

# 基于BMI的人体内脏脂肪影响因素的回归分析

梁 翌\*, 朱恩文#, 洪圣光, 法 焱

长沙理工大学数学与统计学院, 湖南 长沙

收稿日期: 2023年8月10日; 录用日期: 2023年10月24日; 发布日期: 2023年10月31日

## 摘 要

体重指数(BMI)是目前全球常用的衡量身体胖瘦与健康的指标。本文收集了国内199名成年人的体检数据, 运用相关分析、主成分回归、logistic回归和Lasso回归研究了身高、体重、腰臀比等12项指标与内脏脂肪含量(VAT)以及肥胖之间的关系, 并使用中国标准的BMI作为肥胖程度的评价依据。基于主成分回归, 对内脏脂肪含量有显著影响( $p = 0.05$ )的因素有BMI、腰臀比、腹部矢状径以及矢状径与左右径之比。体重指数与内脏脂肪含量之间存在显著的正相关性( $\gamma = 0.785$ ), 基于二分变量的logistic回归分析得到, 对BMI有显著影响的因素有腰臀比、腹部左右径和内脏脂肪含量。体重指数的波动表明内脏脂肪含量可能产生变化, 有可能影响身体健康, 需要及时评估和干预; 内脏脂肪含量的增减也可能导致BMI变动, 分析表明内脏脂肪含量是判断一个人是否肥胖的更优指标。

## 关键词

体重指数, 内脏脂肪含量, 相关分析, 主成分回归, logistic回归, Lasso回归

# Regression Analysis of Factors Influencing Human Visceral Fat Based on BMI

Zhao Liang\*, Enwen Zhu#, Shengguang Hong, Zhan Fa

School of Mathematics and Statistics, Changsha University of Science and Technology, Changsha Hunan

Received: Aug. 10<sup>th</sup>, 2023; accepted: Oct. 24<sup>th</sup>, 2023; published: Oct. 31<sup>st</sup>, 2023

## Abstract

Body mass index (BMI) is currently a commonly used global measure of body weight and health situation. This article used correlation analysis, principal component regression, logistic regres-

\*通讯作者。

#第一作者。

sion, and Lasso regression to study the relationship between height, weight, waist-to-hip ratio, visceral adipose tissue (VAT), obesity, etc., totaling 12 indicators based on the physical examination data collected from 199 adults in China, under the Chinese standard BMI as the evaluation basis for obesity level. Based on principal component regression, the factors that have a significant impact ( $\rho = 0.05$ ) on visceral adipose tissue include BMI, waist-hip ratio, abdominal sagittal diameter, and the ratio of sagittal diameter to right and left diameter. There is a significant positive correlation ( $\gamma = 0.785$ ) between body mass index and visceral adipose tissue. Based on logistic regression with dichotomous variables, the principal factors influencing body mass index were waist-hip ratio, abdominal right to left diameter, and visceral adipose tissue. The fluctuation of body mass index indicates that there may be changes in visceral adipose tissue, which may cause health risk and require timely evaluation and intervention; the increase or decrease in visceral adipose tissue may also lead to changes in BMI, and analysis shows that visceral adipose tissue is a better indicator to judge the obesity of a person.

## Keywords

Body Mass Index, Visceral Fat Content, Correlation Analysis, Principal Component Regression, Logistic Regression, Lasso Regression

