Statistical Frameworks for Setting Cutoff Points of Metabolic Syndrome Criteria

[1]Tien-Mu Hsiao, [2]Huifen Chen,[3]Chien-Hua Wu *

[1]Department of Community Medicine

Landseed Hospital

Taoyuan, Taiwan

[2]Department of Industrial and Systems Engineering

Chung-Yuan University

Taoyuan, Taiwan 32023

[3]Department of Applied Mathematics

Chung-Yuan University

Taoyuan, Taiwan 32023

March 17, 2016

*Corresponding author: Chien-Hua Wu, Tel:+886-3-2653130; Fax:886-3-265-3199. E-mail address: cwu@cycu.edu.tw

Abstract

We propose three statistical frameworks for determining the cutoff points of metabolic syndrome (MetS) criteria, consisting of six components that are the same as in widely used MetS definitions, e.g., the 2004 updated NCEP-ATPIII criteria. Several international organizations have proposed MetS definitions; no literature indicates that any of these definitions is based on statistical frameworks. For all the three frameworks, the cutoff points are set to maximize the observed prevalence rate of stroke and DM. The three frameworks differ in assumptions on the joint distribution of the six components. Using the cohort data from a regional hospital in Taiwan, we illustrate applications of the three frameworks and compare them with the updated NCEP-ATPIII definition and the 2009 consensus definition of IDF and AHA/NHLBI. The performance measure is the odds ratio, the odds of getting stroke or DM within subjects with MetS divided by the analogous odds for subjects without MetS. Our numerical results show that the odds ratios of the three frameworks are higher than those of the updated-NCEP and consensus definitions, showing that the proposed frameworks seem to provide a better association of stroke and DM.

Keywords: cut-off point, metabolic syndrome, stroke, diabetes mellitus

1 Introduction

Metabolic syndrome (MetS) was first termed as "syndrome X" for a cluster of various metabolic abnormalities by Reaven [1] in 1988. Such metabolic abnormalities have been observed in patients by Kylin [2] in 1923 and discussed later by Vague [3] in 1956. MetS consists of a cluster of risk factors for the development of cardiovascular diseases (CVD) and diabetes mellitus (DM) type 2. A meta-analysis according to the National Cholesterol Education Program (NCEP) Adult Treatment Panel (ATP) III definition reveals that metabolic syndrome is associated with a twofold increase

in cardiovascular outcomes [4]. The MetS prevalence rate among U.S. adults is more than 25% [5].

We propose here statistical frameworks for determining the cutoff points of MetS criteria that are similar to most widely accepted MetS definitions (e.g., NCEP-ATPIII [6]). Since 1998 many organizations have proposed MetS definitions (composed of criteria) for medical diagnoses, e.g., the definitions of World Health Organization (WHO) in 1998 [7], European Group for the Study of Insulin Resistance in 1999 [8], US NCEP-ATPIII in 2001 [6], American Association of Clinical Endocrinology in 2003 [9], American Heart Association/National Heart, Lung, and Blood Institute (AHA/NHLBI) in 2004 ([10], [11]), International Diabetes Federation (IDF) in 2005 [12], and the consensus definition incorporating IDF and AHA/NHLBI definitions in 2009 [13]. Since the AHA/NHLBI definition updates the NCEP-ATPIII definition [6] with minor modifications, it is also called the updated NCEP-ATPIII definition. Among these definitions, the NCEP-ATPIII definition has emerged as the most widely used definition, primarily because the NCEP-ATPIII criteria are easy to use in clinical practices ([6], [14]). However, Kassi et al. [15] suggest that the consensus definition of IDF and AHA/NHLBI is most suitable for practical use in clinical medicine. Nevertheless, none of these MetS definitions are based on statistical models.

These existing MetS criteria are based on cutoff points of a cluster of risk factors of CVD since MetS was originally proposed as a common pathway to CVD. Most MetS definitions use six components (i.e., risk factors of CVD): the central-obesity waist circumference(WC), triglycerides(TG), systolic-blood-pressure (SBP), diastolic-blood-pressure (DBP), high-density lipoprotein-cholesterol (HDLC), and fasting glucose(FG). For example, the updated NCEP-ATPIII (i.e., AHA/NHLBI) definition is any three of the five criteria: (i) TG $\geq$ 150 mg/dL (or use drugs for elevated TG), (ii)

FG $\geq$ 100 mg/dL (or use drugs for elevated glucose), (iii) WC $\geq$ 102 cm in men and $\geq$ 88 cm in women, (iv) HDLC $<$ 40 mg/dL for males and $<$ 50 mg/dL for females (or use drugs for reduced HDLC), and (v) SBP $\geq$ 130 mmHg or DBP $\geq$ 85 mmHg (or use drugs with a history of hypertension). The consensus definition of IDF and AHA/NHLBI has the same cutoff points of TG, FG, HDLC, SBP, and DBP as those of the updated NCEP-ATPIII definition, but uses geography-specific cutoff points for WC (e.g., for Asians: $\geq$ 90 cm in men and $\geq$ 80 cm in women; for US: $\geq$ 102 cm in men and $\geq$ 88 cm in women). See also Table 1.

Existing MetS criteria may be updated occasionally, usually by changing the cutoff points. If a component has significantly increased the CVD morbidity and mortality (with significant increased odds ratios compared among different categorized groups), its cutoff point may be changed. Two components with frequently updated cutoff points are FG and WC. For example, the cutoff point of FG was reduced from 110 mg/dl in 2001 (by NCEP) to 100 mg/dl in 2004 (by AHA/NHLBI) because the CVD morbidity and mortality for individuals with FG in the range of 100 to 110 mg/dl has increased.

There has been no generally accepted pathogenic mechanism for MetS. Debates of MetS definitions arose even after the consensus definition proposed in 2009 ([16], [17]). Simmons et al. [18] pointed out that MetS may be useful as an educational concept but has limited practical utility as a diagnostic or management tool. Reaven (first proposing insulin resistance as underlying pathophysiology mechanism of syndrome X) criticized central obesity as the main pathogenesis factor of MetS ([19]-[21]). Parikh et al. [22] and Parikh and Mohan [23] proposed that WC be replaced by the index of central obesity and lipid accumulation product, respectively. Johnson et al. [24] proposed the fat storage as main pathogenesis factor of MetS. The proposed mechanisms or indices by [22], [23], and [24] are different from widely used definitions such

as the NCEP-ATPIII definition and consensus definition of IDF and AHA/NHLBI.

Four shortcomings of existing MetS definitions motivate this research: (i) None of the existing definitions establishes a statistical model based on the CVD morbidity or mortality risk; (ii) None of the existing definitions sets the criteria's cutoff points by considering all criteria simultaneously; (iii) The cutoff-points modifications for an existing MetS definition are usually based on an increase of CVD morbidity or mortality for a subset of the criteria (rather than all criteria); and (iv) The cutoff points of MetS criteria (excluding WC) are based on western-country data and hence, may not be suitable for Asian population.

