

# 四种宫颈癌筛查方法的预测值比较

王慧 王彤 胡尚英 赵方辉 张询 潘秦镜 章文华 李凌 乔友林

**【导读】** 以病理组织学诊断为金标准,比较宫颈癌筛查中5%醋酸染色肉眼观察法(VIA)、薄层液基细胞学(LBC)、人乳头瘤病毒(HPV)自体采样法和HPV医生采样法的预测值。首先采用联合假设检验方法比较四种筛检方法的阳性预测值(PPV)和阴性预测值(NPV),并对两预测值进行组间两两比较,再采用边际回归法和加权最小二乘法分别比较该两预测值,多重性调整采用Hochberg法。结果表明HPV医生采样法的NPV最高,LBC的PPV最高,均为有效的宫颈癌初筛方法,但后者需要有经验的细胞病理学家读片。HPV自体采样法的PPV和NPV均高于VIA,其PPV与HPV医生采样法相当,NPV与LBC相当,因该方法易得到受检者配合而有可能提高筛查依从性,成为一种简便、客观、易于重复和适宜在资源贫乏地区使用的宫颈癌初筛方法。

**【关键词】** 宫颈肿瘤; 筛检; 人乳头瘤病毒; 诊断试验; 预测值

## Comparison on the predictive values of four screening methods regarding cervical cancer

WANG Hui<sup>1</sup>, WANG Tong<sup>1</sup>, HU Shang-ying<sup>2</sup>, ZHAO Fang-hui<sup>2</sup>, ZHANG Xun<sup>2</sup>, PAN Qin-jing<sup>2</sup>, ZHANG Wen-hua<sup>2</sup>, LI Lin<sup>2</sup>, QIAO You-lin<sup>2</sup>. 1 Department of Health Statistics, Shanxi Medical University, Taiyuan 030001, China; 2 The Cancer Institute/Hospital, Chinese Academy of Medical Sciences

Corresponding authors: WANG Tong, Email: wtstat1@sina.com; QIAO You-lin, Email: qiaoy@cicams.ac.cn

This work was supported by grants from the National Natural Science Foundation of China (No. 81072385) and the National Key Statistical Research Project (No. 2009LZ033).

**【Introduction】** To compare the predictive values of 5% acetic acid stain and visual inspection, human papillomavirus (HPV) Self test, ThinPrep Pap and HPV direct test in screening for cervical cancer with biopsy as gold standard. Positive predictive values and negative predictive values were compared simultaneously by joint hypothesis tests and then either positive predictive values or negative predictive values of the any two screening tests were compared by marginal regression based on both GEE and weighted least square methods. Hochberg method was used for multiplicity adjustment. It was showed that HPV direct test had the highest negative predictive value and ThinPrep Pap the highest positive predictive value. 5% acetic acid stain and visual inspection had both the lowest positive predictive value and negative predictive value. Both HPV direct test and ThinPrep Pap were efficient, but the latter required compatible infrastructure and skilled caregivers to go with. Both 5% acetic acid stain and visual inspection were inexpensive, and their positive predictive value and negative predictive value were lower than HPV self-test. They also had similar positive predictive value with HPV direct test and similar negative predictive value with ThinPrep Pap. HPV self-test appeared to be efficient, suggesting that it had significant potential for screening program to be implemented in the rural areas of China since the test could be performed without speculum examination in low-resource regions.