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

体重指数(BMI)是目前国际常用的衡量人体胖瘦程度以及是否健康的标准,定义为体重(千克)除以身高的平方(平方米)。世界卫生组织标准: BMI 低于 18.5 为体重过轻; 18.5~24.9 为体重正常; 大于等于 25 为超重: 其中偏胖 25.0~29.9、肥胖 30.0~34.9、重度肥胖 35.0~39.9、极重度肥胖大于等于 40。体重指数无论是升高还是降低都应该进行评估,如果出现异常,则应积极干预。肥胖已成为一个主要的健康问题,肥胖人群容易出现各种代谢疾病[1]。在一些国家,肥胖的患病率不断上升[2] [3]。根据世界卫生组织(WHO)的统计,1975 年至 2016 年间,全球肥胖流行率几乎翻了三倍[4] [5]。中国的专家制定了针对国人肥胖的标准,认为以 BMI 值 24 为中国成人超重的界限, BMI 28 为肥胖的界限; 男性腰围  $\geq 85$  cm, 女性腰围  $\geq 80$  cm 为腹部脂肪蓄积的界限[6] [7]。最新发布的《中国居民营养与慢性病状况报告(2020 年)》显示,成人( $\geq 18$  岁)的超重率为 34.3%、肥胖率为 16.4%, 超重/肥胖成年人已过半[8]。在 11 省(市)中,超重和肥胖的现患率男性是北京最高,女性是山东最高;肥胖的现患率呈现出北方高于南方,城市高于农村的流行特点[9] [10]。

一般来说,肥胖是由 BMI 确定的,特定的身体脂肪分布和胰岛素反作用会增加肥胖相关疾病的风险[11] [12]。内脏脂肪含量(VAT)的增加会导致腹部肥胖的异常表现,并增加代谢并发症的风险。大量研究表明, BMI、内脏脂肪含量(VAT)和肥胖之间存在着密切的关系,使用 BMI 作为相关指标可以及时评估内脏脂肪含量的变化,以降低健康风险。然而, BMI 作为肥胖和疾病风险的指标有其局限性,它不能代表皮下和内脏脂肪的情况,从而导致肥胖测定的不准确[13]。换句话说,并非所有 BMI 在健康可接受范围内的成年人都拥有健康的体重,相反,他们可能有超标的内脏脂肪,也可能有更高的心血管疾病风险。

目前,在人体各项健康指标与肥胖程度关系的研究中,大部分研究都有一定局限性: Verma 的研究表明 BMI 易于计算并反映了个体身高和体重之间的关系,但它不能预测由高内脏脂肪含量(VAT)引起相关疾病的风险[13]; Tukaram 的研究表明腰臀比(WHR)是预测内脏脂肪含量(VAT)相对较好的指标[14], Oliveira 的研究则认为虽然该参数与人体脂肪含量之间存在强相关性,但这一指标更多地代表皮下脂肪,

而非内脏脂肪[15], 不能有效反应个体的健康情况。

本文根据湖南省某三甲医院体检中心收集的体检数据, 运用相关性分析、多元线性回归、主成分回归等方法, 分析人体内脏脂肪含量(VAT)的主要影响因素。根据中国标准下的体重指数(BMI)对体检对象进行分类, 用 logistic 回归分析内脏脂肪含量(VAT)及相关因素与肥胖的关系, 并最终建立内脏脂肪体积的预测模型。

## 2. 数据分析

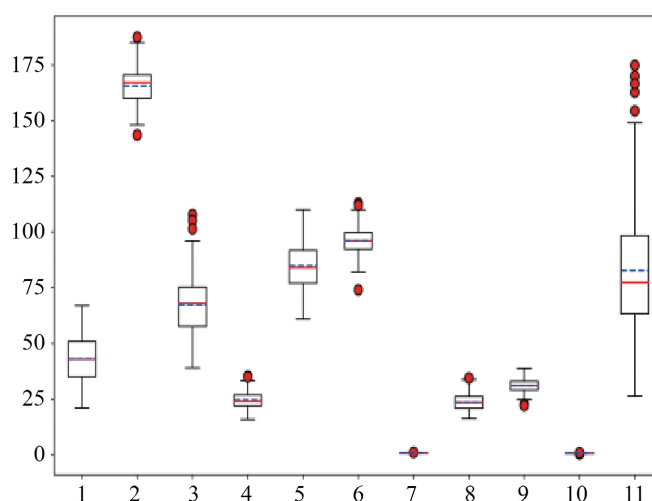
借助统计软件 SPSS 程序 12.0 版(SPSS Inc)、Python 3.8.0 和 R 语言 3.6.1 对数据进行分析。Python 清理数据后, 使用直方图查看变量的数据分布特征, 使用皮尔逊相关系数描述变量之间的线性相关性并检验显著性, 显著性水平为  $\rho = 0.05$ 。

