

· 现场调查 ·

台湾地区35~74岁健康体检人群代谢综合征发病风险预测模型的建立

杨兴华 陶秋山 孙凤 曹纯铿 詹思延

【摘要】目的 构建台湾地区35~74岁健康体检人群代谢综合征5年发病风险(个体化)预测模型。**方法** 在1997—2006年初次参加台湾美兆自动化健康体检机构(美兆健检)的35~74岁人群中,将随访满5年基线时无代谢综合征13 973人作为随访队列,并分为建模队列(用于建立5年发病预测模型)和验证队列(用于评估模型外部效度),采用logistic回归构建预测模型。以ROC曲线下面积(AUC)评价拟合优度,并将人群的预测风险概率进行风险等级划分。**结果** 去除基线患者后研究人群5年代谢综合征患病率为11.7%。纳入发病风险预测模型变量有年龄、糖尿病家族史、收缩压、空腹血糖、甘油三酯、高密度脂蛋白胆固醇、低密度脂蛋白胆固醇、总胆固醇、体重指数和血尿酸,建模队列建立预测模型的AUC为0.827(95%CI:0.814~0.839),验证队列的AUC分别为0.813(0.789~0.837)、0.826(0.800~0.852)、0.794(0.768~0.820)。将建模队列划分为4个风险等级后,提示个体发病概率≥17.6%者为中危人群,发病概率≥59.0%者为高危人群。**结论** 由美兆健检纵向数据库建立的5年代代谢综合征个体风险预测模型有较高的验证效度,对于体检人群5年代代谢综合征发病预测具有实用、可行的特点,预测模型对评估代谢综合征个体发病和群体监测均有较高应用价值。

【关键词】 代谢综合征; 风险预测模型; 纵向数据

Setting up a risk prediction model on metabolic syndrome among 35–74 year-olds based on the Taiwan MJ Health-checkup Database YANG Xing-hua¹, TAO Qiu-shan², SUN Feng², CAO Chun-keng³, ZHAN Si-yan². 1 Department of Epidemiology and Biostatistics, School of Public Health, Capital Medical University, Beijing 100069, China; 2 Department of Epidemiology and Biostatistics, School of Public Health, Health Science Center of Peking University; 3 MJ Health Management Organization, Taiwan

Corresponding author: ZHAN Si-yan, Email: siyan-zhan@bjmu.edu.cn

[Abstract] **Objective** This study aimed to provide an epidemiological modeling method to evaluate the risk of metabolic syndrome (MS) development in the coming 5 years among 35–74 year-olds from Taiwan. **Methods** A cohort of 13 973 subjects aged 35–74 years who did not have metabolic syndrome but took the initial testing during 1997–2006 was formed to derive a risk score which tended to predict the incidence of MS. Multivariate logistic regression was used to derive the risk functions and using the ‘check-up center’ (Taipei training cohort) as the overall cohort. Rules based on these risk functions were evaluated in the remaining three centers (as testing cohort). Risk functions were produced to detect the MS on a training sample using the multivariate logistic regression models. Started with those variables that could predict the MS through univariate models, we then constructed multivariable logistic regression models in a stepwise manner which eventually could include all the variables. The predictability of the model was evaluated by areas under curve (AUC) the receiver-operating characteristic (ROC) followed by the testification of its diagnostic property on the testing sample. Once the final model was defined, the next step was to establish rules to characterize 4 different degrees of risks based on the cut points of these probabilities, after being transformed into normal distribution by log-transformation. **Results** At baseline, the range of the proportion of MS was 23.9% and the incidence of MS in 5-years was 11.7% in the non-MS cohort. The final multivariable logistic regression model would include ten risk factors as: age, history of diabetes, contractive pressure, fasting blood-glucose, triglyceride, high density lipoprotein cholesterol,

DOI: 10.3760/cma.j.issn.0254-6450.2013.09.004

作者单位:100069 北京,首都医科大学公共卫生学院流行病与卫生统计学系(杨兴华);北京大学公共卫生学院流行病与卫生统计学系(陶秋山、孙凤、詹思延);台湾美兆健康管理机构(曹纯铿)

通信作者:詹思延,Email: siyan-zhan@bjmu.edu.cn

low density lipoprotein cholesterol, body mass index and blood uric acid. AUC was 0.827 (95% CI: 0.814–0.839) that could predict the development of MS within the next 5 years. The curve also showed adequate performance in the three tested samples, with the AUC and 95% CI as 0.813 (0.789–0.837), 0.826 (0.800–0.852) and 0.794 (0.768–0.820), respectively. After labeling the degrees of the four risks, it was showed that over 17.6% of the incidence probability was in the population under mediate risk while over 59.0% of them was in the high risk group, respectively. Conclusion Both predictability and reliability of our Metabolic Syndrome Risk Score Model, derived based on Taiwan MJ Longitudinal Health-checkup-based Population Database, were relatively satisfactory in the testing cohort. This model was simple, with practicable predictive variables and feasible form on degrees of risk. This model not only could help individuals to assess the situation of their own risk on MS but could also provide guidance on the group surveillance programs in the community regarding the development of MS.