**【Key words】** Cervical neoplasia; Screening; Human papillomavirus; Binary diagnostic test; Predictive value

宫颈癌是女性第三大恶性肿瘤,2008年全球有

近53万宫颈癌新发病例和27.5万宫颈癌死亡病例,其中分别约85%和88%发生在发展中国家<sup>[1-3]</sup>。2011年中国卫生统计年鉴公布数据显示<sup>[4]</sup>,我国宫颈癌患病率由1999年的8.2/10万升高至2010年的15.1/10万,其中山西省2010年宫颈癌患病率为40.7/10万,位居全国之首。由于宫颈癌存在较长、可逆转的癌前病变期,且早期治疗预后良好,患者5年生存率为92%<sup>[5,6]</sup>,因此宫颈癌的筛查研究一直备受关注。

诊断或筛检试验中通常采用灵敏度和特异度描

DOI: 10.3760/cma.j.issn.0254-6450.2013.02.020

基金项目:国家自然科学基金(81072385); 全国统计科研计划重点项目(2009LZ033)

作者单位:030001 太原,山西医科大学卫生统计学教研室(王慧、王彤); 中国医学科学院北京协和医学院肿瘤医院研究所医院流行病学研究室(胡尚英、赵方辉、乔友林), 肿瘤医院病理科(张询、李凌), 细胞学室(潘秦镜), 妇瘤科(章文华)

通信作者:王彤, Email: wtstat1@sina.com; 乔友林, Email: qiaoy@cicams.ac.cn

述试验的准确度<sup>[7]</sup>,但在临床实践中多采用阳性预测值(PPV)和阴性预测值(NPV),两者是描述诊断或筛检试验效益的指标。有金标准时在同一人群中应用两种或多种方法进行筛查,可采用 McNemar 检验比较其中两种方法的灵敏度和特异度,常用的统计软件如 SPSS 和 SAS 等均可实现,但 PPV 和 NPV 通常只是报告数值,其总体参数的比较方法还鲜有文献报道。Leisenring 等<sup>[8]</sup>曾提出基于广义估计方程的边际回归(marginal regression)方法,Wang 等<sup>[9]</sup>提出加权最小二乘法用于 PPV 或 NPV 中一个指标的比较。如要同时比较该两指标,由于同一项诊断或筛检试验的 PPV 和 NPV 均是该试验的灵敏度、特异度和人群患病率的函数,因此两个预测值具有一定的相关性,如忽略可能增大 I 类错误,据此 Nofuentes 等<sup>[10]</sup>提出了联合假设检验的方法。以下简要介绍在同一人群中采用不同诊断或筛检方法获得 PPV 和/或 NPV 总体参数的比较及分析策略,及其应用于多种宫颈癌筛查方法的比较。

### 基本原理

利用 McNemar 检验比较两种筛检方法的灵敏度,即仅在金标准证实确有发病的受试对象中比较:试验 1 阳性且试验 2 阴性的例数与试验 1 阴性且试验 2 阳性的例数是否相等。两种筛检方法的特异度较类似,即仅在金标准证实确无发病的受试对象中比较:试验 1 阳性且试验 2 阴性的例数与试验 1 阴性且试验 2 阳性的例数是否相等。灵敏度和特异度的假设检验相同,只是比较的人群不同,PPV 与 NPV 的比较也相似。以 PPV 为例说明单个指标的比较。

Leisenring 等<sup>[8]</sup>建立边际回归模型:

$$\text{logit}[P(D_i=1 | Z_i=1, x_i)] = \alpha + \beta x_i$$

式中, $D_i$ 为第  $i$  个受检对象的金标准结果, $D_i=0$  表示非病例, $D_i=1$  表示病例; $Z_i$ 为第  $i$  个受检对象的筛检结果, $Z_i=0$  表示阴性, $Z_i=1$  表示阳性; $x_i$ 为第  $i$  个受检对象采用的筛检方法, $x_i=0$  表示采用筛检方法 1, $x_i=1$  表示采用筛检方法 2。因此

$$PPV_1 = P(D_i=1 | Z_i=1, x_i=0)$$

$$PPV_2 = P(D_i=1 | Z_i=1, x_i=1)$$

$\beta=0$  表示  $PPV_1=PPV_2$ ,比较两个筛检试验的 PPV 相等,即 $\beta=0$ ,可采用广义估计方程构造广义得分统计量,其服从自由度为 1 的卡方分布。

