Role of Physical Activity and Fitness in the Characterization and Prognosis of the Metabolically Healthy Obesity Phenotype: a Systematic Review and Meta-Analysis

Francisco B. Ortega[1,2*]. Cristina Cadenas-Sanchez[1*], Jairo H. Migueles[1], Idoia Labayen[3], Jonatan R. Ruiz[1,2], Xuemei Sui[4], Steven N Blair[4], Vicente Martínez-Vizcaino[5] and Carl J. Lavie[6]

* Co-First authors.

1. PROFITH "PROmoting FITness and Health through physical activity" research group, Department of Physical Education and Sports, Faculty of Sport Sciences, University of Granada, Spain.

2. Department of Biosciences and Nutrition at NOVUM, Karolinska Institutet, Huddinge,
Sweden

3. Department of Health Sciences, Public University of Navarra, Pamplona, Spain

4. Department of Exercise Science, University of South Carolina, Columbia, SC, U.S.A.

5. Universidad de Castilla-La Mancha: Health and Social Research Center, Cuenca, Castilla-La Mancha, Cuenca, Spain

6. Ochsner Heart and Vascular Institute, The University of Queensland School of Medicine, New Orleans, LA, U.S.A.

Correspondence to:

Francisco B Ortega, Department of Physical Education and Sports, Faculty of Sport Sciences, University of Granada, Carretera de Alfacar s/n, Granada 18071, Spain.

Tel. +34 958 244374. Fax. +34 958244369

E-mail: ortegaf@ugr.es

Funding: Current research activities of Dr. Ortega and Dr. Ruiz are supported by the Spanish Ministry of Economy and Competitiveness –MINECO/FEDER (DEP2016-79512-R); the University of Granada, Plan Propio de Investigación 2016, Excellence actions: Units of Excellence, Unit of Excellence on Exercise and Health (UCEES); the European Union's Horizon 2020 research and innovation programme under grant agreement. No 667302; the SAMID III network, RETICS, funded by the PN I+D+I 2017-2021 (Spain), ISCIII- Sub-Directorate General for Research Assessment and Promotion, the European Regional Development Fund (ERDF) (Ref. RD16/0022) and the EXERNET Research Network on Exercise and Health in Special Populations (DEP2005-00046/ACTI). Dr. Cadenas-Sanchez is supported by a grant from the Spanish Ministry of Economy and Competitiveness (BES-2014-068829) and Jairo H. Migueles by a grant from the Spanish Ministry of Education, Culture and Sport (FPU15/02645). Support for the ACLS study was provided by the U.S. National Institutes of Health (AG06945, HL62508, DK088195, and HL133069).

Conflict of interest

The authors declare no conflict of interest related to this work. Prof. Blair has served as consultants for weight loss and fitness companies and for The Coca-Cola Company, which has also provided them unrestricted research grants.

Abstract

The aims of the present article are to systematically review and meta-analyze the existing evidence on: 1) differences in physical activity (PA), sedentary behavior (SB), cardiorespiratory fitness (CRF) and muscular strength (MST) between metabolically healthy obesity (MHO) and metabolically unhealthy obesity (MUO); and 2) the prognosis of all-cause mortality and cardiovascular disease (CVD) mortality/morbidity in MHO individuals, compared with the best scenario possible, i.e., metabolically healthy normal-weight (MHNW), after adjusting for PA, SB, CRF or MST. Our systematic review identified 67 cross-sectional studies to address aim 1, and 11 longitudinal studies to address aim 2. The major findings and conclusions from the current meta-analysis are: 1) MHO individuals are more active, spend less time in SB, and have a higher level of CRF (yet no differences in MST) than MUO individuals, suggesting that their healthier metabolic profile could be at least partially due to these healthier lifestyle factors and attributes. 2) The meta-analysis of cohort studies which accounted for PA (N=10 unique cohorts, 100% scored as high-quality) support the notion that MHO individuals have a 24-33% higher risk of all-cause mortality and CVD mortality/morbidity compared to MHNW individuals. This risk was borderline significant/non-significant, independent of the length of the follow-up and lower than that reported in previous meta-analyses in this topic including all type of studies, which could be indicating a modest reduction in the risk estimates as a consequence of accounting for PA. 3) Only one study has examined the role of CRF in the prognosis of MHO individuals. This study suggests that the differences in the risk of all-cause mortality and CVD mortality/morbidity between MHO and MHNW are largely explained by differences in CRF between these two phenotypes.

Keywords

Obesity; Metabolically Healthy Obesity; Metabolically Unhealthy Obesity; Metabolically Health Normal-Weight; Mortality; Cardiovascular Disease; Cardiorespiratory Fitness, Muscular Strength, Physical Activity, Exercise; Sedentary Behaviors.

List of abbreviations

ACLS – Aerobics Center Longitudinal Study

BMI – Body mass index

CI – Confidence intervals

CRF – Cardiorespiratory fitness

CVD – Cardiovascular disease

HR – Hazard ratio

HTN-Hypertension

MeSH – Medical Subject Heading

MetS– Metabolic Syndrome

MHNW – Metabolically healthy normal weight

MHO – Metabolically healthy obesity

MST – Muscular strength

MUO – Metabolically unhealthy obesity

MVPA – Moderate-to-vigorous physical activity

NHS – The Nurses' Health Study

NOS – Newcastle Ottawa Scale

OR – Odds ratio

PA – Physical activity

PRISMA – Preferred Reporting Items for Systematic Reviews and Meta-Analysis

PROSPERO – International Prospective Register of Systematic Reviews

RR – Relative risk

SB – Sedentary behaviors

SMD – Standardized mean difference

T2DM – Type 2 Diabetes Mellitus

THIN – The Health Improvement Network

WC – Waist Circumference

Obesity has increased markedly over the last several decades in most developed and developing countries (1-3), and it is clear today that obesity is a pathological condition associated with higher risk of suffering a myriad of physical, psychological and social problems. Among these problems are the metabolic alterations, which are clearly more frequent in obese than in normal-weight individuals (4). However, nearly 2 decades ago the existence of a subset of individuals that, despite being obese, had otherwise a healthy metabolic profile (5,6), and this was later named as the metabolically healthy obesity (MHO) phenotype. The concept of MHO refers to those individuals who are obese but do not have dyslipidemia, hyperglycemia/type 2 diabetes mellitus (T2DM), or hypertension (HTN). A recent meta-analysis has reported that one in every three (more precisely 35%) individuals with obesity is MHO (7), yet this percentage might change depending on the definition of MHO used, as previously discussed elsewhere (8). Although several specific definitions of MHO have been used in the literature, there seems to be more and more consensus that MHO should be defined as being obese according to the standard definition of a body mass index (BMI) equal or higher than 30kg/m^2 and having 0 of the metabolic syndrome (MetS) criteria (waist circumference -WC-, excluded). More details about the proposed harmonized definition of MHO and the scientific base for it is provided in the original publication (see Table 2 and 3 in Ortega et al. (8)). The opposite condition among obese individuals is most commonly named as metabolically unhealthy obesity (MUO) and defined as an obese individual who meets at least 1 of the 4 MetS criteria (WC excluded). Both, MHO and MUO are often compared regarding prognosis and metabolic characteristics to the normal weight/metabolic group, i.e., the metabolically healthy normal weight (MHNW) phenotype.

From the landmark studies of Drs. Morris, Paffenbarger and Blair (9-12) to date, vast, consistent and accumulating evidence supports the health benefits of high levels of physical activity (PA), low levels of sedentarism (i.e., sedentary behaviors, SB) and high levels of cardiorespiratory fitness (CRF) and more recently also muscular strength (MST) (13-24). However, to the best of our knowledge, the specific role of PA, SB, CRF and MST in the characterization and prognosis of the MHO has not yet been systematically reviewed and meta-analyzed.

Over the last several years, the amount of studies focused on the MHO intriguing phenotype has been overwhelming. **Figure 1** shows the number of articles published in PubMed about the MHO concept since it was first described in 2001. Some of these studies have addressed two key questions: 1) What are the characteristics of the MHO individuals? and 2) How is the prognosis of the MHO individuals compared with MUO and MHNW individuals? Concerning the first question, the early reviews on this topic did not consider a higher level of PA and/or CRF as a characteristic of the MHO when compared with MUO (25), except more recent ones (26), given recent evidence supporting this notion (27-33). However, to the best of our knowledge, the differences in PA and/or CRF between MHO and MUO have not been systematically reviewed and quantified using meta-analysis methods. Similarly, whether there are systematic differences between MHO and MUO in time spent in SB, and in other components of physical fitness, such as MST, is also currently unknown.

Concerning the prognosis of MHO, a number of systematic reviews and meta-analyses have examined the risk of fatal and non-fatal cardiovascular disease (CVD), as well as all-cause mortality in MHO compared to MUO and MHNW (34-39). In addition to these systematic reviews and meta-analyses, some powerful studies have recently been published on this topic, such as The Health Improvement Network (THIN) cohort in 3.5

million participants (40), and the Nurses' Health Study (NHS) in 90,257 women (32,41). Although controversy about this topic has existed in the near past, we believe that the current evidence coming from the systematic reviews/meta-analyses and the latest and more powerful original research studies mentioned above, clearly supports that there is no benign obesity. Thus, MHO individuals, even they have a markedly lower risk of future disease and death than MUO individuals, they still have a higher risk of mortality and morbidity than MHNW individuals (32). However, it is of utmost importance to note that most of the existing cohorts lack information about probably one of the most powerful predictors, and therefore potential confounder, of current and future metabolic health, all-cause mortality and CVD risk; that is CRF (24,28,30-33). Moreover, few studies have accounted for PA in their analyses, which is closely related to CRF, although always less objectively and accurately measured (42). Therefore, there is a gap in the current knowledge about the prognosis of MHO individuals, whether or not the higher risk reported in MHO, compared to MHNW, in most of existing studies, is explained by differences in CRF or PA. To the best of our knowledge, the role of CRF and/or PA in the prognosis of the MHO has not yet been systematically reviewed and meta-analyzed.

Thus, the aims of the present article are to systematically review and meta-analyze the existing evidence on: 1) differences in PA, SB, CRF and MST between MHO and MUO; and 2) the prognosis of future all-cause and CVD mortality and morbidity (i.e., non-fatal CVD events) in MHO individuals, compared with the best scenario possible, i.e., MHNW, once PA, SB, CRF and/or MST have been taken into account.

METHODS

Protocol and Registration

The present systematic review and meta-analysis was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement (43). The study protocol was registered in the International Prospective Register of Systematic Reviews (PROSPERO reference number: CRD42018093955).

Data Sources and Search Strategies

The search was conducted for studies published from inception to March 21, 2018 in two major electronic databases: PubMed and Web of Science. The keywords used in search strategy were related to the following topics: PA (motor activity, exercise, etc.), physical fitness (CRF, functional capacity, etc.), metabolic phenotypes (metabolically benign, metabolically healthy, etc.) and weight status (overweight and obesity). The connectors 'OR' and 'AND' were used to combine the search terms. Specifically, for PubMed search, we used Medical Subject Heading (MeSH) terms combined with the tag for searching in title, abstract and keywords. As an example, for CRF we introduced: ("cardiorespiratory fitness" [Title/Abstract] OR cardiorespiratory fitness [MeSH]). Search terms and search strategies were adapted to each database (for more information, see **Supplemental Material, Table S1**). Additionally, the reference lists of retrieved studies were examined for identifying potential interesting articles.

Once the search strategies for both databases were executed, we imported all the references found into Covidence software (44). All the process of duplicates, screening, data extraction and risk of bias analysis were performed by this web-based systematic review tool.

Eligibility Criteria

Studies were considered eligible for the inclusion if they met the following criteria: 1) provided results that allowed comparing the MHO and MUO in regards to PA, SB, CRF and MST (aim 1) or provided risk estimates on longitudinal prognosis of MHNW and MHO for all-cause mortality, non-fatal CVD or CVD mortality after adjustments for PA, SB, CRF or MST (aim 2); 2) healthy participants without any illness that could influence their metabolic profile; and 3) original studies written in English or Spanish (excluding letters, meeting abstracts, systematic reviews and/or meta-analyses, etc.). To be included in the meta-analysis, studies additionally had to provide, for the aim 1, mean and standards deviations (or 95% confidence intervals, CI) and sample sizes for MHO and MUO group; and for the aim 2, hazard ratio (HR), odds ratio (OR) or relative risk (RR) and their 95% CI for all-cause mortality, non-fatal CVD or CVD mortality. When a cohort was represented in more than one study, we included only the study with the largest sample size. In **Supplemental Material**, **Table S2** shows the studies that were excluded after revision of these inclusion criteria and the reasons for exclusion.

Study Selection Process

Two reviewers (CC-S and JHM) independently performed the study selection process. Firstly, the reviewers examined title and abstract of each article in order to identify those studies that could be included for the next step of the selection process (inter-reviewer agreement = 99%). Then, studies that appeared eligible based on the first screening were read full-text against the inclusion criteria for their final inclusion or exclusion in the systematic review (inter-reviewer agreement = 89%). Disagreements on the study selection were solved by reaching consensus between reviewers. When the inclusion of a study was ambiguous, a third reviewer (FBO) was included for resolving by discussion and a consensus decision was made.

Of the articles included in the systematic review, we meta-analyzed those studies that provided cross-sectional information (aim 1) on moderate-to-vigorous PA (MVPA), SB, CRF and/or MST, since these components were the most reported and most health-related. Although the meta-analysis was conducted using MVPA data, we refer to PA in the text as general concept and for simplicity. Moreover, for MST, we analyzed the relative MST because previous studies have shown to be a stronger predictor of CVD risk factors than indicators of absolute strength (45,46), yet in the text we refer to just MST as general concept and for simplicity. Concerning the longitudinal studies (aim 2), we meta-analyzed those which adjusted for PA or CRF and presented all-cause mortality, non-fatal CVD or CVD mortality as outcomes. A minimum of three studies was requested to perform the meta-analysis.

Data Collection Process

Data extraction from eligible studies was performed by two reviewers (CC-S and JHM) using a consensual template with a third reviewer (FBO). The extracted data vary depending on the aim of the study: for the aim 1, we extracted 1) first author's name and year of publication of the study, 2) Setting, study design and sample size, 3) age, 4) metabolic criteria used, 5) definition for classifying MHO or MUO, 6) outcomes of interest (measurement and unit), 7) statistical analysis used and adjustments, and 8) main findings of the studies included; for the aim 2, we additionally included years of follow-up and number of all-cause mortality, non-fatal CVD and/or CVD mortality cases in the template.

Quality of the Studies and Risk of Bias

Study quality was assessed by two reviewers (CC-S and JHM) independently and using two different tools depending on the aim. Cross-sectional studies included in the aim 1

were evaluated following the Critical Appraisal Checklist for Analytical Cross-sectional studies from the Joanna Briggs Institute (47). The checklist consists of 8 items regarding inclusion criteria, study sample and setting, exposure measured, standard criteria for measurement, confounding factors, strategies to deal with confounders, outcomes measured and statistical analysis (**Supplementary material 1**). Each item was assessed as "yes", "no", "unclear" or "not applicable". For standardization, we considered "yes" as low risk of bias, and "no" and "unclear" as high risk of bias. Overall, low risk of bias (i.e., high quality study) was considered when a study accumulated at least 5 items answered as "yes". Studies assessed as 'yes' in less than 5 items were categorized as high risk of bias (i.e., low quality study).

Longitudinal studies were assessed by Newcastle Ottawa Scale (NOS) risk of bias tool (48). The NOS tool assesses the study quality of cohort studies in 8 items grouped in three categories: selection, comparability and outcome (**Supplementary material 2**). A study can be awarded a maximum of one star for each numbered item within the selection (i.e., items 1-4) and outcome categories (i.e., items 6-8), and a maximum of two stars in the comparability domain (i.e., item 5). Thus, each study can be awarded with a maximum of 9 stars. Overall, low risk of bias (i.e., high quality study) was defined when a study got at least 6 stars. High risk of bias (i.e., low quality study) was then defined as studies awarded with less than 6 stars.

Meta-Analysis

The meta-analysis was performed in: 1) MHO vs. MUO groups for cross-sectional studies (aim 1), since the difference in fitness between obese individuals (both MHO and MUO) and normal-weight individuals is well-known (27); and 2) MHNW vs. MHO for longitudinal studies in the prognosis of all-cause mortality, non-fatal CVD and CVD mortality adjusting for CRF or PA (aim 2), since the differences in risk of disease between

MHO and MUO are well-known and was therefore not needed to be tested in this meta-analysis. In order to answer the aim 1, we first calculated the mean difference between metabolic phenotypes' groups (MHO *minus* MUO) and, then, its standardized mean difference (SMD, d-Cohen) and 95% confidence intervals (CI) on PA, SB, CRF and MST. The main analyses are presented with overweight and/or obese participants. Sensitivity analysis was performed with only the obese sample, and results are presented as supplementary materials. The pooled SMD in the outcomes analyzed was obtained using fixed or random effects models depending on the heterogeneity level detected (I^2; the larger the value, the greater the heterogeneity). Heterogeneity was evaluated by the percentage of total variability attributed to between-study heterogeneity (I^2 statistics). Low, moderate and high degrees of heterogeneity were identified as I^2 values of 25, 50 and 75%, respectively (49).

For the aim 2, we extracted the HR and its 95% CI of all the studies included. All the studies had as reference group the MHNW, except the one from Ortega et al.(27) who presented the data having the MHO group as reference. For comparison purposes, the results extracted from Ortega et al.(27) were inverted. Likewise, most of the studies included provided the HR estimate except Appleton et al. (50) which provided OR. In this case, as the OR was close to 1 and the prevalence was lower than 5%, the OR and HR values are assumed to be similar and, thus, we included it in the analysis as HR as previously done in another meta-analysis (35). Therefore, pooled HR estimates were obtained for all-cause mortality, non-fatal CVD and CVD mortality outcomes using fixed or random models (depending on the I^2 value).

Funnel plots were also examined for assessing risk of potential publication bias. We also calculated the P value of the Egger's intercept. The leave-one-out analysis helped to examine the influence of an article excluded on the combined SMD. If the evidence is

consistent on a certain finding, the leave-one-out analysis should not change the conclusions.

All the statistical analyses were performed with the Comprehensive Meta-Analysis software version 2 (Englewood, NJ: Biostat, USA) and the level of significance was set at $p<0.05$.

RESULTS

Literature Search

In total, 70 unique studies were included in the systematic review; 67 of them were included for the systematic review for aim 1. Among them, 55 studies (51-105) were focused on the differences between MHO and MUO in PA (N=53) and/or SB (N=9), while only 19 (5,27,55,72,81,86,89,94,97,106-115) compared CRF (N=19) and MST (N=6) between phenotypes. Seven studies presented both PA and SB data, while another 7 studies examined both PA/SB and fitness data. In regards to the aim 2, 11 studies were included in the systematic review, 10 studies (50,60,70,77,82,91-93,116,117) adjusted for PA in the association of MHO vs. MHNW with risk of all-cause mortality, CVD mortality or non-fatal CVD, while only one study adjusted for fitness (27). **Figure 2** shows the flowchart of the study selection process.

For meta-analysis purposes, a total of 25 unique studies that provided comparable data (or data that could be transformed into comparable data) were included for the aim 1, where most of the information was provided for CRF (N=19) (5,27,55,72,81,86,89,94,97,106-115). PA (N=6) (55-58,72,87), SB (N=5) (55,58,75,87,102) and relative MST (N=6) (55,81,106,107,110,115), were also analyzed (**Figure 2**). For the aim 2, a total of 10 unique studies adjusting for PA were included in the meta-analysis (50,60,70,77,82,91-93,116,117). We did not perform a meta-analysis with studies adjusting for CRF, since we found only one study doing this adjustment (27). The meta-analysis on studies adjusting for PA examined the following outcomes: all-cause mortality (N=4) (60,92,93,117), non-fatal CVD and CVD mortality (N=7) (50,70,77,82,91,92,116) compared with MHO individuals after adjusting for PA (**Figure 2**).

Characteristics of the Study Sample

The characteristics of the 67 unique studies included in the systematic review for the aim 1 are shown in **supplementary material Table S3**. Studies reporting estimates for PA and/or SB ranged in sample size from 8 to 342,442 participants. The mean age for these studies ranged from 11.8 to 72 years-old. PA and SB were mainly self-reported (N=47 and N=6, respectively). Few studies examined PA and SB by objective measurements such as accelerometers or pedometers (N=9 and N=4, respectively). For fitness, sample size ranged from 8 to 3,911 (mean age ranged from 14.1 to 61.1 years-old). CRF was the most reported fitness component (N=19), followed by muscular strength (N=6). Less articles were found studying other components of fitness, i.e., speed-agility (N=2), flexibility (N=2) and/or balance (N=1).

Table 1 shows the characteristics of the study samples of 11 unique studies included in the systematic review for aim 2. Sample sizes of the studies included ranged from 72 to 65,175 participants. The duration of the follow-up ranged from 8 to 14 years. Of the studies that examined all-cause mortality, the number of cases ranged from 9 to 449. The number of cases from non-fatal CVD and CVD mortality ranged from 2 to 261.

Quality of the Studies and Risk of Bias

The analysis of the quality of the studies, i.e., risk of bias, included in the systematic review and meta-analysis can be found in **Figure 3** for cross-sectional and longitudinal studies. The specific information with the scoring of each study in each item and total is shown in **supplementary material Table S4** for cross-sectional studies and **Table S5** for longitudinal studies. Overall, 80% of the cross-sectional studies included for the aim 1 were scored as high-quality studies, indicating a low risk of bias in most of the studies

included. Likewise, all of the longitudinal studies (i.e., 100%) included for the aim 2 were scored as high-quality studies, indicating a low risk of bias in the studies included.

Differences in PA, SB, CRF and MST Between MHO and MUO

Figure 4 depicts the meta-analyzed differences between MHO and MUO for PA (SMD= 0.267, 95% CI: 0.090, 0.444, P=0.003, I^2= 48.3%, total N=5539) and SB (SMD= -0.199, 95% CI: -0.317, -0.081, P=0.001, I^2= 44.6%, total N=5290). There was no publication bias for any of the outcomes studied (PA, Egger's test, p=0.534; SB, Egger's test, p=0.944; **supplementary material Figure S1**). Sensitivity analyses also showed significant differences between MHO and MUO when only obese sample were analyzed (PA, SMD= 0.635, 95% CI: 0.178, 1.091, P=0.006, I^2= 77.7%; and SB, SMD= -0.175, 95% CI: -0.281, -0.070, P=0.001, I^2= 30.7%; **supplementary material Figure S2**). Egger's tests indicated no significant publication bias (all P≥0.473, **supplementary material Figure S3**). The leave-one-out analysis did not alter the results (data not shown).

The meta-analysis of the differences between MHO and MUO showed a significant difference in CRF in favor of MHO (SMD=0.317, 95% CI: 0.232, 0.402, P<0.001, I^2= 49.1%, total N=11758) (**Figure 5**). In regards to MST, no significant difference was observed between MHO and MUO (SMD= -0.049, 95% CI: -0.241, 0.143, P=0.618, I^2= 0%, total N=851). There was no significant publication bias according to Egger's test both for CRF (p=0.224, **supplementary material Figure S4a**) and MST (p=0.393, **supplementary material Figure S4b**). In sensitivity analyses, we observed that these findings persisted when the analyses were restricted to obese participants (i.e., excluding overweight participants from the analyses) (CRF, SMD= 0.276, 95% CI: 0.206, 0.346, P<0.001, I^2= 29.2%; and MST SMD= -0.126, 95% CI: 0.389, 0.138,

P=0.351, I^2= 0%; **supplementary material Figure S5**). No publication bias was observed (Egger's test ≥0.729; **supplementary material Figure S6**).

Prognosis of MHO After Considering PA

After adjustment for PA, no significant differences were observed between MHO and MHNW in the risk of all-cause mortality (HR= 1.32, 95% CI: 0.833, 2.108, P=0.235, I^2= 73.0%, total N=93.561; **Figure 6a**). No significant publication bias was observed (Egger's test, P=0.601, **supplementary material Figure S7a**). However, the leave-one-out analysis by omitting one study (i.e., Sung et al. (92)) turned the HR from non-significant to significant (HR= 1.58, 95% CI: 1.205, 2.076, P=0.001, data not shown in figures). The meta-analysis focused on CVD showed that MHO individuals presented a 24% higher risk of non-fatal CVD and CVD mortality than MHNW individuals (HR= 1.24, 95% CI: 1.071, 1.444, P=0.004, I^2= 0%; **Figure 6b**). However, in the leave-one-out analysis (i.e., omitting Lassale et al. (77)), we observed that the HR became non-significant (HR= 1.21, 95% CI: 0.983, 1.492, P=0.073, data not shown in figures). No publication bias was observed (**supplementary material Figure S7b**). A sensitivity analysis was performed considering only those studies that examined non-fatal CVD (data not shown). The result was similar, showing that MHO presented 26% higher risk of non-fatal CVD compared to MHNW (HR= 1.26, 95% CI: 1.084, 1.474, P=0.003, I^2= 0%). The Eggers' test did not show significant publication bias either.

Exploratory analyses stratifying the analysis by the studies with a follow-up <10 years and ≥10 years (**supplementary material Figure S8**), showed that there were not differences in the pooled HR for CVD mortality and non-fatal CVD, i.e., HR= 1.24 for <10 years (N=5) and HR= 1.24 for ≥10 years (N=2).

Prognosis of MHO After Considering CRF

After excluding one study using a smaller sample (118) from a cohort already included (27), only one study examined the prognosis of MHO after adjustment for CRF (27), not being therefore applicable for use in the meta-analysis. Nevertheless, for the purpose of this article, we re-analyzed (changed the reference group from MHO in the original paper to MHNW in this review) the original data used in this study and crafted **Figure 7**, which shows the role of CRF in the prognosis of MHO. We observed a markedly higher risk in the MHO group in all-cause mortality, non-fatal CVD and CVD mortality compared with the MHNW group, independent of a set of potential confounders. However, this risk was strongly attenuated and became non-significant after additional adjustment for CRF for the 3 outcomes studied. These findings were consistent when obesity was defined based on BMI or body fat percentage.