【Key words】 Metabolic syndrome; Risk predictive model; Longitudinal data

代谢综合征(MS)是肥胖、高血压、血糖代谢异常和血脂异常等多种心血管病危险因素同一个体聚集的现象^[1],并已成为新的慢性病和公共卫生问题^[2-4]。目前MS多为横断面研究,欧美国家开展的前瞻性研究探讨了肥胖、胰岛素抵抗等对MS发病影响,但在中国人群中的发病因素还有待证实,迄今尚缺乏大规模、全国性MS发病资料^[5]。为此本研究应用台湾美兆自动化健康体检机构(美兆健检)提供的1997—2006年台湾地区大样本纵向数据资料,分析基线时无MS人群的5年内MS发病情况,并探讨发病因素,构建该人群发病预测模型,为中国人群MS个体发病的预测提供可借鉴的评估工具。

对象与方法

1. 研究对象:本研究数据来自美兆健检。美兆健检是连锁会员制专业健康体检机构,1994年起会员资料实施全面电子化管理,1996年起平均每年约有3万至7万人次资料,服务人口覆盖全台湾,岛内有4家诊所。本研究队列人群是在1996—2006年438 693名健康体检者记录数据中,选取随访刚满5年的45 870人作为基线人群,再从中选取35~74岁人群,并剔除有MS诊断及正在服用降血压药、降血糖药和降血脂药者,再剔除1996年体检者记录(1997年1月才有腰围指标测量数据),以最后得到的13 973名(35~74岁且基线无MS者)作为研究队列人群。将该队列人群按4个体检中心分为4个发病队列,其中台北队列为建模队列,另3个队列(桃园、台中和高雄体检中心)作为验证队列。

2. 检测内容及方法:由4个体检中心按统一的筛查程序和采用同型号的仪器设备完成体检测量及资料收集、管理和保存。检查项目包括体格检查、血压测量和营养状况评估。每名受检者在完成检查的同时填写一份自填式问卷以收集社会经济状况和生

活习惯(如吸烟、饮酒、饮食、运动状况),以及主要慢性疾病的个人和家族史、服药史及过敏史。以上信息由体检医师核查后扫描录入。采用全自动测量仪(KN-5000A, Nakamura, Japan)测量受试者身高、体重,并计算BMI(kg/m^2);腰围测量取腋中线肋骨下缘与髂嵴连线中点水平;血压测量采用自动水银血压计(CH-5000, Citizen, Japan)测量右手臂位2次,取均值。采集空腹8 h以上受检者血清样本,通过日本OlympusAu-1000型全自动生物化学分析仪集中测定FPG、TC、TG、HDL-C、LDL-C。

3. 诊断标准:MS诊断标准采用2005年美国NCEP-ATP III。即高血糖判定标准为FPG $\geq 100 \text{ mg/dl}$ 或服用降糖药;血压升高判定标准为SBP $\geq 130 \text{ mm Hg}$ (1 mm Hg=0.133 kPa)和/或DBP $\geq 85 \text{ mm Hg}$,或服用降压药;TG升高判定标准为TG $\geq 150 \text{ mg/dl}$ 或服用调脂药;HDL-C降低的判定标准为男性HDL-C $< 40 \text{ mg/dl}$,女性 $< 50 \text{ mg/dl}$,或服用调脂药;中心性肥胖判定标准(亚洲人群)为男性腰围 $\geq 90 \text{ cm}$,女性 $\geq 80 \text{ cm}$ 。以上5个组分中有3个达到标准即诊断为MS。高血压定义为SBP $\geq 140 \text{ mm Hg}$ 和/或DBP $\geq 90 \text{ mm Hg}$,或正在服降压药。