This proposed statistical frameworks for determining the cutoff points of MetS criteria aim to improve MetS's association of stroke and DM. We use the same six components (WC, TG, FG, HDLC, SBP, and DBP) as in the NCEP-ATPIII definition and consensus definition of IDF and AHA/NHLBI. Three statistical frameworks are proposed, depending on assumptions of the joint distribution of the six components. All the three frameworks set the cutoff points to match the prevalence rate of stroke and DM for better association of CVD and DM morbidity. In this work, we assumed that the prevalence rate of stroke and DM for a certain population is given. Using the Li-Shin Outreach Neighborhood Screening (LIONS) data from Landseed Hospital in Taiwan, we illustrate applications of the proposed statistical frameworks for computing the cutoff points of MetS criteria. We also compare these MetS criteria with those of the updated NCEP-ATPIII (i.e., AHA/NHLBI) definition and the consensus definition of IDF and AHA/NHLBI, where the latter two are called the updated-NCEP and consensus definitions, respectively, in the rest of the paper for simplicity.

The rest of this paper is organized as follows. Section 2 proposes three statistical frameworks for computing the cutoff points of MetS criteria similar to the updated-

NCEP and consensus criteria. Section 3 uses LIONS data to demonstrate applications of these three statistical frameworks and compare them with the updated-NCEP and consensus definitions. Section 4 is our conclusions.

2 Statistical Frameworks

We propose here three statistical frameworks for computing cutoff points of MetS criteria: (i) ellipsoidal boundary, (ii) Bonferroni-type cutoff points, and (iii) modification of current popular definitions. Like most existing MetS criteria, all the three frameworks consider six components WC, TG, FG, HDLC, SBP, and DBP; all are risk factors of stroke and DM. Furthermore, the components' cutoff points are set as the values that maximize the observed prevalence rate of stroke and DM.

The three frameworks differ in assumptions on the joint distribution of the six components. The first framework assumes that the six components follow a multivariate normal distribution and hence provides a simultaneous ellipsoidal boundary for defining MetS criteria. Framework 2 does not require the normality assumption and computes the cutoff points by implementing the Boole's inequality to bound the probability of having MetS. Framework 3, like most popular MetS definitions (e.g., NCEP), identifies an individual as having MetS if three or more out of five criteria are satisfied. For Frameworks 1 and 2, however, an individual with only an extremely unusual measurement on a single component is prone to be identified as having MetS. Such MetS definitions may be reasonable. As mentioned in Alberti et al. [13], most patients with type-2 DM (i.e., FG $\geq$ 126 mg/dl) have MetS under the 2009 consensus definition.

Like updated-NCEP and consensus definitions, Frameworks 2 and 3 view SBP and DBP together as one blood-type criteria and hence have only 5 criteria; Frameworks

2 and 3 also set the cutoff points of WC and HDLC differently for males and females for considering the gender effect. Values of these gender-specific cutoff points are evaluated based on each gender's prevalence rate.

All three frameworks use the maximum likelihood approach to maximize the conformity of the computed boundary of MetS with the observed stroke/DM prevalence rate. The reason is that MetS is associated with risk factors for stroke and DM. Specifically, suppose that there are n_1 persons without stroke/DM among a population of size n. The MetS cutoff boundaries of Frameworks 1 to 3 are set so that the MetS prevalence rate matches the stroke/DM prevalence rate, which is $1 - n_1/n$. A case study with LIONS data in Section 3 illustrates the calculation of the prevalence rate $(1 - n_1/n)$. The maximum likelihood approach is detailed in Appendix A.

We describe Frameworks 1 to 3 in Sections 2.1 to 2.3, respectively.

2.1 Framework 1: Simultaneous ellipsoidal boundary

The first statistical framework constructs an ellipsoidal boundary by considering all the components simultaneously. Framework 1 assumes that the random vector of these components after proper transformations follows a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. A typical way (e.g., confidence region) to construct a boundary on a multivariate normal support is to use the relation between multivariate normal and chi-squared distributions and obtain an ellipsoidal boundary whose size depends on a chi-squared quantile.

Specifically Framework 1 defines MetS as the region outside $R(B_1)$, where

$$R(B_1) = \{\boldsymbol{x} : \boldsymbol{(x - \hat{\mu})'\hat{\Sigma}^{-1}(x - \hat{\mu})} \leq \chi^2_6(n_1/n)\}.$$

Here $x_1 =$ WC, $x_2 = \ln$(TG), $x_3 = \ln$(FG), $x_4 = \ln$(HDLC), $x_5 =$ SBP, and $x_6 =$DBP. The estimators $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, respectively, are the sample mean vector and

sample covariance matrix of $\boldsymbol{X} = (X_1, X_2, \cdots, X_6)$, computed from a random sample of $\boldsymbol{X}$. The constants n_1 and n are defined earlier in Section 2, and $\chi^2_6(\alpha)$ is the 100αth percentile of a chi-squared distribution with 6 degrees of freedom. In Section 3, we provide a case study to illustrate computations of n_1, n, $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$. We describe the structure of Framework 1 in Appendix B.

The ellipsoidal boundary has the advantage of having an exact probability that values of the six components of a subject falls inside the boundary. However, it has five disadvantages: (i) The normality assumption is essential; (ii) Any easy interpretation of MetS criteria would usually contain statements of individual components but such statements are difficult to obtain from the ellipsoidal boundary; (iii) The ellipsoidal region $R(B_1)$ is a multivariate generalization of a one-dimensional two-sided interval (with length depending on a Student t percentile) but an "one-sided" region is more appropriate for MetS criteria because only one-sided extreme values (too big or too small) of MetS components are associated with stroke or DM, (iv) The ellipsoidal boundary considers SBP and DBP as two components, but they are usually considered together as a single risk factor (even with different cutoff points) for stroke and DM; furthermore the cutoff points of WC and HDLC for males and females are the same; and (v) When an individual has only one component with an extreme value and the others normal, such individual may be classified to have MetS. These disadvantages may cause difficulties in MetS diagnoses and explanations.

2.2 Framework 2: Bonferroni-type cutoff points

To overcome the first four disadvantages of Framework 1, we propose here Framework 2 to compute the cutoff points of MetS criteria. Specifically, to avoid the fourth disadvantage of Framework 1, Framework 2 views SBP and DBP together as one

blood-pressure-type criterion but each with a different cutoff point. Furthermore, Framework 2 sets the cutoff points of WC and HDLC differently for males and females due to the gender effect. To avoid the second and third disadvantages of Framework 1, Framework 2 consists of five criteria corresponding to five simultaneous one-sided intervals for the six components—WC, TG, FG, HDLC, SBP, and DBP. (Since SBP and DBP are considered together as one criterion, there are only five criteria for the six components.) To avoid the first disadvantage of Framework 1, Framework 2 uses Bonferroni corrections to compute the cutoff points of the six components. In Framework 2, a subject is classified as having MetS if the subject's measurement has any component lying beyond the associated cutoff point. Although Framework 2 eliminates the first four disadvantages of Framework 1, it still contains the last disadvantage and lose the advantage of Framework 1.