为了分析人体内脏脂肪含量(VAT)的主要影响因素, 考虑性别、年龄、身高、体重、BMI、腰围(WC)、臀围(HC)、腰臀比(WHR)、腹部矢状径(ASD)、腹部左右径(ARL)、矢状径/左右径(SDLR), 共计 11 个变量作为独立变量(或预测变量), 并分别使用多元线性回归、逐步回归、主成分回归和 Lasso 回归来确定对内脏脂肪含量有显著影响的变量并对其进行分析。

首先根据中国标准对体检数据进行分类, 将肥胖定义为一个二分变量, BMI 28 及以上为肥胖, 值为 1; 28 以下为非肥胖, 值为 0。以 BMI 作为因变量, 其余 11 个人体测量参数作为自变量。通过主成分分析对 11 个独立变量进行了降维, 通过 logistic 回归对降维变量进行建模, 并使用 Wald 检验验证模型的统计显著性。

收集共计 199 份体检数据, 排除一例数据丢失, 最终 198 份数据构成了研究样本。每个样本包含 12 个指标值。平均年龄为 43.1 岁, 从 21 岁到 67 岁不等, 主要为男性; BMI 在 15.86 至 35.35 之间变化, 其中 2.86%的女性和 28.13%的男性肥胖。

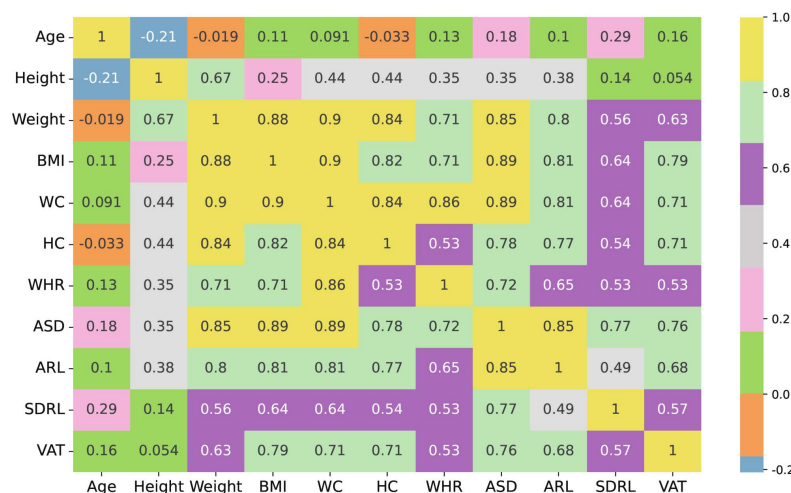
箱线(图 1)显示, 年龄分布轻微偏右; BMI 数据的中线正好位于上下底部的中心, 显示为标准的正态分布; 内脏脂肪含量(VAT)数据为左偏分布, 有一些异常值, 其中值低于上下四分位数的平均值。



**Figure 1.** Box plot: From left to right, age, height, weight, BMI, WC (waist circumference), HC (hip circumference), WHR (waist to hip ratio), ASD (sagittal abdominal diameter), ARL (right to left abdominal diameter), SDRL (sagittal diameter/right to left diameter), VAT (visceral fat content)

