Risk models and scores for metabolic syndrome: a systematic review

Musa Saulawa Ibrahim (✉ musa.ibrahim@study.beds.ac.uk)
University of Bedfordshire - Putteridge Bury Campus  https://orcid.org/0000-0002-3902-6895

Dong Pang
University of Bedfordshire - Putteridge Bury Campus

Yannis Pappas
University of Bedfordshire - Putteridge Bury Campus

Gurch Randhawa
University of Bedfordshire - Putteridge Bury Campus

Research article

Keywords: Metabolic syndrome, Syndrome X, Prediction, Risk score, Model, Prognosis, Performance, Validity

Posted Date: November 9th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-24184/v3

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

Background: Metabolic syndrome - ‘a clustering of risk factors which includes hypertension, central obesity, impaired glucose metabolism with insulin resistance, and dyslipidaemia’ is linked with increased risk of CVD, T2DM and all-cause mortality. Despite the high number of models and scores for assessing the risk of developing MetS, there is hardly any used in research or practical setting. Hence, we conducted a systematic review to determine the performance of risk models and scores for predicting MetS.

Methods: Following the methods proposed by EPPI-Centre, we systematically searched MEDLINE, CINAHL, PUBMED and Web of Science to identify studies that either derive or validate risk models or scores for predicting the risk of MetS. Search was originally done in September 2018 and updated in September 2020. Data concerning models’ statistical properties as well as details of internal or external validations were extracted. Tables were used to compare various components of models and statistical properties. Finally, PROBAST was used to assess the methodological quality (risk of bias) of included studies.

Results: A total of 27 studies reporting about the development, validation or both of MetS risk models were included. There is significant heterogeneity between studies in terms of geography/demographics, data type and methodological approach. Majority of the models or risk scores were developed or validated using data from cross-sectional studies, or routine data. Various combinations of risk factors (predictors) were considered significant in the respective final model. Similarly, different criteria were used in the diagnosis of MetS, but, NCEP criteria including its modified versions were by far the most widely used (32.5%). There is generally poor reporting quality across the studies, especially concerning statistical data. Any form of internal validation is either not conducted, or not reported in nearly a fifth of the studies. Only two (2) risk models or scores were externally validated

Conclusions: There is an abundance of MetS models in the literature. But, their usefulness is doubtful, due to limitations in methodology, poor reporting and lack of external validation and impact studies. Therefore, researchers in the future should focus more on externally validating/ applying such models in a different setting.

Protocol: The protocol of this study can be found at https://bmjopen.bmj.com/content/9/9/e027326

PROSPERO registration number CRD42019139326

Background

The prevalence of Metabolic syndrome (MetS)- ‘a clustering of risk factors which includes hypertension, central obesity, impaired glucose metabolism with insulin resistance, and dyslipidaemia’ has increased significantly over the last three decades [1]. In 2015, the International Diabetic Federation (IDF) suggested that approximately a quarter of the global adult population (>20 years) have MetS [1]. This figure is disturbing, especially because individuals with MetS have increased risk of Cardiovascular diseases
(CVD) (2 to 3-fold)[2], and type 2 diabetes (T2DM) (up to 5-fold)[3]. Additionally, the number of those with MetS may likely reach a frightening level in the nearest future given the current worldwide trends of rising prevalence of hypertension, obesity and diabetes [1,4]. Therefore, it is necessary to adopt effective preventive strategies and work in a systematic way to reduce the rising burden of morbidity and mortality related to MetS.

However, studies of MetS are often challenging and problematic. First, there is lack of a universally acceptable definition of MetS [5]. This lack of definition that is generally acceptable makes comparison between studies difficult. Second, the exact mechanism that brings about MetS remains unclear despite advances in pathophysiology and risk factor identification. Certainly, the widely observed difference in terms of susceptibility and age of onset is highly suggestive of a major interplay between genetic and environmental factors [6]. Similarly, the application of MetS in clinical setting is still debatable [7]. But, its presence could be used by physicians to improve the care they provide or the outcome of their patients [8]. Furthermore, abdominal obesity is generally presumed to be central to the diagnosis of MetS, but, not all individuals with obesity have MetS and vice versa [9]. Without a valid test, it may not be possible to detect or suspect MetS in nonobese individuals [10]. Identifying the subset of individuals (both obese and nonobese) with MetS is necessary and could have implications on subsequent management approach [11].