DISCUSSION

Main Findings

In the present systematic review and meta-analysis, we focused on two relevant topics related to the MHO phenotype, its characterization (aim 1) and prognosis (aim 2). The studies included in this review were mostly (80-100%) scored as high-quality studies indicating a low risk of bias in the findings obtained. We found 67 studies examining differences between MHO and MUO in PA, SB, CRF and MST. Our meta-analysis showed that MHO individuals have significantly higher levels of PA and CRF and lower levels of SB than MHO individuals, without differences in MST between these phenotypes. In addition, we found 10 unique longitudinal studies examining the prognosis of MHO compared to MHNW after accounting for PA, only one study (after excluding another one from the same cohort that used a smaller sample) accounting for CRF, and none accounting for SB or MST. The meta-analysis of 10 studies accounting for PA showed marginally non-significant higher (33%, CI=0.83-2.11) risk in MHO in the risk of all-cause mortality, and a marginally significant higher (24%, CI=1.07-1.44) risk of CVD mortality and non-fatal CVD compared to MHNW. However, the leave-one-out analysis showed that the effect sizes observed were slightly decreased or increased, changing the pooled effect from non-significant to significant for all-cause mortality and *vice versa* for CVD outcomes. In addition, we did not observe a different effect size in studies with shorter (<10 years) or longer (≥10 years) of follow-up, yet the number of studies was limited. Altogether, the findings from the present meta-analysis in studies adjusting for PA support the notion that the differences between MHO and MHNW in the risk of future all-cause mortality and CVD mortality/morbidity are borderline significant/non-significant, independently of the length of the follow-up.

The present systematic review identified only one unique study which adjusted for the potential confounding effect of CRF (27). This study observed that after adjustment for a set of potential confounders except for CRF, MHO had a significantly increased risk of all-cause mortality, non-fatal CVD and CVD mortality. However, this risk became markedly attenuated and non-significant, after additional adjustment for CRF.

Differences in PA, SB, CRF and MST Between MHO and MUO

Previous narrative reviews have pointed out that a higher level of PA and CRF seemed to be a characteristic of MHO when compared to MUO individuals (8,26,28,29,119,120). The present systematic review and meta-analysis has quantified, for the first time, the existing evidence in this regard and statistically tested whether these differences between MHO and MUO are significant. In addition, since longer time spent in SB and low levels of MST have also shown to predict a higher risk of metabolic disorders and CVD risk (15-17,20-24), we included them as well in our systematic search. The results from the present meta-analysis support that MHO have higher levels of PA, lower levels of SB and higher levels of CRF, without differences in MST, suggesting that these factors could be contributing to the better metabolic profile of the MHO phenotype. The largest differences were observed in CRF, followed by PA and then by SB.

Prognosis of MHO After Considering PA

Even if the present meta-analysis restricted the systematic search only to studies adjusting for PA, we were able to retrieve 5 new cohort studies adjusting for PA (59,60,77,117,121) not included in the previous meta-analyses (35,37-39). The pooled risk observed was higher for all-cause mortality (33%) and for CVD mortality/morbidity (24%) in MHO compared to MHNW. This risk was borderline

significant/non-significant and seems to be lower than the 45 to 60% higher risk of CVD morbidity reported in the meta-analysis by Eckel et al. (37) and Zheng et al. (35) respectively, which included all studies focused on MHO, not only those adjusting for PA as ours. The lower risk observed in our analysis of studies adjusting for PA seems to support that the well-known differences in PA existing between obese and normal-weight individuals (and therefore also between MHO and MHNW) could explain, at least partially, the difference between the risk observed in our meta-analysis and that observed in most of the literature (not adjusting for PA) for CVD mortality/morbidity. However, this should be interpreted cautiously, and to definitively test this hypothesis, more studies are needed showing the effect sizes before and after additional adjustment for PA, and then meta-analyze them.

Kramer et al. (39), and Fan et al.(38), observed a distinct risk between studies with shorter or longer follow-ups, i.e., HR=1.19 (all studies) vs. 1.24 (longer follow-up) and HR=1.05 (shorter follow-up) vs. 1.60 (longer follow-up), respectively; however, we did not observe that difference in our study, i.e., HR= 1.24 vs. 1.24 for studies with less than 10 years vs. equal or more than 10 years of follow-up, respectively (using the 10 years cut-point, as in the study by Kramer et al. (39)).

Different definitions to classify metabolic phenotypes have been used, being the most commonly used in relation to the number of MetS criteria met, i.e., ≤2 criteria (few studies), ≤1 criteria (most commonly used in a near past) and 0 criteria (new and most currently accepted definition) (8). It has been argued that a person should not be classified as "metabolically healthy" if that person has T2DM or HTN, as an example. Therefore, based on this sound reasoning, meeting ≤1 or ≤2 criteria of MetS are not recommended any longer to define MHO. However, the problem is that most of existing evidence on the prognosis of MHO is based in studies that used these definitions. In our

present meta-analysis 3 (92,93,117) out of cohorts available used this newer and more accepted definition of meeting 0 criteria for all-cause mortality, and 1 out of 7 available for CVD mortality/morbidity. The pooled effect for all-cause mortality outcome would remain non-significant if conducted only in these 3 cohort studies, and the only 1 study for CVD outcomes does not allow a meaningful meta-analysis. Nevertheless, it is expected that the prognosis of MHO will be better with this new/current definition (37), since individuals classified as MHO will be "healthier" than with the older definitions, and so seems to support our systematic review, since the study by Sung et al. (92) showed by far the lowest risk of CVD mortality (HR=0.4) in MHO compared to MHNW, yet the sample size of this study was small, and we found no other studies to test this hypothesis in CVD mortality/morbidity.

Prognosis of MHO After Considering CRF

Given the strong and consistent association of CRF with metabolic risk, all-cause mortality and CVD mortality/morbidity (14,18,19,24), and given also the large difference existing in CRF between obese and normal-weight individuals, which as expected translate to differences in CRF between MHO and MHNW (27,28), there is a good rationale that adjusting for CRF could potentially change the conclusions about the prognosis of the MHO. In this context, the only study accounting for the potential influence CRF is our study using the Aerobics Center Longitudinal Study (ACLS) data (27). This study supports the notion that the differences in risk of all-cause mortality and CVD mortality/morbidity observed in the literature could be explained by the differences in CRF between MHO and MHNW. However, future cohort studies accounting for CRF in their models will confirm or contrast these findings. Since BMI is often criticized as a marker of adiposity, we tested the same hypothesis using accurate methods (hydrostatic weighing and skinfold thicknesses) to define obesity based on

body fat percentage and the results were consistent (27), making this conclusion stronger. Compared to PA, we expected CRF to show a stronger reduction of risk in MHO vs. MHNW (122), and so seems to support the findings of this systematic review, yet given the limited information including CRF found, this hypothesis on MHO prognosis needs to be confirmed in future cohort studies.

Limitations and Strengths

This study presents several limitations. First, although two major databases (i.e., PubMed and Web of Science) were used for the search, the no inclusion of other electronic databases such as EMBASE should be acknowledged as a limitation. Second, the use of different definitions of MHO and different ways of measuring PA or SB, could explain the high degree of heterogeneity observed among the studies analyzed. Thus, the use of the harmonized definition of the MHO and the use of objective methods for PA and SB, will allow more solid conclusions on this topic. Third, the low number of studies reporting the results without and with additional adjustment for PA or CRF did not allow us to meta-analyze how this can influence the prognosis of all-cause mortality and CVD mortality/morbidity in the MHO compared to MHNW. Fourth, to increase the number of studies and power, we included studies that included overweight plus obesity in definition of MHO, but our sensitivity analyses (reported in supplementary material) showed consistent findings for all meta-analysis conducted, suggesting that the conclusions of the present meta-analysis are valid for MHO when including or excluding overweight.

Despite of these limitations, to the best of our knowledge, this is the first meta-analysis focused on the differences between MHO and MUO in PA, SB, CRF and MST outcomes, as well as, examining the role of PA and CRF in the prognosis of all-cause mortality and non-fatal CVD. The thorough and complete methods used in this meta-

analysis should be acknowledged, e.g., double examination of the selection process by two independent reviewers, the assessment of the studies' quality and risk of bias (Joanna Briggs Institute and NOS checklists), funnel plots and Egger's tests for testing a potential publication bias, the leave-one-out analysis, etc.

Practical Implications

Based on the findings presented in this meta-analysis, future description and characterization of the MHO phenotype, when compared with the rest of obese patients (i.e., MUO), should include a higher CRF as a significant factor, but also a longer time spent in PA, particularly MVPA (that is what was analyzed in the present meta-analysis), and shorter time spent in SB. The fact that MHO and MHNW individuals differed in these lifestyle factors and attribute, does not imply causality. Future intervention studies aiming to increase the time spent in PA, reduce the time spent in SB and enhance CRF will test the effectiveness of such intervention in turning MUO into MHO. Of course, in an ideal world, the best scenario possible would be a transition from MUO to MHNW, which would have a huge benefit on multiple health indicators. However, reality shows that most of existing lifestyle interventions have failed to maintain the weight losses in the long term, with bariatric surgery being the most effective long-term treatment for patients with more severe obesity (123). In this context, Stefan and colleagues (123) have suggested that for people with obesity, to become normal-weight, should be a longer-term goal, but they should be encouraged to follow a healthy lifestyle in order to improve their metabolic profile, i.e., become (or remain) MHO as perhaps a more achievable and shorter term goal (what they named as the "low-hanging fruit"). Building on this idea, our meta-analysis suggests that PA and CRF seems to play a protective role in the prognosis of MHO, and they should be targeted in lifestyle interventions, together with other lifestyle factors, such as diet, to

treat MUO and hopefully transition them into people with a healthier metabolic profile, leaner bodies and, ultimately, lower risk of CVD mortality/morbidity and all-cause mortality.

CONCLUSIONS

Our systematic review and meta-analysis provides novel insights in the characterization and prognosis of the MHO phenotype. First, the meta-analysis on cross-sectional studies supports that MHO individuals are more active, spent less time in SB, and have a higher level of CRF (yet no differences in MST) than MUO individuals, suggesting that their healthier metabolic profile could be, at least partially, due to these healthier lifestyle factors and attributes. Second, the findings from the present meta-analysis of cohort studies, which accounted for PA, support the notion that MHO individuals have a 24-33% higher risk of all-cause mortality and CVD mortality/morbidity compared to MHNW individuals. This risk seems to be borderline significant/non-significant and independent of the length of the follow-up in our meta-analysis. This risk was lower than that reported in previous meta-analyses which included all studies focused on MHO participants (i.e., 45-60% higher risk in MHO) (35,37), that could be indicating a modest reduction in the risk estimates once PA is accounted for. Third, only one study (27) has examined the role of CRF in the prognosis of MHO individuals. This study suggests that the differences in the risk of all-cause mortality and CVD mortality/morbidity between MHO and MHNW are largely explained by differences in CRF between these two groups, supporting the idea that gathering information about the metabolic and CRF status in clinical settings could improve risk stratification in obese patients. Collectively, from a clinical and public health point of view, our meta-analysis indirectly support that obesity treatment efforts should be targeted not only on losing

weight, but also on improving CRF through lifestyle programs focused on increasing PA levels along with other healthy lifestyle factors.

REFERENCES

1. NCD Risk Factor Collaboration. Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants. The Lancet 2016;387:1377-1396.
2. Ng M, Fleming T, Robinson M et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2014;384:766-81.
3. NCD Risk Factor Collaboration. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128.9 million children, adolescents, and adults. Lancet 2017;390:2627-2642.
4. Bhupathiraju SN, Hu FB. Epidemiology of Obesity and Diabetes and Their Cardiovascular Complications. Circ Res 2016;118:1723-35.
5. Brochu M, Tchernof A, Dionne IJ et al. What are the physical characteristics associated with a normal metabolic profile despite a high level of obesity in postmenopausal women? J Clin Endocrinol Metab 2001;86:1020-5.
6. Sims EA. Are there persons who are obese, but metabolically healthy? Metabolism 2001;50:1499-504.
7. Lin H, Zhang L, Zheng R, Zheng Y. The prevalence, metabolic risk and effects of lifestyle intervention for metabolically healthy obesity: a systematic review and meta-analysis: A PRISMA-compliant article. Medicine (Baltimore) 2017;96:e8838.
8. Ortega FB, Lavie CJ, Blair SN. Obesity and Cardiovascular Disease. Circ Res 2016;118:1752-70.
9. Morris JN, Heady JA, Raffle PA, Roberts CG, Parks JW. Coronary heart-disease and physical activity of work. Lancet 1953;265:1053-7; contd.
10. Paffenbarger RS, Jr., Hyde RT, Wing AL, Hsieh CC. Physical activity, all-cause mortality, and longevity of college alumni. N Engl J Med 1986;314:605-13.
11. Paffenbarger RS, Jr., Wing AL, Hyde RT. Physical activity as an index of heart attack risk in college alumni. Am J Epidemiol 1978;108:161-75.
12. Blair SN, Kohl 3rd HW, Paffenbarger Jr. RS, Clark DG, Cooper KH, Gibbons LW. Physical fitness and all-cause mortality. A prospective study of healthy men and women. JAMA 1989;262:2395-2401.
13. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 2012;380:219-29.
14. Kodama S, Saito K, Tanaka S et al. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis. JAMA 2009;301:2024-35.
15. Ruiz JR, Sui X, Lobelo F et al. Association between muscular strength and mortality in men: prospective cohort study. BMJ 2008;337:a439.
16. Artero EG, Lee D-c, Ruiz JR et al. A prospective study of muscular strength and all-cause mortality in men with hypertension. J Am Coll Cardiol 2011;57:1831-7.
17. Ortega FB, Silventoinen K, Tynelius P, Rasmussen F. Muscular strength in male adolescents and premature death: cohort study of one million participants. BMJ 2012;345:e7279.
18. Myers J, McAuley P, Lavie CJ, Despres J-P, Arena R, Kokkinos P. Physical Activity and Cardiorespiratory Fitness as Major Markers of Cardiovascular Risk: Their Independent and Interwoven Importance to Health Status. Prog Cardiovasc Dis 2015;57:306-314.
19. Ross R, Blair SN, Arena R et al. Importance of Assessing Cardiorespiratory Fitness in Clinical Practice: A Case for Fitness as a Clinical Vital Sign: A Scientific Statement From the American Heart Association. Circulation 2016;134:e653-e699.

20. Henriksson H, Henriksson P, Tynelius P, Ortega FB. Muscular strength in adolescence and later disability: a population based cohort study of one million men. Br J Sports Med In Press.
21. Ekelund U, Steene-Johannessen J, Brown WJ et al. Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. Lancet 2016;388:1302-10.
22. Garcia-Hermoso A, Cavero-Redondo I, Ramirez-Velez R et al. Muscular Strength as a Predictor of All-Cause Mortality in an Apparently Healthy Population: A Systematic Review and Meta-Analysis of Data From Approximately 2 Million Men and Women. Arch Phys Med Rehabil 2018.
23. Bakker EA, Lee DC, Sui X et al. Association of Resistance Exercise With the Incidence of Hypercholesterolemia in Men. Mayo Clin Proc 2018;93:419-428.
24. Ortega FB, Cadenas-Sanchez C, Lee D-c, Ruiz JR, Blair SN, Sui X. Fitness and Fatness as Health Markers through the Lifespan: An Overview of Current Knowledge. Progress in Preventive Medicine 2018;3:e0013.
25. Primeau V, Coderre L, Karelis AD et al. Characterizing the profile of obese patients who are metabolically healthy. Int J Obes (Lond) 2011;35:971-81.
26. Bluher S, Schwarz P. Metabolically healthy obesity from childhood to adulthood - Does weight status alone matter? Metabolism 2014;63:1084-92.
27. Ortega FB, Lee DC, Katzmarzyk PT et al. The intriguing metabolically healthy but obese phenotype: cardiovascular prognosis and role of fitness. Eur Heart J 2013;34:389-97.
28. Ortega FB, Cadenas-Sanchez C, Sui X, Blair SN, Lavie CJ. Role of Fitness in the Metabolically Healthy But Obese Phenotype: A Review and Update. Prog Cardiovasc Dis 2015.
29. Stefan N, Haring HU, Hu FB, Schulze MB. Metabolically healthy obesity: epidemiology, mechanisms, and clinical implications. The lancet Diabetes & endocrinology 2013;1:152-62.
30. Deedwania P, Lavie CJ. Dangers and Long-Term Outcomes in Metabolically Healthy Obesity: The Impact of the Missing Fitness Component. J Am Coll Cardiol 2018;71:1866-1868.
31. Kennedy AB, Lavie CJ, Blair SN. Fitness or Fatness: Which Is More Important? JAMA 2018;319:231-232.
32. Lavie CJ, Deedwania P, Ortega FB. Obesity is rarely healthy. The lancet Diabetes & endocrinology 2018.
33. Lavie CJ, Ortega FB, Kokkinos P. Impact of Physical Activity and Fitness in Metabolically Healthy Obesity. J Am Coll Cardiol 2018;71:812-813.
34. Roberson LL, Aneni EC, Maziak W et al. Beyond BMI: The "Metabolically healthy obese" phenotype & its association with clinical/subclinical cardiovascular disease and all-cause mortality -- a systematic review. BMC Public Health 2014;14:14.
35. Zheng R, Zhou D, Zhu Y. The long-term prognosis of cardiovascular disease and all-cause mortality for metabolically healthy obesity: a systematic review and meta-analysis. J Epidemiol Community Health 2016;70:1024-31.
36. Brant LC, Wang N, Ojeda FM et al. Relations of Metabolically Healthy and Unhealthy Obesity to Digital Vascular Function in Three Community-Based Cohorts: A Meta-Analysis. Journal of the American Heart Association 2017;6.
37. Eckel N, Meidtner K, Kalle-Uhlmann T, Stefan N, Schulze MB. Metabolically healthy obesity and cardiovascular events: A systematic review and meta-analysis. European journal of preventive cardiology 2016;23:956-66.
38. Fan J, Song Y, Chen Y, Hui R, Zhang W. Combined effect of obesity and cardio-metabolic abnormality on the risk of cardiovascular disease: a meta-analysis of prospective cohort studies. Int J Cardiol 2013;168:4761-8.
39. Kramer CK, Zinman B, Retnakaran R. Are metabolically healthy overweight and obesity benign conditions?: A systematic review and meta-analysis. Ann Intern Med 2013;159:758-69.
40. Caleyachetty R, Thomas GN, Toulis KA et al. Metabolically Healthy Obese and Incident Cardiovascular Disease Events Among 3.5 Million Men and Women. PLoS One 2017;70:1429-1437.
41. Eckel N, Li Y, Kuxhaus O, Stefan N, Hu FB, Schulze MB. Transition from metabolic healthy to unhealthy phenotypes and association with cardiovascular disease risk across BMI categories in 90 257 women (the Nurses' Health Study): 30 year follow-up from a prospective cohort study. The lancet Diabetes & endocrinology 2018.
42. Blair SN, Cheng Y, Holder JS. Is physical activity or physical fitness more important in defining health benefits? Med Sci Sports Exerc 2001;33:S379-99; discussion S419-20.
43. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 2009;62:1006-12.

44. Veritas Health Innovation Ltd. Covidence systematic review software. Melbourne, Australia, 2013.
45. Artero EG, Ruiz JR, Ortega FB et al. Muscular and cardiorespiratory fitness are independently associated with metabolic risk in adolescents: the HELENA study. Pediatr Diabetes 2011;12:704-12.
46. Steene-Johannessen J, Anderssen SA, Kolle E, Andersen LB. Low muscle fitness is associated with metabolic risk in youth. Med Sci Sports Exerc 2009;41:1361-7.
47. The Joanna Briggs Institute. Joanna Briggs Institute Reviewers' Manual. Australia: The Joanna Briggs Institute, 2016.
48. Wells G, Shea B, O'Connell D et al. The Newcastle–Ottawa Scale (NOS) for Assessing the Quality of Non-Randomized Studies in Meta-Analysis, 2000.
49. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539-58.
50. Appleton SL, Seaborn CJ, Visvanathan R et al. Diabetes and cardiovascular disease outcomes in the metabolically healthy obese phenotype: a cohort study. Diabetes Care 2013;36:2388-94.
51. Aparicio VA, Soriano-Maldonado A, Buitrago F, Felix-Redondo FJ, Fernandez-Berges D. The Role of Sex and Domestic Physical Activity on the Metabolically Healthy and Unhealthy Obesity. The HERMEX Study. Rev Esp Cardiol (Engl Ed) 2016;69:983-986.
52. Bell JA, Hamer M, van Hees VT, Singh-Manoux A, Kivimaki M, Sabia S. Healthy obesity and objective physical activity. Am J Clin Nutr 2015;102:268-75.
53. Berezina A, Belyaeva O, Berkovich O et al. Prevalence, Risk Factors, and Genetic Traits in Metabolically Healthy and Unhealthy Obese Individuals. BioMed research international 2015;2015:548734.
54. Bouchard DR, Langlois MF, Brochu M, Dionne IJ, Baillargeon JP. Metabolically healthy obese women and functional capacity. Metab Syndr Relat Disord 2011;9:225-9.
55. Cadenas-Sanchez C, Ruiz JR, Labayen I et al. Prevalence of Metabolically Healthy but Overweight/Obese Phenotype and Its Association With Sedentary Time, Physical Activity, and Fitness. J Adolesc Health 2017;61:107-114.
56. Camhi SM, Crouter SE, Hayman LL, Must A, Lichtenstein AH. Lifestyle Behaviors in Metabolically Healthy and Unhealthy Overweight and Obese Women: A Preliminary Study. PLoS One 2015;10:e0138548.
57. Camhi SM, Waring ME, Sisson SB, Hayman LL, Must A. Physical activity and screen time in metabolically healthy obese phenotypes in adolescents and adults. J Obes 2013;2013:984613.
58. de Winter M, Rioux BV, Boudreau JG, Bouchard DR, Senechal M. Physical Activity and Sedentary Patterns among Metabolically Healthy Individuals Living with Obesity. 2018;2018:7496768.
59. Ding WQ, Yan YK, Zhang MX et al. Hypertension outcomes in metabolically unhealthy normal-weight and metabolically healthy obese children and adolescents. J Hum Hypertens 2015;29:548-54.
60. Doustmohamadian S, Serahati S, Barzin M, Keihani S, Azizi F, Hosseinpanah F. Risk of all-cause mortality in abdominal obesity phenotypes: Tehran Lipid and Glucose Study. Nutr Metab Cardiovasc Dis 2017;27:241-248.
61. Elmaogullari S, Demirel F, Hatipoglu N. Risk factors that affect metabolic health status in obese children. J Pediatr Endocrinol Metab 2017;30:49-55.
62. Gao X, Zhang W, Wang Y et al. Serum metabolic biomarkers distinguish metabolically healthy peripherally obese from unhealthy centrally obese individuals. Nutr Metab (Lond) 2016;13:33.
63. Goday A, Calvo E, Vazquez LA et al. Prevalence and clinical characteristics of metabolically healthy obese individuals and other obese/non-obese metabolic phenotypes in a working population: results from the Icaria study. BMC Public Health 2016;16:248.
64. Gutierrez-Repiso C, Soriguer F, Rojo-Martinez G et al. Variable patterns of obesity and cardiometabolic phenotypes and their association with lifestyle factors in the Di@bet.es study. Nutr Metab Cardiovasc Dis 2014;24:947-55.
65. Hankinson AL, Daviglus ML, Van Horn L et al. Diet composition and activity level of at risk and metabolically healthy obese American adults. Obesity (Silver Spring) 2013;21:637-43.
66. Hashimoto Y, Hamaguchi M. Fatty liver as a risk factor for progression from metabolically healthy to metabolically abnormal in non-overweight individuals. 2017;57:89-97.
67. Hayes L, Pearce MS, Firbank MJ, Walker M, Taylor R, Unwin NC. Do obese but metabolically normal women differ in intra-abdominal fat and physical activity levels from those with the expected metabolic abnormalities? A cross-sectional study. BMC Public Health 2010;10:723.