The MetS criteria by Framework 2 require at least one of the following:

- $WC \geq b_{1f}$ (female), $\geq b_{1m}$ (male)
- $TG \geq b_2$
- $FG \geq b_3$
- $HDLC < b_{4f}$ (female), $< b_{4m}$ (male)
- $SBP \geq b_{5S}$ or $DBP \geq b_{5D}$

The cutoff points b_{1f}, b_{1m}, b_2, and b_3 are the $100(1-(n-n_1)(5n)^{-1})$th percentiles of WC for the female population, WC for the male population, TG, and FG, respectively. Cutoff points b_{4f} and b_{4m} are the $100(n-n_1)(5n)^{-1}$th percentiles of HDLC for the female and male populations, respectively. Finally, b_{5S} and b_{5D} are the $100(1-(n-n_1)(10n)^{-1})$th percentile of SBP and DBP, respectively. These percentiles can be estimated using observed data as shown in Section 3. The derivation of Framework 2 is illustrated in Appendix C.

This boundary can overcome the first four disadvantages of Framework 1. Specifically, the MetS criteria of Framework 2 consist of five one-sided-interval criteria with cutoff points for the six components. Statements for each individual component can

be made. Using the Bonferroni correction, the value of cutoff point only depends on the marginal distribution of component and hence, the joint distribution of these components is not restricted to multivariate normal. Each interval is one-sided, either upper or lower to match the component's support that is ordinarily considered "normal". However, this region has the disadvantage of possibly being too big because the probability of measurement lying inside of region might be much higher than the target probability. Furthermore, Framework 2 still retains the last disadvantage of Framework 1 (a subject is classified as having MetS even if only one MetS criterion is met).

2.3 Framework 3: modification of current popular definitions

Framework 3 modifies the updated-NCEP and consensus definitions by considering all criteria simultaneously. Framework 3 is designed to avoid the five disadvantages but keep the advantage of Framework 1. Like Framework 2, Framework 3 views SBP and DBP as one blood-pressure-type criterion, consists of five criteria corresponding to six simultaneous one-sided intervals for the six components, sets the cutoff points of WC and HDLC by considering the gender effect, and does not assume normality. However, in Framework 3, a subject is classified as having MetS if three or more out of the five criteria are satisfied.

Most current MetS definitions, including the updated-NCEP and consensus definitions, assume that the six components—WC, TG, FG, HDLC, SBP and DBP—are independent, meaning that a component with an extremely "bad" value has no influence on the other components. We found no literature indicating that any of the current MetS definitions considers the dependence among the six components. In practice, the six components are dependent. For example, the obesity usually results

in low HDLC and large WC, TG, FG, SBP, and DBP. Framework 3 is a modification of the updated-NCEP and consensus definitions and hence, the independence assumption is also kept. Nevertheless, Framework 3 determines the cutoff points by considering all criteria simultaneously (rather than one at a time) and matching the stroke/DM prevalence rate for better association.

The MetS criteria by Framework 3 have the same form as those by Framework 2 but with different cutoff points. Specifically, Framework-3 criteria require at least *three* of the following:

- $WC \geq b_{1f}$ (female), $\geq b_{1m}$ (male)
- $TG \geq b_2$
- $FG \geq b_3$
- $HDLC < b_{4f}$ (female), $< b_{4m}$ (male)
- $SBP \geq b_{5S}$ or $DBP \geq b_{5D}$

The cutoff points b_{1f}, b_{1m}, b_2, and b_3 are the $100(1-q)$th percentiles of WC for the female population, WC for the male population, TG, and FG, respectively. The constant q satisfies the following equation:

$$\sum_{i=0}^{2} \binom{5}{i} q^i (1-q)^{5-i} = n_1/n.$$

Cutoff points b_{4f} and b_{4m} are the $100q$th percentiles of HDLC for the female and male populations, respectively. Finally, b_{5S} and b_{5D} are the $100\sqrt{1-q}$th percentile of SBP and DBP, respectively. These percentiles can be estimated using observed data as shown in Section 3. The derivations of Framework 3 is illustrated in Appendix D.

3 Case Study with LIONS Data

This section evaluates the MetS criteria defined by the three statistical frameworks in Section 2 and compares them with the updated-NCEP and consensus criteria.

The study population includes subjects that participated in the Li-Shin Outreaching Neighborhood Screening (LIONS) program during 2006 and 2011. The LI-

ONS program has been conducted by Taiwan Landseed Hospital since 2006 and is a community-based (in Pingjen City of Taoyuan County) cancer and chronic-disease-oriented health screening project. In the study population, 6003 subjects with complete data associated with MetS risk factors at their first visits are included, such as gender, age, height, body weight, WC, TG, HDLC, FG, SBP, DBP, body mass index (BMI), the history of having stroke attacks and/or DM, and the history of drug treatments for hypertension, elevated glucose, and/or hyperlipidemia. However, due to lack of information on the drug treatment of hyperlipidemia, we are not sure whether the drug treatment is for elevated TG or cholesterol. There are 4877 (out of 6003) subjects that experienced no stroke and no DM before their first visits in the LIONS program.

Excluding subjects having only one visit during 2006 and 2011, 3915 subjects are included in this case study. (That is, $n = 3915$.) The numbers of these 3915 subjects having two, three, four, five, and six visits during 2006 and 2011 are 750, 1682, 1428, 54, and 1, respectively. The mean (standard deviation) of a subject's age (year old) at his/her first visit is 52.5 (12.6). Among these 3915 subjects, 121 subjects have experienced stroke or DM during the study period and hence, $n_1 = 3794$ and $n_2 = n - n_1 = 121$. The observed prevalence rate for having stroke or DM is 3.09%.

Six components of MetS are considered: WC, TG, HDLC, FG, SBP, and DBP, as mentioned in Section 2. SBP and DBP are considered separately in Framework 1 but considered together as one blood-pressure-type criterion in Frameworks 2 and 3. Frameworks 2 and 3 also consider the gender effect for WC and HDLC. Furthermore, since Framework 1 requires the multivariate normal assumption, the log transformation is used for TG, FG, and HDLC data for achieving the normality. That is, in Framework 1, $X_1, \ldots, X_6$ denote ln(TG), ln(FG), WC, ln(HDLC), SBP, and DBP, respectively. No transformation is needed for Frameworks 2 and 3.