Risk prediction models are of great significance in supporting decision making both in clinical and public health practice, and are increasingly being incorporated in guidelines [12,13]. For instance, in CVD (where the application of models is more advanced), several prediction models have been developed and are currently in use, e.g. QRISK (qrisk.org), the Framingham risk score (framinghamheartstudy.org), Assign score (assign-score.com) etc. Furthermore, prediction modelling is becoming more popular in chronic disease research due to the increase in availability of large datasets, advanced statistical methods and computational power [13,14]. This may have a crucial role to play in informing how the rising burden of MetS on public health can be reduced. An accurate and reliable MetS prediction model could be used in screening individuals at increased risk of MetS, both in children and adults. Furthermore, lifestyle interventions have been shown to be effective in modifying the risk of MetS [15,16]. Focusing on assessing risk of MetS provides a single unifying theme which could enable clinicians to identify an individual's global risk of NCD (holistic approach) [8]. Therefore, both from clinical and public health perspective, early identification and control of MetS is of great significance as it could result in the reduction of T2DM and CVD related morbidity and mortality. Similarly, CVD and T2DM share specific risk factors and are the target of many prevention guidelines [17]. But, so far, no perfect risk model for CVD and T2DM exists. Since the purpose of prevention programmes is to address specific issues before they progress into chronic disease, MetS may become an important assessment tool [11].

In recent years, there is proliferation of models and scores for assessing the risk of developing MetS, however, to the best of our knowledge, none is in routine use either in clinical or in public health setting, and there is no available systematic review in the academic literature. This we believe may present a confusing picture for GPs, public health specialists and policymakers, who would be potentially faced
with very complex literature, multiple different methodologies, and probably very few studies of use in real life.

Therefore, the purpose of this review is to determine the performance of risk models and scores for predicting MetS.

**Methods**

This review follows standard methodology for systematic reviews proposed by Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) in “Systematic Reviews: Centre for Reviews and Dissemination guidance for undertaking reviews in health care” [18].

**Search strategy**

Mixed search strategy involving both electronic and manual databases [19,20] was adopted in this study. The search strategy was designed with the help of IHR specialist librarian (DA), and relevant guidance was drawn from “Systematic Reviews: Centre for Reviews and Dissemination guidance for undertaking reviews in health care”, and “Systematic Reviews to Support Evidence-Based Medicine” [18,21] to identify any relevant studies of MetS risk models and scores. The final search strategy was implemented by MSI and was double-checked by DP, GR and YP. The initial search was conducted in September 2018. However, as it has been nearly two years since the original search was conducted, a full update search was done between 01 and 12 September 2020.

The literature was searched using keywords which includes: predict, screen, risk, score, metabolic syndrome, insulin resistance syndrome, model, regression, risk assessment, risk factor, calculator, analysis, sensitivity and specificity, ROC and odds ratio. Both MESH terms and text words were used. Articles were searched using titles and abstracts, the search was limited to studies conducted in English language, but no date restriction was applied.

The details of the search strategies used can be found in supplementary material 1.0.

The literature search was conducted in databases such as MEDLINE, CINAHL, Web of Science and PUBMED.

**Eligibility criteria**

We included peer-reviewed studies that either combine two or more known risk factors to derive a MetS risk model or score, or validated a pre-existing model on a different population or conducted both. Furthermore, the main outcome of this review is MetS, and the secondary outcomes are any related predictive outcomes (discrimination and calibration inclusive). Finally, this review only includes studies published in English.
We excluded studies on screening and early detection, genetic mutation models, conducted on animals, investigating one or more single risk factors which are not connected to build a model or score, studies that applied other disease model or score to predict MetS. Also, studies whose main outcome is not MetS, studies that did not report any related predictive outcomes (either discrimination or calibration).

Selection of studies

All identified articles were transferred into the electronic reference software Endnote version 9 (endnote.com), and duplicates were removed automatically. The duplicate titles that were not removed automatically by endnote were removed manually.

After de-duplication, the titles and abstracts of the records were screened by MSI. Furthermore, the full paper review was conducted by applying the inclusion/exclusion criteria to the retrieved articles. Full paper review was conducted by MSI while DP double-checked ten (10) of the reviewed full papers.