68. Hayes L, Pearce MS, Unwin NC. Lifecourse predictors of normal metabolic parameters in overweight and obese adults. Int J Obes (Lond) 2006;30:970-6.
69. Heinzle S, Ball GD. Variations in the prevalence and predictors of prevalent metabolically healthy obesity in adolescents. 2016;11:425-33.
70. Hosseinpanah F, Barzin M, Sheikholeslami F, Azizi F. Effect of different obesity phenotypes on cardiovascular events in Tehran Lipid and Glucose Study (TLGS). Am J Cardiol 2011;107:412-6.
71. Hwang YC, Cho IJ, Jeong IK, Ahn KJ, Chung HY. Differential association between sarcopenia and metabolic phenotype in Korean young and older adults with and without obesity. Obesity (Silver Spring) 2017;25:244-251.
72. Ingle L, Swainson M, Brodie D, Sandercock GR. Characterization of the metabolically healthy phenotype in overweight and obese British men. Prev Med 2017;94:7-11.
73. Jennings CL, Lambert EV, Collins M, Joffe Y, Levitt NS, Goedecke JH. Determinants of insulin-resistant phenotypes in normal-weight and obese Black African women. Obesity (Silver Spring) 2008;16:1602-9.
74. Kanagasabai T, Thakkar NA, Kuk JL, Churilla JR, Ardern CI. Differences in physical activity domains, guideline adherence, and weight history between metabolically healthy and metabolically abnormal obese adults: a cross-sectional study. Int J Behav Nutr Phys Act 2015;12:64.
75. Korhonen PE, Korsoff P, Vahlberg T, Kaaja R. Lifestyle of metabolically healthy obese individuals. Prim Care Diabetes 2015;9:179-83.
76. Kuzik N. Physical Activity and Sedentary Time Associations with Metabolic Health Across Weight Statuses in Children and Adolescents. PLoS One 2017;25:1762-1769.
77. Lassale C, Tzoulaki I, Moons KGM et al. Separate and combined associations of obesity and metabolic health with coronary heart disease: a pan-European case-cohort analysis. Eur Heart J 2018;39:397-406.
78. Lee SK, Kim SH, Cho GY et al. Obesity phenotype and incident hypertension: a prospective community-based cohort study. J Hypertens 2013;31:145-51.
79. Lim HH. Sleep duration independently influences metabolic body size phenotype in children and adolescents: a population-based study. Journal of diabetes research 2018;42:47-52.
80. Matta J, Nasreddine L, Jomaa L et al. Metabolically Healthy Overweight and Obesity Is Associated with Higher Adherence to a Traditional Dietary Pattern: A Cross-Sectional Study among Adults in Lebanon. Nutrients 2016;8.
81. McElroy JA, Gilbert T, Hair EC, Mathews KJ, Redman SD, Williams A. Obese But Fit: The Relationship of Fitness to Metabolically Healthy But Obese Status among Sexual Minority Women. Womens Health Issues 2016;26 Suppl 1:S81-6.
82. Moon S, Oh CM, Choi MK et al. The influence of physical activity on risk of cardiovascular disease in people who are obese but metabolically healthy. J Korean Med Sci 2017;12:e0185127.
83. Oh CM, Jun JK, Suh M. Risk of cancer mortality according to the metabolic health status and degree of obesity. Asian Pac J Cancer Prev 2014;15:10027-31.
84. Park YM, Steck SE, Fung TT et al. Mediterranean diet and mortality risk in metabolically healthy obese and metabolically unhealthy obese phenotypes. Int J Obes (Lond) 2016;40:1541-1549.
85. Phillips CM, Dillon C, Harrington JM et al. Defining metabolically healthy obesity: role of dietary and lifestyle factors. PLoS One 2013;8:e76188.
86. Poelkens F, Eijsvogels TM, Brussee P, Verheggen RJ, Tack CJ, Hopman MT. Physical fitness can partly explain the metabolically healthy obese phenotype in women. Exp Clin Endocrinol Diabetes 2014;122:87-91.
87. Prince RL, Kuk JL, Ambler KA, Dhaliwal J, Ball GD. Predictors of metabolically healthy obesity in children. Diabetes Care 2014;37:1462-8.
88. Schroder H, Ramos R, Baena-Diez JM et al. Determinants of the transition from a cardiometabolic normal to abnormal overweight/obese phenotype in a Spanish population. Eur J Nutr 2014;53:1345-53.
89. Senechal M, Wicklow B, Wittmeier K et al. Cardiorespiratory fitness and adiposity in metabolically healthy overweight and obese youth. Pediatrics 2013;132:e85-92.
90. Slagter SN, Corpeleijn E, van der Klauw MM et al. Dietary patterns and physical activity in the metabolically (un)healthy obese: the Dutch Lifelines cohort study. Nutrition journal 2018;17:18.
91. Song Y, Manson JE, Meigs JB, Ridker PM, Buring JE, Liu S. Comparison of usefulness of body mass index versus metabolic risk factors in predicting 10-year risk of cardiovascular events in women. Am J Cardiol 2007;100:1654-8.

92. Sung KC, Ryu S, Cheong ES et al. All-Cause and Cardiovascular Mortality Among Koreans: Effects of Obesity and Metabolic Health. Am J Prev Med 2015;49:62-71.
93. van der A DL, Nooyens AC, van Duijnhoven FJ, Verschuren MM, Boer JM. All-cause mortality risk of metabolically healthy abdominal obese individuals: the EPIC-MORGEN study. Obesity (Silver Spring) 2014;22:557-64.
94. Wiklund PK, Pekkala S, Autio R et al. Serum metabolic profiles in overweight and obese women with and without metabolic syndrome. Diabetol Metab Syndr 2014;6:40.
95. Wildman RP, Muntner P, Reynolds K et al. The obese without cardiometabolic risk factor clustering and the normal weight with cardiometabolic risk factor clustering: prevalence and correlates of 2 phenotypes among the US population (NHANES 1999-2004). Arch Intern Med 2008;168:1617-24.
96. Yoon DY, Lee YA. Prevalence and Clinical Characteristics of Metabolically Healthy Obesity in Korean Children and Adolescents: Data from the Korea National Health and Nutrition Examination Survey. 2017;32:1840-1847.
97. Yu R, Yau F, Ho SC, Woo J. Associations of cardiorespiratory fitness, physical activity, and obesity with metabolic syndrome in Hong Kong Chinese midlife women. BMC Public Health 2013;13:614.
98. Zhang N, Chen Y, Guo X, Sun G, Dai D, Sun Y. Metabolic Abnormalities, But Not Metabolically Healthy Obesity, Are Associated with Left Ventricular Hypertrophy. Heart Lung Circ 2017;26:251-257.
99. Zhang N, Chen Y, Guo X, Sun G, Sun Y. The prevalence and predictors of metabolically healthy obesity in obese rural population of China: a cross-sectional study. Psychol Health Med 2017;22:303-309.
100. Donini LM, Merola G, Poggiogalle E et al. Disability, Physical Inactivity, and Impaired Health-Related Quality of Life Are Not Different in Metabolically Healthy vs. Unhealthy Obese Subjects. Nutrients 2016;8.
101. Bell JA, Kivimaki M, Batty GD, Hamer M. Metabolically healthy obesity: what is the role of sedentary behaviour? Prev Med 2014;62:35-7.
102. de Rooij BH, van der Berg JD, van der Kallen CJ et al. Physical Activity and Sedentary Behavior in Metabolically Healthy versus Unhealthy Obese and Non-Obese Individuals - The Maastricht Study. PLoS One 2016;11:e0154358.
103. Zheng R, Yang M, Bao Y et al. Prevalence and Determinants of Metabolic Health in Subjects with Obesity in Chinese Population. Int J Environ Res Public Health 2015;12:13662-77.
104. Saghafi-Asl M, Amiri P, Naghizadeh M, Ghavami SM, Karamzad N. Association of endotoxaemia with serum free fatty acids in metabolically healthy and unhealthy abdominally obese individuals: a case-control study in northwest of Iran. BMJ Open 2017;7:e015910.
105. Chang SH, Chen MC, Chien NH, Wu LY. CE: Original Research: Examining the Links Between Lifestyle Factors and Metabolic Syndrome. Am J Nurs 2016;116:26-36.
106. Aparicio VA, Carbonell-Baeza A, Senhaji M, Martin S, Camiletti-Moiron D, Aranda P. Usefulness of fitness testing to establish metabolic syndrome in perimenopausal Moroccan women. Eur J Cardiovasc Nurs 2014;13:524-31.
107. Carvalho LP, Basso-Vanelli RP, Di Thommazo-Luporini L et al. Myostatin and adipokines: The role of the metabolically unhealthy obese phenotype in muscle function and aerobic capacity in young adults. Cytokine 2018;107:118-124.
108. Dalzill C, Nigam A, Juneau M et al. Intensive lifestyle intervention improves cardiometabolic and exercise parameters in metabolically healthy obese and metabolically unhealthy obese individuals. Can J Cardiol 2014;30:434-40.
109. Dobson R, Burgess MI, Sprung VS et al. Metabolically healthy and unhealthy obesity: differential effects on myocardial function according to metabolic syndrome, rather than obesity. Int J Obes (Lond) 2016;40:153-61.
110. Gregorio-Arenas E, Ruiz-Cabello P, Camiletti-Moiron D et al. The associations between physical fitness and cardiometabolic risk and body-size phenotypes in perimenopausal women. Maturitas 2016;92:162-167.
111. Jae SY, Franklin B, Choi YH, Fernhall B. Metabolically Healthy Obesity and Carotid Intima-Media Thickness: Effects of Cardiorespiratory Fitness. Mayo Clin Proc 2015;90:1217-24.
112. Jae SY, Babu AS, Yoon ES et al. Impact of Cardiorespiratory Fitness and Risk of Systemic Hypertension in Nonobese Versus Obese Men Who Are Metabolically Healthy or Unhealthy. Am J Cardiol 2017;120:765-768.
113. Karelis AD, Faraj M, Bastard JP et al. The metabolically healthy but obese individual presents a favorable inflammation profile. J Clin Endocrinol Metab 2005;90:4145-50.

114. Messier V, Karelis AD, Prud'homme D, Primeau V, Brochu M, Rabasa-Lhoret R. Identifying metabolically healthy but obese individuals in sedentary postmenopausal women. Obesity (Silver Spring) 2010;18:911-7.
115. Messier V, Malita FM, Rabasa-Lhoret R, Brochu M, Karelis AD. Association of cardiorespiratory fitness with insulin sensitivity in overweight and obese postmenopausal women: a Montreal Ottawa New Emerging Team study. Metabolism 2008;57:1293-8.
116. Bo S, Musso G, Gambino R et al. Prognostic implications for insulin-sensitive and insulin-resistant normal-weight and obese individuals from a population-based cohort. Am J Clin Nutr 2012;96:962-9.
117. Loprinzi PD, Frith E. Cardiometabolic healthy obesity paradigm and all-cause mortality risk. Eur J Intern Med 2017;43:42-45.
118. Katzmarzyk PT, Church TS, Janssen I, Ross R, Blair SN. Metabolic syndrome, obesity, and mortality: impact of cardiorespiratory fitness. Diabetes Care 2005;28:391-7.
119. Blüher M. The distinction of metabolically 'healthy' from 'unhealthy' obese individuals. Curr Opin Lipidol 2010;21:38-43.
120. Stefan N, Kantartzis K, Machann J et al. Identification and characterization of metabolically benign obesity in humans. Arch Intern Med 2008;168:1609-16.
121. Moon S, Oh CM, Choi MK et al. The influence of physical activity on risk of cardiovascular disease in people who are obese but metabolically healthy. 2017;12:e0185127.
122. Lee DC, Sui X, Ortega FB et al. Comparisons of leisure-time physical activity and cardiorespiratory fitness as predictors of all-cause mortality in men and women. Br J Sports Med 2011;45:504-10.
123. Stefan N, Haring HU, Schulze MB. Metabolically healthy obesity: the low-hanging fruit in obesity treatment? The lancet Diabetes & endocrinology 2018;6:249-258.

Table 1. Longitudinal studies examining the differences between MHNW and MHO after further adjustment of PA OR physical fitness (N=11).

Author, year	Setting Study design and years of follow-up Sample size	Age at baseline (SD)	Metabolic criteria	Subgroup definition	Outcomes of interest	Number of cases	Statistical analysis and adjustments	Main findings
Appleton et al, 2013	NWAHS study Longitudinal, follow-up: 8.2 years Random sample N = 2315 **Subgroups:** N_{MHNW} = 636 N_{MHO} = 281	Not specified	- HDL< 1.0♂ / 1.3♀mmol/l - TG ≥ 1.7 mmol/l - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 5.6 mmol/l	*MHO/MHNW*: ≤1 of the criteria met	CVD/stroke cases	*CVD/stroke cases:* N_{MHNW} = 25 N_{MHO} = 15	OR Model 1: age, sex, smoking, household income, highest education level, LDL and **PA**	Compared with the MHNW group, MHO did not show an increase in CVD/stroke cases after adjustment for confounders (OR: 1.16, p>0.05).
Bo et al, 2012	Asti (Italy) Longitudinal, follow-up: 9 years Not specified N = 1658 **Subgroups:** N_{MHNW} = 540 N_{MHO} = 72	53.6 (5.6)	- WC ≥ 94♂ / 80♀ cm - TG ≥ 1.7 mmol/L - HDL < 1.0♂ / 1.3♀mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 5.6 mmol/L	*MHO/MHNW*: ≤2 of the criteria met	-Hyperglycemia cases -CVD cases	Not specified	HR -Model 1: adjusted for age, sex, fiber intake, waist circumference and **PA**	MHO showed higher incident in hyperglycemia and CVD cases than in MHNW (HR:2.16 and 2.76, respectively)
Doustmohamadian et al, 2017	*Tehran Lipid and Glucose Study* Longitudinal, follow-up: 8 years Random sample N = 8804 **Subgroups:** N_{MHNW} = 2086 N_{MHOO} = 1125	47.7 (12.6)	- HDL < 40♂ / 50♀ mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 100 mg/dL - 2-h blood glucose ≥ 140 mg/dL	*MHOO/MHNW*: ≤1 of the criteria met Participants were classified as abdominal obese (i.e. ≥ 89 cm for men and ≥ 91 cm for women)	All-cause mortality	*-All-cause mortality:* N_{MHNW} = 65 N_{MHOO} = 55	HR -Model 1: Unadjusted -Model 2: adjusted for age and sex -Model 3: adjusted for model 1 plus	In unadjusted model, MHO presented higher risk of mortality for all-cause than MHNW (HR: 1.65, p<0.05). However, after adjustment for either model 2 or 3, they did not find significant differences between MHO and MHNW (Model 1, HR: 1.21; model 2, HR: 1.35; all p>0.05).

				healthy and unhealthy			smoking, educational level and **PA**	
Hosseinpanah et al, 2011	*TLGS study* Longitudinal, follow-up: 8.1 years Random sample N = 6215 **Subgroups:***:* N_{MHNW} = 1555 N_{MHO} = 408	47.4 (0.2)	-WC > 89♂ / 91♀ cm - HDL < 40♂ / 50♀ mg/dL - TG ≥ 150 mg/dL - SBP ≥ 140 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 100 mg/dL	*MHO/MHNW*: ≤1 of the criteria met (excluding WC)	CVD cases	*CVD cases:* N_{MHNW} = 64 N_{MHO} = 13	HR -Model 1: adjusted for age and sex -Model 2: adjusted for model 1 plus smoking, family history of premature CAD, high TC and **VPA**	No significant differences were observed in the incidents for CVD between MHO and MHNW after adjusted for confounders (Model 1, HR: 1.01; model 2, HR: 1.07, p>0.05)
Ortega et al, 2013	ACLS United States Longitudinal, follow-up: 14.3 (for mortality), 7.9 (non-fatal CVD) years Random sample N=43265 **Subgroups:** *All-cause and CVD mortality* N_{MHNW} = 16002 N_{MHO} = 1738 *Non-fatal CVD cases* N_{MHNW} = 7001 N_{MHO} = 544	44.2 (9.9)	- HDL≤ 40♂ / 50♀ mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥100 mg/dL *Note: in the paper they provided results from obese categorized based on their BF%. For more information see the paper*	*MHO/MHNW*: ≤1 of the criteria met	-All-cause mortality -CVD mortality -Non-fatal CVD cases	-*All-cause mortality:* N_{MHNW} = 449 N_{MHO} = 52 -*CVD mortality:* N_{MHNW} = 98 N_{MHO} = 17 -*Non-fatal CVD cases:* N_{MHNW} = 261 N_{MHO} = 30	HR -Model 1: adjusted for age, sex, examination year, smoking, alcohol consumption and parental history of CVD (this last confounder were removed for cancer mortality analyses) -Model 2: Model 1 plus **CRF**	In all-cause mortality, no significant difference was observed between MHO and MHNW participant (Model 2, HR: 0.91). This result was consistent for CVD mortality and incident non-fatal CVD, as well as for cancer mortality after adjustment for the set of confounders including fitness (HR: 0.73, 0.78 and 0.61, respectively).
Song et al, 2007	Women's Health study United States	54.3 (6.5)	- HDL < 50 mg/dL - TG ≥ 150 mg/dL - SBP ≥ 135 mmHg	*MHO/ MHNW*: ≤2 of the criteria met	- CVD cases - CHD cases -Stroke cases	-*CVD cases* N_{MHNW} = 278 N_{MHO} = 77	RR	MHO showed similar risk of CVD as compared with the MHNW after

	Longitudinal, follow-up: 10.2 years Random sample N=25626 **Subgroups:** N_{MHNW} = 12943 N_{MHO} = 2925		- DBP ≥ 85 mmHg - Glucose ≥ 110 mg/dL			- *CHD cases* N_{MHNW} = 149 N_{MHO} = 55 - *Stroke cases* N_{MHNW} = 129 N_{MHO} = 22	-Model 1: adjusted for age and vitamin E and aspirin. -Model 2: Model 1 with additional adjustment for smoking, alcohol intake, total calorie intake, postmenopausal hormone use, multivitamin use, parental history of myocardial infarction, BMI and **PA**	adjusted for model 2 (RR: 1.05). The RR of MHO in model 1 was 1.36. In regards to CHD, the MHO showed higher RR than MHNW (Model 1, RR: 1.80; model 2, RR: 1.25). MHO showed lower RR of stroke cases than MHNW in both models examined (Model 1, RR: 0.85; model 2, RR: 0.82).
Sung et al, 2015	Korea (Asia) Longitudinal, follow-up: not reported Random sample N=275867 N_{men}=156252 N_{women}=119615 **Subgroups:** *Men:* N_{MHO} = 12731 N_{MUO} = 49269 *Women:* **Subgroups:** N_{MHO}= 5730 N_{MUO} = 15906	40.2 (10)	- SBP ≥ 140 mmHg - DBP ≥ 90 mmHg - Glucose ≥ 126 mg/Dl	*MHO/MHNW*: none of the criteria met	-All-cause mortality -CVD mortality	-*All-cause mortality:* N_{MHNW} = 212 N_{MHO} = 39 -*CVD mortality:* N_{MHNW} = 21 N_{MHO} = 2	HR -Model 1: adjusted for age and sex -Model 2: Model 1 plus smoking status, alcohol intake, education level among participants without diabetes, HTN and a history of CVD and **PA**	For both all-cause and CVD mortality, there was no significant difference between MHO and MHNW (HR~0.70 for all-cause mortality and HR~0.40 for CVD mortality for both models, p>0.05) in any of the models examined.

Lassale et al, 2018	*EPIC-CVD study* 10 European countries Longitudinal, follow-up: 12.2 years (median) Random sample N=10474 **Subgroups:** N_{MHNW}= 6165 N_{MHO} = 1103	52.4 (not specified)	- WC ≥ 94♂ / 80♀ cm - HDL < 40♂ / 50♀ mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥100 mg/dL	*MHO/MHNW*: ≤2 of the criteria met	CHD cases	*-CHD cases:* N_{MHNW} = 1978 N_{MHO} = 360	HR adjusted for age, smoking, educational level, Mediterranean diet score, energy, alcohol intake and **PA**	MHO presented higher risk of CHD cases than MHNW (HR: 1.28, p=0.02)
Loprinzi et al, 2017	NHANES United States Longitudinal, follow-up: 8.58 years (median) Random sample N=7579 **Subgroups:** N_{MHNW} = 879 N_{MHO} = 299	46.9 (not specified)	- HDL< 40♂ / 50♀ mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥100 mg/dL	*MHO/MHNW*: none of the criteria met	All-cause mortality	*-All-cause mortality:* N_{MHNW} = 26 N_{MHO} = 9	HR -Model 1: adjusted for age, sex, race-ethnicity and smoking status -Model 2: Model 1 plus **meeting PA guidelines**	MHO presented higher risk of mortality for all-cause than MHNW (HR: 2.48, p=0.02) after adjustment for model 1. When adding meeting MVPA guidelines, result was unchanged (HR: 2.41).
Moon et al, 2017	*KoGES study* Korea (Asia) Longitudinal, follow-up: 8.4 years Random sample N=8144 **Subgroups:** N_{MHNW} = 2929 N_{MHO} = 1262	50.6 (not specified)	- HDL ≤ 1.03♂ / 1.30♀ mmol/L - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 100 mg/dL	*MHO/MHNW*: ≤1 of the criteria met	CVD cases	*CVD cases:* N_{MHNW} = 135 N_{MHO} = 70	HR -Model 1: adjusted for age and sex -Model 2: adjusted for model 1 plus alcohol drinking, smoking, medication of HTN, diabetes mellitus,	MHO participants were not at elevated risk of CVD compared with their MHNW counterparts after adjustment for confounders (Model 1, HR: 1.28; model 2, HR: 1.288, p≥0.088).

							dyslipidemia and **PA**	
Van der A et al, 2014	*The Dutch EPIC-MORGEN study* Amsterdam, Maastrich, Doetinchem (The Netherlands) Longitudinal, follow-up: 13.4 years Random samples N=20299 **Subgroups:** N_{MHNW} = 4799 N_{MHO} = 737	42.8 (10.5)	- WC ≥ 102♂ / 88♀ cm - HDL < 1.0♂ / 1.3♀ mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 5.6 mmol/L - TC ≥ 6.5 mmol/L	*MHO/MHNW*: none of the criteria met (including WC)	All-cause mortality	*All-cause mortality:* N_{MHNW} = 82 N_{MHO} = 35	HR -Model 1: Age and sex -Model 2: Model 1 plus smoking, education, total energy intake, protein intake, carbohydrate intake and **PA**	MHO had higher risk of all-cause mortality than MHNW individuals (model 1, HR: 1.62; model 2, HR: 1.66).

♂: males, ♀: females, BMI: body mass index, WC: waist circumference, TG: triglycerides, TC: total cholesterol, HDL: high-density lipoprotein cholesterol, LDL: low-density lipoprotein cholesterol, SBP: systolic blood pressure, DBP: diastolic blood pressure, HTN: hypertension, MHNW: metabolically healthy normal weight, MHO: metabolically healthy obese, CRF: cardiorespiratory fitness, PA: physical activity, CVD: cardiovascular disease, CHD: coronary heart disease, CAD: coronary artery disease, HR: hazard ratio, OR: odds ratio.

Note:

-N sample include all participants in the study (not only those stated in the subgroups), thus, in some cases the first sample size provided in the setting-study design column is higher than the sum of those stated in the subgroups subsection.

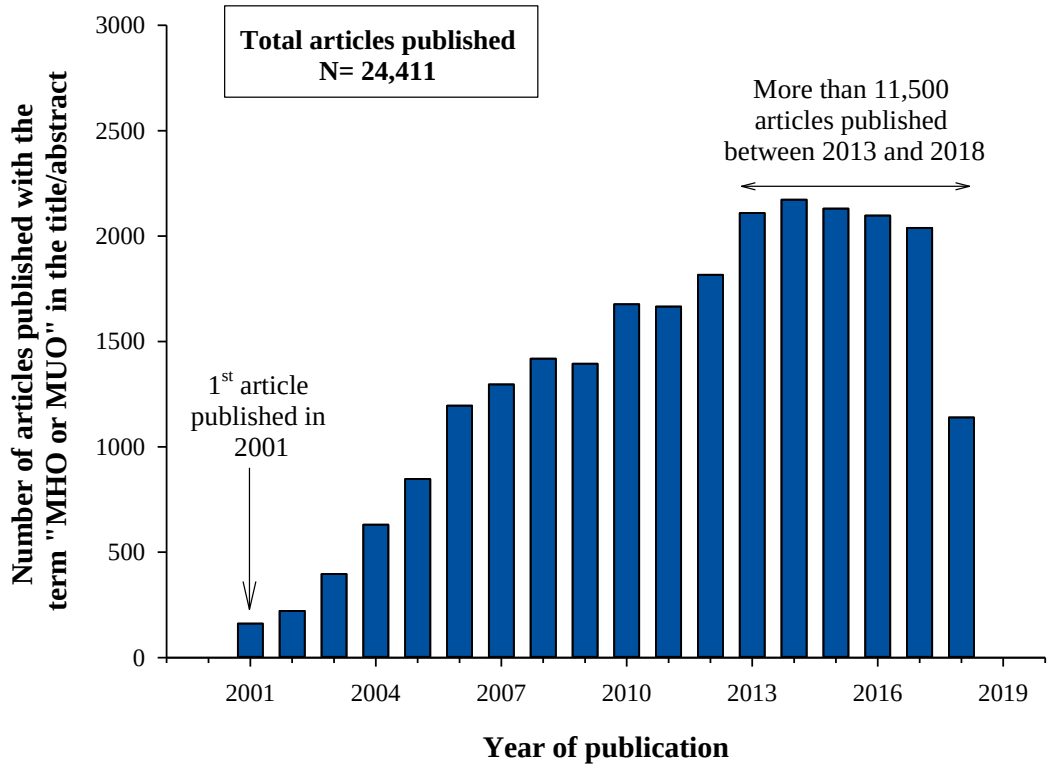


Figure 1. Number of publications focused on MHO/MUO during recent years.

MHO: Metabolically healthy obesity. MUO: Metabolically unhealthy obesity. Terms included for this search are provided as Supplementary material Table S1.

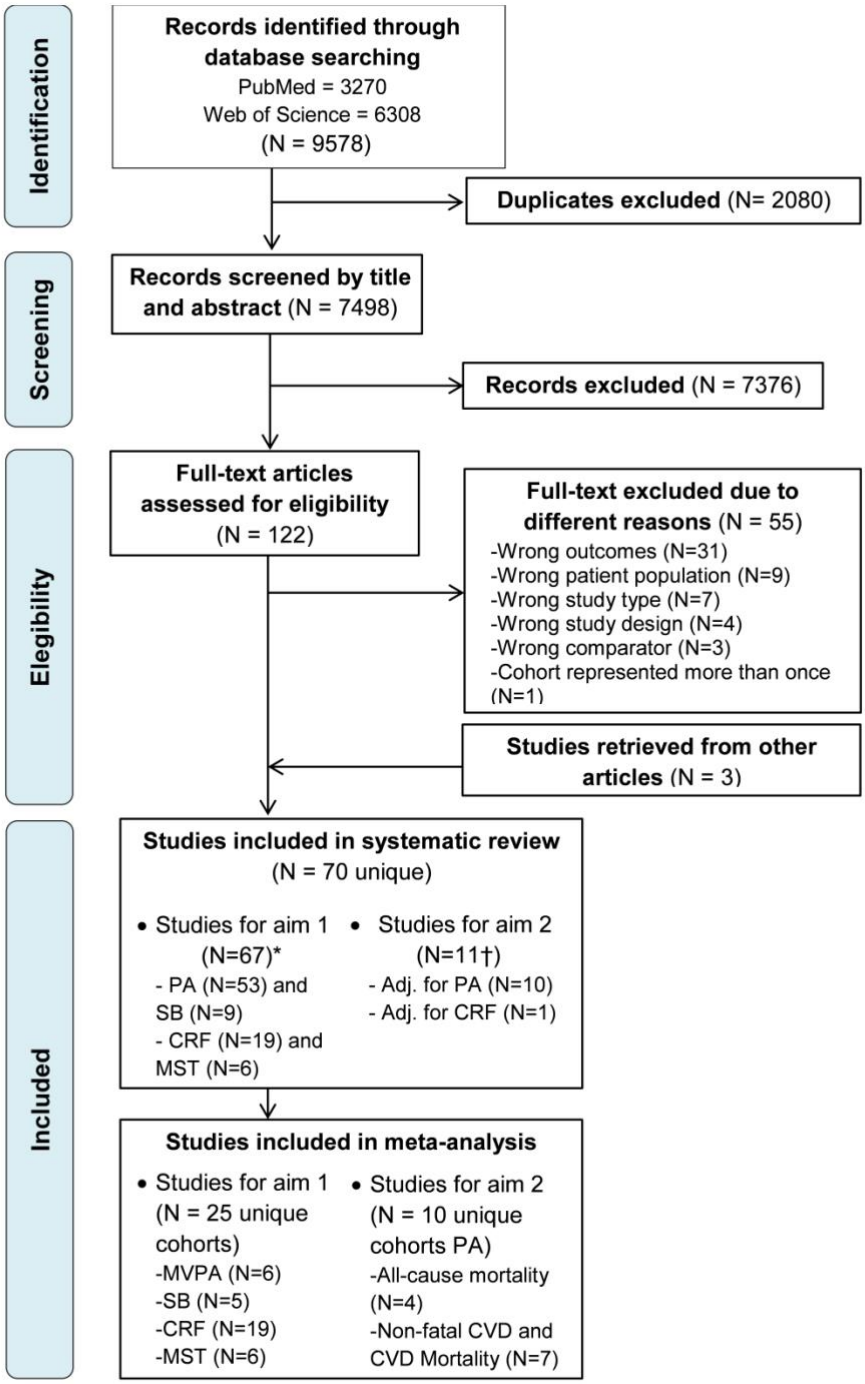


Figure 2. Flow diagram of studies included through the review process according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)(43).