In this study, the MetS cutoff points for Frameworks 1 to 3 are computed using measurements of the 3915 subjects joining the LIONS program. Table 1 list the ellipsoid boundary of Framework 1 and cutoff points of Frameworks 2 and 3. The cutoff points for the updated-NCEP and consensus definitions are also listed for comparisons. Each of the 3915 subjects are then identified to have MetS or not using these MetS criteria. Numerical results are listed in Tables 2 and 3.

Table 2 shows the cross classification of the stroke/DM and MetS defined by the updated-NCEP and consensus definitions. Table 2 also shows the odds ratio and its 95% confidence interval (using the method by [25]). The odds ratio is the odds of getting stroke or DM within subjects with MetS divided by the same odds but within subjects without MetS. MetS definition with a larger odds ratio implies its higher association of stroke and DM. For the stroke/DM classification, a subject is classified in the *Yes* row (5th row) if this subject has stroke or DM and classified in the *No* row (4th row), otherwise. Since there are 3794 subjects have no stroke and no DM, the row sum of the *No* category for each definition is 3794; similarly, the row sum of the *Yes* category for each definition is 121. For the MetS classification, if a subject satisfies the updated-NCEP definition, the count under the *Yes* category in Column 4 increases by 1; otherwise, the count under the *No* category in Column 3 increases by 1. The row number where the count increment occurs depends on whether this subject has stroke/DM or not. Columns 5 and 6 are the same as Columns 3 and 4 but are for the consensus definition. Notice that due to the LIONS data limitation, we do not use the drug treatments as an alternative cutoff point to check TG and HDLC for the updated-NCEP and consensus definitions.

For the updated-NCEP definition, Table 2 shows that among the 3794 subjects having neither stroke or DM, 358 subjects are identified to have MetS, while among the 121 subjects with stroke or DM, 49 subjects are identified to have MetS. The odds

ratio is 6.5 (= (49/358)/(72/3436)), with 95% confidence interval $(4.5, 9.5)$. That is, based on the updated-NCEP definition, the estimated odds of getting stroke/DM for the population with MetS is 6.5 times as large as the estimated odds for the population without MetS. For the consensus definition, among the 3794 subjects without stroke or DM, 617 subjects are identified to have MetS, while among the 121 subjects with stroke or DM, 60 subjects are identified to have MetS. The odds ratio is 5.1 (= (60/617)/(61/3177)), with 95% confidence interval $(3.5, 7.3)$. The odds ratio of the consensus definition is lower than that of the updated-NCEP definition. Notice that the consensus definition has looser MetS criteria (with lower WC cutoff points), and hence, more subjects are classified to have MetS. Since there are only 121 subjects have stroke/DM and 3794 haven't, a MetS definition with more subjects classified to have MetS tends to result in a lower odds ratio.

Table 3 provides the odds-ratios results for Frameworks 1 to 3. The structure of Table 3 is the same as that of Table 2. Notice that all the three frameworks do not consider the drug treatments effect on TG, FG, HDLC, or blood pressure (BP). Although some subjects joining the LIONS program were on drug treatments for elevated TG, BP, or glucose, or for reduced HDLC, their measurements of TG, FG, HDLC, and BP are still used for MetS identification. The odds ratio (confidence interval) for Frameworks 1 to 3 are 11.2 (7.5, 16.7), 7.9 (4.6, 13.7), and 7.3 (4.9, 11.0), respectively. Framework 1 has the highest odds ratio and the odds ratio of Framework 2 is a little bit higher than that of Framework 3. Framework 1 performs best might be because it considers dependence among the six components. Comparing Frameworks 1 and 3, we see that the number of subjects classified to have MetS for Framework 1 is lower than that of Framework 3, but the number of subjects having stroke/DM among the MetS subjects is higher. This result shows that Framework 1 has higher association with stroke/DM. Framework 2 has the lowest number of subjects classified

to have MetS because the Bonferroni correction guarantees that the probability of without MetS for an individual is at least n_1/n and hence, results in the tightest criteria among the three frameworks. Although Framework 2 classifies fewest MetS subjects, the number of stroke/DM subjects within this group is also much lower and hence, the odds ratio of Framework 2 is not higher than that of Framework 1.

4 Conclusions and Discussions

MetS is composed of interrelated risk factors of CVD. Various diagnostic criteria for metabolic syndrome have been proposed by different organizations over the past two decades. However, the prediction ability on the prevalence, incidence, and outcome measurement risk of CVD for a MetS definition depends on the definition itself and various populations and hence, good criteria with proper cutoff points of MetS are important.

This is a pioneer work to propose statistical frameworks to compute cutoff points of MetS criteria for a certain population. Six components (WC, TG, FG, HDLC, SBP, and DBP) are considered simultaneously, rather than one at a time. Based on the proposed statistical frameworks, rational comparisons of risk of CVD for different populations (with different cutoff points) can be made.

The proposed three frameworks are compared with the two most widely used definitions, the updated-NCEP and consensus definitions, using the LIONS program data. Our numerical results show that the odds ratios of the three proposed frameworks are higher than those of the updated-NCEP and consensus definitions; hence, the proposed frameworks seem to show higher association between MetS and DM/stroke. The reason might be that the three frameworks set the cutoff points so that the MetS prevalence rate matches the stroke/DM prevalence rate, and hence the three

frameworks provide higher association of stroke and DM. Despite lack of evidence on prediction ability, we conjecture that the proposed frameworks could provide better prediction of DM/stroke because of its using DM/stroke prevalence rate. Furthermore, the odds ratios of the updated-NCEP and consensus definitions are not much lower than those of Frameworks 2 and 3, meaning that the updated-NCEP and consensus definitions are useful in practice.

We further compare these methods by considering eight desired properties of a MetS definition, in terms of MetS's rationality and real-world applications. Table 4 illustrates these eight properties. The properties that the three frameworks, updated NCEP and consensus definitions satisfy are also shown. Framework 3, as well as the updated NCEP and consensus definitions, satisfy seven out of the eight properties. This indicates that Framework 3 with statistical structure is easy to implement as the two most popular MetS definitions.

The logic of the proposed statistical frameworks may be applied on defining other syndromes or diseases—e.g., toxic shock syndrome, Kawasaki disease, rheumatoid arthritis, and systemic lupus erythematosus—to provide more rational criteria based on the associated outcome morbidity and mortality. Details (e.g., the chosen components and number of criteria) might change but the main ideas would still hold.

Acknowledgments

We thank Taiwan Landseed Hospital for providing LIONS data. We also thank referees for comments that improved the presentation of this paper. This research was supported by Landseed Hospital under Contract LH-2013-09. The second author is partially supported by Taiwan Ministry of Science and Technology under Contract NSC 102-410-H-033-035.

References

[1] Reaven GM. Role of insulin resistance in human disease. *Diabetes* 1988; **37**: 1595–1607.

[2] Kylin E. Studien ueber das Hypertonie-Hyperglykamie Hyperurikamiesyndrom. *Zentralbl Innere Medizin* 1923; **44**: 105–127.