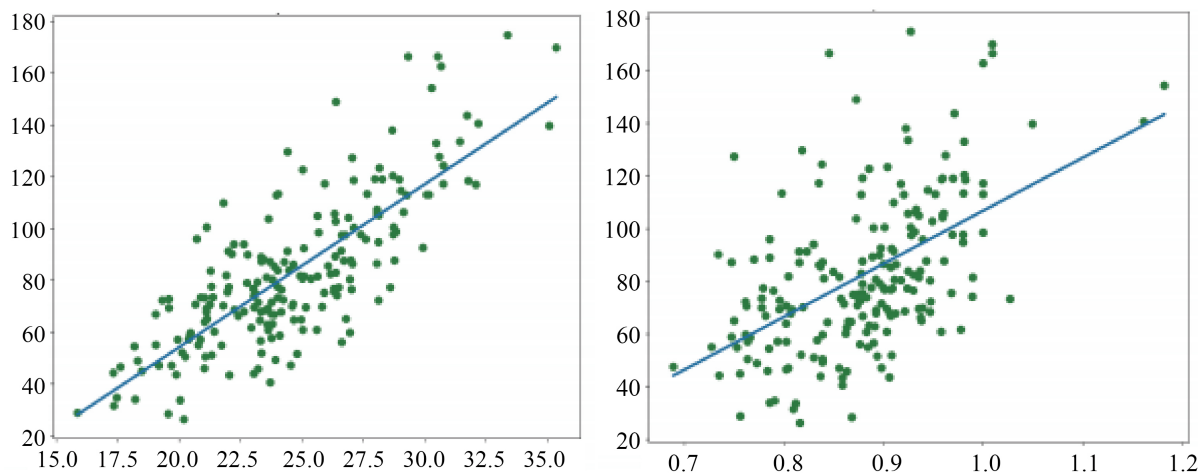
**图 1.** 箱线图: X 轴从左至右依次为年龄、身高、体重、BMI、WC (腰围)、HC (臀围)、WHR (腰臀比)、ASD (腹部矢状径)、ARL (腹部右至左径)、SDRL (矢状径/右至左直径)和 VAT (内脏脂肪含量), Y 轴-取值

相关系数的热力图(图 2)显示, 年龄、身高和 BMI 之间没有明显的线性关系, 体重、腰围、腰臀比、内脏脂肪含量和 BMI 之间有更明显的正相关。年龄与身高和内脏脂肪的相关程度较低, 相关系数分别只有 0.16 与 0.054; 而 BMI、腰围以及臀围与内脏脂肪的线性相关程度较高, 相关系数分别达到了 0.79、0.71 与 0.71, 这说明内脏脂肪的主要影响因素可能是 BMI、腰围以及臀围。基于此进一步作相关系数的显著性检验, 绘制散点图, 查看 BMI、腰臀比(WHR)和内脏脂肪含量(VAT)之间的线性相关性(图 3)。



**Figure 2.** Thermogram: BMI: Body Mass Index; WC: Waist circumference; HC: Hip circumference; WHR: waist to hip ratio; ASD: Abdominal sagittal diameter; ARL: Abdominal left and right diameter; SDRL: sagittal diameter/left right diameter; VAT: visceral adipose tissue

**图 2.** 热力图: BMI: 体重指数; WC: 腰围; HC: 臀围; WHR: 腰臀比; ASD: 腹部矢状径; ARL: 腹部左右径; SDRL: 矢状径/左右径; VAT: 内脏脂肪组织



**Figure 3.** Scatterplot showing the correlation between VAT and BMI, WHR

**图 3.** 散点图显示 VAT 与 BMI、WHR 之间的相关性

### 3. 人体内脏脂肪主要影响因素分析

使用内脏脂肪含量(VAT)作为解释变量, 分别通过逐步回归和主成分回归提取具有较高 VIF 的变量。逐步回归(表 1)最终保留了 BMI、性别、腹部矢状径(ASD)和腰围(WC)为解释变量, 建立拟合优度为 0.799

的线性模型(1)，其显著性水平  $\rho$  为 0.05，模型具有统计学意义。

$$\text{VAT} = -120.417 - 31.241(\text{Gender}) + 3.974(\text{BMI}) + 0.693(\text{WC}) + 2.835(\text{ASD}) \quad (1)$$

**Table 1.** Stepwise regression estimates of visceral fat content

**表 1.** 逐步回归估算内脏脂肪含量

方程	常数	标准差	性别	BMI	WC	ASD	R <sup>2</sup> (%)	R <sup>2</sup> (%) Adjust.
1	-71.652	18.1811	-	6.295	-	-	61.6	61.4
2	-95.947	14.6715	-25.813	7.966	-	-	75.1	74.9
3	105.309	13.4970	-28.680	4.913	-	3.654	79.1	78.8
4	120.417	13.2524	-31.241	3.974	0.693	2.853	79.9	79.5