*Studies were classified both in fitness and PA/SB outcomes, N=7.
†Studies were classified both for aim 1 and 2, N=8.
PA: Physical activity. SB: Sedentary behaviors. Adj.: adjusting. CRF: cardiorespiratory fitness. MST: relative muscular strength. MVPA: moderate-to-vigorous physical. CVD: cardiovascular diseases.

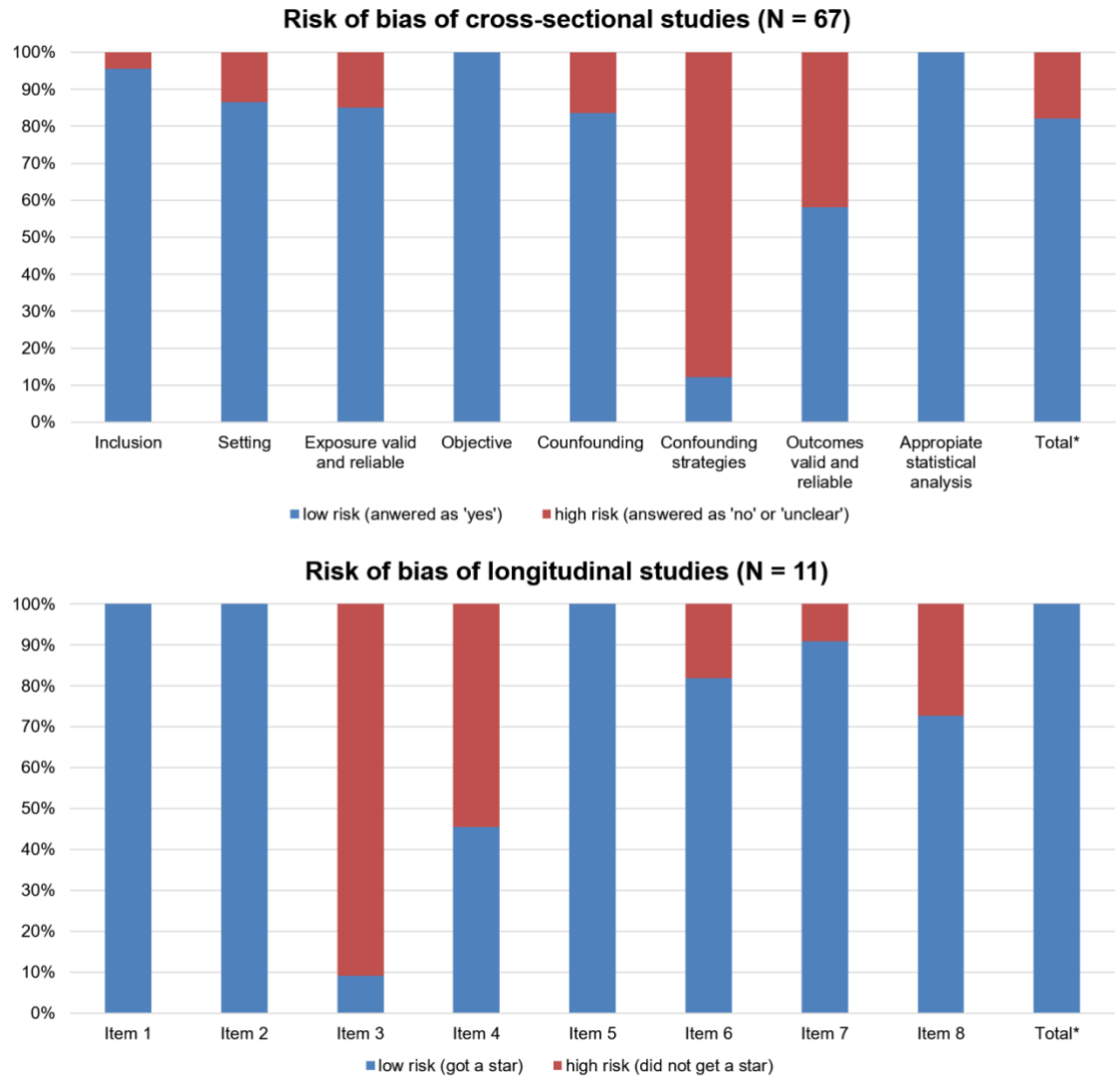


Figure 3. Stacked bar plot showing the risk of bias assessed in all cross-sectional and longitudinal studies included in the systematic review and meta-analysis.

The checklist used for risk of bias assessment for cross-sectional and longitudinal studies can be found in Supplementary material 1 and 2, respectively. *Total risk of bias was considered 'low' if cross-sectional studies accumulated at least 5 items (total = 8 items) answered as 'yes' or longitudinal studies awarded at least 6 stars (total = 9 stars). The specific information with the scoring of each study in each item and total is shown in Table S4 for cross-sectional studies and Table S5 for longitudinal studies.

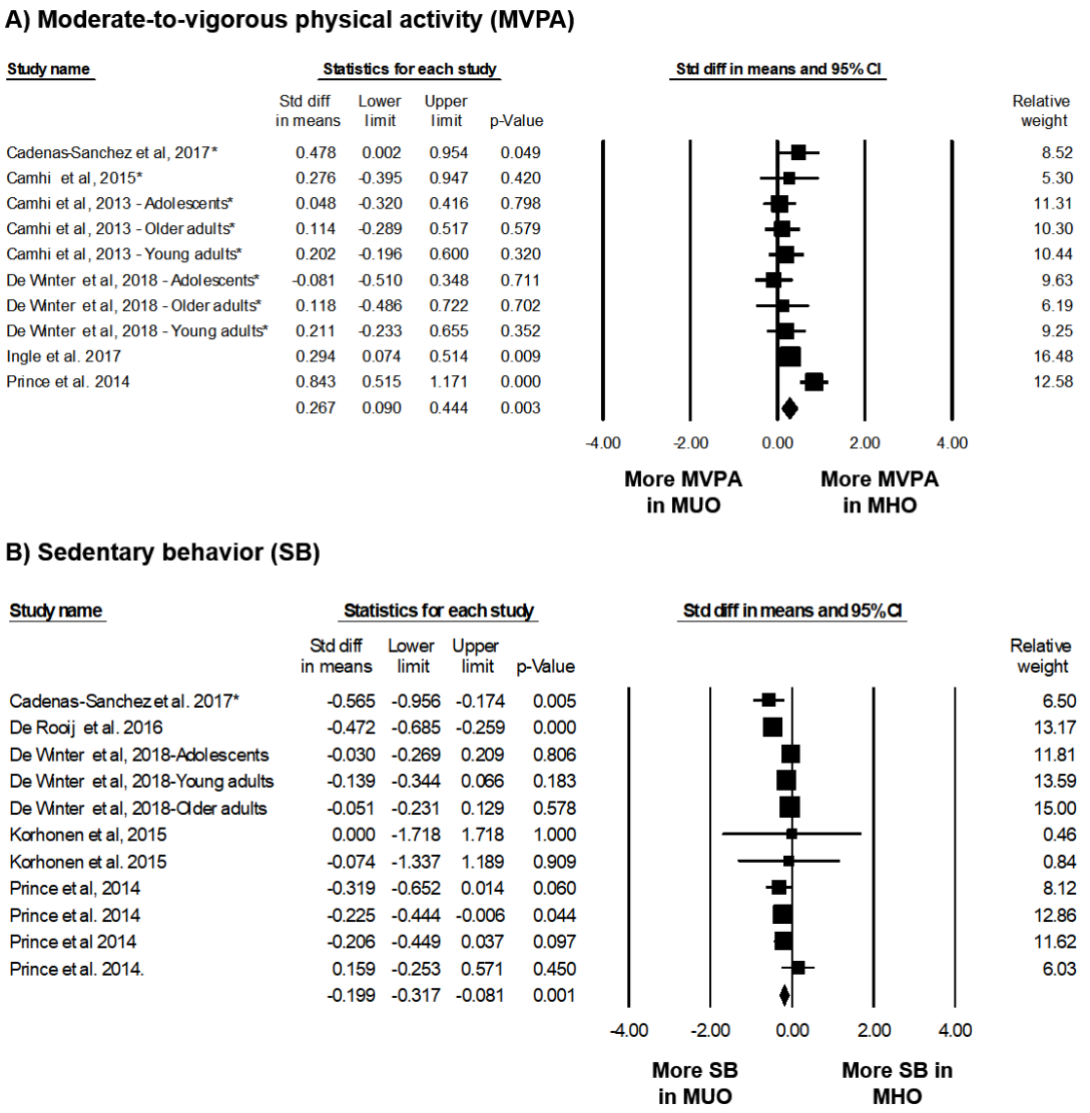


Figure 4. Forest plot of moderate-to-vigorous physical activity (Panel A) and sedentary behavior (Panel B) differences (pooled standardized mean difference) between metabolic phenotypes in overweight/obese participants.

*Studies combining overweight and obese participants. Standardized mean difference are expressed as MHO *minus* MUO, thus, positive values represents that MHO showed higher values than MUO. The leave-one-out analysis did not alter the pooled estimate. MHO: metabolically healthy obesity. MUO: Metabolically unhealthy obesity. MVPA: moderate-to-vigorous physical activity. SB: sedentary behavior. CI: confidence interval.

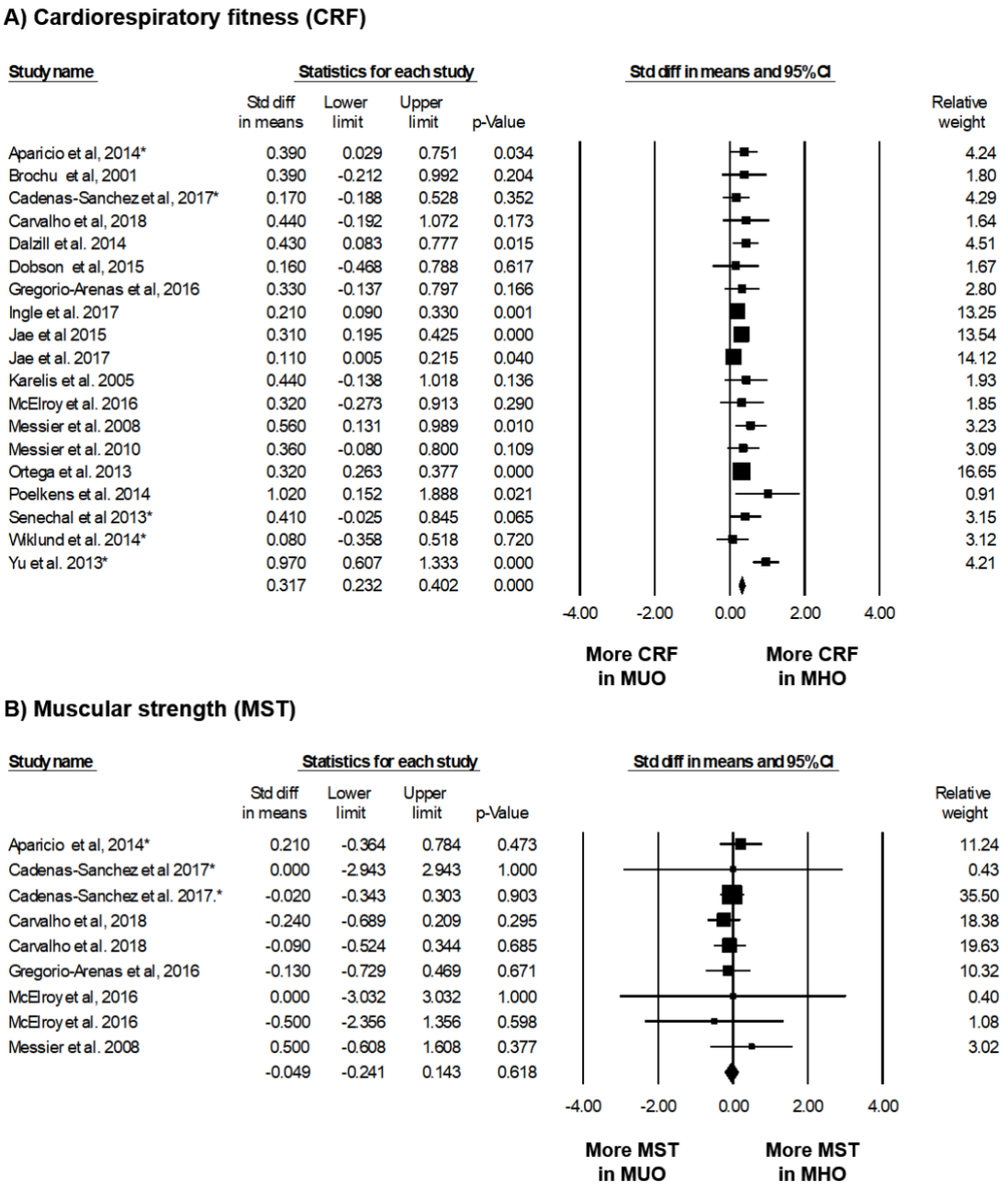


Figure 5. Forest plot of cardiorespiratory fitness (Panel A) and muscular strength (Panel B) differences (pooled standardized mean difference) between metabolic phenotypes in overweight/obese participants.

*Studies combining overweight and obese participants. Standard differences in means are expressed as MHO *minus* MUO, thus, positive values represent that MHO showed higher values than MUO. The leave-one-out analysis did not alter the pooled estimate. MHO: metabolically healthy obesity. MUO: Metabolically unhealthy obesity. CRF: cardiorespiratory fitness. MST: muscular strength. CI: confidence interval

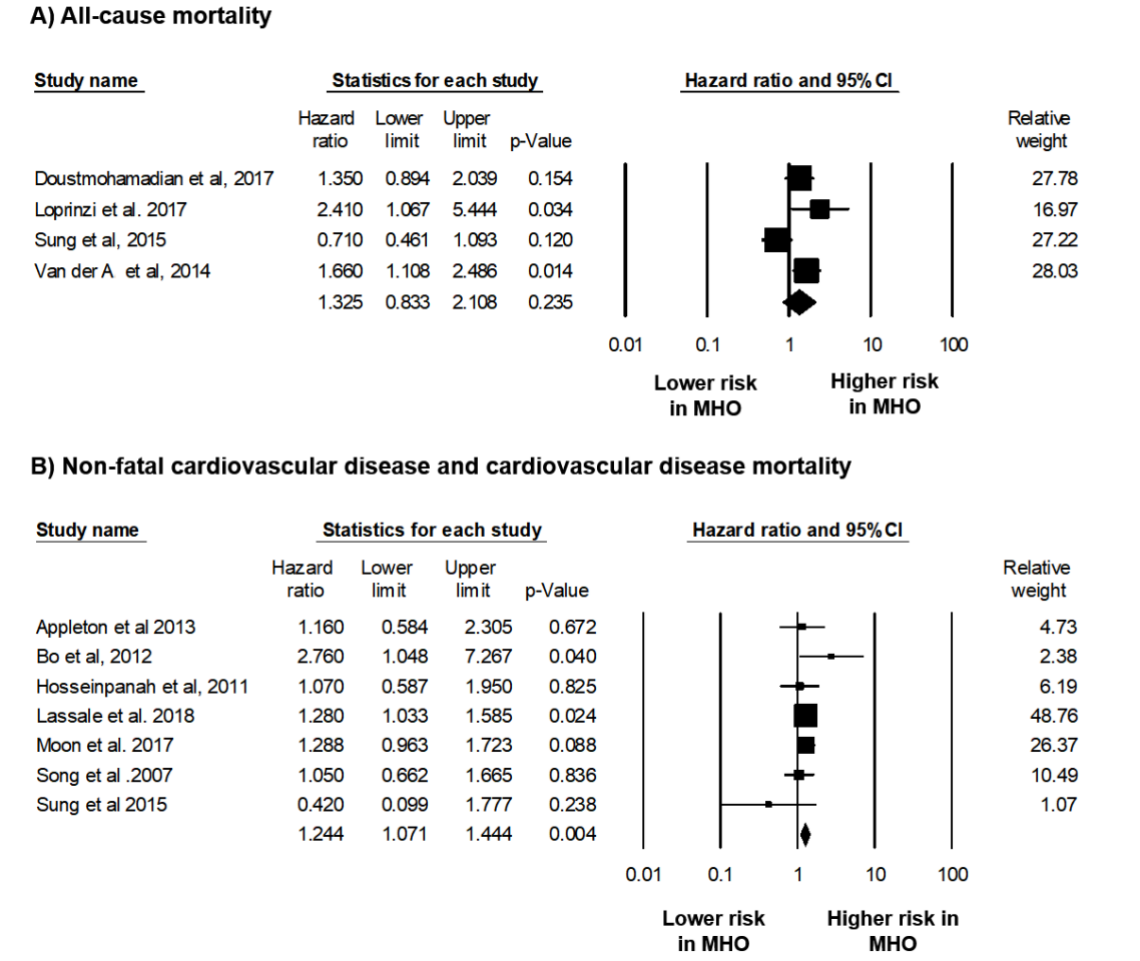


Figure 6. Forest plot of metabolically healthy normal weight participants for the risk of all-cause mortality (Panel A) and non-fatal cardiovascular disease and cardiovascular disease mortality (Panel B) compared with metabolically healthy obese individuals after adjusting for physical activity.

Hazard ratio is presented having metabolic healthy normal weight phenotype as reference (1). The leave-one-out analysis showed that the effect sizes observed were slightly decreased or increased, changing the pooled effect from non-significant to significant for all-cause mortality and *vice versa* for CVD outcomes (see exact estimates in the text, Results). MHO: metabolically healthy obesity. CI: confidence interval.

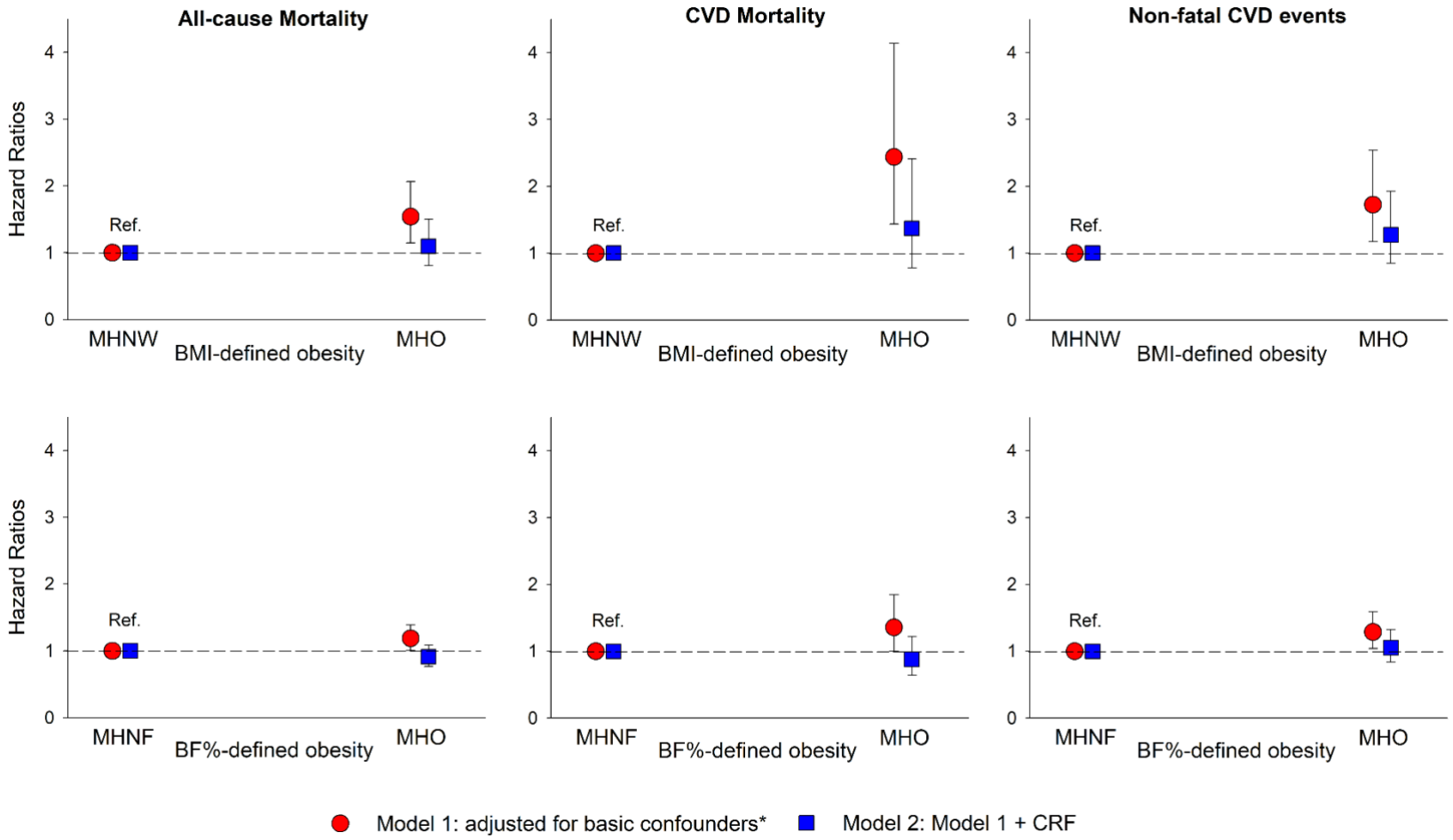


Figure 7. Role of cardiorespiratory fitness (CRF) on the prognosis of all-cause mortality and cardiovascular disease (CVD) of metabolically healthy obesity (MHO) men and women compared to metabolically healthy normal-weight (MHNW) or normal-fat (MHNF) men and women from the Aerobics Center Longitudinal Study (ACLS, N= 43,265 adults).

This figure has been newly created for the purpose of this review, using the data presented in tables as well as methods and models described in detail by Ortega et al., 2013 (32). The major change from the original article is that here we set the MHNW/MHNF group as reference in the models instead of the MHO group as in the original paper, for an easier interpretation of the role of fitness. *Model 1 is adjusted for age, sex, examination year, smoking, alcohol consumption and parental history of CVD. Non-fatal CVD events include: myocardial infarction, stroke, and coronary revascularization (i.e. bypass, coronary angioplasty); data available in a subsample of 18430 participants.

SUPPLEMENTARY MATERIAL

Supplementary material 1. Tool used to assess the study quality and risk of bias of cross-sectional studies.

Checklist for analytical cross-sectional studies from The Joanna Briggs Institute:

1. Were the criteria for inclusion in the sample clearly defined?
2. Were the study subjects and the setting described in detail?
3. Was the exposure measured in a valid and reliable way?
4. Were objective, standard criteria used for measurement of the condition?
5. Were confounding factors identified?
6. Were strategies to deal with confounding factors stated?
7. Were the outcomes measured in a valid and reliable way?
8. Was appropriate statistical analysis used?

Supplementary material 2. Tool used to assess the study quality and risk of bias of longitudinal studies.

Newcastle Ottawa Scale for cohort studies:

Selection

1) Representativeness of the exposed cohort

* a) Truly representative of the average _____ (describe) in the community
* b) Somewhat representative of the average _____ in the community

c) Selected group of users eg nurses, volunteers

d) No description of the derivation of the cohort

2) Selection of the non-exposed cohort

* a) Drawn from the same community as the exposed cohort

b) Drawn from a different source

c) Selected group of users eg nurses, volunteers

d) No description of the derivation of the non-exposed cohort

3) Ascertainment of exposure

* a) Directly measured physical activity by accelerometers or cardiorespiratory fitness by gas exchange measurement
* b) Physical activity measured by pedometers or physical fitness by field-based tests

c) Physical activity or fitness measured with self-reports

d) No description

4) Demonstration that outcome of interest was not present at start of study

* a) Yes

b) No

Comparability

5) Comparability of cohorts on the basis of the design or analysis

* a) Study controls for physical activity/physical fitness
* b) Study controls for any additional factor

Outcome

6) Assessment of outcome

* a) Independent blind assessment
* b) Record linkage

c) Self-report

d) No description

7) Was follow-up long enough for outcomes to occur

* a) Yes (≥ 5 years)

b) No

8) Adequacy of follow up of cohorts

* a) Complete follow up - all subjects accounted for
* b) Subjects lost to follow up unlikely to introduce bias - small number lost ≥ 80 % follow up, or description provided of those lost.

c) Follow up rate < 80% and no description of those lost

d) No statement

Table S1. Search terms combined used for this systematic review in each database and number of studies found, and search term used for the purpose of Figure 1.