[3] Vague J. The degree of masculine differentiation of obesities: a factor determining predisposition to diabetes, atherosclerosis, gout, and uric calculous disease. *Am J Clin Nutrition* 1956; **4**: 20–34.

[4] Mottillo S, Filion KB, Genest J, Joseph L, Pilote L, Poirier P, Rinfret S, Schiffrin EL, Eisenberg MJ. The metabolic syndrome and cardiovascular risk: A systematic review and meta-analysis. *Journal of the American College of Cardiology* 2010; **56**(14): 1113–1132.

[5] Ford ES, Giles WH, Dietz WH. Prevalence of metabolic syndrome among US adults: Findings from the third National Health and Nutrition Examination Survey. *JAMA* 2002; **287**(3): 356–359.

[6] Expert Panel on Detection and Evaluation of Treatment of High Blood Cholesterol in Adults. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). *JAMA* 2001; **285**: 2486–2497.

[7] Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med* 1998; **15**: 539–553.

[8] Balkau B, Charles MA. Comment on the provisional report from the WHO consultation. European Group for the Study of Insulin Resistance (EGIR). *Diabet Med* 1999; **16**: 442–443.

[9] Einhorn D, Reaven GM, Cobin RH, Ford E, Ganda OP, Handelsman Y, Hellman R, Jellinger PS, Kendall D, Krauss RM, Neufeld ND, Petak SM, Rodbard HW, Seibel JA, Smith DA, Wilson PW. American College of Endocrinology position statement on the insulin resistance syndrome. *Endocr Pract* 2003; **9**: 237–252.

[10] Grundy SM, Brewer HB Jr, Cleeman JI, Smith SC Jr, Lenfant C. Definition of metabolic syndrome: Report of the National Heart, Lung, and Blood Institute/American Heart Association conference on scientific issues related to definition. *Arterioscler Thromb Vasc Biol* 2004; **24**: e13–e18.

[11] Grundy SM, Cleeman JI, Daniels SR, Donato KA, Eckel RH, Franklin BA, Gordon DJ, Krauss RM, Savage PJ, Smith SC Jr, Spertus JA, Costa F. American Heart Association; National Heart, Lung, and Blood Institute: Diagnosis and management of the metabolic syndrome. An American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. Executive summary. *Cardiology in Review* 2005; **13**: 322–327.

[12] Alberti KG, Zimmet P, Shaw. IDF Epidemiology Task Force Consensus Group. The metabolic syndrome: a new worldwide definition. *Lancet* 2005; **366**: 1059–1062.

[13] Alberti KG, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, Fruchart JC, James WP, Loria CM, Smith SC Jr. International Diabetes Federation Task Force on Epidemiology and Prevention, National Heart, Lung, and

Blood Institute, American Heart Association, World Heart Federation, International Atherosclerosis Society, International Association for the Study of Obesity: Harmonizing the metabolic syndrome: A joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation* 2009; **120**: 1640–1645.

[14] Moebus S, Hanisch JU, Neuhäuser M., Aidelsburger P, Wasem J, Jöckel KH. Assessing the prevalence of the Metabolic Syndrome according to NCEP ATP III in Germany: feasibility and quality aspects of a two step approach in 1550 randomly selected primary health care practices. *German Medical Science* 2006; **4**: Doc07.

[15] Kassi E, Pervanidou P, Kaltsas G, Chrousos G. Metabolic syndrome: Definitions and controversies. *BMC Medicine* 2011; **9**: 48–60.

[16] Thaman RG, Arora GP. Metabolic syndrome: definition and pathophysiology - the discussion goes on!. *Journal of Physiology and Pharmacology Advances* 2013; **3**(3): 48–56.

[17] Prabhakaran D, Reddy KS. The metabolic syndrome: looking beyond the debates. *Clinical Pharmacology and Therapeutics* 2011; **90**: 19–21.

[18] Simmons RK, Alberti KGMM, Gale EAM, Colagiuri S, Tuomilehto J, Qiao Q, Ramachandran A, Tajima N, Brajkovich Mirchov I, Ben-Nakhi A, Reaven G, Hama Sambo B, Mendis S, Roglic G. The metabolic syndrome: useful concept or clinical tool? Report of a WHO Expert Consultation. *Diabetologia* 2010; **53**: 600–605.

[19] Reaven GM. The metabolic syndrome: requiescat in pace. *Clinical Chemistry* 2005; **51**(6): 931–938.

[20] Reaven GM. The metabolic syndrome: is this diagnosis necessary? *American Journal of Clinical Nutrition* 2006; **83**: 1237–1247.

[21] Reaven GM. The metabolic syndrome: time to get off the merry-go-round? *Journal of internal medicine* 2011; **269**: 127–136.

[22] Parikh RM, Joshi SR, Menon PS, Shah NS. Index of central obesity–A novel parameter. *Med Hypotheses* 2007; **68**: 1272–1275.

[23] Parikh RM, Mohan V. Changing definitions of metabolic syndrome. *Indian Journal of Endocrinology and Metabolism* 2012; **16**(1): 7–12.

[24] Johnson RJ, Stenvinkel P, Martin SL, Jani A, Sanchez-Lozada LG, Hill JO, Lanaspa MA. Redefining metabolic syndrome as a fat storage condition based on studies of comparative physiology. *Obesity (Silver Spring)* 2013; **21**(4): 659–664.

[25] Morris JA, Gardner MJ. Calculating confidence intervals for relative risks (odds ratios) and standardised ratios and rates. *Statistics in Medicine* 1988; **296**: 1313–1316.

[26] Sen, P. K.; Singer, J. M. *Large Sample Methods in Statistics: An Introduction with Applications* 1993; Chapman & Hall, New York-London.

Appendices

A Maximum Likelihood Approach

Let $\boldsymbol{X}$, a 6×1 random vector, denote the measurement of the six components, WC, TG, FG, HDLC, SBP, and DBP, for an individual in a certain population. Furthermore, let Ω be the sample space of $\boldsymbol{X}$, i.e., the collection of all possible observations $\boldsymbol{x}$ of $\boldsymbol{X}$.

The cutoff boundary (called B) divides the sample space Ω into two exclusive and exhaustive regions: (i) $R(B)$: the set of $\boldsymbol{x}$ values for which we classify the corresponding subjects as normal (i.e., without MetS), and (ii) $\overline{R}(B)$: the set of $\boldsymbol{x}$ values for which we classify the corresponding subjects as abnormal (i.e., with MetS). Any subject with an observed $\boldsymbol{x}$ value falling outside $R(B)$ is considered to have MetS.