In order to identify more relevant articles, a manual search of the reference lists of all the selected articles was conducted. Furthermore, relevant "grey literature" was searched for in the following: The Grey Literature Report (www.greylit.org/), OpenGrey (www.opengrey.eu/) and OAISTER (www.oaister.org/).

Data extraction

Data extraction was conducted using a standard form adopted from a similar study [22]. The extracted data were on those variables relevant to the review question and which satisfied the conditions for the narrative synthesis conducted. It is noteworthy that, some of the studies presented several models with each model composed of different risk factors. However, it is beyond the capacity of theses researchers to study in details each of those models. Furthermore, the relevant authors themselves often conclude that one of their reported models is obviously better than the others in terms of performance. Therefore, where this is the case, data from the authors’ preferred model(s) or (if no clear preference was stated in the article), the one judged to be more detailed or robust statistically was extracted.

The primary data extraction was conducted by MSI while DP, GR and YP each double-checked the extracted data and discrepancies were resolved by discussions.

Assessment of methodological quality

The PROBAST (Prediction study Risk of Bias Assessment Tool) [23], a tool for assessing the risk of bias and applicability of prognostic model studies, was used to assess the quality (risk of bias and applicability) of included studies. Briefly, the PROBAST is a tool recently developed to assess the quality of primary studies included in a systematic review. It evaluates both risk of bias and issues concerning applicability of studies that develop, validate or update a multivariable model (both diagnostic and prognostic).
Furthermore, PROBAST comprises of 4 domains covering 20 signalling questions to enable risk of bias assessment and applicability. These domains are concerned about participants such as (the study design used, whether appropriate inclusion/ exclusion criteria were used), the predictors used, outcome, and how the analysis was conducted.

Aside from its specific purpose of appraising studies in systematic reviews of prediction models, PROBAST can also be utilised in the general critical appraisal of primary prediction model studies. Noteworthy, PROBAST is not meant for generating summary “quality score” due to the documented drawbacks related to such scores [24]. Therefore, the effect of problems observed within each domain should be discussed by users [23].

Results

After removing duplicates, a total of 17543 records were screened, out of which 77 full-text articles were reviewed (see figure 1.0). A total of 27 studies (including 3 studies identified during the update search) met the inclusion criteria. Other full-text studies were excluded because of the following: predicting genetics (n=1), investigating one or more single risk factors which are not connected to build a model or score (n=22), used unconventional predictors (complimentary/alternative medicine) (n=3), applied other disease model or score to predict MetS (CVD, T2DM) (n=4) main outcome is not MetS (n=8) predicts MetS risk reduction (n=1), did not report any related predictive outcome (either discrimination or calibration) (n=9) conducted in languages other than English (n=2).

Furthermore, out of the 27 included papers, 25 report about the development of one or more risk model or score [25-36,38,39,41-51], and 2 studies report about the development and validation of one or more risk model or scores on an external population [37,40]. Overall, the 27 studies reported 43 models, out of which 27 models were selected for full data extraction. The rest (16 models) were not selected, either because they were judged to be minimally different from the reported ones or they were not the preferred models by the authors or they were significantly deficient in details or statistical reporting.

Table 1.0 provides a summary of the quality assessment of the included studies. In summary, the quality assessment revealed that in the entire included studies there is moderate-to-high risk of bias, primarily due to the use of inappropriate study design and absence of external validation. Further look at the studies, it was observed that majority of the models suffered a high risk of bias and significant methodological deficiencies arising from poor choice of model analyses, significantly underpowered analyses, dichotomisation of continuous variables, lack of adjustment for optimisation, poor handling of missing data and overall poor model presentation.
Table 2.0 provides the summary of the included studies. In summary, there is high heterogeneity in the studies. Studies were conducted in 17 countries with sample size ranging from 62 to 36944 participants. Overall, risk models were tested on 155,027 individuals. Furthermore, studies differ in terms of the target population (children/young vs adults) as well as purpose (diagnosis vs prognosis of MetS). Due to the heterogeneity of data, difference in methodological approach and presentations, it is challenging to make comparisons across studies.
MetS risk models or scores for children