Wang 等<sup>[9]</sup>提出加权最小二乘法。即假定一次筛检试验是多项分布  $\{N=T, \pi=(\pi_1, \dots, \pi_8)\}$  的一个结果,类似于掷一次八面体得到的结果,故筛查  $n$  个

受检对象也就等于独立掷  $n$  次八面体。设每个面朝上出现的概率分别为  $\pi_1, \dots, \pi_8$ ,即对应于 8 种可能结果的理论概率。记  $g(\pi) = PPV_1 - PPV_2$  为两个筛检试验 PPV 的差值。拟合  $g(\pi) = \beta_0$ ,即仅含有截距项的回归方程,当  $PPV_1=PPV_2$  时  $\beta_0=0$ ,因此比较两个筛检试验的 PPV 相等即检验  $\beta_0=0$ ,可构造自由度为 1 的卡方统计量。 $g(\pi)$  是连接函数,在 SAS 的 Proc CATMOD 过程中可通过 response 语句自定义,参数  $\beta_0$  的估计采用加权最小二乘法,其中权重为  $g(\pi)$  的方差倒数。

Nofuentes 等<sup>[10]</sup>提出联合假设检验,即构造一个统计量同时考虑 PPV 和 NPV 的比较:将一个或多个筛检试验的 PPV 和 NPV 用一个向量  $\eta$  表示,由多元中心极限定理得到  $\eta$  渐近服从多元正态分布,多个假设检验用一个矩阵  $\varphi$  表示,构造的卡方统计量是  $\eta$  和  $\varphi$  的函数,且自由度为  $\varphi$  的行秩。

对于多个筛检试验的 PPV 和 NPV 的比较,可采用以下分析策略。第一步:采用联合假设检验方法进行多组 PPV 和 NPV 的总体比较。第二步:若第一步得到有统计学意义的结果,则采用联合假设检验方法对 PPV 和 NPV 同时进行组间两两比较,并采用 Hochberg 或 Holm 等方法进行多重性调整<sup>[11]</sup>,以控制其总 I 类错误在名义水准  $\alpha=0.05$  以内;否则,结束检验。第三步:对于第二步中两两比较有统计学意义的筛检组,采用边际回归法或者加权最小二乘法对 PPV、NPV 分别进行两两比较,并采用 Hochberg 或 Holm 等方法进行多重性调整,以控制其总 I 类错误在名义水准  $\alpha=0.05$  以内;否则,结束检验。

边际回归方法可以推导出统计量的表达式,采用 SAS 简单编程可求解;加权最小二乘法采用 SAS 中的 Proc CATMOD 可实现,联合检验方法采用 R 编程求解。此外,分析中如有交叉格子数为 0,采用以上方法都将得不出结果,此时可对每一个格子的频数加上格子个数的倒数<sup>[12]</sup>。

### 实例分析

为评价现行宫颈癌及癌前病变筛查方法的诊断效果,乔友林和 Belinson 等<sup>[5,13]</sup>于 1999 年 5—7 月在宫颈癌高发地区山西省襄垣县,对 35~45 岁、无子宫切除史和盆腔放射治疗史、当时未怀孕的妇女筛查宫颈癌及癌前病变,以病理组织学为金标准,对同一人群同时采用 6 种常用的宫颈癌筛查方法[5%醋酸染色肉眼观察法(5% acetic acid stain and visual inspection, VIA)、薄层液基细胞学(liquid-based

cytology, LBC)、人乳头瘤病毒(HPV)自体采样法(HPV self-test, Self-HPV)、阴道镜检法(colposcopy)、HPV 医生采样法(HPV direct test, Physician-HPV)、荧光分光镜检法(fluorescence spectroscopy)]。本研究仅分析其中最常用的 4 种筛查方法,其中病理组织学结果以  $\geq$ CIN II 为阳性,LBC 以  $\geq$ LGSIL 为阳性,Physician-HPV 和 Self-HPV 以  $\geq 1.0$  pg/ml 为阳性<sup>[13]</sup>。1997 例合格筛查对象中剔除 69 例采样量不足、无报告和结果不满意者<sup>[13]</sup>,本研究采用 1928 例有完整检测结果的数据进行分析,其中病理组织诊断  $\geq$ CIN II 者 83 例。