4. 统计学分析:采用SAS 9.1.3软件(北京大学公共卫生学院)。测量数据由体检中心核查和清理,并建立用于分析的随访数据库。连续性变量以 $\bar{x}\pm s$ 表示,分类变量以百分数(%)表示。分类变量其他比较应用 χ^2 检验, $P<0.05$ 为差异有统计学意义。采用多元logistic回归构建预测模型,根据logistic回归方程推导出概率预测公式

$$\text{logitP} = \text{logit}\left(\frac{p}{1-p}\right) = \beta_0 + \sum_{i=1}^n \beta_i x_i \quad (1)$$

$$p = \frac{\exp(\beta_0 + \sum_{i=1}^n \beta_i x_i)}{1 + \exp(\beta_0 + \sum_{i=1}^n \beta_i x_i)} = \frac{e^{\text{logitP}}}{1 + e^{\text{logitP}}} \quad (2)$$

表2 MS的单因素logistic回归分析

因 素	β	s_{β}	P值	OR值(95%CI) ^a
性别				
男	0.729	0.073	0.000	2.074(1.798~2.392)
年龄	0.040	0.004	0.000	1.041(1.033~1.048)
糖尿病家族史	0.429	0.084	0.000	1.536(1.303~1.812)
受教育水平	-0.129	0.034	0.000	0.879(0.822~0.940)
吸烟				
现在吸	0.356	0.095	0.000	1.428(1.185~1.721)
已戒	-0.065	0.148	0.659	0.937(0.701~1.252)
饮酒				
现在饮	0.168	0.094	0.073	1.182(0.984~1.420)
已戒	0.740	0.181	0.000	2.096(1.471~2.987)
运动				
偶尔	-0.282	0.098	0.004	0.754(0.623~0.913)
经常	-0.233	0.111	0.036	0.792(0.638~0.985)
每天	-0.104	0.104	0.318	0.901(0.734~1.106)
BMI(kg/m ²)				
超重(24~28)	1.205	0.078	0.000	3.336(2.862~3.887)
肥胖(>28)	2.277	0.152	0.000	9.746(7.230~13.137)
中心性肥胖				
是	1.196	0.095	0.000	3.308(2.747~3.983)
腰围	0.027	0.004	0.000	1.027(1.020~1.034)
身高	0.017	0.007	0.008	1.017(1.005~1.030)
体重	0.098	0.005	0.000	1.102(1.092~1.113)
FPG	0.011	0.002	0.000	1.011(1.007~1.015)
SBP	0.022	0.002	0.000	1.022(1.017~1.026)
DBP	0.030	0.003	0.000	1.031(1.024~1.038)
TG	0.010	0.001	0.000	1.010(1.009~1.011)
HDL-C	-0.027	0.002	0.000	0.973(0.969~0.978)
GPT	0.007	0.001	0.000	1.010(1.000~1.010)
UA	0.320	0.027	0.000	1.377(1.307~1.451)

注:性别中女性、吸烟中从不吸、饮酒中从不饮、运动中很少运动、BMI<24、非中心性肥胖为参照组,OR=1.0;^a为调整年龄和性别后的数值

资料,对本研究建立的MJ-MSRSM进行3次独立效度检验,结果其AUC分别为0.813(95%CI:0.789~0.837)、0.826(95%CI:0.800~0.852)、0.794(95%CI:0.768~0.820)(图1B、C、D),说明曲线的拟合度较高且有较好的外部效度。

4. 预测MS个体5年发病风险的等级及其划分:受检人群5年MS预测发病概率呈偏态分布,危险度

表3 MS的多因素logistic回归

因 素	β	Wald χ^2 值	P值	OR值(95%CI)
常数项	-13.486	555.050	0.000	-
年龄	0.030	43.973	0.000	1.030(1.021~1.040)
糖尿病家族史	0.410	19.025	0.000	1.507(1.253~1.812)
SBP	0.021	72.496	0.000	1.021(1.016~1.026)
FPG	0.014	36.864	0.000	1.014(1.010~1.019)
TG	0.009	116.976	0.000	1.009(1.008~1.010)
TC	-0.011	17.630	0.000	0.990(0.985~0.994)
HDL-C	-0.019	29.405	0.000	0.981(0.974~0.988)
LDL-C	0.011	23.548	0.000	1.011(1.007~1.016)
BMI	0.257	253.443	0.000	1.293(1.252~1.334)
UA	0.0429	34.2264	<0.0001	1.044(1.029~1.059)

等级划分时将预测概率进行正态转换(取自然对数)。划分标准: $\bar{x}-s$ 以下为低风险, $\bar{x}-s$ 至 $\bar{x}+s$ 为一般风险, $\bar{x}+s$ 至 $\bar{x}+2s$ 为中度风险, $\bar{x}+2s$ 以上为高度风险。再将相应界值点取反自然对数后得到预测概率与风险等级对应关系,见表4(分别以1、2、3、4代表低风险、一般风险、中度风险和高度风险)。