Consider a population of size n. Given a cutoff boundary B, let N_1 denote the number of subjects whose measurements $\boldsymbol{X}$ lying in Region $R(B)$ (i.e., the number of normal subjects in a population of size n). If each subject in the population is independently and equally likely to be a normal subject with probability $p = \mathrm{P}\{\boldsymbol{X} \in R(B)\}$, then N_1 follows a binomial distribution with probability p of success. The probability mass function of N_1 is therefore

$$\mathrm{P}\{N_1 = i\} = \binom{n}{i} p^i (1-p)^{n-i}, \quad i = 0, 1, \cdots, n. \tag{1}$$

Suppose that n_1 out of n subjects in the population are observed to have no stroke and no DM. Given n_1, the likelihood function of B is

$$L(B|n_1) = \mathrm{P}\{N_1 = n_1\} = \binom{n}{n_1} p^{n_1} (1-p)^{n-n_1}.$$

The value of B that maximizes the likelihood $L(B|n_1)$ satisfies the equation $\partial L(B|n_1)/\partial B = 0$, or equivalently $\partial \ln L(B|n_1)/\partial B = 0$. Since the logarithm of $L(B|n_1)$ is

$$\ln L(B|n_1) = \ln \binom{n}{n_1} + n_1 \ln(p) + (n - n_1)\ln(1-p),$$

we have

$$\frac{\partial \ln L(B|n_1)}{\partial B} = \frac{n_1(\frac{\partial p}{\partial B})}{p} + \frac{(n-n_1)(\frac{\partial (1-p)}{\partial B})}{1-p}.$$

By setting $\partial \ln L(B|n_1)/\partial B = 0$, we obtain

$$n_1(1-p)\Big(\frac{\partial p}{\partial B}\Big) - (n-n_1)p\Big(\frac{\partial p}{\partial B}\Big) = \Big(\frac{\partial p}{\partial B}\Big)(n_1 - np) = 0.$$

Hence,

$$p = n_1/n. \tag{2}$$

Since MetS defines an assemblage of abnormality associated with risk factors for stroke and DM, the maximum likelihood approach is to maximize the conformity of the computed boundary with the observed stroke/DM prevalence rate $(1 - n_1/n)$, by assuming that the observed stroke/DM prevalence rate is equivalent to the MetS prevalence rate. The objective of this article is to proposed a solution for the cutoff boundary B (given $p = n_1/n$) by the three proposed frameworks. Frameworks 1 to 3 use different assumptions on the joint distribution of $\boldsymbol{X}$ to determine the boundary B as described in Sections 2.1 to 2.3, respectively.

B Ellipsoidal Boundary

Since the random vector $\boldsymbol{X}^{6\times 1}$ follows a multivariate normal distribution, the quantity $\boldsymbol{(X-\mu)'\Sigma^{-1}(X-\mu)}$ follows a chi-squared distribution with 6 degrees of freedom. Using this relation, Framework 1 defines the boundary and hence, normal region as

$$B_1 = \{\boldsymbol{x} : \boldsymbol{(x-\mu)'\Sigma^{-1}(x-\mu)} = c\} \tag{3}$$

and

$$R(B_1) = \{\boldsymbol{x} : (\boldsymbol{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) \leq c\}$$

for a constant c. Since c determines the size of the boundary B_1, finding B_1 is the same as finding c.

Recall that the boundary B_1, and therefore, c, are set to satisfy $p = \mathrm{P}\{\boldsymbol{X} \in R(B_1)\} = n_1/n$ as shown in Equation (2). By setting

$$P\{(\boldsymbol{X} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{X} - \boldsymbol{\mu}) \leq c\} = n_1/n,$$

we then have

$$c = \chi^2_6(n_1/n),$$

where $\chi^2_6(\alpha)$ is the 100αth percentile of a chi-squared distribution with 6 degrees of freedom. Therefore, Framework 1 sets the boundary B_1 to be

$$B_1 = \{\boldsymbol{x} : (\boldsymbol{x} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) = \chi^2_6(n_1/n)\}.$$

The shape of B_1 is an ellipsoid. An individual with an observed measurement $\boldsymbol{x}$ lying inside the ellipsoid is considered normal (without MetS); otherwise, the subject is considered abnormal (with MetS).

In practice, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are unknown and need to be estimated. Suppose that m observations $\{\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_m\}$ are randomly selected from the target population. We can estimate $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ by unbiased estimators

$$\hat{\boldsymbol{\mu}} = \sum_{i=1}^{m} \boldsymbol{X}_i/m$$

and

$$\hat{\boldsymbol{\Sigma}} = \sum_{i=1}^{m} (\boldsymbol{X_i} - \hat{\boldsymbol{\mu}})(\boldsymbol{X_i} - \hat{\boldsymbol{\mu}})'/(m-1),$$

respectively. If $\boldsymbol{X}$ follows a multivariate normal distribution and independent of $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_m$, then $(\boldsymbol{X} - \hat{\boldsymbol{\mu}})'\hat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{X} - \hat{\boldsymbol{\mu}})$ is approximately chi-squared distributed

with 6 degrees of freedom by Slutsky Theorem [26]. The ellipsoidal boundary B_1 and the normal region are therefore

$$B_1 = \{\boldsymbol{x} : \boldsymbol{(x-\hat{\mu})'\hat{\Sigma}^{-1}(x-\hat{\mu})} = \chi^2_6(n_1/n)\},$$

and

$$R(B_1) = \{\boldsymbol{x} : \boldsymbol{(x-\hat{\mu})'\hat{\Sigma}^{-1}(x-\hat{\mu})} \leq \chi^2_6(n_1/n)\}.$$

C Bonferroni-type cutoff points

For convenience, let $X_1, X_2, \cdots, X_6$ denote the measurements of WC, TG, FG, HDLC, SBP and DBP, respectively. Let L_i denote the one-sided interval indicating the "normal" range for the ith component, $i = 1, \ldots, 6$. If a large value of the ith components (e.g., WC, TG, FG, SBP, or DBP) predisposes a subject to CVD or DM, L_i is a lower one-sided interval, said $(0, b)$ for some b value. Similarly, if a small value of the ith component (e.g., HDLC) predisposes a subject to CVD or DM, L_i is an upper one-sided interval, said (b, ∞) for some b value. Given cutoff points, the normal region for Framework 2 is

$$R(B_2) = \{\boldsymbol{x} : \ x_1 \in L_1, x_2 \in L_2, \cdots, x_6 \in L_6\}.$$

A subject is considered to have MetS if the subject's measurement $\boldsymbol{X}$ has one or more X_i values falling outside the interval L_i (or equivalently, beyond the cutoff point b_i), $i = 1, \ldots, 6$.

Using the maximum likelihood approach, the cutoff points are set so that the probability of having MetS is the stroke/DM prevalence rate $(1 - n_1/n)$ for a target population of size n. That is,

$$\mathrm{P}\{X_1 \in L_1, X_2 \in L_2, \cdots, X_6 \in L_6\} = n_1/n. \tag{4}$$

Solving Equation (4) for cutoff points is difficult because $X_1, \ldots, X_6$ are not independent.