采用联合假设检验方法,对 4 种筛检试验的 PPV 和 NPV 进行总体比较,得到有统计学意义的结果后,再对 PPV 和 NPV 进行组间两两比较。对于得到差异有统计学意义的结果采用边际回归法和加权最小二乘法分别比较两种筛检方法的 PPV、NPV。对此 3 种方法在多组间或多指标间比较所涉及的多重性问题均采用 Hochberg 方法进行调整。

4 种筛查方法的 PPV 和 NPV 结果见表 1。采用联合假设检验的方法同时比较 4 种筛检试验的 PPV 和 NPV,取

$$\varphi = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$

$$\eta = (PPV_1, \dots, PPV_4, NPV_1, \dots, NPV_4)^T$$

$$H_0: \varphi\eta = 0$$

表示原假设为检验相邻的两个 PPV 相等且相邻的两个 NPV 相等,等价于检验

$$H_0: PPV_1 - PPV_2 = 0 \text{ 且 } PPV_2 - PPV_3 = 0 \text{ 且 } PPV_3 - PPV_4 = 0$$

$$\text{且 } NPV_1 - NPV_2 = 0 \text{ 且 } NPV_2 - NPV_3 = 0 \text{ 且 } NPV_3 - NPV_4 = 0$$

由于存在频数为 0 的格子,对所有的格子频数加上 1/32,计算得到  $\chi^2 = 110.8649, \nu = 6, P < 0.0001$ ,表明至少有一对相邻的 PPV 或 NPV 不相等。进一步采用联合假设检验的方法对四组间 PPV 和 NPV 进行两两比较,取

$$\varphi = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

表 2 任意两种宫颈癌筛查方法的 PPV 和 NPV 同时比较的结果

统计学分析	Physician-HPV vs. LBC	Physician-HPV vs. Self-HPV	Physician-HPV vs. VIA	Self-HPV vs. LBC	Self-HPV vs. VIA	LBC vs. VIA
$\chi^2$ 值	70.135 49	14.718 35	57.900 78	55.017 03	48.294	93.009 29
P 值	<0.0001	0.0006	<0.0001	<0.0001	<0.0001	<0.0001

$$\eta = (PPV_1, PPV_2, NPV_1, NPV_2)^T$$

$$H_0: \varphi\eta = 0$$

表示原假设为两种筛检方法的 PPV 相等且 NPV 相等,等价于检验

$$H_0: PPV_1 - PPV_2 = 0 \text{ 且 } NPV_1 - NPV_2 = 0$$

其结果见表 2,采用 Hochberg 法进行多重性调整后拒绝所有原假设,表明任意两种筛检方法的 PPV 或 NPV 不相等。

表 1 4 种宫颈癌筛查方法的 PPV 和 NPV

预测值(%)	Physician-HPV	Self-HPV	LBC	VIA
PPV	22.3	21.0	37.1	11.0
NPV	99.9	99.1	99.4	98.3

为进一步确定四组间差异是源于 PPV 不同还是 NPV 不同,采用边际回归方法和加权最小二乘法两两比较四组间的 PPV、NPV,两种方法得到的结果一致(表 3)。当存在频数为 0 的格子时,对所有的格子频数加上 1/8。采用 Hochberg 法进行多重性调整后的结果表明:Physician-HPV 的 NPV 高于其他 3 种方法;LBC 的 PPV 高于其他 3 种方法;VIA 的 PPV 和 NPV 均小于其他 3 种方法。