表4 MJ-MSRSM预测概率与风险等级的划分

项目	低风险	一般风险	中度风险	高度风险
预测概率(%)	<(0~1.6)	<(1.6~17.6)	<(17.6~59.0)	59.0~100.0
对应评价等级	1	2	3	4

根据表3参数,列出logitP公式

$$\text{logitP} = \beta_0 + \sum_{i=1}^n \beta_i x_i \\ = -13.486 + 0.03x_1 + 0.410x_2 + 0.021x_3 \\ + 0.014x_4 + 0.009x_5 - 0.011x_6 - 0.019x_7 \\ + 0.011x_8 + 0.257x_9 + 0.429x_{10}$$

式中 x_1 ~ x_{10} 分别为年龄、糖尿病家族史、SBP、FPG、TG、TC、HDL-C、LDL-C、BMI和UA共计10项预测指标。若一名47岁男性受检者,初次参加体检时体重76.10 kg,身高178.0 cm,有糖尿病家族史,SBP 123 mm Hg、FPG 114 mg/dl、TG 345 mg/dl、HDL-C 44 mg/dl、LDL-C 177 mg/dl、TC 290 mg/dl、UA 8 mg/dl,经计算其5年MS发病风险概率为31.46%,其风险等级为第3级(中度风险)。

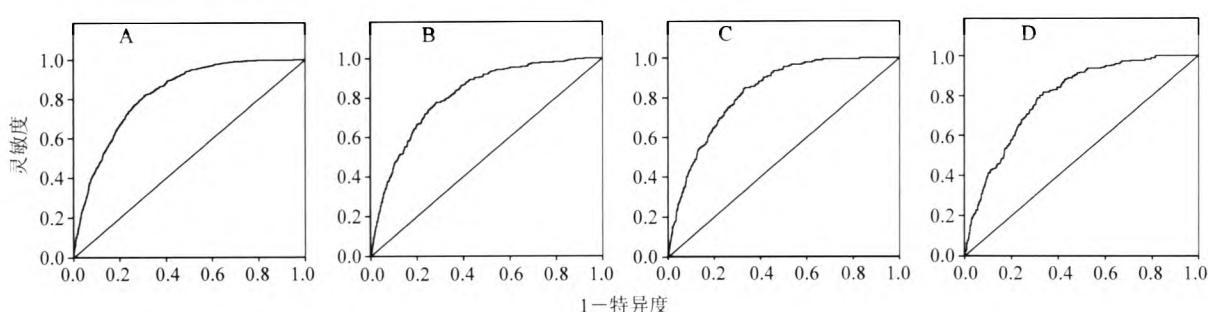


图1 MJ-MSRSM预测能力及其验证

讨 论

国内外研究均已表明,MS 可增加 2 型糖尿病和心血管病的发病风险^[6-9],并与全死因心血管病死亡率的增加相关联^[10-12]。与欧美国家相比较,亚洲地区人群易患 MS^[13]。MS 在不同地区和人群其流行特征有所不同,美国患病率为 23%,中国为 15.1%^[14]。本研究美兆健检 35~74 岁人群中 MS 粗患病率约为 23.7%。无 MS 人群 5 年累积患病率为 11.7%,与 2007 年北京市自然人群 5 年 MS 发病率调查结果(12.7%)相近^[5]。

本研究由逐步 logistic 回归筛选建立的 5 年 MS 预测模型中 10 个变量(年龄、糖尿病家族史、SBP、FPG、TG、HDL-C、LDL-C、TC、BMI 和 UA),与 MS 的关系同国内外既往病因学研究结果一致,其中除 UA 指标外,其他均已得到前瞻性研究证实。而受教育水平、经济情况、锻炼强度和是否吸烟、饮酒以及饮食习惯等均未进入模型,其他如生化检测指标、糖尿病家族史、年龄等与预测有关。其中糖尿病家族史、BMI、UA 对应 OR 值较大,而性别、腰围均未进入模型,提示“性别”在预测中作用不大,“腰围”主要与中心性肥胖有关,该因素未进入模型可能与研究人群体脂率高而中心性肥胖不明显有关^[15]。