Systematic review-PubMed (N=3270)
(("cardiorespiratory fitness”[Title/Abstract] OR cardiorespiratory fitness [MeSH] OR "aerobic capacity”[Title/Abstract] OR "aerobic fitness” [Title/Abstract] OR "physical fitness”[Title/Abstract] OR "functional capacity”[Title/Abstract] OR "physical characteristic*”[Title/Abstract] OR "oxygen consumption”[Title/Abstract] OR "vo2max” [Title/Abstract] OR "vo2peak” [Title/Abstract] OR fitness[Title/Abstract] OR “physical activity”[Title/Abstract] OR “motor activity” [Title/Abstract] OR motor activity [MeSH] OR “exercise” [Title/Abstract] OR exercise [MeSH]) AND ("metabolically healthy"[Title/Abstract] OR "metabolically abnormal"[Title/Abstract] OR "metabolically unhealthy"[Title/Abstract] OR "metabolic profile"[Title/Abstract] OR "metabolic syndrome"[Title/Abstract] OR "metabolic syndrome"[MeSH] OR "metabolically phenotype"[Title/Abstract] OR “metabolically benign” [Title/Abstract] OR “metabolic* benign” [Title/Abstract] OR “Obesity, Metabolically Benign”[Mesh] OR "metabolic* healthy"[Title/Abstract] OR "metabolic* abnormal"[Title/Abstract] OR "metabolic* unhealthy"[Title/Abstract] OR "metabolic* phenotype"[Title/Abstract] OR mho[Title/Abstract] OR muo[Title/Abstract] OR muho[Title/Abstract] OR mao[Title/Abstract]) AND (“overweight” [Title/Abstract] OR “overweight” [Mesh] OR “Obese” [Title/Abstract] OR obesity[Title/Abstract] OR “obesity”[Mesh]))
Systematic review-Web of science (N=6308)
(("cardiorespiratory fitness” OR "aerobic capacity” OR "aerobic fitness” OR "physical fitness” OR "functional capacity” OR "physical characteristic*” OR "oxygen consumption” OR "vo2max” OR "vo2peak” OR fitness OR “physical activity” OR “motor activity” OR “motor activities” OR “exercise” OR “exercises” OR “physical exercise” OR “acute exercise” OR “isometric exercise” OR “aerobic exercise” OR “exercise training”) AND ("metabolically healthy" OR "metabolically abnormal" OR "metabolically unhealthy" OR "metabolic profile" OR "metabolic syndrome" OR "metabolic syndrome" OR “insulin resistance syndrome X” or “metabolic syndrome X” OR “dysmetabolic syndrome X” OR “Reaven syndrome X” OR “metabolic cardiovascular syndrome” OR "metabolically phenotype" OR “metabolically benign” OR “metabolic* benign” OR "metabolic* healthy" OR "metabolic* abnormal" OR "metabolic* unhealthy" OR "metabolic* phenotype" OR mho OR muo OR muho OR mao) AND (overweight OR Obese OR obesity))
Figure 1-Search in Pubmed
("metabolically healthy"[Title/Abstract] OR "metabolically abnormal"[Title/Abstract] OR "metabolically unhealthy"[Title/Abstract] OR "metabolic profile"[Title/Abstract] OR "metabolic syndrome"[Title/Abstract] OR "metabolic syndrome"[MeSH] OR "metabolically benign"[Title/Abstract] OR "metabolic* benign"[Title/Abstract] OR

"Obesity, Metabolically Benign"[Mesh] OR "metabolic* phenotype"[Title/Abstract] OR mho[Title/Abstract] OR muo[Title/Abstract] OR muho[Title/Abstract] OR mao[Title/Abstract]) AND ("overweight"[Title/Abstract] OR "overweight"[Mesh] OR "Obese"[Title/Abstract] OR obesity[Title/Abstract] OR "obesity"[Mesh]) AND ("2001/01/01"[PDAT] : "3000/12/31"[PDAT])

Table S2. Exclusion reasons for the studies not included in the systematic review and meta-analysis (N=55).

Author, year	Exclusion reason
Amouzegar et al., 2015	Wrong outcomes
Aparicio et al., 2014	Wrong study type
Arsenault et al., 2009	Wrong study design
Babu et al., 2014	Wrong study type
Barbat-Artigas et al., 2012	Wrong outcomes
Bell et al., 2015	Wrong outcomes
Bjelakovic et al., 2017	Wrong outcomes
Bradshaw et al., 2013	Wrong outcomes
Brandon et al., 2017	Wrong outcomes
Camhi et al., 2015	Wrong study design
Chang et al., 2016	Wrong outcomes
Chang et al., 2012	Wrong outcomes
Choi et al., 2013	Wrong patient population
Conus et al., 2004	Wrong patient population
Corona et al., 2014	Wrong outcomes
Dalleck et al., 2014	Wrong study type
De Lorenzo et al., 2017	Wrong outcomes
Delgado-Floody et al., 2018	Wrong outcomes
Durward et al., 2012	Wrong study type
Fung et al., 2015	Wrong comparator
Gao et al., 2016	Wrong patient population
Gayda et al., 2013	Wrong study type
Gordon-Larsen et al., 2013	Wrong outcomes
Gorostegi-Anduaga et al., 2018	Wrong patient population
Guo et al., 2015	Wrong patient population
Hamer et al., 2015	Wrong outcomes
Hamer et al., 2012	Wrong comparator
Henriques et al., 2015	Wrong outcomes
Jae et al., 2014	Wrong study type
Janiszewski et al., 2010	Wrong outcomes
Kantartzis et al., 2011	Wrong outcomes
Katzmarzyk et al., 2005	Cohort represented more than once
Kelishadi et al., 2010	Wrong outcomes
Khan et al., 2014	Wrong outcomes
Kim et al., 2014	Wrong outcomes
Kimokoti et al., 2014	Wrong outcomes
Lee et al., 2009	Wrong outcomes
Lee et al., 2018	Wrong outcomes
Lopez-Garcia et al., 2013	Wrong study type
Lwow et al., 2011	Wrong outcomes
Manu et al., 2012	Wrong outcomes
Messier et al., 2010	Wrong outcomes
Nogueira et al., 2016	Wrong patient population

Rodriguez-Garcia et al., 2018	Wrong outcomes
Rodriguez-Garcia et al., 2017	Wrong outcomes
Ruhunuhewa et al., 2017	Wrong outcomes
Ryu et al., 2015	Wrong study design
St-Onge et al., 2004	Wrong patient population
Stefan et al., 2008	Wrong outcomes
Taylor et al., 2018	Wrong patient population
Tomiyama et al., 2016	Wrong outcomes
Twig et al., 2014	Wrong comparator
Velho et al., 2010	Wrong study design
Wedell-Neergaard et al., 2018	Wrong outcomes
Yun et al., 2013	Wrong patient population

Table S3. Studies examining the differences between MHO and MUO on physical fitness and PA and sedentary behavior (N=67).

Author, year	Setting Study design Sample size	Age (SD)	Metabolic criteria	Subgroup definition	Outcomes of interest (measurement, *unit*)	Statistical analysis and adjustments	Main findings
Physical fitness (N=19)							
Aparicio et al, 2014	North of Morocco Cross-sectional Convenience sample N = 151 **Subgroups:** N_{MHOO} = 50 N_{MUOO} = 71	52.5 (3.8)	- WC ≥ 88 cm - HDL < 50 mg/dl - TG ≥ 150 mg/dl - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 100 mg/l *Note: in the paper they provided results from another definition. For more information see the paper*	*MHOO*: ≤ 1 of the criteria met (excluding WC) *MUOO*: ≥ 2 of the criteria met (excluding WC)	- CRF (6-min walk test, *m*) - Upper strength (handgrip, *kg*) - Lower strength (30-second chair stand, *no. stands*) - Upper flexibility (back scratch, *cm*) - Lower flexibility (chair sit-and-reach, *cm*) - Static balance (blind flamingo, *no. fails*) - Dynamic balance (8-foot up-and-go, *s*)	ANCOVA adjusted for age	The MHOO group scored better than the MUOO group in CRF, ($p<0.05$), static balance and dynamic balance/agility ($p<0.05$ and $p=0.004$, respectively). No differences between groups were observed in muscular fitness and flexibility (all, $p>0.05$)
Brochu et al, 2001	Not specified Cross-sectional Not specified N = 43 **Subgroups:** N_{MHOO} = 17 N_{MUOO} = 26	58.0 (6.0)	- Glucose disposal rate > 8.0 mg/min·kg	*MHOO*: criterion not met *MUOO*: criterion met	- CRF (VO_2peak in treadmill effort test, *ml/kg·min*) - PAEE (TEE·0.9 – RMR, *Cal/day*)	T-test	MHOO and MUOO presented not significantly different values of VO_2peak and energy expenditure

Cadenas-Sanchez et al, 2017*	*HELENA study* European countries Random sample Cross-sectional N=237 **Subgroups:** N_{MHOO} = 190 N_{MUOO} = 47 N_{MHO} = 24 N_{MUO} = 8	14.7 (1.3)	Age and gender- specific cut-off points by Jolliffe and Janssen (see article for more information)	*MHOO*: ≤1 of the criteria met (excluding WC) *MUOO*: ≥ 2 of the criteria met (excluding WC) *Sensitivity analyses:* *MHOO*: none of the criteria met (excluding WC) *MUOO*: ≥ 1 of the criteria met (excluding WC) *Note: The same criteria were applied when the analyses were done for MHO and MUO*	- CRF (VO_2max in 20m shuttle run test, *ml/kg-min*) - Lower and upper strength (relative handgrip and standing long jump test, *kg and cm respectively*) - Speed-agility (4x10m shuttle run test, *s*)	ANCOVA adjusted for age, sex, socioeconomic status, and maternal education level	No significant differences were observed in fitness in any of the analyses done with overweight/obese and only obese (including also the sensitivity analyses) (p>0.149)
Carvalho et al, 2018	Brazil Cross-sectional Not specified N = 61 **Subgroups:** N_{MHO} = 16 N_{MUO} = 21	33.1 (5.3)	- WC > 102♂ / 88♀ cm - HDL < 1.04♂ / 1.3♀ mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 6.1 mmol/L - HOMA > 3.46	*MHO*: ≤ 2 of the criteria met *MUO*: ≥ 3 of the criteria met	- CRF (VO_2peak in treadmill effort test, *ml/kg·min*) - Lower strength (isometric extensor peak torque and extensor isokinetic total work, *Nm/kg and J/kg*), - Resting energy expenditure (indirect calorimetry, VCO_2/VO_2)	Kruskall-Wallis with Mann-Whitney post-hoc	Both MHO and MUO groups had lower VO_2peak than normal weight participants, and lower muscle strength and endurance, expressed by isometric extensor peak torque and extensor isokinetic total work. All groups presented similar resting energy expenditure

Dalzill et al, 2014	*Montreal Heart Institute* Montreal (Canada) Intervention study *Convenience sample* N=134 **Subgroups:** N_{MHO} = 55 N_{MUO} = 79	52.8 (10.9)	- HDL < 1.0♂ / 1.30♀ mmol/L - TG ≥ 1.70 mmol/L - SBP ≥ 130 mm Hg - DBP ≥ 85 mm Hg - Glucose ≥5.6 mmol/L	*MHO*: ≤ 1 of the criteria met *MUO*: ≥ 2 of the criteria met	-CRF (VO_2peak in incremental treadmill test, *METs*)	Two-way ANOVA	MHO had significantly higher VO_2peak compared with their MUO peers (p<0.01)
Dobson et al, 2016	Liverpool (England) Cross-sectional Convenience sample N = 67 **Subgroups:** N_{MHO} = 21 N_{MUO}= 16	49 (11)	- WC > 102♂ / 88♀ cm - HDL < 1.04♂ / 1.3♀ mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 6.1 mmol/L	*MHO*: ≤ 2 of the criteria met *MUO*: ≥ 3 of the criteria met	- CRF (VO_2max in treadmill effort test, *ml/FFM·min*)	ANOVA	There were no significant differences between the four groups in terms of CRF
Gregorio-Arenas et al, 2016	Granada (Spain) Cross-sectional Random sample N = 228 **Subgroups:** N_{MHO} = 30 N_{MUO} = 40	53 (5)	- HDL < 50 mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 100 mg/dL	*MHO*: ≤ 1 of the criteria met *MUO*: ≥ 2 of the criteria met	- CRF (6-min walk test, *m*) - Upper strength (handgrip, *kg*) - Lower strength (30-second chair stand, *no. stands*) - Upper flexibility (back scratch, *cm*) - Lower flexibility (chair sit-and-reach, *cm*) - Agility (timed up-and-go, *s*).	ANCOVA adjusted for age, smoker, marital status and educational level.	The 6-min walk and the back-scratch tests, for the measurement of CRF and upper flexibility, respectively, presented the most robust differences across metabolic phenotypes (both p<0.001)

Ingle et al, 2017*	England Cross-sectional Random sample N=9117 **Subgroups:** N_{MHO} = 1631 N_{MUO} = 336	48.9 (7.4)	- HDL < 1.036 mmol/L - TG > 1.695 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose > 6.1 mmol/L	*MHO*: ≤1 of the criteria met *MUO*: ≥ 2 of the criteria met	- CRF (VO_2peak in treadmill effort test, *ml/kg·min*)	ANOVA Chi-square	MHO showed higher CRF compared to those with a metabolic unhealthy profile (all p≤0.05)
Jae et al, 2017	Korea (Asia) Cross-sectional Random sample N=3800 **Subgroups:** N_{MHO} = 803 N_{MUO} = 594	47.8 (6.3)	- WC > 90 cm - HDL < 40 mg/dL - TG > 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose > 100 mg/ml	*MHO*: none of the criteria met *MUO*: ≥ 1 of the criteria met	- CRF (VO_2peak in treadmill effort test, *ml/kg·min*)	ANOVA	No significant difference in CRF was observed between MHO and MUO (p>0.05)
Jae et al, 2015	Korea (Asia) Cross-sectional Random sample N=3838 **Subgroups:** N_{MHO} = 958 N_{MUO} = 435	50.8 (5.7)	- WC > 90 cm - HDL < 40 mg/dL - TG > 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose >100 mg/ml	*MHO*: none of the criteria met *MUO*: ≥ 1 of the criteria met	- CRF (VO_2peak in treadmill effort test, *ml/kg·min*)	ANOVA	MHO showed higher CRF than MUO (p<0.001)
Karelis et al, 2005	Montreal (Canada) Cross-sectional Not specified N=44 **Subgroups:** N_{MHO} = 22 N_{MUO} = 22	57.6 (5.9)	- Glucose disposal rates/FFM ≤ 9.29	*MHO*: ≥ 12.62 *MUO*: ≤ 9.29	- CRF (VO_2peak in ergocycle effort test, *ml/kg·min*)	ANOVA	No differences between MHO and MUO were found in CRF (p>0.05)

McElroy et al, 2016*	*LOLA project* Chicago (United States) Cross-sectional Random sample N = 45 **Subgroups:** N_{MHO} = 17 N_{MUO} = 28	53 (Not specified)	- HDL < 40 mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 110 mg/dL	*MHO*: none of the criteria met *MUO*: ≥ 1 of the criteria met	- CRF (VO_2max and 1-min heart rate recovery in treadmill effort test or recumbent bicycle ergometer test, *not specified*) - Strength (single maximal repetition lift, *kg/weight*).	Logistic linear regression	MHO presented better 1-min heart rate recovery after peak exercise performance compared with MUO (p=0.02)
Messier et al, 2008	*Montreal Ottawa New Emerging Team weight loss project* Montreal (Canada) Convenience sample Cross-sectional N=127 **Subgroups:** N_{MHO} = 42 N_{MUO} = 42	57.7 (4.8)	- Hyperinsulinemic-euglycemic clamp < 10.9	*MHO:* belonging to tertile 3rd of HE clamp *MUO:* belonging to tertile 1st of HE clamp	-CRF (VO_2max in cycle ergometer test, *ml/kg·min*) - Lower strength (1-repetitium maximum in maximal leg-press, *kg*)	One-way ANOVA	MHO had significantly higher VO_2max compared with their MUO peers (p=0.028) Relative lower strength was not significant between MHO and MUO (p>0.05)
Messier et al, 2010	Montreal (Canada) Cross-sectional Not specified N=113 **Subgroups:** N_{MHO} = 26 N_{MUO} = 84	58.1(5.0)	*Wildman:* - HDL < 1.3 mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 5.6 mmol/L - HOMA > 5.13 - hs-CRP > 0.1 mg/l *Note: in the paper they provided results from different definitions (i.e.*	MHO: ≤ 1 of the criteria met MUO: ≥ 2 of the criteria met	- CRF (VO_2peak in graded exercise test on an ergocycle, *ml/min·FFM*)	T-test	Using the Clamp categorization, MHO presented higher VO_2peak than those MUO (p<0.05). No significant differences were observed in terms of VO_2peak between phenotypes and different categorizations applied

			Clamp, Matsuda index, HOMA and Karelis). For more information see the paper				
Ortega et al, 2013	ACLS United States Cross-sectional and longitudinal Random sample N=43265 **Subgroups:** N_{MHO} = 1738 N_{MUO} = 3911	44.2 (9.9)	- HDL≤ 40♂ / 50♀ mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥100 mg/dL *Note: in the paper they provided results from obese categorized based on their BF%. For more information see the paper*	*MHO*: ≤ 1 of the criteria met *MUO*: ≥ 2 of the criteria met	- CRF (VO_2max in maximal treadmill exercise test, *ml/kg·min*)	ANCOVA adjusted for age, sex, examination year, smoking, and alcohol consumption	MHO participants had a better baseline fitness level than MUO participants, after adjustment for the set of confounders (p<0.001)
Poelkens et al, 2014*	The Netherlands Cross-sectional Not specified N=20 **Subgroups:** N_{MHO} = 10 N_{MUO} = 10	51 (6)	- WC ≥ 88 cm - HDL < 1.3 mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg	*MHO*: none of the criteria met (excluding WC) *MUO*: ≥ 1 of the criteria met (excluding WC)	- CRF (VO_2max maximal exercise test on ergocycle, *ml/kg·min*)	T-test	MHO showed higher CRF than those MUO (p=0.04)
Senechal et al, 2013*	Canada Cross-sectional Random sample N=108 **Subgroups:** N_{MHOO} = 27 N_{MUOO} = 81	15.2 (1.6)	Age and gender- specific cut-off points by Jolliffe and Janssen (see article for more information)	*MHOO*: none of the criteria met *MUOO*: ≥ 1 of the criteria met	-CRF (VO_2peak in graded maximal cycle ergometer test, *ml/kg·min*)	T-test	CRF were not significantly different between MHOO and MUOO (p>0.05)
Wiklund et al, 2014*	*EWI study* Jyväskylä (Finland) Convenience sample	41.7 (6.9)	- WC ≥ 88 cm - HDL < 1.30 mmol/L - TG ≥ 1.7 mmol/L	*MHOO*: none of the criteria met (excluding WC)	-CRF (VO_2max in maximum bicycle	Not specified	MHOO and MUOO did not differ in CRF (all p>0.05)

	Cross-sectional N= 78 **Subgroups:** N_{MHOO} = 42 N_{MUOO} = 36		- SBP ≥ 130 mm Hg - DBP ≥ 85 mm Hg - Glucose ≥5.6 mmol/L	*MUOO*: ≥ 3 of the criteria met (excluding WC)	ergometer test, *ml/kg·min*)		
Yu et al, 2013*	Hong Kong (China) Random sample Cross-sectional N=518 **Subgroups:** N_{MHOO} = 144 N_{MUOO} = 40	61.1 (3.1)	- WC ≥ 80 cm - HDL < 1.30 mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mm Hg - DBP ≥ 85 mm Hg - Glucose ≥6.1 mmol/L	*MHOO: ≤2* of the criteria met *MUOO:* ≥ 3 of the criteria met	-CRF (VO_2max in maximal exercise ergometer test, *ml/kg·min*)	ANCOVA adjusted for age	MHOO participants showed higher CRF than those with MUOO ($p<0.001$).
Physical activity and sedentary behavior (N=55)							
Aparicio et al, 2016	Badajoz (Spain) Cross-sectional Random sample N = 2833 **Subgroups:** N_{MHOO} = 332 N_{MUOO} = 580	51.2 (14.7)	- WC ≥ 88 cm - HDL < 50 mg/dl - TG ≥ 150 mg/dl - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 100 mg/l	*MHOO*: ≤ 1 of the criteria met (excluding WC) *MUOO*: ≥ 2 of the criteria met (excluding WC)	- Reported PA and energy expenditure (Minnesota Leisure Time Physical Activity Questionnaire, *MET/week*)	ANCOVA adjusted for age, sex, smoker, alcohol consumption, educational level	MHOO participants showed higher light, moderate and vigorous PA than MUOO, but not statistically significant. In regards to energy expenditure, similar patterns were observed with the exception of the activity energy expenditure including housing hold activities for MHOO was significantly higher than for MUOO
Bell et al, 2015	*Whitehall II cohort study* Cross-sectional Convenience sample N = 3457	69.2 (5.6)	- HDL < 1.03♂ / 1.29♀ mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 5.6 mmol/L	*MHO*: ≤ 1 of the criteria met *MUO*: ≥ 2 of the criteria met	- Reported PA (Minnesota Leisure Time Physical Activity Questionnaire, *MET-h/week and*	Linear and logistic regressions adjusted for demographic factors (model	Higher total PA in MHO than in MUO is evident only when measured objectively. Likewise, associations of total PA with the different phenotypes were stronger when measured objectively

	Subgroups: N_{MHO} = 580 N_{MUO} = 580		- HOMA > 5.12		*no. participants above 2.5 h/week of MVPA*) - Objective Total PA (GENEActiv accelerometer, *mg and no. participants above 2.5 h/week of MVPA*)	1) and additionally for socioeconomic position, health behaviors and presence of an illness that limits MVPA	
Bell et al, 2014	*English Longitudinal Study of Ageing* Cross-sectional Random sample N = 4931 **Subgroups:** N_{MHO} = 299 N_{MUO} = 1135	65.1 (8.9)	- HDL < 1.03♂ / 1.29♀ mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - hs-CRP ≥ 3 mg/L - HbA1c ≥ 6%	*MHO*: ≤ 1 of the criteria met *MUO*: ≥ 2 of the criteria met	- Reported television viewing (how many hours during weekdays and weekend days, *h*)	ANCOVA adjusted for age and sex and additionally adjusted for marital status, occupational class, self-reported presence of any long-standing illness which limits activities, limitations in basic and instrumental activities of daily living, depressive symptoms and health behaviors including smoking status.	Results of this study of older adults indicate that a common type of leisure-time sedentary behavior varies across metabolic and obesity phenotypes. However, the higher television viewing time observed in MHO was not significantly different with the observed in MUO
Berezina et al, 2015	Saint Petersburg Cross-sectional Not specified N = 503	45.8 (0.3)	- WC > 94♂ / 80♀ cm - HDL < 1.03♂ / 1.29♀ mmol/L - TG ≥ 1.7 mmol/L	*MHOO*: ≤ 2 of the criteria met	- Reported physical training (*no. sessions and time per session*)	ANOVA	Frequency and duration of physical training were greater in MHOO than MUOO patients ($p < 0.05$)

	Subgroups: N_{MHOO} = 44 N_{MUOO} = 459		- SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 5.6 mmol/L	*MUOO*: ≥ 3 of the criteria met (including WC)			
Bouchard et al, 2011	Not specified Cross-sectional Convenience sample N = 86 **Subgroups:** N_{MHOO} = 18 N_{MUOO} = 68	58.7 (1.3)	- WC > 88 cm - HDL < 1.3 mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 6.1 mmol/L	*MHOO*: ≤ 1 of the criteria met *MUOO*: ≥ 2 of the criteria met	- Reported PA (Physical Activity Scale for the Elderly, *z-score*) - Objective PA (CALTRAC accelerometer, *z-score*)	T-test or Mann-Whitney tests	Although MHOO presented higher mean PA level than MUOO, the difference was no significant
Cadenas-Sanchez et al, 2017*	*HELENA study* European countries Random sample Cross-sectional N=237 **Subgroups:** N_{MHOO} = 190 N_{MUOO} = 47 N_{MHO} = 24 N_{MUO} = 8	14.7 (1.3)	Age and gender- specific cut-off points by Jolliffe and Janssen (see article for more information)	*MHOO*: ≤1 of the criteria met (excluding WC) *MUOO*: ≥ 2 of the criteria met (excluding WC) *Sensitivity analyses:* *MHOO*: none of the criteria met (excluding WC) *MUOO*: ≥ 1 of the criteria met (excluding WC) *Note: The same criteria were applied when the analyses were done for MHO and MUO*	- Objective SB (Accelerometry, *min/day*) - Objective PA (accelerometry, *min/day*)	ANCOVA adjusted for age, sex, socioeconomic status, and maternal education level	MHOO showed lower time spent in SB and higher in MVPA compared to those peers MUOO ($p<0.05$). Overall, these results persisted when only obese adolescents were included in the analyses In sensitivity analyses, MHOO presented lower SB time ($p=0.004$) and borderline non-significant higher MVPA compared to MUOO ($p=0.071$). Taking into account only obese adolescents, no significant results were found neither in SB nor PA ($p>0.149$)