Regarding the target population, more risk scores were reported in adults than in children or adolescents. Of the 27 included models, 17 focused on adult subjects [25,26,32-39,42-45,48,51], while 10 targets children and adolescents [27-29,31,40,41,46,47,49,50].
The main purpose of developing most MetS models targeting children and adolescents is diagnosing MetS. Only one model was built for the purpose of predicting MetS prospectively [40,34]. Similarly, most of the children MetS models were developed using cross-sectional data [27-29,31,41,46,47,49,50] and they often report incidence/prevalence of MetS at the end of the study, which ranges from as little as 1.2% [47] to as high as 54.8% [50]. In all the 10 risk scores aimed at children, various combinations of risk factors (predictors) were considered significant in the respective final model. The average number of predictors utilised in a single risk score is five. Similarly, majority of the studies used a continuous criteria in the diagnosis of MetS [27-29,31,41,46,47,49,50] (see table 3.0).
Poor reporting quality is observed across the studies, especially concerning statistical data. For instance, only one study reported calibration (of any statistic form) [40] and two studies reported positive and negative predictive values [28,40]. The commonest measures reported are sensitivity/specificity (all but one study) [47] and AUROC (all studies). AUROC ranged from 0.87 to 0.98. In the same vein, the final model equation was reported in only 5 studies [27-29,47]. Also, one study reported having an online risk calculator [46].
Regarding the validity of the risk models or scores, any form of internal validation is either not conducted, or not reported in nearly a fifth of the studies. Similarly, only one risk model was externally validated [40]; by same authors and, reported in the same paper (with corresponding model development). However, there is some form of uniformity when it comes to the biomarkers used to capture the assumption of MetS with majority of studies using a combination of WC, MAP, HDL-C, TG and FBG. However, 3 studies used HOMA-IR instead of FBG [27,46,50], while a study employed salivary biomarkers (as against blood biomarkers) [47].

**MetS risk models or scores for adults**

Both diagnostic and prognostic MetS models exist for adults. But, majority (70.6%) are diagnostic developed using data from cross-sectional studies [25,26,30,32,33,37-39,43,45,51]. Surprisingly, the observed incidence/prevalence of MetS at the end of the study is not reported in over forty-one percent of the studies [25,26,32,38,42,43]. For those that reported it, the prevalence/ incidence of MetS ranged from as little as 7.9% [37] to as high as 57.8% [51]. In all the 17 risk scores targeting adults, various combinations of risk factors (predictors) were considered significant in the respective final model. Again, different weights were assigned to different components in the various models. The number of predictors utilised in a single risk score ranged from 2 to 11 (mean 5.6, SD 1.95). Similarly, NCEP including its modified versions were by far the most widely used criteria (32.5%) in the diagnosis of MetS (see table 3.0).

Similar to the children studies, there is generally poor reporting quality across the studies, especially concerning statistical data. Only two studies reported calibration (of any statistic form) [36,48]. Similarly, more than one in five studies did not report sensitivity/specificity[25,26,36,42]; four-fifth did not report positive and negative predictive values[25,26,30,32-36,38,39,42,45,48]. Similarly, Area Under Receiver Operating Curve (AUROC) ranged from 63.0 to 95.0. One study did not report any discrimination measure [33].

Regarding the validity of the risk models or scores, any form of internal validation is either not conducted, or not reported in nearly a fifth of the studies. Similarly, only one risk model was externally validated [37]. More so, the external validation was done by the same authors and, reported in the same paper with corresponding model development. In addition to the traditional biomarkers/ predictors of MetS (i.e. abdominal obesity, blood pressure, blood glucose, triglyceride and HDL-cholesterol), some studies employed other (novel) biomarkers, such as phenotypic biomarkers (double chin, buffalo hump) [38], Quadriceps muscle peak torque/body mass, (Nm/kg) [39], and lifestyle factors (alcohol, moderate physical activity, smoking, food insecurity, habit of eating less salt, dairy consumption) [37].