BMI: 体重指数; WC: 腰围; ASD: 腹部矢状径; VAT: 内脏脂肪含量。

通过主成分分析(PCA)，保留负荷系数最大的前四个主要成分，分别为 BMI、腰臀比(WHR)、腹部矢状径(ASD)和矢状径/左右径(SDLR)，选择这四个因子进行主成分分析，累积方差贡献率为 89.057%。主成分回归模型的回归方程为式(2)，对应的各项参数见表 2。

$$\begin{aligned} \text{VAT} = & -96.494 - 9.04(\text{Gender}) - 0.126(\text{Age}) - 0.79(\text{Height}) + 0.244(\text{Weight}) + 2.055(\text{BMI}) \\ & + 0.355(\text{WC}) + 1.6937(\text{HC}) + 37.581(\text{WHR}) + 0.5121(\text{ASD}) - 0.452(\text{ARL}) + 38.947(\text{SDRL}) \end{aligned} \quad (2)$$

**Table 2.** Estimation of visceral fat content by principal component regression

**表 2.** 主成分回归估算内脏脂肪含量

模型	非标准化系数		标准系数	t	Sig.
	B	标准误差			
常数	82.715	1.111	-	74.419	0.000
成分 1	22.235	1.114	0.759	19.954	0.000
1 成分 2	7.290	1.114	0.249	6.542	0.000
成分 3	-2.003	1.114	-0.068	-1.798	0.074
成分 4	8.135	1.114	0.278	7.301	0.000

VAT: 内脏脂肪组织; 因子表示相应的主成分。

在第十次迭代中获得的模型最小 Cp 统计量为 9.2258 (表 3)，Lasso 回归最终保留了 9 个变量。计算具有非零系数的自变量相对应的系数值，得到 Lasso 回归方程为方程(3)。

**Table 3.** Cp statistics for Lasso regression

**表 3.** Lasso 回归的 Cp 统计量

DF	RSS	Gp	DF	RSS	Cp
1	168899	785.1768	7	33810	12.4132
2	108253	434.8680	8	33314	13.3093
3	81202	279.7241	9	33620	13.5315

Continued

4	70271	218.2204	10	32229	9.2258
5	40674	48.2875	11	32190	11.0000
6	38152	35.6357			

DF: 自由度; RSS: 残差平方和; Cp 统计量: 最优子集的选择标准。

$$\begin{aligned} \text{VAT} = & -134.2155 - 33.66(\text{Gender}) + 0.236(\text{Age}) + 3.675(\text{BMI}) + 0.314(\text{Weight}) \\ & + 0.401(\text{WC}) + 0.286(\text{HC}) + 24.098(\text{WHR}) + 3.07(\text{ASD}) - 1.22(\text{ARL}) \end{aligned} \quad (3)$$

#### 4. 肥胖主要影响因素分析

如图 3 回归结果表明, 所有变量的 Wald 统计量都对应非常大的  $p$  值, 因此没有一个变量是重要的。这可能是因为变量之间存在严重的多重共线性, 因此考虑使用主成分分析和因子分析去除多重共线性, 同时选择一些可解释性指标, 为二分变量的 logistic 回归做准备。