Camhi et al, 2015	Massachusetts (US) Cross-sectional Convenience sample N = 46 **Subgroups:** N_{MHOO} = 37 N_{MUOO} = 9	26.7 (4.7)	- HDL < 50 mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - HOMA > 5.49 - hs-CRP > 14.4 mg/L	*MHOO*: ≤ 2 of the criteria met *MUOO*: ≥ 3 of the criteria met	- Objective SB and PA (ActiGraph GT3X+ accelerometer, *min/day and steps/day*) - Reported SB and PA (7-Day PA recall, *h/week*)	Mixed linear regression models or logistic regressions adjusted for age, race, BMI, smoker and accelerometer wear time	Compared to MUOO, MHOO young women demonstrate healthier lifestyle habits with less sedentary behavior and more time in light PA
Camhi et al, 2013	*NHANES* United States Cross-sectional Random sample N = 766 **Subgroups:** *Adolescents (12-18 y)* N_{MHOO} = 163 N_{MUOO} = 62 *Young adults (19-44 y)* N_{MHOO} = 152 N_{MUOO} = 118 *Older adults (45-85 y)* N_{MHOO} = 64 N_{MUOO} = 207	*Adolescents* 14.7 (0.3) *Young adults* 32.5 (0.7) *Older adults* 56.3 (0.9)	*Adolescents* - HDL < 40 mg/dL - TG ≥ 110 mg/dL - SBP ≥ 90th percentile - DBP ≥ 90th percentile - Glucose > 100 mg/dL *Young and older adults* - HDL < 40♂ / 50♀mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose > 100 mg/dL	*MHOO*: ≤ 1 of the criteria met *MUOO*: ≥ 2 of the criteria met	- Reported PA (Interview, *min/day, MET * min/day and no. training sessions and no. and time in active commuting*) - Reported SB (Interview, *no. participants spending X h/day in sitting, television viewing and computer*)	T-tests Logistic regression adjusted for age, gender, BMI, race/ethnicity, menopausal status and NHANES cycle	Among adolescents, PA was not associated with MHOO. In contrast, MHOO adults 19–44 years were 85% more likely to engage in active transportation and 2.7 times more likely to be involved in light intensity usual daily activity versus sitting. For each minute per day, adults 45–85 years were 36%more likely to have the MHOO phenotype with higher levels of moderate PA. SB was not associated with metabolic phenotypes in adolescents or adults
Chang et al, 2016	Taiwan Cross-sectional Random sample N = 734 **Subgroups:** N_{MHO} = 91 N_{MUO} = 146	56.5 (13.6)	- WC > 90♂ / 80♀ cm - HDL < 1.04♂ / 1.3♀ mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 6.1 mmol/L - HOMA > 3.46	*MHO*: ≤ 2 of the criteria met *MUO*: ≥ 3 of the criteria met	- Regular exerciser (*no. participants exercising more than once a week*)	Chi-squared T-test	Compared with obese participants who had metabolic syndrome, those who were obese but did not have metabolic syndrome, were more likely to exercise regularly

De Rooij et al, 2016	*The Maastricht Study* Cross-sectional study Not specified N = 2449 **Subgroups:** N_{MHO} = 107 N_{MUO} = 440	60.0 (8.1)	- WC > 102♂ / 88♀ cm - HDL < 1.04♂ / 1.3♀ mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 6.1 mmol/L	*MHO*: ≤ 2 of the criteria met *MUO*: ≥ 3 of the criteria met	- Objective SB, standing and stepping time (activPAL acceleromter, *min/day*)	ANCOVA adjusted for age, sex, educational level, smoking, alcohol use and waking time. Additional adjustments for T2DM status, history of CVD and mobility limitation	After adjustments for age, sex, educational level, smoking, alcohol use, waking time, type 2 diabetes mellitus, history of cardiovascular disease and mobility limitation, MHO spent, per day, more time stepping (118.2 versus 105.2 min; p<0.01) and less time sedentary (563.5 versus 593.0 min., p = 0.02) than MUO (n = 440)
De Winter et al, 2018	*NHANES* *United States* Cross-sectional Convenience sample N = 1446 **Subgroups:** *12-18 years:* N_{MHO} = 335 N_{MUO} = 79 *19-44 years:* N_{MHO} = 691 N_{MUO} = 111 *45-85 years:* N_{MHO} = 1082 N_{MUO} = 148	*12-18 years* 14.9 (2.0) *19-44 years* 32.4 (7.5) *45-85 years* 61.9 (10.9)	*12-18 years* Age and gender- specific cut-off points by Jolliffe and Janssen (see article for more information) *19-44 / 45-85 years* - HDL < 1♂ / 1.3♀ mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 5.6 mmol/L	*12-18 years:* *MHO*: ≥ 1 of the criteria met *MUO*: ≥ 2 of the criteria met *19-44 years:* *MHO*: none of the criteria met *MUO*: ≥ 1 of the criteria met *45-85 years:* *MHO*: none of the criteria met *MUO*: ≥ 1 of the criteria met	- Objective PA (ActiGraph 7164 accelerometer, *min/day*)	T-tests Logistic linear regressions	MHO adults displayed a higher 1-minute bout of MVPA per day compared to non-MHO (p = 0 02), but no difference was observed for MVPA and sedentary behavior patterns for youth and older adults. When adjusted for confounders, all bouts of sedentary behavior patterns in youth were significantly associated with being classified as MHO
Ding et al, 2015	*Beijing Child and Adolescent Metabolic Syndrome*	10.0 (1.9)	- WC > 90th percentile for age and sex - HDL < 1.03 mmol/L - TG ≥ 1.24 mmol/L	*MHO*: none of the criteria met *MUO*: ≥ 1 of the criteria met	- Reported PA (frequency and duration, *no. participants*	Chi-square	No significant differences were found between healthy and unhealthy phenotypes in obese participants (p>0.05)

	Cross-sectional and longitudinal study Not specified N = 1149 **Subgroups:** N_{MHO} = 172 N_{MUO} = 297		- SBP ≥ 90th percentile for age and sex - DBP ≥ 90th percentile for age and sex - Glucose ≥ 5.6 mmol/L		*reporting at least exercise for once biweekly and over 30 min per time*)		
Donini et al, 2016	*University of Rome* Cross-sectional Convenience sample N = 253 **Subgroups:** N_{MHO} = 151 N_{MUO} = 102	50.6 (11.9)	- WC > 102♂ / 88♀ cm - HDL < 1.04♂ / 1.3♀ mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 6.1 mmol/L	*MHO*: none of the criteria met *MUO*: ≥ 1 of the criteria met	- Reported PA (IPAQ, *METs-min/week*)	T-test ANCOVA adjusted for age, sex and BMI	No physical activity differences between MHO and MUO was observed (p≥0.360)
Doustmohamadian et al, 2017	*Tehran Lipid and Glucose Study* Cross-sectional and longitudinal study Random sample N = 8804 **Subgroups:** N_{MHOO} = 1125 N_{MUOO} = 3686	47.7 (12.6)	- HDL < 40♂ / 50♀ mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 100 mg/dL - 2-h blood glucose ≥ 140 mg/dL	*MHOO*: ≤ 1 of the criteria met *MUOO*: ≥ 2 of the criteria met Participants were classified as abdominal obese (i.e. ≥ 89 cm for men and ≥ 91 cm for women) healthy and unhealthy	- Reported PA (Lipid research clinic questionnaire, *no. participants reporting light moderate or heavy lifestyle*)	Chi-Square	Between MHOO and MUOO, no significant differences were observed in PA
Elmaogullari et al, 2017	Not specified Cross-sectional study Convenience sample N = 876	11.9 (2.9)	- HDL < 40 mg/dL - TG > 150 mg/dL - LDL > 130 mg/dL - Glucose ≥ 100 mg/dL - TC > 200 mg/Dl	*MHOO*: none of the criteria met *MUOO*: ≥ 1 of the criteria met	- Reported PA (frequency and duration in different behaviors, *no. participants*	Chi-square Logistic linear regression.	A sedentary lifestyle was more common among MUOO children compared to the MHOO children (63.2% vs. 55.1%, respectively); and the negative effects of a sedentary

	Subgroups: N_{MHOO} = 363 N_{MUOO} = 513		- HOMA > 2.5 (prepubertal) > 4 (pubertal) - TSH ≥ 0.34 - Free thyroxine 4 ≥ 0.6 - Hepatosteatosis (not defined the cut-off) - HTN (not defined the cut-off)	Also separated in pre-pubertal and pubertal.	*classified as sedentary, moderate or active*)		lifestyle on metabolic health were found to be statistically significant (p = 0.048).
Gao et al, 2016	*CODING study* *Newfoundland (*Canada*)* Cross-sectional Not specified N=20 **Subgroups:** N_{MHO} = 10 N_{MUO} = 10	46.4 (8.25)	- WC > 102♂ / 88♀ cm - HDL < 1.03♂ / 1.30♀ mmol/L - TG ≥ 150 mg/dL - Glucose ≥ 5.6 mmol/L - HOMA > 4.27	*MHO*: none of the criteria met *MUO*: ≥ 1 of the criteria met	- Reported PA (ARIC-Baecke Questionnaire, *PA level*)	T-test	There was no significant differences between MHO and MUO in physical activity levels (p=0.875)
Goday et al, 2016	*ICARIA study* Spain Cross-sectional *Random sample* N=241420 **Subgroups:** N_{MHO} = 38600 N_{MUO} = 31452	38.7 (10.6)	- WC > 102♂ / 88♀ cm - HDL < 40♂ / 50♀ mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mm Hg - DBP ≥ 85 mm Hg - Glucose ≥100 mg/dL	*MHO*: ≤ 2 of the criteria met *MUO*: ≥ 3 of the criteria met	- Reported PA (Questionnaire on frequency and duration, *no. participants classified as regular exercisers, ≥2h/week of exercise, <2h/week of exercise or no exercise*)	Chi-square	In obese participants, those with a metabolically healthy phenotype showed higher levels of PA than those metabolically unhealthy (p<0.001)
Gutiérrez-Repiso et al, 2014	*The diabet.es study* Spain Cross-sectional Not specified	50.4 (0.2)	- HDL < 40♂ / 50♀ mg/dL - TG ≥ 150 mg/dL - SPB ≥ 130 mmHg - DBP ≥ 85 mmHg	*MHO*: ≤ 1 of the criteria met *MUO*: ≥ 2 of the criteria met	- Reported PA (frequency and duration, *no. participants*	Logistic linear regression	Moderate PA was associated with being metabolically healthy in all the weight groups

	N = 5728 **Subgroups:** N_{MHO} = 338 N_{MUO} = 1089		- Glucose ≥ 100 mg/dL - HOMA > 3.1 (90th percentile) - hs-CRP level > 6.0 mg/L (90th percentile)		*reporting never, once a week, 1-3 times/week or >3 times/week*)		
Hankinson et al, 2013	*The Internations Population Study on Macro/Micronutrients and Blood Pressure* Cross-Sectional Random sample N = 775 **Subgroups:** N_{MHO} = 75 (men) N_{MHO} = 74 (women) N_{MUO} = 323 (men) N_{MUO} = 303 (women)	49.1 (5.2)	- SPB ≥ 120 mmHg - DBP ≥ 80 mmHg - Presence of cardiovascular disease - Physician diagnosis, medication or special diet for diabetes or dyslipidemia	*MHO*: none of the criteria met *MUO*: ≥ 1 of the criteria met	- Reported PA (time sitting or doing light and MVPA, *h/day*)	T-tests of means adjusted for age, race and education.	Diet composition and most activity behaviors were similar between obesity phenotypes. These results do not support hypotheses that diet composition and/or physical activity account for the absence of cardiometabolic abnormalities in metabolically healthy obese
Hashimoto et al, 2017	*Gifu (Japan)* Cross-sectional and longitudinal Random sample N=3290 **Subgroups:** N_{MHO} = 1255 N_{MUO} = 536	45.9 (8.9)	- HDL < 1.03♂ / 1.29♀ mmol/L - TG ≥ 1.70 mmol/L - SBP ≥ 130 mm Hg - DBP ≥ 85 mm Hg - Glucose ≥5.6 mmol/L	*MHO*: ≤ 1 of the criteria met *MUO*: ≥ 2 of the criteria met	- Regular exerciser (*no. participants exercising more than once a week*)	Chi-square	No significant differences in the percentage of regular exercisers were observed between metabolic healthy and unhealthy phenotype (all p>0.05)
Hayes et al, 2010	*The Internations Population Study on Macro/Micronutrients and Blood Pressure* Cross-Sectional Random sample N = 39	46.8 (9.7)	- HDL ≤ 1.1 mmol/L - TG ≥ 2.2 mmol/L - SBP ≥ 160 mmHg - DBP ≥ 95 mmHg - Glucose ≥ 6.1 mmol/L - 2-h blood glucose ≥ 7.8 mmol/L	*MHO*: none of the criteria met *MUO*: ≥ 1 of the criteria met	- Reported PA (Modifiable Activity Questionnaire,) - Objective PA (pedometer, *steps/day*).	T-tests Mann-Whitney tests	A higher proportion of those with a healthy compared to a less healthy metabolic profile met current physical activity guidelines (70% vs 25%). Intra-abdominal fat, insulin resistance and physical activity make independent contributions to

	Subgroups: N_{MHO} = 20 N_{MUO} = 19						metabolic status in very obese women but explain only around a third of the variance
Hayes et al, 2006	*The Newcastle Thousand Families cohort* Cross-sectional and Longitudinal Random sample N = 223 **Subgroups:** N_{MHOO} = 40 (men) N_{MHOO} = 70 (women) N_{MUOO} = 48 (men) N_{MUOO} = 65 (women)	≥50	- TG ≥ 1.7 mmol/L - HDL ≤ 0.9♂ / 1.1♀ mmol/L - SBP ≥ 140 mmHg - DBP ≥ 90 mmHg - Glucose ≥ 6.1 mmol/L - 2-h blood glucose ≥ 7.8 mmol/L	*MHOO*: none of the criteria met *MUOO*: ≥ 1 of the criteria met	- Reported PA (Medical Research Council's National Survey of Health and Development, *no. participants classified as low, medium or high PA*)	Logistic linear regression	After adjusting for BMI, higher levels of physical activity were independently associated with being MHOO in men
Heinzle et al, 2016	*NHANES* *United States* Cross-sectional Random sample N = 223 **Subgroups:***:* N_{MHOO} = 21 (men) N_{MHOO} = 27 (women) N_{MUOO} = 295 (men) N_{MUOO} = 289 (women)	15.1 (0.3)	- HDL: diagnosed Hypoalphalipoproteinemia - TG ≥ 1.47 mmol/L - SBP ≥ 140 mmHg - DBP ≥ 90 mmHg - Glucose ≥ 7.0 mmol/L Note: *in the paper they provided results from another different definition. For more information see the paper.*	*MHOO*: none of the criteria met *MUOO*: ≥ 1 of the criteria met	- Reported PA (questionnaire on frequency and duration engaged in moderate and vigorous activities, *min/day*)	T-tests Logistic linear regressions adjusted for age and socioeconomic status.	Physical activity and inflammation were not associated with MHOO in male and female adolescents (p>0.05)
Hosseinpanah et al, 2011	*TLGS study* Cross-sectional and longitudinal Random sample N = 6215	47.4 (0.2)	-WC > 89♂ / 91♀ cm - HDL < 40♂ / 50♀ mg/dL - TG ≥ 150 mg/dL - SBP ≥ 140 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 100 mg/dL	*MHO*: ≤ 1 of the criteria met (excluding WC) *MUO*: ≥ 2 of the criteria met (excluding WC)	- Reported PA (Lipid research clinic questionnaire, *no. participants who performed LPA, MPA or VPA*)	Chi-square	Participants classified as MHO showed higher percentage of VPA and lower percentages in LPA than those peers MUO

	Subgroups:*:* N_{MHO} = 408 N_{MHO} = 1294						
Hwang et al, 2017	*KNHANES* Korea (Asia) Cross-sectional Random Sample N= 6021 **Subgroups:** N_{MHO} = 309 N_{MUO} = 1781	53.01 (12.9)	- WC > 90♂ / 80♀ cm - HDL < 40♂ / 50♀ mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 100 mg/dL	*MHO*: ≤ 1 of the criteria met *MUO*: ≥ 2 of the criteria met	- Reported PA (Frequency and duration, *no. participants reporting 5 days of 30 min/day of moderate PA or 3 days with 20 min/day of vigorous PA*)	Chi-square	MHO showed higher regular PA than MUO (p=0.038).
Ingle et al, 2017*	England Cross-sectional Random sample N=9117 **Subgroups:** N_{MHO} = 1631 N_{MUO} = 336	48.9 (7.4)	- HDL < 1.036 mmol/L - TG > 1.695 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose > 6.1 mmol/L	*MHO*: ≤1 of the criteria met *MUO*: ≥ 2 of the criteria met	- Reported PA (questionnaire on frequency and duration, *no. participants reporting 150 min/week of MVPA or 75 min/week of vigorous PA*)	ANOVA Chi-square	MHO showed higher PA compared to those with a metabolic unhealthy profile (all p≤0.05)
Jennings et al, 2008	South Africa Cross-sectional Convenience N=120 **Subgroups:** N_{MHO} = 40 N_{MUO} = 82	29.6 (6.4)	- HOMA ≥ 1.95	*MHO:* do not met HOMA criteria *MUO:* met HOMA criteria	- Reported PA (Global Physical Activity Questionnaire, *METs*)	Chi-square	MHO participants did not show any significant difference in PA energy expenditure compare to MUO (all p>0.05)
Kanagasabai et al, 2015	*NHANES* *United States* Cross-sectional	48.8 (0.6)	- WC ≥ 102♂ / 88♀ cm - HDL < 1.04♂ / 1.29♀ mmol/L	*MHO*: ≤ 2of the criteria met	- Reported PA (questionnaire on frequency and	Chi-square	MHO engaged in higher levels of total PA compared to MUO (p<0.05)

	Random sample N=2753 **Subgroups:** N_{MHO} = 816 N_{MUO} = 1937		- TG ≥ 1.69 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 5.6 mmol/l	*MUO*: ≥ 3 of the criteria met	duration, *no. participants classified as inactive, somewhat active or active*)		
Korhonen et al, 2015	*Harmonica project* Harjavalta and Kokemäki (Finland) Cross-sectional Random sample N=2752 **Subgroups:** N_{MHO} = 269 N_{MUO} = 632	59.1 (6.8)	- WC ≥ 94♂ / 80♀ cm - HDL < 1.0♂ / 1.3♀ mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 5.6 mmol/l	*MHO*: none of the criteria met *MUO*: ≥ 1 of the criteria met	- Reported PA (questionnaire on frequency and duration, *no. participants classified as no PA, functional, occasional or regular PA*) - Reported SB (Time spent TV watching, reading, *h/24h*)	Chi-square T-test	MHO exercised more regularly than MUO participants (p = 0.007) while no differences were found in sedentary behaviors (p ≥ 0.14)
Kuzik et al, 2017	*ICAD study* Denmark, Estonia, Portugal and United States Random sample Cross-sectional N=1039 **Subgroups:** N_{MHO} = 394 N_{MUO} = 258	11.8 (2.8)	- HDL ≤ 1.17 mmol/L - TG ≥ 0.85 (age≤9) / 1.02 (age>9) mmol/L - SBP ≥ 90th percentile - DBP ≥ 90th percentile - Glucose ≥5.56 mmol/L - HOMA ≥ 2.22♀ (age ≤10), 2.67♂ (age ≤11), 3.82♀ (age >10), and 5.22♂ (age >11)	*MHO*: none of the criteria met *MUO*: ≥ 1 of the criteria met	- Objective PA (ActiGraph accelerometer, *min/day*) - Objective SB (ActiGraph accelerometer, *min/day*)	Logistic regression adjusted for age, sex, study and accelerometer wear time	In sedentary time, no significant differences were observed between groups (p>0.05) Each additional 10 min of MVPA was associated with lower odds of MUO compared with their metabolic healthy group (p≤0.05)
Lassale et al, 2018	*EPIC-CVD study* 10 European countries Cross-sectional and longitudinal	52.4 (not specified)	- WC ≥ 94♂ / 80♀ cm - HDL < 40♂ / 50♀ mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg	*MHO*: ≤2 of the criteria met *MUO*: ≥ 3 of the criteria met	- Reported PA (Cambridge index, *no. participants classified as*	ANCOVA adjusted for center, sex, age, education,	MUO were more moderately inactive than MHO (p=0.04). Likewise, MHO showed higher activity than MUO (p=0.07)

	Random sample N=10474 **Subgroups:** N_{MHO} = 751 N_{MUO} = 909		- DBP ≥ 85 mmHg - Glucose ≥100 mg/dL		*inactive, moderately inactive, moderately active and active*)	smoking status, and energy intake	
Lee et al, 2013	*KoGES study* Korea (Asia) Cross-sectional and longitudinal Random sample N=2352 **Subgroups:** N_{MHO} = 135 N_{MUO} = 610	48.7 (7.5)	- WC ≥ 90♂ / 80♀ cm - HDL < 40♂ / 50♀ mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥100 mg/dL	*MHO*: none of the criteria met *MUO*: ≥ 1 of the criteria met	- Reported PA (not specified the method, METs h/day)	ANOVA	No significant differences were observed between metabolic healthy phenotypes compared to those peers metabolically unhealthy (p=0.820)
Lim et al, 2018	*KNHANES study* Korea (Asia) Random Sample Cross-sectional N= 3650 **Subgroups:** N_{MHOO} = 451 N_{MUOO} = 247	14.8 (2.0)	- WC ≥ 90th percentile - HDL ≤ 40 mg/dL - TG ≥ 110 mg/dL - SBP ≥ 90th percentile - DBP ≥ 90th percentile - Glucose ≥100 mg/dL	*MHOO*: none of the criteria met *MUOO*: ≥ 1 of the criteria met	- Reported PA (questionnaire on frequency and duration, *no. days per week of MVPA*)	ANOVA	MHOO presented higher physical activity levels than MUOO (p=0.015)
Matta et al, 2016	*National Nutrition and Non-Communicable Disease Risk Factor Survey* Cross-sectional Random sample N = 196 **Subgroup:** N_{MHOO} = 73	41.4 (14.7)	- HDL ≤ 1.03♂ / 1.30♀ mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥5.6 mmol/L	*MHOO*: ≥ 1 of the criteria met *MUOO*: ≥ 2 of the criteria met	- Reported PA (IPAQ, *no. participants classified as low, medium or high PA*)	Chi-square Logistic linear regression	MHOO presented higher odds of high PA in comparison to MUOO (p<0.001)

	N_{MUOO} = 123						
McElroy et al, 2016*	*LOLA project* Chicago (United States) Cross-sectional Random sample N = 45 **Subgroups:** N_{MHO} = 17 N_{MUO} = 28	53 (Not specified)	- HDL < 40 mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 110 mg/dL	*MHO*: none of the criteria met *MUO*: ≥ 1 of the criteria met	- Reported PA (IPAQ, *min/day*)	Logistic linear regression	There was no significant difference in meeting the PA guidelines between MHO and MUO (p=0.80)
Moon et al, 2017	*KoGES study* Korea (Asia) Cross-sectional and longitudinal Random sample N=8144 **Subgroups:** N_{MHO} = 1262 N_{MUO} = 1006	50.6 (not specified)	- HDL ≤ 1.03♂ / 1.30♀ mmol/L - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 100 mg/dL	*MHO*: ≤ 1 of the criteria met *MUO*: ≥ 2 of the criteria met	- Reported PA (questionnaire on frequency and duration, *no. participants reporting at least 30 min/day*)	Chi-square	MHO presented higher physical activity engagement than those peers with a metabolically unhealthy profile (all p<0.001)
Oh et al, 2014	*National Health Insurance Corporation* Korea (Asia) Cross-sectional and longitudinal Random sample N=363881 **Subgroups:** N_{MHO} = 342442 N_{MUO} = 21439	≥20	- Glucose ≥ 126 mg/dL	*MHO*: criterion not met *MUO*: criterion met	- Reported PA (questionnaire on frequency and duration, *times/week*)	Chi-square	MHO showed higher PA levels than those peers MUO (p<0.001)

Park et al, 2016	*NHANES* *United States* Cross-sectional and longitudinal Random sample N=1739 **Subgroups:** N_{MHO} = 598 N_{MUO} = 1141	44.9 (0.7)	- HDL < 40♂ / 50♀ mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 100 mg/Dl - HOMA > 90th percentile - hs-CRP > 90th percentile	*MHO*: ≤ 1of the criteria met *MUO*: ≥ 2 of the criteria met	- Reported PA (IPAQ, *no. participants classified as meeting and not meeting PA recommendations; and inactivity)*	Chi-square	Participants categorized as MHO did not show any significant difference in meeting or not meeting the PA recommendations and inactivity in comparison with those peers MUO (p=0.60).
Phillips et al, 2013	*The Cork and Kerry Diabetes and Heart Disease Study* Cork (Irland) Cross-sectional Random sample N=2047 **Subgroups:** N_{MHO} = 158 N_{MUO} = 512	60.1 (0.3)	- HDL < 1.04♂ / 1.30♀ mmol/L - TG ≥ 1.70 mmol/L - SBP ≥ 130 mm Hg - DBP ≥ 85 mm Hg - Glucose ≥ 5.55 mmol/L - HOMA > 90th percentile - CRP > 90th percentile	*MHO*: ≤ 1 of the criteria met *MUO*: ≥ 2 of the criteria met	- Reported PA (IPAQ, *no. participants classified as low, moderate, and high PA; meeting PA recommendations; and total PA in min/day*)	Chi-square T-test	MHO did not show any significant difference in the PA outcomes examined, including the total PA performed during the day compared to those peers MUO (all p>0.05)
Poelkens et al, 2014*	The Netherlands Cross-sectional Not specified N=20 **Subgroups:** N_{MHO} = 10 N_{MUO} = 10	51 (6)	- WC ≥ 88 cm - HDL < 1.3 mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg	*MHO*: none of the criteria met (excluding WC) *MUO*: ≥ 1 of the criteria met (excluding WC)	- Objective PA (SenseWear Armband, *METs and steps/day*)	T-test	Borderline non-significant between MHO and MUO in favor of a healthy phenotype was observed in the number of steps (p=0.06)
Prince et al, 2014	Canada Cross-sectional Random sample N=181	12.9 (2.7)	- HDL < 1.02♂ / 1.29♀ mmol/L - TG ≥ 1.25 mmol/L	*MHO*: none of the criteria met *MUO*: ≥ 1 of the criteria met	- Reported PA (questionnaire on frequency and duration in	T-test	MHO presented higher moderate PA, hard PA, and MVPA and lower computer time compared to MUO (all p≤0.04)