Framework 2 uses Bonferroni corrections to set the values of cutoff points. By Boole's inequality, we have

$$\begin{aligned} & \mathrm{P}\{X_1 \in L_1, X_2 \in L_2, \cdots, X_6 \in L_6\} \\ = \; & 1 - \mathrm{P}\{X_1 \notin L_1 \quad \text{or} \quad X_2 \notin L_2 \quad \text{or} \cdots \quad \text{or} \quad X_6 \notin L_6\} \\ \geq \; & 1 - \Big(\sum_{i=1}^{4} \mathrm{P}\{X_i \notin L_i\} \Big) - \mathrm{P}\{X_5 \notin L_5 \text{ or } X_6 \notin L_6\} \end{aligned}$$

By assuming equal probabilities for the five events $(X_1 \notin L_1)$,..., $(X_4 \notin L_4)$, and $[(X_5 \notin L_5)$ or $(X_6 \notin L_6)]$, we have

$$\mathrm{P}\{X_1 \in L_1, X_2 \in L_2, \cdots, X_6 \in L_6\} \geq 1 - 5\mathrm{P}\{X_1 \notin L_1\}.$$

Furthermore, by setting

$$\mathrm{P}\{X_i \notin L_i\} = \frac{n - n_1}{5n} \text{ for } i = 1, \ldots, 4, \text{ and } \mathrm{P}\{X_5 \notin L_5 \text{ or } X_6 \notin L_6\} = \frac{n - n_1}{5n} \quad (5)$$

(or equivalently, $\mathrm{P}\{X_i \in L_i\} = 1 - (n - n_1)(5n)^{-1}$, for $i = 1, \ldots, 4$, and $\mathrm{P}\{X_5 \in L_5, X_6 \in L_6\} = 1 - (n - n_1)(5n)^{-1}$), we have

$$\mathrm{P}\{X_1 \in L_1, X_2 \in L_2, \cdots, X_6 \in L_6\} \geq n_1/n.$$

The cutoff points can be computed using Equation (5). Let b_{1f}, b_{1m}, b_2, b_3, b_{4f}, b_{4m}, b_{5S}, and b_{5D} denote the cutoff points of WC for the female population, WC for the male population, TG, and FG, HDLC for the female population, HDLC for the male population, SBP, and DBP, respectively. Equation (5) shows that b_{1f}, b_{1m}, b_2, and b_3 are the $100(1 - (n - n_1)(5n)^{-1})$th percentiles of WC for the female population, WC for the male population, TG, and FG, respectively; b_{4f} and b_{4m} are the $100(n - n_1)(5n)^{-1}$th percentiles of HDLC for the female and male populations, respectively.

To determine the cutoff points b_{5S} and b_{5D} for X_5 (SBP) and X_6 (DBP), respectively, we further assume that $\mathrm{P}\{X_5 \notin L_5\} = \mathrm{P}\{X_6 \notin L_6\}$. Since the cutoff points b_5 and b_6 need to satisfy

$$\frac{n - n_1}{5n} = \mathrm{P}\{X_5 \notin L_5 \text{ or } X_6 \notin L_6\} \leq \mathrm{P}\{X_5 \notin L_5\} + \mathrm{P}\{X_6 \notin L_6\} \ , \tag{6}$$

we let

$$\mathrm{P}\{X_5 \notin L_5\} = \mathrm{P}\{X_6 \notin L_6\} = \frac{n - n_1}{10n} \ .$$

Therefore, the cutoff point b_{5S} and b_{5D} are the $100(1-(n-n_1)(10n)^{-1})$th percentiles of SBP and DBP, respectively.

D Modification of current popular definitions

As in Appendix C, let $X_1, \cdots, X_6$ denote the measurements of WC, TG, FG, HDLC, SBP and DBP, respectively. Let b_{1f}, b_{1m}, b_2, b_3, b_{4f}, b_{4m}, b_{5S}, and b_{5D} denote the cutoff points of WC for the female population, WC for the male population, TG, and FG, HDLC for the female population, HDLC for the male population, SBP, and DBP, respectively. Let L_i denote the one-sided intervals indicating the "normal" range for the ith component. Recall that L_i has form (b, ∞) for $i = 4$ (i.e., HDLC) and $(0, b)$, otherwise, for some value b.

If a subject's measurements of the six components satisfy k or more out of the five criteria, this subject is classified to have MetS. Here, $k = 3$. Specifically, the abnormal region is

$$\begin{aligned} \bar{R}(B_3) \ = \ & \{\boldsymbol{x} : \text{the number of the following five criteria} \\ & (x_1 \notin L_1), \cdots, (x_4 \notin L_4), \text{ and } [(x_5 \notin L_5) \text{ or } (x_6 \notin L_6)] \\ & \text{satisfied is at least } k\} \ . \end{aligned} \tag{7}$$

As in Framework 2, the boundary B_3 depends on the cutoff points.

Let the indicator variable I_j, $j = 1, \ldots, 5$, denote whether the jth criteria is satisfied, i.e.,

$$I_j = \begin{cases} 1\ , & \text{if } X_j \notin L_j \\ 0\ , & \text{if } X_j \in L_j \end{cases} \quad \text{for } \ j = 1, ..., 4,$$

and

$$I_5 = \begin{cases} 1\ , & \text{if } \ x_5 \notin L_5 \text{ or } \ x_6 \notin L_6 \\ 0\ , & \text{else} \end{cases} .$$

A subject is considered to have MetS if $\sum_{j=1}^{5} I_j \geq k$.

Given cutoff points $\boldsymbol{b} = (b_{1f}, b_{1m}, b_2, b_3, b_{4f}, b_{4m}, b_{5S}, b_{5D})$, the probability that a subject has MetS defined by Framework 3 is then

$$\mathrm{P}\Big\{ \sum_{j=1}^{5} I_j \geq k \Big\} = \sum_{i=k}^{5} \mathrm{P}\Big\{ \sum_{j=1}^{5} I_j = i \Big\}.$$

To determine values of the cutoff points $\boldsymbol{b}$, Framework 3 further makes the assumption that the probabilities of satisfying each of the five criteria are equal, i.e., the cutoff points $\boldsymbol{b}$ satisfying $q_1 = \cdots = q_5 \equiv q$, where $q_j = \mathrm{P}(I_j = 1)$, $j = 1, \ldots, 5$. All i-combinations of the 5 criteria have a same probability of being abnormal, $i = k, \ldots, 5$. That is, if $\{j_1, \ \ldots, \ j_i\}$ is a subset of $\{1, \ldots, 5\}$ with size i and $j_1 \neq \ \ldots \ \neq j_i$, $\mathrm{P}\{I_j = 1 \text{ for all } j \in \{j_1, \ldots j_i\} \text{ and } I_j = 0 \text{ for all } j \notin \{j_1, \ldots j_i\}\}$ is the same for all $\binom{5}{i}$ possible subsets $\{j_1, \ \ldots, \ j_i\}$ of $\{1, \ldots, 5\}$.