本研究提供的预测模型可精确计算个体 5 年后 MS 的发病概率,并针对受检者具体情况提出相应的健康干预建议;如受检者采纳建议,可重新计算个体 MS 发病概率,直观可见采纳建议后发病风险的降低情况,对受检者开展健康教育和健康促进提供简便的评估工具。模型可预测体检者 5 年内新发 MS 的风险,并根据模型了解各风险因素对应的 OR 值。

本研究也存在局限性。首先体检人群为自愿参加者,而长期参加体检和很少参加体检者可能存在社会人口学特征的差异;其次 MS 发病人群在 55 岁前是男性多于女性,55 岁后男女性发病数量接近,如模型中分段评估应更精确;最后应强调该模型仅用于 5 年发病预测,不可直接外推至 35~74 岁以外人群。此外模型中糖尿病家族史信息为体检者自填也可能影响信息准确性。

参 考 文 献

- [1] Expert Panel on Detection, Evaluation, and 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(19):2486-2497.
- [2] Grundy SM, Cleeman JI, Daniels SR, et al. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. Circulation, 2005, 112(17):2735-2752.
- [3] Palaniappan L, Carnethon MR, Wang Y, et al. Predictors of the incident metabolic syndrome in adults: the Insulin Resistance Atherosclerosis Study. Diabetes Care, 2004, 27(3):788-793.
- [4] Gu DF, Reynolds K, Yang WJ. The prevalence of metabolic syndrome in the general adult population aged 35-74 years in China. Chin J Diabetes, 2005, 13(3):181-186. (in Chinese)
顾东风, Reynolds K, 杨文杰. 中国成年人代谢综合征的患病率. 中华糖尿病杂志, 2005, 13(3):181-186.
- [5] Liu J, Zhao D, Wang W. Incidence of the metabolic syndrome and its risk factors. J Cardiovasc Pulm Dis, 2007, 26(2):65-68. (in Chinese)
刘静, 赵冬, 王薇. 北京自然人群代谢综合征发病率及影响因素的研究. 心肺血管病杂志, 2007, 26(2):65-68.
- [6] Hanson RL, Imperatore G, Bennett PH, et al. Components of the "metabolic syndrome" and the incidence of type 2 diabetes. Diabetes, 2002, 51(10):3120-3127.
- [7] Isomaa B, Almgren P, Tuomi T, et al. Cardiovascular morbidity and mortality associated with the metabolic syndrome. Diabetes Care, 2001, 24(4):683-689.
- [8] Wannamethee SG, Shaper AG, Lennon L, et al. Metabolic syndrome vs Framingham Risk Score for prediction of coronary heart disease, stroke, and type 2 diabetes mellitus. Arch Intern Med, 2005, 165(22):2644-2650.
- [9] He Y, Jiang B, Wang J, et al. Prevalence of the metabolic syndrome and its relation to cardiovascular disease in an elderly Chinese population. J Am Coll Cardiol, 2006, 47(8):1588-1594.
- [10] Lakka HM, Laaksonen DE, Lakka TA, et al. The metabolic syndrome and total and cardiovascular disease mortality in middle-aged men. JAMA, 2002, 288(21):2709-2716.
- [11] Ford ES. Risks for all-cause mortality, cardiovascular disease, and diabetes associated with the metabolic syndrome: a summary of the evidence. Diabetes Care, 2005, 28(7):1769-1778.
- [12] Tsai SP, Wen CP, Chan HT, et al. The effects of pre-disease risk factors within metabolic syndrome on all-cause and cardiovascular disease mortality. Diabetes Res Clin Pract, 2008, 82(1):148-156.
- [13] Yusuf S, Reddy S, Ounpuu S, et al. Global burden of cardiovascular diseases. Part II: variations in cardiovascular disease by specific ethnic groups and geographic regions and prevention strategies. Circulation, 2001, 104(23):2855-2864.
- [14] Sheu WH, Chuang SY, Lee WJ, et al. Predictors of incident diabetes, metabolic syndrome in middle-aged adults: a 10-year follow-up study from Kinmen, Taiwan. Diabetes Res Clin Pract, 2006, 74(2):162-168.
- [15] Sun F, Tao QM, Tao QS, et al. Estimation on the risk of 5-years obesity development among adults aged 30-59, based on the Taiwan MJ Health-checkup Database. Chin J Epidemiol, 2012, 33(9):921-925. (in Chinese)
孙凤, 陶庆梅, 陶秋山, 等. 台湾地区 30~59 岁健康体检人群肥胖 5 年发病风险预测模型. 中华流行病学杂志, 2012, 33(9):921-925.

(收稿日期:2013-03-22)

(本文编辑:张林东)