Reporting the model equation can add to the reporting quality and replicability of the study. However, the final model equation was reported in only 10 studies [25,26,30,32,33,38,42,43,45,48]. Also, one study reported having a risk estimation chart [48] and another has mobile application and excel risk calculators [25].
Diagnostic vs prognostic MetS risk models or scores

The diagnostic and prognostic MetS models differ in certain characteristics. On the one hand, MetS diagnostic models are usually developed using data from cross-sectional studies, or routine data that were often assembled for other reasons [25-33,37-39,41,43,45-47,49-51]; have smaller sample size; and mostly targeted at young individuals. On the other hand, prognostic MetS models tend to be developed using cohort data (both general population and primary care) [34-36,40,42,44,48]; with relatively larger sample size; target mainly adults [34-36,40,42,44,48]; have prediction horizon ranging from 2 years [48] to 10 years [40] (see table 3.0).

Discussion

This systematic review revealed that there are numerous MetS prediction models or scores in the literature. This finding is similar to what is seen in other chronic diseases such as CVD [52,53,54] and T2DM [22,55,56]. Similarly, studies of MetS risk models vary significantly in terms of methodological approach, presentation, purpose (diagnosis vs prognosis of MetS) and target population (children vs adult). However, despite their variations, they share common characteristics.

Poor conduct and reporting of prediction model studies is a common finding reported across most similar reviews [22,52-56], and often, this leads to missing of vital information. This finding is also observed in most of the studies included in this review. The lack of standardised guideline both for conduct and reporting of prediction model studies is believed to be largely responsible for this occurrence [15]. However, with tools such as TRIPOD [57] and PROBAST [23] being developed and validated, the situation is likely to change in the nearest future.

Assessing the overall performance of prediction models is necessary before translating research findings into the real-world setting [58]. However, model performance can be affected in several ways. One of these ways is how continuous variables are handled; are they retained as continuous measures or are they categorised into two or more categories [59]. Categorisation of variables is frequently observed in risk prediction model studies [60-63]. In this review, categorisation of some, or all variables was conducted in (65%) of the studies. This finding is in keeping with what is reported previously in similar reviews [52,54,55]. However, it is recommended that while developing a model, continuous risk variables (predictors) should be retained as continuous variables, or rather, splines or fractional polynomial functions should be used if the relationship between the predictor and the outcome is nonlinear [64].

Another way that the performance of models is affected is through missing values. Missing values are a common occurrence in most datasets [55]. Nearly half of the studies in this review failed to report information regarding how they treated missing values. This finding is in keeping with other similar reviews [52,55,60,63]. Potent ways of minimising the effect of missing values such as the multiple imputation technique have been recommended [65,66]. Therefore, researchers should always report the completeness of the overall data and how the missing values are dealt with so that the representativeness and quality of the data can be judged by readers.
Again, there is a lack of consistency in studies of MetS prediction as they used different predictors and statistical methods. To the very least, discrimination and calibration measures are recommended to be reported [67]. Although nearly all studies reported some form of discrimination, however, calibration is rarely reported. In this review, only three studies reported any form of calibration measure [36,40,48]. This is similar to other relevant reviews [52,55,60,63]. This makes it difficult to make comparison across studies, e.g. meta-analysis and to assess the generalizability of the studies [68].

Furthermore, majority of the studies used common biomarkers (blood pressure, fasting blood sugar, cholesterol, triglycerides and waist circumference) as predictors in building their models. In addition to these, however, other novel predictors/biomarkers have been used once or twice by some researchers. However, none of the models that reported using novel biomarkers has been used elsewhere or externally validated. A similar observation is made in CVD models studies [52]. This shows that researchers in the field give more significance to the process of identifying new predictors and new model building as against validating and applying existing ones.

Regarding the definition/criteria of MetS, there is significant heterogeneity amongst the studies. But, the NCEP criteria [69] or its modified versions are the most commonly used. This further makes it difficult to compare between studies because different definitions of outcome result in difference in predictor effect and resultant model performance [52]. Having a more uniform definition/criteria would help significantly in mitigating the above (thereby making it easy to compare between studies and eventually translate research findings into clinical setting) [57].

Furthermore, most of the studies included in this review described developing MetS prediction models, but, the external validation of such models is seriously lacking. Certainly, the ultimate aim of any multivariable model study is to show that the model in question works [55]. It is, therefore; of paramount importance that the model performance is assessed once it is developed [68,70]. Only two models in this review were externally validated. Lack of external validations is a common problem of most prediction model studies [52,55,60,63].