	Subgroups: N_{MHO} = 39 N_{MUO} = 142		- SBP ≥ 90th percentile for age, sex and height - DBP ≥ 90th percentile for age, sex and height -Glucose ≥ 5.6 mmol/L. *Note: in the paper they provided results from another definition (i.e. HOMA). For more information see the paper.*		moderate, hard and very hard PA, *min/day*) - Objective PA (pedometer, *steps/day*) - Reported SB (questionnaire on television viewing, computer time, video game time and total screen time, *min/day*)		
Saghafi-Asl et al, 2017	Iran Cross-sectional Random sample N=164 **Subgroups:** N_{MHO} = 83 N_{MUO} = 81	37.7 (8.6)	-WC ≥ 95 cm - HDL≤ 40♂ / 50♀ mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥100 mg/dL	*MHO*: ≤ 2 of the criteria met (including WC) *MUO*: ≥ 3 of the criteria met (including WC)	- Reported PA (IPAQ, *PA score*)	T-test	No significant differences between MHO and MUO in the PA score
Schroder et al, 2014	Girona (Spain) Cross-sectional and longitudinal Random sample N=1445 **Subgroups:** N_{MHOO} = 301 N_{MUOO} = 1144	51.6 (11.8)	- HDL≤ 40♂ / 50♀ mg/dL - TG ≥ 200 mg/dL - SBP ≥ 140 mmHg - DBP ≥ 90 mmHg - Glucose >125 mg/dL - TC ≥ 240 mg/dL - LDL ≥ 160 mg/dL	*MHOO*: none of the criteria met *MUOO*: ≥ 1 of the criteria met	- Reported PA (Minnesota Leisure time PA questionnaire, *METs·min/day*)	T-test	No significant difference were observed between MHOO and MAOO in leisure time PA (p=0.902)
Senechal et al, 2013*	Canada Cross-sectional Random sample N=108	15.2 (1.6)	Age and gender- specific cut-off points by Jolliffe and Janssen (see article for more information)	*MHOO*: none of the criteria met *MUOO*: ≥ 1 of the criteria met	- Objective PA (pedometer, *steps/day*)	T-test	PA were not significantly different between MHOO and MUOO (p>0.05)

	Subgroups: N_{MHOO} = 27 N_{MUOO} = 81						
Slagter et al, 2018	Lifelines study The Netherlands Cross-sectional Random sample N=9270 **Subgroups:** N_{MHO} = 1774 N_{MUO-1} = 3479 N_{MUO-2} = 4017	48.6 (9.3)	- HDL < 40 mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 100 mg/dL	*MHO*: none of the criteria met (excluding WC) *MUO-1:* 1 of the criteria met (excluding WC) *MUO-2*: ≥ 2 of the criteria met (excluding WC)	- Reported PA (SQUASH questionnaire, *min/week*)	T-test	MHO were more moderate-vigorous physically active than participants with MUO (p≤0.01)
Song et al, 2007	Women's Health study United States Cross-sectional and longitudinal Random sample N=25626 **Subgroups:** N_{MHO} = 2925 N_{MUO} = 1341	54.3 (6.5)	- HDL < 50 mg/dL - TG ≥ 150 mg/dL - SBP ≥ 135 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 110 mg/dL	*MHO*: ≤ 2 of the criteria met *MUO:* ≥ 3 of the criteria met	- Reported PA (questionnaire, *kcal/week*)	Not specified	MHO presented higher PA than MUO (919 vs. 781 kcal/week, p value not given)
Sung et al, 2015	Korea (Asia) Cross-sectional and longitudinal Random sample N=275867 N_{men}=156252 N_{women}=119615 **Subgroups:**	40.2 (10)	- SBP ≥ 140 mmHg - DBP ≥ 90 mmHg - Glucose ≥ 126 mg/Dl	*MHO*: none of the criteria met *MUO*: ≥ 1 of the criteria met	- Regular exerciser (*no. participants exercising more than once a week*)	Chi-square	MHO showed higher levels of regular exercise than those peers MUO (all p<0.001)

	Men: N_{MHO} = 12731 N_{MUO} = 49269 *Women:* **Subgroups:** N_{MHO}= 5730 N_{MUO} = 15906						
Van der Ar et al, 2014	*The Dutch EPIC-MORGEN study* Amsterdam, Maastrich, Doetinchem (The Netherlands) Cross-sectional and longitudinal Random samples N=20299 **Subgroups:** N_{MHOO} = 7660 (non-abdominal obese) N_{MHOO} = 895 (abdominal obese) N_{MUOO} = 8060 (non-abdominal obese) N_{MUOO} = 3684 (abdominal obese)	42.8 (10.5)	- WC ≥ 102♂ / 88♀ cm - HDL < 1.0♂ / 1.3♀ mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 5.6 mmol/L - TC ≥ 6.5 mmol/L	*Abdominal obese:* *MHOO*: none of the criteria met (including WC) *MUOO*: ≥ 1 of the criteria met (including WC) *Nonabdominal obese:* *MHOO*: none of the criteria met (excluding WC) *MUOO*: ≥ 1 of the criteria met (excluding WC)	- Reported PA (Cambridge PA Index, *no. participants classified as inactive, moderately inactive, moderately active and active*)	Cochran-Mantel-Haenszel chi-square statistics adjusted for age and sex	MHOO with abdominal obesity was more physically active than those peers metabolically unhealthy (with abdominal obesity). The same pattern is presented in those categorized as metabolically healthy and nonabdominal obese (all p<0.001)
Wiklund et al, 2014*	*EWI study* Jyväskylä (Finland) Convenience sample Cross-sectional N= 78	41.7 (6.9)	- WC ≥ 88 cm - HDL < 1.30 mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mm Hg - DBP ≥ 85 mm Hg - Glucose ≥5.6 mmol/L	*MHOO*: none of the criteria met (excluding WC) *MUOO*: ≥ 3 of the criteria met (excluding WC)	-Reported PA (questionnaire)	Not specified	MHOO and MUOO did not differ in PA (all p>0.05)

	Subgroups: N_{MHOO} = 42 N_{MUOO} = 36						
Wildman et al, 2008	United States Cross-sectional Random sample N = 5440 **Subgroups:** N_{MHO} = 528 N_{MUO} = 1137	45.0 (0.4)	- HDL < 40♂ / 50♀ mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 100 mg/dL - HOMA >5.13 - hs-CRP >0.1 mg/L	*MHO*: ≤ 1 of the criteria met *MUO*: ≥ 2 of the criteria met	- Reported PA (questionnaire on frequency and duration, *no. participants classified as 0, 1-49-9, 50-131.9, 132-279.9, ≥280 METs/d*)	Chi-square	No significant differences were found between MHO and MUO (all p>0.05).
Yoon et al, 2017	Korea (Asia) Cross-sectional Random sample N = 530 **Subgroups:** N_{MHO} = 197 N_{MUO} = 333	14.9 (0.2)	- HDL < 40♂ / 50♀ mg/dL - TG ≥ 150 mg/dL - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 100 mg/dL	*MHO*: none criteria met *MUO*: ≥ 1 of the criteria met	- Reported PA (questionnaire on frequency and duration in moderate, vigorous PA and walking, *min/day*)	Not specified	No significant differences were observed between MHO and MUO in PA (all p≥0.077)
Yu et al, 2013*	Hong Kong (China) Random sample Cross-sectional N=518 **Subgroups:** N_{MHOO} = 144 N_{MUOO} = 40	61.1 (3.1)	- WC ≥ 80 cm - HDL < 1.30 mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mm Hg - DBP ≥ 85 mm Hg - Glucose ≥6.1 mmol/L	*MHOO: ≤2* of the criteria met *MUOO:* ≥ 3 of the criteria met	- Reported PA (ARIC-Baecke Questionnaire, *PA level*)	ANCOVA adjusted for age	No significant difference was found between groups in PA (p=0.364)
Zhang et al, 2017	Liaoning (China) Cross-sectional Random sample N = 10804 **Subgroups:**	53.8 (10.3)	- WC ≥ 90♂ / 80♀ cm - HDL < 1.0♂ / 1.3♀ mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg	*MHOO*: none criteria met *MUOO*: ≥ 1 of the criteria met	- Reported PA (questionnaire on frequency and duration, *no. participants classified as light,*	Chi-square	Compared to MHOO participants, MUOO had significantly higher PA (p<0.001)

	N_{MHOO} = 112 N_{MUOO} = 4712		- Glucose ≥ 5.6 mmol/L		*moderate and severe PA*)		
Zhang et al, 2017	Liaoning Province (China) Cross-sectional Random sample N=2037 **Subgroups:** N_{MHO} = 470 N_{MUO} = 1567	52.9 (10.0)	- WC > 90♂ / 80♀ cm - HDL < 1.0♂ / 1.30♀ mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mm Hg - DBP ≥ 85 mm Hg - Glucose ≥5.6 mmol/L	*MHO*: ≤ 1 of the criteria met *MUO*: ≥ 2 of the criteria met	- Reported PA (questionnaire, *no. participants reporting low, moderate or high leisure PA*)	Chi-square	MHO showed higher levels of PA compared to MUO peers (p<0.001)
Zheng et al, 2015	Shanghai, Hangzhou, Beijing, Shenyang, Taiyuan, Chengdu and Guangzhou (China) Cross-sectional Random sample N = 5013 **Subgroups:** N_{MHO} = 196 N_{MUO} = 506	35-72	- HDL < 1.03♂ / 1.29♀ mmol/L - TG ≥ 1.7 mmol/L - SBP ≥ 130 mmHg - DBP ≥ 85 mmHg - Glucose ≥ 5.6 mmol/L	*MHO*: ≤ 1 of the criteria met (excluding WC) *MUO*: ≥ 2 of the criteria met (excluding WC)	- Reported PA (IPAQ, *no. participants classified as short or moderate-high PA*) - Reported SB (IPAQ, *no. participants classified as short or moderate-high SB*)	Logistic linear regression adjusted for gender, age and other factors	Obese individuals with higher moderate or long SB (p=0.039) and lower moderate or long PA (p=0.002) were significantly associated with metabolic abnormalities

♂: males, ♀: females, WC: waist circumference, TG: triglycerides, HDL: high-density lipoprotein cholesterol, SBP: systolic blood pressure, DBP: diastolic blood pressure, HTN: hypertension, MHOO: metabolically healthy overweight/obese, MUOO: metabolically unhealthy overweight/obese, MHO: metabolically healthy obese, MUO: metabolically unhealthy obese, CRF: cardiorespiratory fitness, PA: physical activity, LPA: light physical activity, MPA: moderate physical activity, VPA: vigorous physical activity, SB: sedentary, Cal: calories, PAEE: physical activity energy expenditure, HOMA: homeostatic model assessment, hs-CRP: high sensitivity c reactive protein, TSH: thyroid-stimulating hormone, MET: metabolic equivalent, TC: total cholesterol, VO_2max: maximum oxygen consumption, VO_2peak: peak oxygen consumption, BF%: body fat percentage, FFM: fat free mass, IPAQ: international physical activity questionnaire; ANCOVA: analysis of covariance, ANOVA: analysis of variance.

*Articles (N=7) included both in physical activity/sedentary behavior and physical fitness sections.

Notes:

-N sample include all participants in the study (not only those stated in the subgroups), thus, in some cases the first sample size provided in the setting-study design column is higher than the sum of those stated in the subgroups subsection.

-When an article presented the data based on different metabolic phenotypes cut-offs, we have shown only that criterion that is closer to the last proposal provided by Ortega et al, 2017.

Table S4. Risk of bias of cross-sectional studies included (N=67).

	Inclusion	Setting	Exposure valid and reliable	Objective	Confounding	Confounding strategies	Outcomes valid and reliable	Appropriate statistical analysis	Number of 'yes'
Aparicio et al, 2014	1	1	1	1	1	0	1	1	7
Aparicio et al, 2016	1	1	0	1	1	0	1	1	6
Bell et al, 2014	1	1	0	1	1	0	1	1	6
Bell et al, 2015	1	1	1	1	1	1	1	1	8
Berezina et al, 2015	1	1	0	1	0	0	1	1	5
Bouchard et al, 2011	1	0	1	1	1	0	1	1	6
Brochu et al, 2001	1	0	1	1	0	0	1	Unclear	4
Cadenas-Sanchez et al, 2017*	1	1	1	1	1	1	1	1	8
Camhi et al, 2013	1	1	0	1	1	0	1	1	6
Camhi et al, 2015	1	1	1	1	1	0	1	1	7
Carvalho et al, 2018	1	0	1	1	1	0	1	1	6
Chang et al, 2016	1	1	0	1	1	0	1	1	6
Dalzill et al, 2014	1	0	1	1	0	0	1	1	5
De Rooij et al, 2016	1	1	1	1	1	0	1	1	7
De Winter et al, 2018	1	1	1	1	1	0	1	1	7
Ding et al, 2015	1	1	0	1	1	0	1	1	6
Dobson et al, 2016	1	0	0	1	1	1	1	1	6
Donini et al, 2016	1	1	1	1	1	0	0	1	6
Doustmohamadian et al, 2017	1	1	1	1	1	0	1	1	7
Elmaogullari et al, 2017	1	1	1	1	1	0	0	1	6

Gao et al, 2016	1	1	1	1	0	Unclear	0	1	5
Goday et al, 2016	1	1	1	1	0	0	0	1	5
Gregorio-Arenas et al, 2016	1	1	1	1	1	0	1	1	7
Gutiérrez-Repiso et al, 2014	1	1	1	1	1	0	0	1	6
Hankinson et al, 2013	1	1	1	1	1	0	0	1	6
Hashimoto et al, 2017	1	1	1	1	1	1	0	1	7
Hayes et al, 2006	1	1	1	1	1	0	0	1	6
Hayes et al, 2010	0	0	1	1	1	0	0	1	4
Heinzle et al, 2016	1	1	1	1	1	1	0	1	7
Hosseinpanah et al, 2011	1	1	1	1	1	0	1	1	7
Hwang et al, 2017	1	1	1	1	1	0	0	1	6
Ingle et al, 2017*	0	0	0	1	1	1	1	1	5
Jae et al, 2015	1	1	1	1	1	0	1	1	7
Jae et al, 2017	1	1	1	1	1	0	1	1	7
Jennings et al, 2008	1	0	1	1	1	0	0	1	5
Kanagasabai et al, 2015	1	1	1	1	1	0	0	1	6
Karelis et al, 2005	1	1	1	1	0	0	1	1	6
Korhonen et al, 2015	1	1	1	1	0	0	0	1	5
Kuzik et al, 2017	1	1	1	1	1	0	1	1	7
Lassale et al, 2018	1	1	1	1	1	0	1	1	7
Lee et al, 2013	1	1	1	1	1	0	0	1	6
Lim et al, 2018	1	1	1	1	1	0	0	1	6
Matta et al, 2016	1	1	1	1	1	1	0	1	7
McElroy et al, 2016*	1	1	0	1	0	0	1	1	5

Messier et al, 2008	1	1	1	1	1	0	1	1	7
Messier et al, 2010	1	1	1	1	0	0	1	1	6
Moon et al, 2017	1	1	1	1	1	0	1	1	7
Oh et al, 2014	1	1	1	1	1	0	0	1	6
Ortega et al, 2013	1	1	1	1	1	0	1	1	7
Park et al, 2016	1	1	1	1	1	0	0	1	6
Phillips et al, 2013	1	1	1	1	1	0	0	1	6
Poelkens et al, 2014*	0	0	1	1	0	0	1	1	4
Prince et al, 2014	1	1	1	1	1	0	1	1	7
Saghafi-Asl et al, 2017	1	1	1	1	0	0	0	1	5
Schroder et al, 2014	1	1	1	1	1	0	0	1	6
Senechal et al, 2013*	1	1	1	1	1	0	1	1	7
Slagter et al, 2018	1	1	1	1	1	0	0	1	6
Song et al, 2007	1	1	1	1	1	0	1	1	7
Sung et al, 2015	1	1	1	1	1	0	1	1	7
Van der A et al, 2014	1	1	1	1	1	0	1	1	7
Wiklund et al, 2014*	1	1	1	1	1	1	1	1	8
Wildman et al, 2008	1	1	1	1	1	0	0	1	6
Yoon et al, 2017	1	1	1	1	1	0	0	1	6
Yu et al, 2013*	1	1	0	1	1	0	1	1	6
Zhang et al, 2017a	1	1	1	1	1	0	0	1	6
Zhang et al, 2017b	1	1	1	1	1	0	0	1	6
Zheng et al, 2015	1	1	1	1	1	0	0	1	6

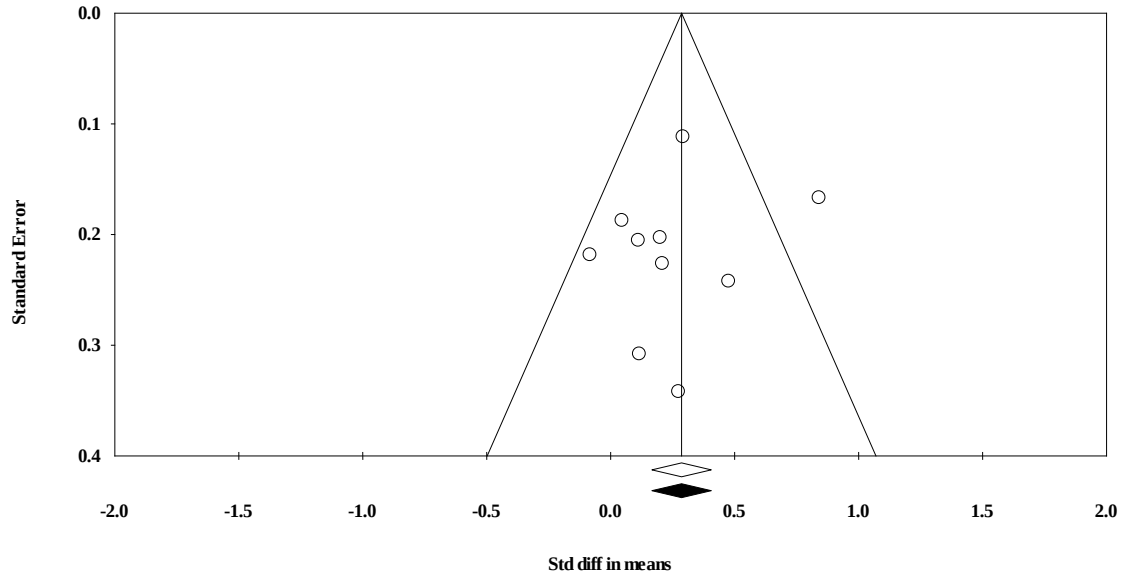
The critical appraisal checklist for analytical cross-sectional studies from The Joanna Briggs Institute can be found in Supplementary material 1.

Table S5. Risk of bias of longitudinal studies included (N=11).

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Number of stars
Appleton et al, 2013	A*	A*	C	A*	A*/ B*	B*	A*	C	7
Bo et al, 2012	A*	A*	C	B	A*/ B*	B*	A*	A*	7
Doustmohamadian et al, 2017	A*	A*	C	B	A*/ B*	B*	A*	B*	7
Hosseinpanah et al, 2011	A*	A*	C	B	A*/ B*	Unclear	A*	A*	6
Lassale et al, 2017	A*	A*	C	A*	A*/ B*	B*	A*	B*	8
Loprinzi et al, 2017	A*	A*	C	B	A*/ B*	B*	A*	A*	7
Moon et al, 2017	A*	A*	C	A*	A*/ B*	C	A*	D	6
Ortega et al, 2013	A*	A*	A*	A*	A*/ B*	B*	A*	C	8
Song et al, 2007	A*	A*	C	B	A*/ B*	B*	A*	A*	7
Sung et al, 2015	A*	A*	C	A*	A*/ B*	B*	Unclear	A*	7
Van der A et al, 2014	A*	A*	C	B	A*/ B*	B*	A*	A*	7

The Newcastle Ottawa Scale for cohort studies can be found in Supplementary material 2. *Indicates that the study gets a star in the item examined.

A) Moderate-to-vigorous physical activity



B) Sedentary behavior

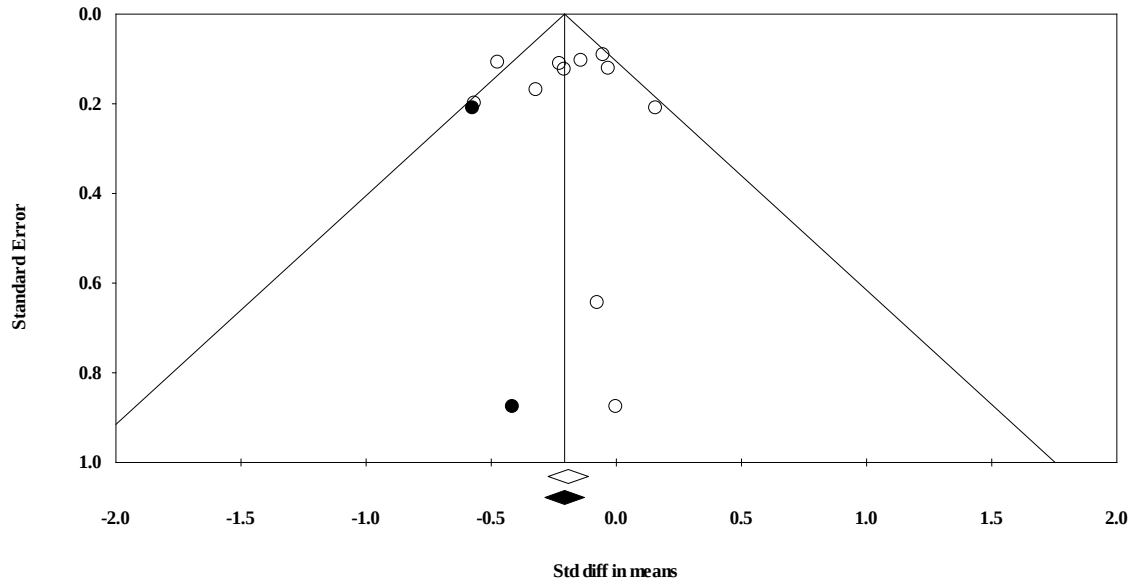


Figure S1. Funnel plots to assess publication bias on moderate-to-vigorous physical activity (Panel A) and sedentary behavior (Panel B) differences (pooled standardized mean difference) between metabolic phenotypes in overweight/obese participants.

White point represents meta-analyzed observed studies whilst the black point shows the imputed studies. Diagonal lines represent pseudo-95% confidence intervals. In reference of Y axis, studies located at the lower part of the graph have a higher standard error (a lower weight in the pooled analysis). The vertical line represents the calculated standardized mean difference.

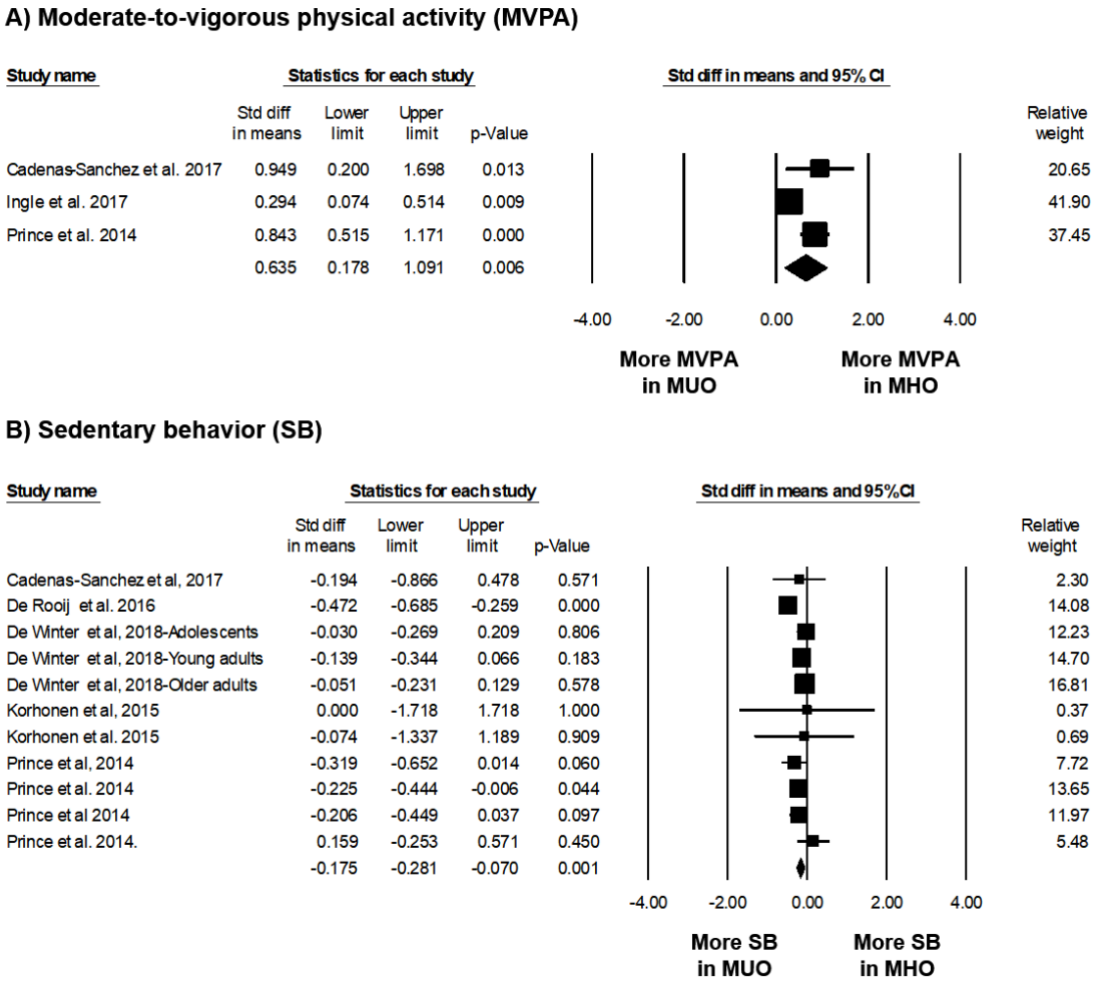
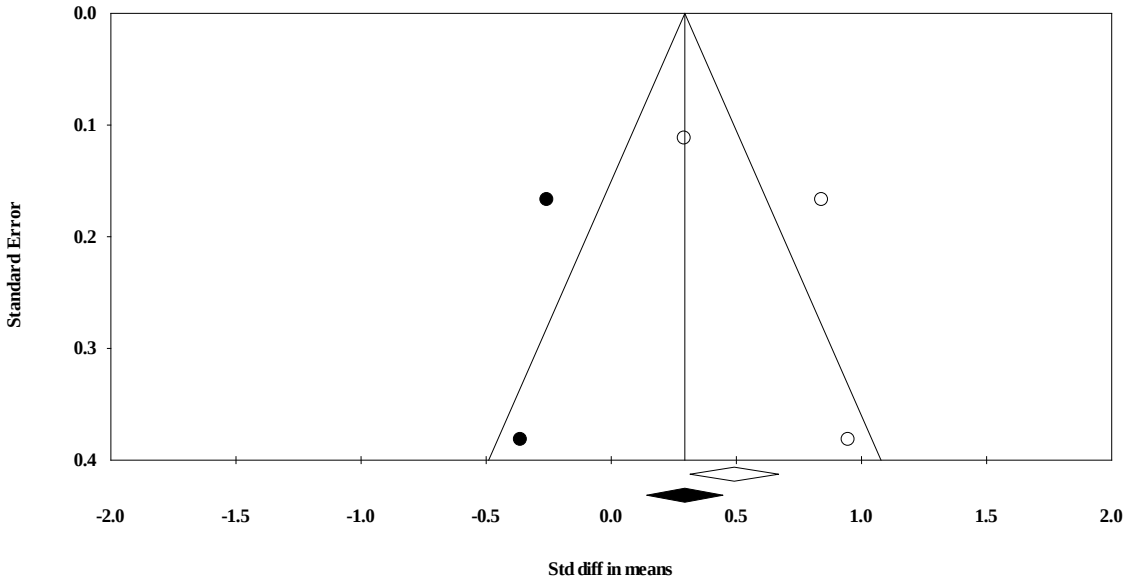


Figure S2. Forest plot of moderate-to-vigorous physical activity (Panel A) and sedentary behavior (Panel B) differences (pooled standardized mean difference) between metabolic phenotypes in obese participants.