Bartlett 球形检验值达到了 2598, 对应的  $p \approx 0.000$ , 在  $\alpha = 0.05$  的显著性水平下拒绝相关系数矩阵为单位阵的原假设, 认为原始变量之间存在相关性, 适合做主成分分析。由 KMO 值也可看出, KMO 统计量的值为 0.780, 比较接近 1, 可认为变量间的相关性强, 原始变量适合做因子分析。选择四个成分因子进行主成分分析, 其累积方差贡献率达到 88.405%。使用最大方差正交旋转法进行因子旋转, 使各因子仍然保持正交的状态, 此时各因子的方差差异达到最大, 即相对载荷平方和达到最大(表 2)。腰臀比在第一主成分中有较大的载荷系数, 因此将第一主成分视为腰臀比; 同理将第二主成分视为腹部左右径(ARL); 第三主成分视为矢状径/左右径(SDLR); 第四主成分视为内脏脂肪含量(VAT)。根据 Wald 统计进行进一步的变量筛选, 保留了最后三个变量: 腰臀比(WHR)、腹部左右径(ARL)和内脏脂肪含量(VAT)。Logistic 回归系数(表 4)表明, 三个变量都是显著的( $\rho = 0.05$ ), 都对 BMI 有积极和显著的影响, 该模型的预测正确率高达 93.4%。由于腰臀比(WHR)对应的比值远大于腹部左右径(ARL)和内脏脂肪含量(VAT)对应的比值, 每个保留变量的 OR 均大于 1, 再次证明这三个变量对肥胖有显著影响。假设腰臀比(WHR)和腹部左右径(ARL)稳定, 则内脏脂肪含量越高, 肥胖的概率越高, 该结论与常识认知一致, 并且参数估计具有统计显著性。

**Table 4.** Rotated component matrix of the four principal components of principal component analysis

**表 4.** 主成分分析四个主要成分的旋转成分矩阵

成分	1	2	3	4	5	6	7	8	9	10	11
Gender	0.228	0.233	0.341	0.123	0.87	0.055	0.03	0.011	0.015	0.003	0.000
Age	0.013	0.043	-0.104	0.125	0.036	0.985	0.018	0.003	0.006	0.001	0.000
Height	0.209	0.097	0.926	0.015	0.263	-0.132	0.047	-0.014	0.008	0.004	-0.002
Weight	0.759	0.265	0.437	0.21	0.279	-0.033	0.056	0.181	0.021	-0.016	0.056
BMI	0.859	0.278	-0.01	0.263	0.226	0.048	0.049	0.239	0.017	-0.029	-0.019
WC	0.757	0.49	0.183	0.254	0.246	0.038	0.039	-0.044	0.029	0.151	-0.001
HC	0.917	0.088	0.232	0.199	0.048	-0.05	-0.023	-0.217	-0.068	-0.022	-0.005
WHR	0.415	0.833	0.131	0.201	0.25	0.066	0.092	0.02	0.017	-0.016	0.001

## Continued

ASD	0.729	0.281	0.119	0.443	0.202	0.101	0.206	0.035	0.289	0.005	0.001
ARL	0.765	0.258	0.167	0.124	0.071	0.071	0.542	0.01	0.028	0.002	0.001
SDRL	0.379	0.177	0.025	0.884	0.109	0.173	0.033	0.005	0.000	0.003	0.001

BMI: 体重指数; WC: 腰围; HC: 臀围; WHR: 腰臀比; ASD: 腹部矢状径; ARL: 腹部左右径; SDRL: 矢状径/左右径。

## 5. 结论

本文根据人体健康指标和成人肥胖程度监测数据,建立了预测人体内脏脂肪含量(VAT)和肥胖程度的回归模型。分析表明:腰臀比(WHR)、腹部左右径(ARL)和内脏脂肪含量(VAT)是预测肥胖与否的主要因素。相比单独使用人体健康指标数据来评估肥胖情况,本文使用复合变量、logistic 和 lasso 回归方程预测内脏脂肪含量(VAT)时效果更好,并可用于筛查由肥胖引起的相关疾病。