Head-to-head comparison of models assists in knowing which models are better in terms of performance. In this review, no such comparative study is observed. This makes it difficult to choose, or advocate amongst the existing studies. Comparative studies (preferably of multiple models in a single study) are recommended in prognostic risk prediction models [52]. However, as significant as the statistical characteristics of a prediction model may be, they do not guarantee its usefulness in a clinical/real-life setting. None of the models in this review is reported to have been applied in clinical setting. Therefore, in the future, more emphasis should be given on impact studies- applying the models in clinical setting and assessing their ability to influence decision making or patients' outcome.

The number of risk models or scores included in the final sample of this review is relatively high. But, we believe none of the existing MetS models or scores in their current state warrants to be applied in real-life setting without being investigated further. For any prediction model or risk score to be considered useful, it should be accurate (statistically significant calibration, and discrimination above 0.70), generalisable
(externally validated by a separate research team on a different population) and usable (has few components that are commonly used in practical setting) [71]. However, MetS prediction discipline is arguably still in its early phase of development; therefore, it is difficult to identify any model or score that fulfils all of the above criteria. Nevertheless, for clarity, it was decided to highlight the risk models or scores we feel should be researched further. To aid in making this decision, the above criteria proposed by Altman et al. [71] was modified. In addition, we also considered the availability of either full model equation, online risk calculator or risk estimate chart in the paper, without which it will not be possible to replicate the study. Based on the above, the following prediction models are recommend to be further tested: For the purpose of long term prediction of MetS in adolescence, we recommend the risk model described in Efstathiou et al. [40]. The model was developed using a cohort data, used few predictors and reported good accuracy (discrimination 0.97 & calibration 6.64). Furthermore, the model has been externally validated. Similarly, for the purpose of predicting risk of developing MetS around mid-life, the model developed by Hsiao et al. [48] should be researched further. The model was developed using a primary care cohort, used few number of common predictors, and reported fairly good performance measures (discrimination 0.77 & calibration 0.20). In addition, both the full model equation as well as the corresponding risk estimation chart are provided, making it easy to replicate. Finally, if a researcher is interested in a risk score for self-assessment by older lay persons, then we suggest the JAMRISC [33]. The score was developed using non-laboratory predictors. Also, the developed questionnaire as well as the final model equation are provided which could be used.

The role of MetS as an effective tool for identifying individuals at increased risk of CVD and T2DM has been recognised [72]. Focusing on identification of individuals at risk of MetS will help in the wider prevention of T2DM, CVD and other NCDs related to MetS. However, despite the clear evidence linking it with the risk of CVD and T2DM, the suitability and practicality of applying MetS in clinical setting is still debatable [7]. The major argument for the use of MetS in clinical practice is that physicians can use its presence to improve the care they provide or the outcome of their patients [8].

Generally, abdominal obesity is presumed to be central in the diagnosis of MetS, but, not all individuals with MetS are obese and vice versa [9]. The consideration of metabolically healthy obese and nonobese MetS is gaining more attention [73]. This is even more relevant with the recognition of other significant risk factors associated with MetS such as stress, poverty and unhealthy diet [74-76]. Distinguishing between these groups of MetS might be necessary because of the potential management implications [11]. But, without a validated special test, it may not be possible to detect or suspect MetS in nonobese individuals and those patients who may not be otherwise suspected to have it [10]. Individuals with nonobese MetS may have relatively low perception of risk of MetS, and may be missed by clinicians (because they may not be considered high risk) [10]. Similarly, in obese patients, MetS can be used as a marker indicating increase need of intervention to reduce subsequent risk of CVD and T2DM.

Through the use of MetS risk models or scores we can identify and focus on individuals with the most risk of MetS without necessarily screening the whole population [77]. As shown in this review, MetS risk models could be deployed for different purposes and various populations. The diagnostic MetS models
are typically built using common clinical biomarkers which are also components of MetS. Therefore, they could be utilised in everyday clinical setting for assessing high risk group (patients highly suspected to have MetS either based on their clinical presentation or significant history) [78]. Arguably, the diagnosis of MetS may motivate patients and their clinicians to take the necessary actions required for reducing their risk of CVD and T2DM [78]. Similarly, some of the prognostic MetS models could be applied in general prevention/screening programmes targeting low risk group (e.g. apparently healthy young people). Also, they could be used to monitor response to treatment in patients diagnosed with MetS and undergoing therapy.