The probability that a subject satisfies the MetS criteria of Framework 3 is

$$\begin{aligned} \mathrm{P}\Big\{ \sum_{j=1}^{5} I_j \geq k \Big\} &= \sum_{i=k}^{5} \mathrm{P}\Big\{ \sum_{j=1}^{5} I_j = i \Big\} \\ &= \sum_{i=k}^{5} \binom{5}{i} \mathrm{P}\{I_1 = 1, \ldots, I_i = 1, I_{i+1} = 0, \ldots, I_5 = 0\} \\ &= \sum_{i=k}^{5} \binom{5}{i} q^i (1-q)^{5-i} \ , \end{aligned} \tag{8}$$

if I's are mutually independent. Since the probability of having MetS is set to be the stroke/DM prevalence rate $(1 - n_1/n)$, using the maximum likelihood approach, the value of q satisfies

$$\sum_{i=k}^{5} \binom{5}{i} q^i (1-q)^{5-i} = 1 - n_1/n$$

or equivalently

$$\sum_{i=0}^{k-1} \binom{5}{i} q^i (1-q)^{5-i} = n_1/n. \tag{9}$$

The common probability q for the five criteria can be obtained by solving Equation (9). For the special case of $k = 1$, solving Equation (9) is easy: $q = 1 - (n_1/n)^{1/5}$. If k is larger than 1, q needs to be computed numerically.

Recall that $\mathrm{P}(X_j \notin L_j) = \mathrm{P}(I_j = 1) = q$ for $j = 1, \ldots, 4$. Hence, cutoff points b_{1f}, b_{1m}, b_2, and b_3 are the $100(1-q)$th percentiles of WC for the female population, WC for the male population, TG, and FG, respectively. Cutoff points b_{4f} and b_{4m} are the $100q$th percentiles of HDLC for the female and male populations, respectively.

The cutoff points b_{5S} and b_{5D} for SBP and DBP, respectively, satisfy

$$\mathrm{P}\{X_5 \notin L_5 \text{ or } X_6 \notin L_6\} = q, \tag{10}$$

i.e.,

$$\mathrm{P}\{X_5 \notin L_5\} + \mathrm{P}\{X_6 \notin L_6\} - \mathrm{P}\{X_5 \notin L_5 \text{ and } X_6 \notin L_6\} = q.$$

If $\mathrm{P}\{X_5 \notin L_5\} = \mathrm{P}\{X_6 \notin L_6\} \equiv r$, and X_5 and X_6 are independent, then

$$r + r - r^2 = q.$$

Therefore,

$$r = 1 - \sqrt{1-q}.$$

Hence, the cutoff points b_{5S} and b_{5D} are the $100(1-r)$th percentiles of X_5 (SBP) and X_6 (DBP), respectively. In practice, the marginal distributions of $X_1, \ldots, X_6$ may be

unknown. Given a random sample of $\boldsymbol{X}$, the values of $\boldsymbol{b}$ can be estimated by sample quantiles.

Table 1: The MetS criteria of the updated-NCEP definition, consensus definition, and Frameworks 1 to 3

Definition	Criteria
Updated NCEP (AHA/NHLBI)	Any three of the following: • $WC \geq 88$ (female), ≥ 102 (male) • SBP ≥ 130 or DBP ≥ 85 (or use drugs with a history of hypertension) • TG ≥ 150 (or use drugs for elevated TG) • FG ≥ 100 (or use drugs for elevated glucose) • HDLC < 50 (female), < 40 (male) (or use drugs for reduced HDLC)
Consensus	Any three of the following: • $WC \geq 80$ (female), ≥ 90 (male) • SBP ≥ 130 or DBP ≥ 85 (or use drugs with a history of hypertension) • TG ≥ 150 (or use drugs for elevated TG) • FG ≥ 100 (or use drugs for elevated glucose) • HDLC < 50 (female), < 40 (male) (or use drugs for reduced HDLC)
Framework 1	Any point $\boldsymbol{x} = (x_1, \cdots, x_6)$ falling outside the ellipsoid $R(B_1)$: $R(B_1) = \{\boldsymbol{x} : \boldsymbol{(x - \hat{\mu})' \hat{\Sigma}^{-1} (x - \hat{\mu})} \leq 13.89\}$, where $x_1 = \ln(\text{HDLC})$, $x_2 = \ln(\text{TG})$, $x_3 = \ln(\text{FG})$, $x_4 = \text{WC}$, $x_5 =$ SBP, $x_6 =$DBP, and $\hat{\mu}$ and $\hat{\Sigma}$ are the sample mean vector and sample covariance matrix of the 3915 observations of $\boldsymbol{x}$.
Framework 2	Any one of the following: • $WC \geq 107$ (female), ≥ 110 (male) • SBP ≥ 194 or DBP ≥ 113 • TG ≥ 676 • FG ≥ 205 • HDLC < 34 (female), < 28 (male)
Framework 3	Any three of the following: • $WC \geq 86$ (female), ≥ 95 (male) • SBP ≥ 153 or DBP ≥ 92 • TG ≥ 184 • FG ≥ 98 • HDLC < 50 (female), < 41 (male)

Table 2: Cross classification of the stroke/DM and MetS based on the updated-NCEP and consensus definitions, as well as the odds ratio and its 95% confidence interval (in parentheses)

		Metabolic Syndrome				
		updated NCEP		Consensus		
		No	Yes	No	Yes	Total
Stroke/DM	No	3436	358	3177	617	3794
	Yes	72	49	61	60	121
	Total	3508	407	3238	677	3915
	odds ratio	6.5, (4.5, 9.5)		5.1, (3.5, 7.3)		

Table 3: Cross classification of stroke/DM and MetS defined by Frameworks 1 to 3, as well as the odds ratio and its 95% confidence interval (in parentheses)

		Metabolic Syndrome						
		Framework 1		Framework 2		Framework 3		
		No	Yes	No	Yes	No	Yes	Total
Stroke/DM	No	3610	184	3712	82	3571	223	3794
	Yes	77	44	103	18	83	38	121
	Total	3687	228	3815	100	3654	261	3915
	odds ratio	11.2, (7.5, 16.7)		7.9, (4.6, 13.7)		7.3, (4.9, 11.0)		

Table 4: Desired properties of a MetS definition

			Framework		
	updated NCEP	Consensus	1	2	3
(1) Needlessness of Normality assumption	√	√		√	√
(2) Diagnosis convenience	√	√		√	√
(3) One-sided cutoff region	√	√		√	√
(4) A combined criterion for SBP and DBP	√	√		√	√
(5) No extreme cut-off points	√	√	√		√
(6) Gender-specific cutoff points	√	√		√	√
(7) Model of dependence among components			√		
(8) Classification with more than one criterion	√	√			√