## 基金项目

国家统计局科学研究项目(2019LY21);长沙理工大学研究生创新基金(CX2020SS86, CX2020SS87)。

## 参考文献

- [1] 李华婷, 贾伟平. 肥胖相关代谢异常诊治现状[J]. 中国实用内科杂志, 2022, 42(2): 89-95. <https://doi.org/10.19538/j.nk2022020101>
- [2] Finucane, M., Stevens, G., Cowan, M., Danaei, G., Lin, J., Paciorek, C., et al. (2011) National, Regional, and Global Trends in Body-Mass Index Since 1980: Systematic Analysis of Health Examination Surveys and Epidemiological Studies with 960 Country-Years and 9.1 Million Participants. *The Lancet Journal*, **377**, 557-567. [https://doi.org/10.1016/S0140-6736\(10\)62037-5](https://doi.org/10.1016/S0140-6736(10)62037-5)
- [3] Stevens, G., Singh, G., Lu, Y., Danaei, G., Lin, J., Finucane, M., et al. (2012) National, Regional, and Global Trends in Adult Overweight and Obesity Prevalences. *Population Health Metrics*, **10**, Article No. 22. <https://doi.org/10.1186/1478-7954-10-22>
- [4] Vasileva, L., Marchev, A. and Georgiev, M. (2018) Causes and Solutions to “Globesity”: The New Fa(s)t Alarming Global Epidemic. *Food and Chemical Toxicology*, **121**, 173-193. <https://doi.org/10.1016/j.fct.2018.08.071>
- [5] Newby, P.K., Muller, D., Hallfrisch, J., Qiao, N., Andres, R. and Tucker, K.L. (2003) Dietary Patterns and Changes in Body Mass Index and Waist Circumference in Adults. *The American Journal of Clinical Nutrition*, **77**, 1417-1425. <https://doi.org/10.1093/ajcn/77.6.1417>
- [6] Stefan, N., Kantartzis, K., Jrgen, M., et al. (2008) Identification and Characterization of Metabolically Benign Obesity in Humans. *Archives of Internal Medicine*, **168**, 1609-1616. <https://doi.org/10.1001/archinte.168.15.1609>
- [7] 中国肥胖问题工作组. 中国成人超重和肥胖症预防与控制指南(节录) [J]. 营养学报, 2004(1): 1-4.
- [8] 刘月姣. 《中国居民营养与慢性病状况报告(2020年)》发布[J]. 中国食物与营养, 2020, 26(12): 2. <https://doi.org/10.19870/j.cnki.11-3716/ts.2020.12.001>
- [9] 王文娟, 王克安, 李天麟, 等. 中国成年人肥胖的流行特点研究: 超重和肥胖的现患率调查[J]. 中华流行病学杂志, 2001, 22(2): 129-132.
- [10] 武阳丰, 马冠生, 胡永华, 等. 中国居民的超重和肥胖流行现状[J]. 中华预防医学杂志, 2005, 39(5): 316-320.
- [11] Després, J. and Lemieux, I. (2006) Abdominal Obesity and Metabolic Syndrome. *Nature*, **444**, 881-887. <https://doi.org/10.1038/nature05488>
- [12] Lee, K., Park, H. and Yum, K. (2012) Cut-Off Values of Visceral Fat Area and Waist-to-Height Ratio: Diagnostic Criteria for Obesity-Related Disorders in Korean Children and Adolescents. *Yonsei Medical Journal*, **53**, 99-105. <https://doi.org/10.3349/ymj.2012.53.1.99>
- [13] Verma, M., Rajput, M., Sahoo, S., Kaur, N. and Rohilla, R. (2016) Correlation between the Percentage of Body Fat and Surrogate Indices of Obesity among Adult Population in a Rural Block of Haryana. *Journal of Family Medicine and*

- Primary Care*, **5**, 154-159. <https://doi.org/10.4103/2249-4863.184642>
- [14] Tukaram, G., Dudeja, P., Ipsita, B., Shruti, V. and Sandip, M. (2020) Correlation of Visceral Body Fat with Waist-Hip Ratio, Waist Circumference and Body Mass Index in Healthy Adults: A Cross-Sectional Study. *Medical Journal Armed Forces India*, **76**, 41-46. <https://doi.org/10.1016/j.mjafi.2017.12.001>
- [15] Oliveira, C.D., Roriz, A., Moreira, P., *et al.* (2014) Anthropometric Indicators Associated with Hypertriglyceridemia in the Prediction of Visceral Fat. *Revista Brasileira de Cineantropometria & Desempenho Humano*, **16**, 485-493. <https://doi.org/10.5007/1980-0037.2014v16n5p485>