Standardized mean differences are expressed as MHO *minus* MUO, thus, positive values represents that MHO showed higher values than MUO. MHO: metabolically healthy obesity. MUO: Metabolically unhealthy obesity. CI: confidence interval. MVPA: moderate-to-vigorous physical activity. SB: sedentary behavior.

A) Moderate-to-vigorous physical activity



B) Sedentary behavior

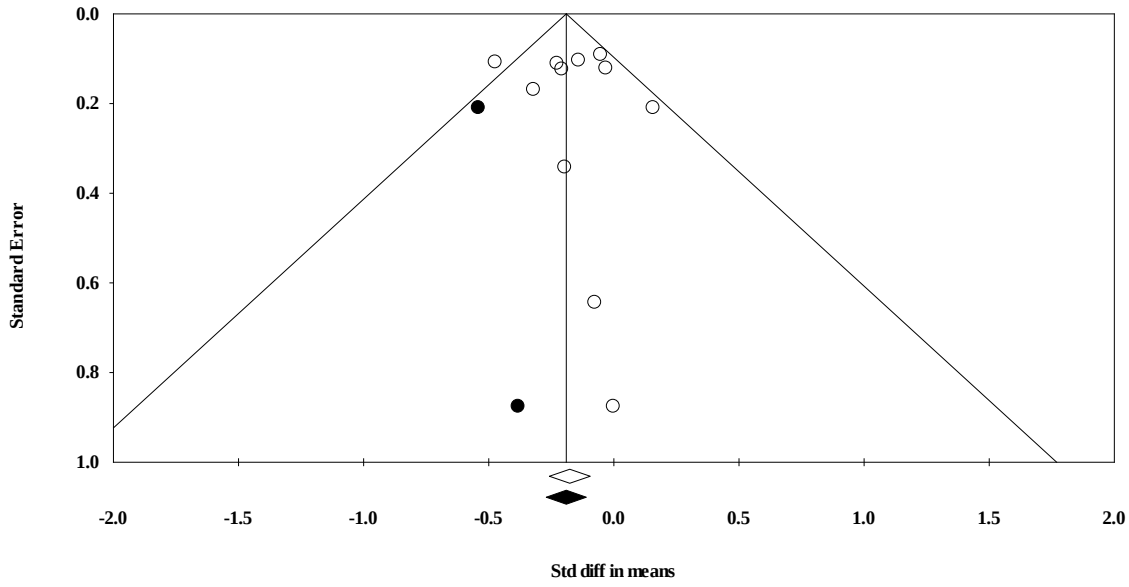
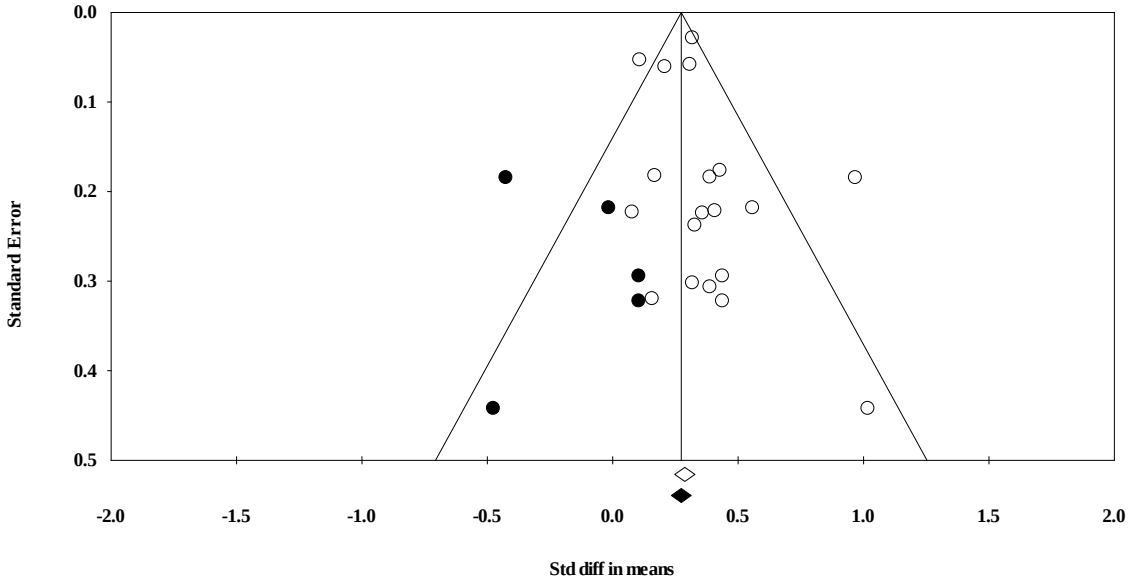


Figure S3. Funnel plots to assess publication bias on moderate-to-vigorous physical activity (Panel A) and sedentary behavior (Panel B) differences (pooled standardized mean difference) between metabolic phenotypes in obese participants.

White point represents meta-analyzed observed studies whilst the black point shows the imputed studies. Diagonal lines represent pseudo-95% confidence intervals. In reference of Y axis, studies located at the lower part of the graph have a higher standard error (a lower weight in the pooled analysis). The vertical line represents the calculated standardized mean difference.

A) Cardiorespiratory fitness



B) Muscular strength

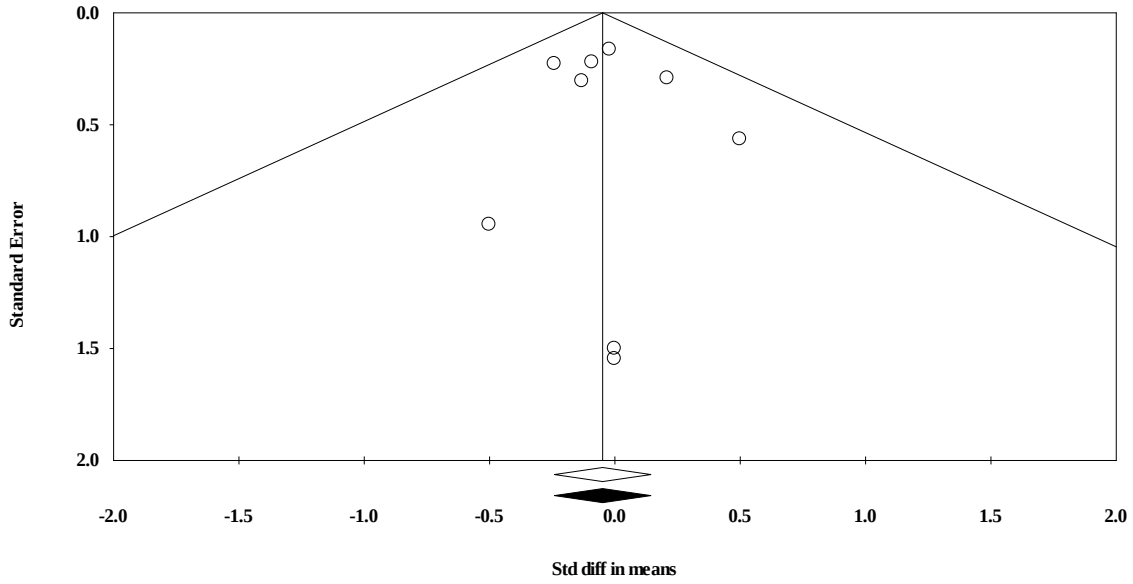


Figure S4. Funnel plots to assess publication bias on cardiorespiratory fitness (Panel A) and muscular strength (Panel B) differences (pooled standardized mean differences) between metabolic phenotypes in overweight/obese participants.

White point represents meta-analyzed observed studies whilst the black point shows the imputed studies. Diagonal lines represent pseudo-95% confidence intervals. In reference of Y axis, studies located at the lower part of the graph have a higher standard error (a lower weight in the pooled analysis). The vertical line represents the calculated standardized mean difference.

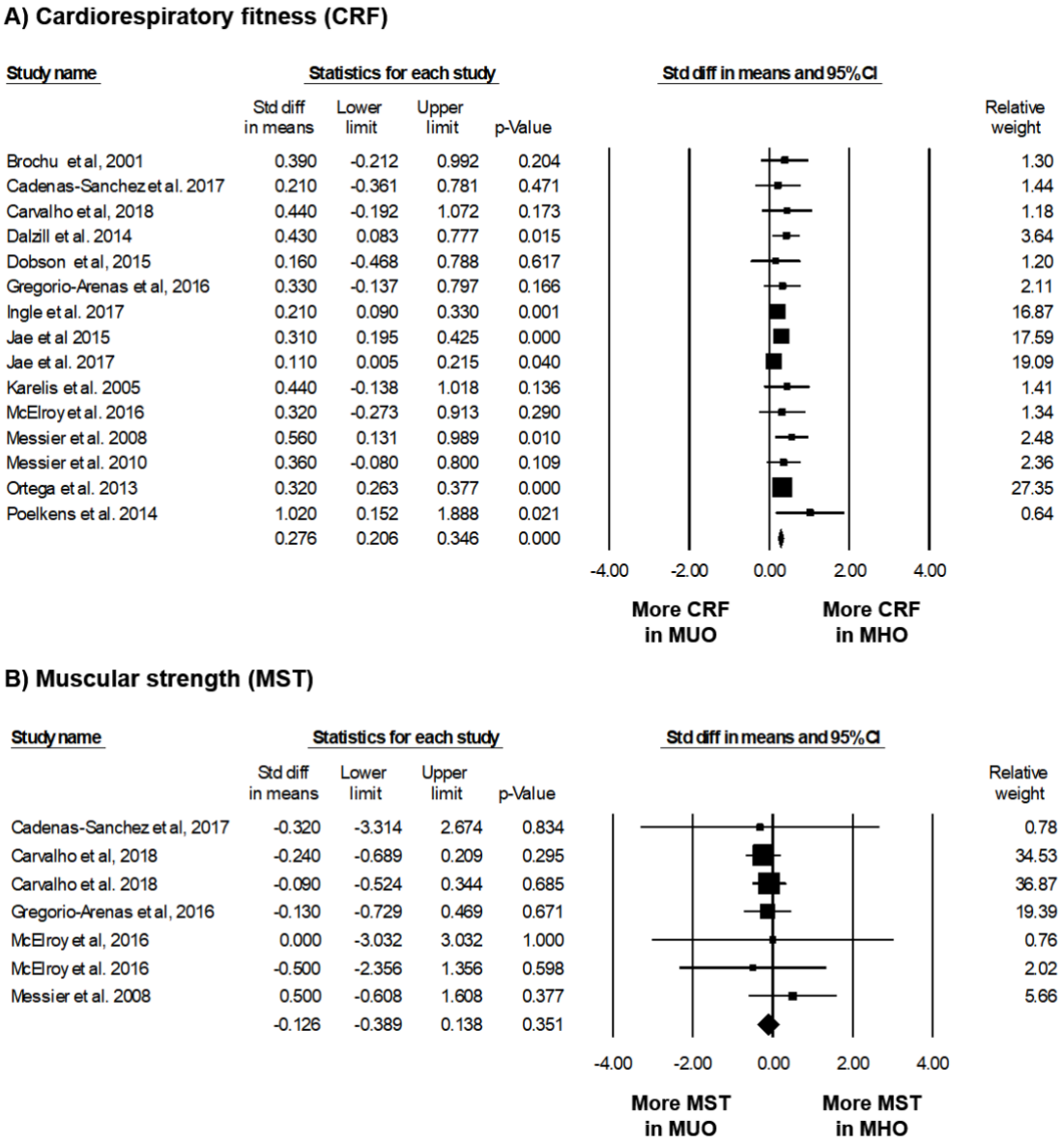
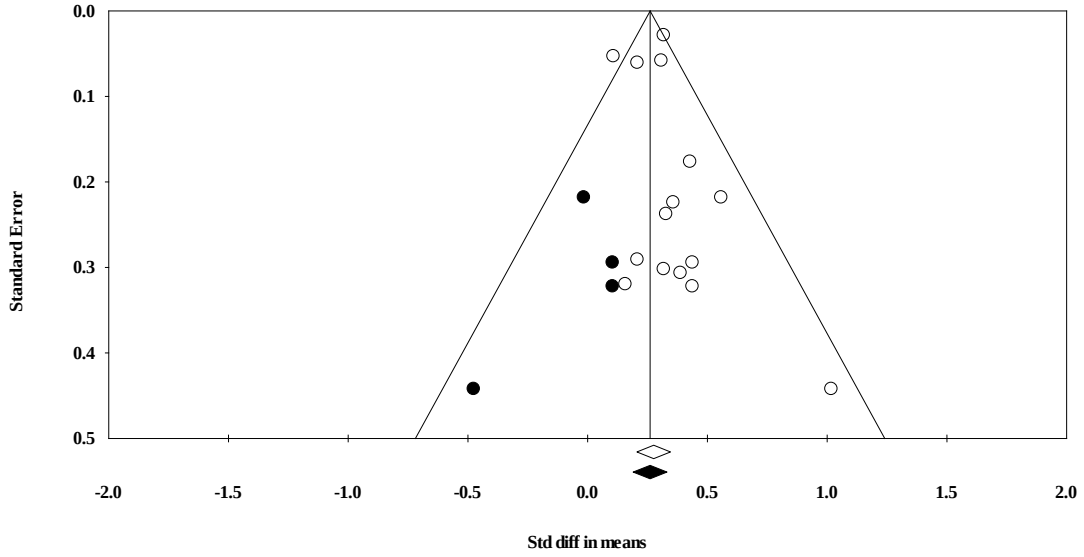


Figure S5. Forest plot of cardiorespiratory fitness (Panel A) and relative muscular strength (Panel B) differences (pooled standardized mean difference) between metabolic phenotypes in obese participants.

Standardized mean differences are expressed as MHO *minus* MUO, thus, positive values represents that MHO showed higher values than MUO. MHO: metabolically healthy obesity. MUO: Metabolically unhealthy obesity. CRF: cardiorespiratory fitness. MST: muscular strength. CI: confidence interval.

A) Cardiorespiratory fitness



B) Relative muscular strength

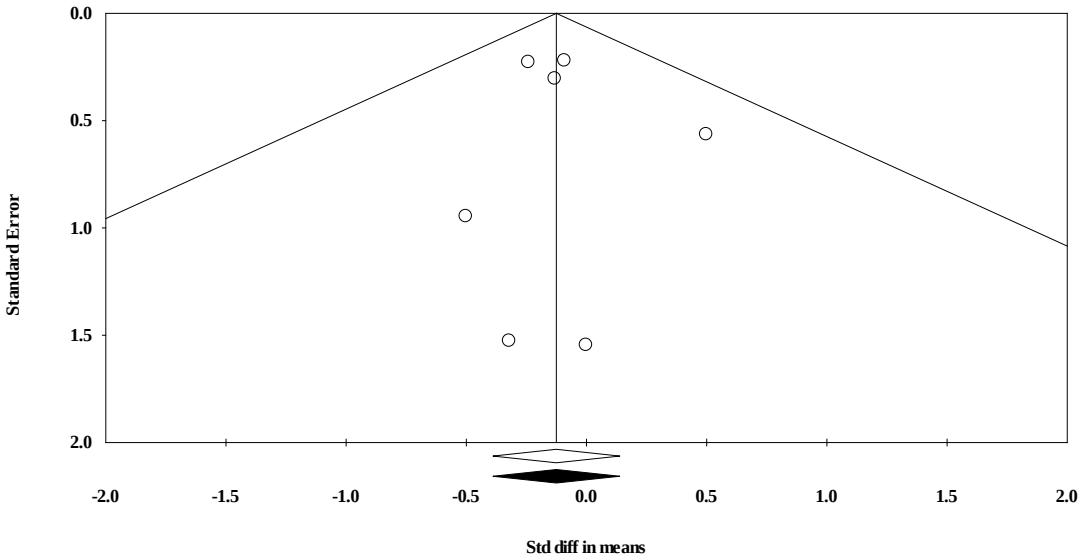
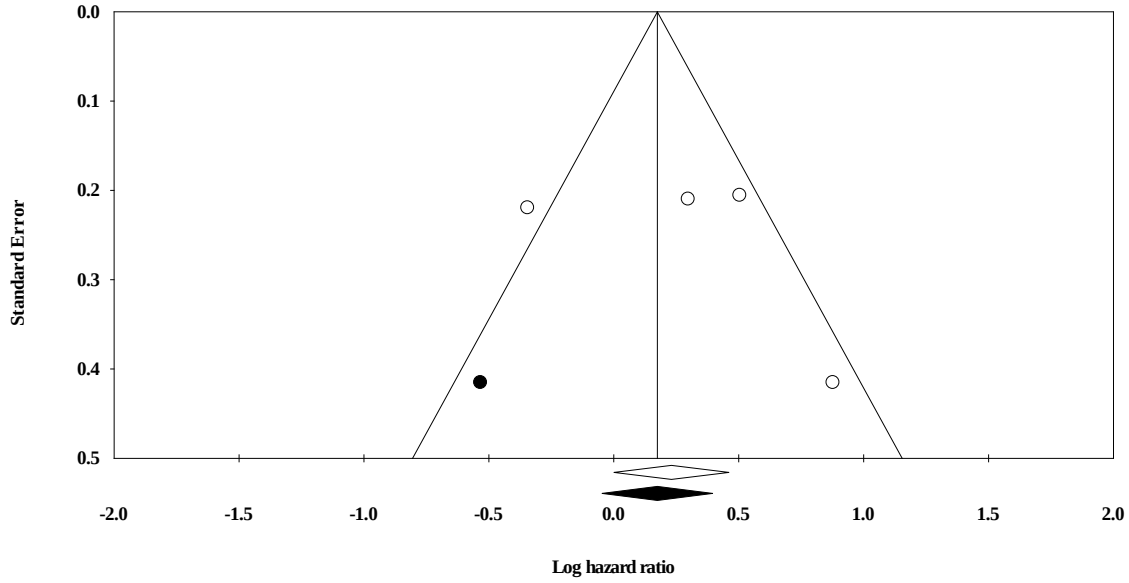


Figure S6. Funnel plots to assess publication bias on cardiorespiratory fitness (Panel A) and relative muscular strength (Panel B) differences (pooled standardized mean difference) between metabolic phenotypes in obese participants.

White point represents meta-analyzed observed studies whilst the black point shows the imputed studies. Diagonal lines represent pseudo-95% confidence intervals. In reference of Y axis, studies located at the lower part of the graph have a higher standard error (a lower weight in the pooled analysis). The vertical line represents the calculated standard differences in means.

A) All-cause mortality



B) Non-fatal cardiovascular disease and cardiovascular disease mortality

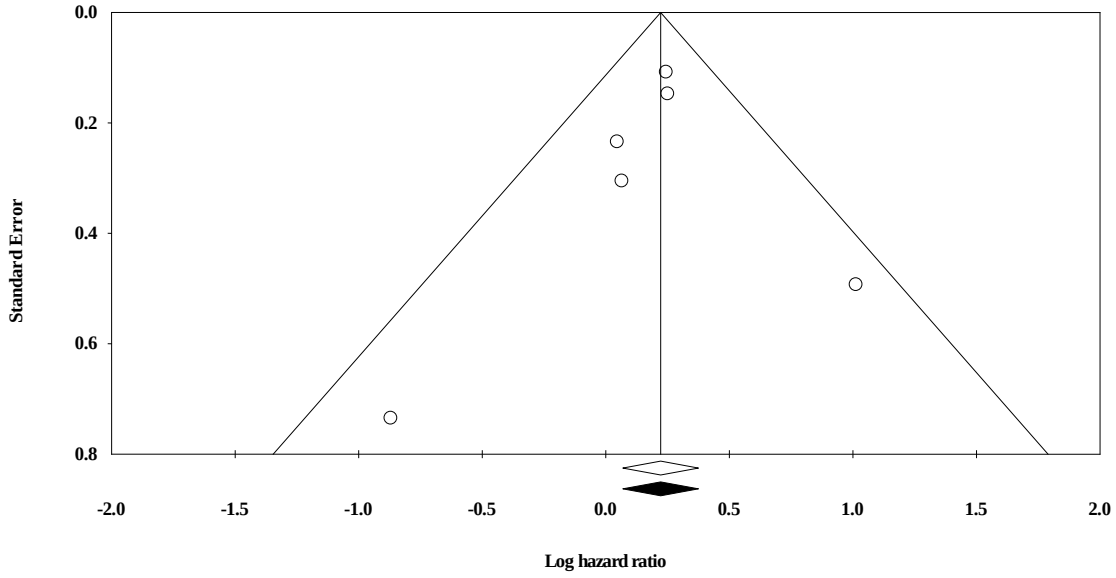


Figure S7. Funnel plot of metabolically healthy normal weight participants for the risk of all-cause mortality (Panel A) and non-fatal cardiovascular disease and cardiovascular disease mortality (Panel B) compared with metabolically healthy obese individuals after adjusting for physical activity.

White point represents meta-analyzed observed studies whilst the black point shows the imputed studies. Diagonal lines represent pseudo-95% confidence intervals. In reference of Y axis, studies located at the lower part of the graph have a higher standard error (a lower weight in the pooled analysis). The vertical line represents the calculated hazard ratio.

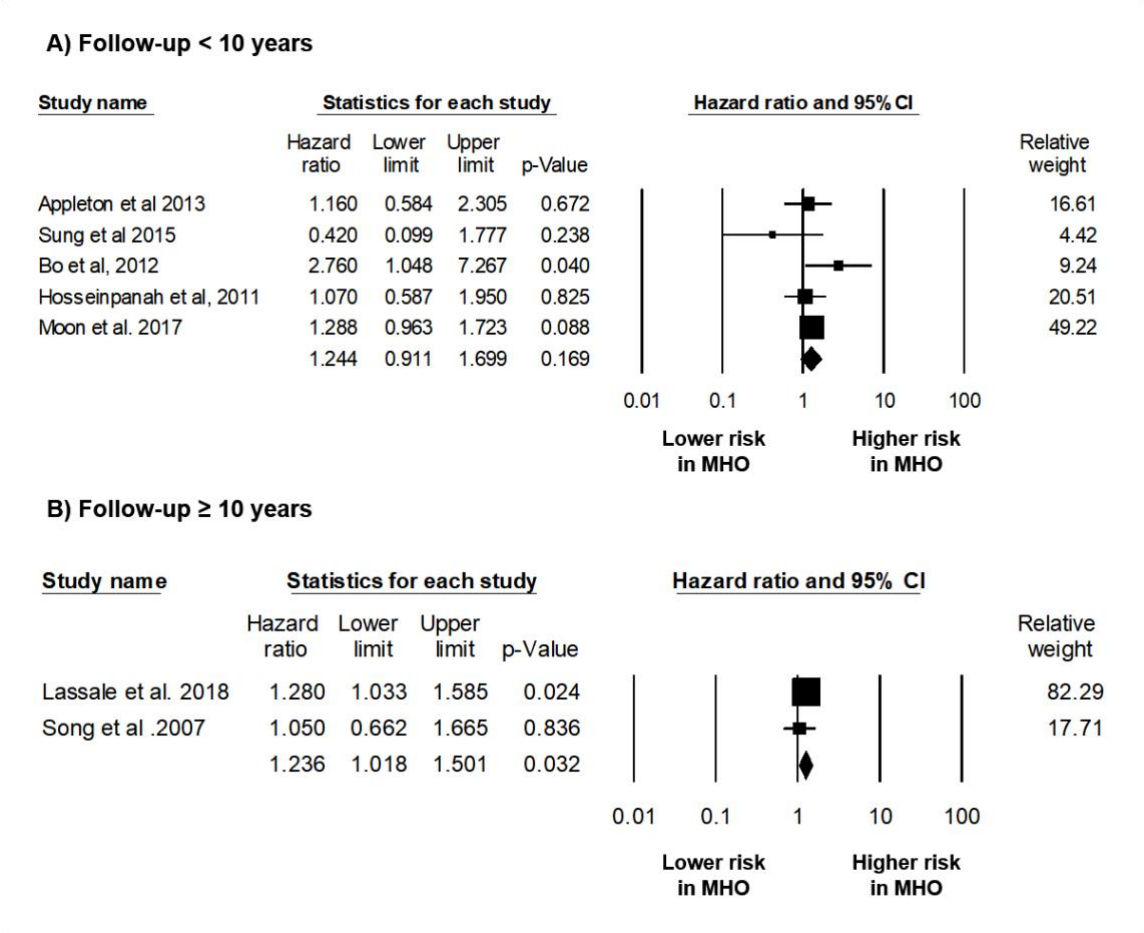


Figure S8. Forest plot of metabolically healthy normal weight participants for the risk of non-fatal cardiovascular disease and cardiovascular disease mortality compared with metabolically healthy obese individuals divided by 10 years of follow-up after adjusting for physical activity.

Hazard ratio is presented having metabolic healthy normal weight phenotype as reference (1). MHO: metabolically healthy obesity. CI: confidence interval.