Identifying and managing MetS is a form of primary or early secondary prevention of CVD and T2DM. Thus, primary care is the most appropriate setting where MetS prediction models should be applied. More so, most of the predictors of the existing MetS models could easily be obtained in such setting. Furthermore, many similar risk models are being utilised in the primary care setting.

Both CVD and T2DM share specific risk factors and are the target of many prevention guidelines [17]. But, so far, no perfect risk model for CVD and T2DM exists. Since the purpose of prevention programmes is to address specific issues before they progress into chronic disease, MetS may become an important assessment tool [11]. The WHO stress that MetS should be viewed as a pre-morbid state of CVD and T2DM, not a tool for diagnosing a distinct clinical disease [72]. Indeed, diagnosing MetS combines several risk factors thereby providing a more holistic picture to the physicians regarding the individual’s risk of CVD and T2DM which is at the core of many primary healthcare systems [11]. However, before this could be achieved, MetS prediction models must be shown to have good discriminative power both in high risk group and general population [56]. Therefore, as a next step, some of the promising MetS models need to be further investigated to ascertain their real worth.

What is already known on the topic

- The findings of this review are in keeping with earlier published reviews of risk prediction models or scores in CVD, T2DM, cancer and stroke.
- Numerous MetS prediction models or scores exist in the literature
- No detailed review has described how these models or scores were developed, their predictive performance and how many were validated (externally).

What this study adds

- Though there are many MetS prediction models or scores, only a few have been externally validated.
- Poor reporting was observed in all aspects of risk prediction models development, specifically, in terms of data description and sufficient detail in all steps taken in building the model.
- This makes them of doubtful value to potential users (practitioners, policy makers and guideline developers).

Strengths
The strengths of the systematic review include:

- This review consists of models or scores that are geographically diverse (across different countries). This is likely to give a broader picture of MetS risk prediction because it has been shown that models performances are affected by race/ethnicity.
- The review used standard methods that apply to reviews of any risk models and scores.

**Limitations**

This review was limited to articles published in English; therefore; some significant additional findings might be missed. Notwithstanding, it is strongly suspected that none of the study findings would be altered by adding more articles. Furthermore, large numbers of titles and abstracts were screened at the initial stages of the review due to lack of a sensitive literature search strategy. Finally, due to significant heterogeneity between the included studies, detailed quantitative analysis and formal meta-analysis could not be performed.

**Conclusions**

This systematic review of 27 published studies highlights numerous methodological deficiencies and a generally poor level of reporting in studies in which risk prediction models were developed for the detection of MetS. Furthermore, there is an abundance of MetS models in the literature. But, their usefulness is doubtful, due to limitations in methodology, poor reporting and lack of external validation and impact studies. Therefore, researchers in the future should focus more on externally validating/applying such models in a different setting.

**Declarations**

**Ethics approval and consent to participate:** The ethics approval for this study was provided by the Institute for Health Research Ethics committee, University of Bedfordshire.

**Consent for publication:** Not applicable

**Availability of data and materials:** Not applicable

**Competing interests:** The authors declare that they have no competing interests

**Funding:** This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors’.

**Authors’ contributions:** MSI conceptualised the review, co-developed the search strategy and ran the search, scanned all titles and abstracts, extracted quantitative data, conducted the quality appraisal and wrote the paper. He is also the guarantor for the paper. DP assisted with double-checking the search strategy, studies selection, data extraction and quality appraisal. YP double-checked the search strategy,
data extraction, read and revise drafts of the paper, and approved the final manuscript. GR checked the search strategy, data extraction, read, provide feedback and approved the final manuscript.

Acknowledgements: The authors will like to acknowledge the contribution of David Abdy, the specialist librarian for Institute for Health Research, University of Bedfordshire for his contribution in designing the search strategy for the study.

References


78. Huang PL. A comprehensive definition for metabolic syndrome. Disease models & mechanisms. 2009 May 1;2(5-6):231-7.

**Figures**
Figure 1

Prisma flow diagram describing the selection of studies
Figure 2

Publication of MetS risk models and scores 2008 to 2018

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- BMC.supplementaryfiles.docx
- BMC.revised.PRISMA2009checklist.pdf