## 讨 论

比较多种筛检方法的 PPV 或 NPV 的总体参数时可采用边际回归法和加权最小二乘法,同时比较一种或多种筛检方法的 PPV 和 NPV 的总体参数时采用联合检验方法。Wang 等<sup>[9]</sup>在不同发病率和样本量下进行模拟研究比较边际回归法和加权最小二乘法,结果显示两种方法得到类似的结果,只是后者可以得到 PPV 或 NPV 总体参数比较的样本量估计公式,与本研究分析的结果一致。

PPV 和 NPV 属后验诊断概率指标,较先验诊断概率指标(灵敏度和特异度)更符合实际临床应用<sup>[14]</sup>。本文宫颈癌筛查是以人群为基础的横断面研究,可估计样本人群的患病率,并直接计算 PPV 和 NPV。

乔友林等<sup>[5]</sup>采用约登指数综合指标比较 6 种筛查方法,结果显示:Physician-HPV 法最好(约登指数为 0.811),LBC(约登指数为 0.807)与之接近,两者差异无统计学意义,但均高于其他方法。本文 4 种

表 3 任意两种宫颈癌筛查方法的 PPV 或 NPV 比较的结果

筛查方法	基于 GEE 的 边际回归		加权最小 二乘法	
	$\chi^2$ 值	P 值	$\chi^2$ 值	P 值
Physician-HPV vs. LBC	PPV	43.49 <0.0001	43.17 <0.0001	
	NPV	6.92 0.0085	6.92 0.0085	
Physician-HPV vs. Self-HPV	PPV	1.58 0.2082	1.59 0.2079	
	NPV	11.77 0.0006	11.77 0.0006	
Physician-HPV vs. VIA	PPV	55.64 <0.0001	56.60 <0.0001	
	NPV	21.07 <0.0001	21.12 <0.0001	
Self-HPV vs. LBC	PPV	47.05 <0.0001	47.36 <0.0001	
	NPV	1.07 0.3015	1.07 0.3015	
Self-HPV vs. VIA	PPV	42.27 <0.0001	42.56 <0.0001	
	NPV	7.51 0.0061	7.52 0.0061	
LBC vs. VIA	PPV	78.80 <0.0001	88.09 <0.0001	
	NPV	10.17 0.0014	10.19 0.0014	

宫颈癌筛查方法比较,以 Physician-HPV 法的 NPV 最高,尽管 PPV 低于 LBC,但灵敏度最高<sup>[5]</sup>,可作为一种高效的宫颈癌初筛方法。LBC 的 PPV 最高,具有制片、阅片效果好等优点,但其准确性受阅片者经验水平的影响,不适合在尚未建立细胞学检查体系的地区推广使用。VIA 虽是一种简便、低廉的检测手段,但 PPV 和 NPV 均低于 Self-HPV 法;而 Self-HPV 的 PPV 与 Physician-HPV 相当, NPV 又与 LBC 相当,能够极大地提高妇女的筛查参与率,节约卫生资源。虽然 HPV 检测的准确性和效益高,但费用较高,目前一些快速的 HPV 检测技术已有新发展<sup>[15]</sup>,将推动 HPV 检测方法成为资源贫乏地区的一种简便、客观、高效、易于重复的宫颈癌初筛手段。

参 考 文 献

[1] Ferlay J, Shin HR, Bray F, et al. Cancer incidence and mortality worldwide: IARC Cancer Base No. 10. <http://globocan.iarc.fr/>. 2010. GLOBOCAN 2008 v2.0.

[2] Bray F, Ren JS, Masuyer E, et al. Estimates of global cancer prevalence for 27 sites in the adult population in 2008. *Int J Cancer*, 2012 [Epub ahead of print].

[3] Ci PW, Wang LH, Zhao FH, et al. Research progress of the transition probabilities of the natural history model of cervical cancer. *Chin J Epidemiol*, 2011, 32 (12): 1292-1295. (in Chinese)  
慈璞娟,王临虹,赵方辉,等.宫颈癌自然史模型转移概率的研究进展. *中华流行病学杂志*, 2011, 32(12): 1292-1295.

[4] Ministry of Health of the People's Republic of China. *Chinese*

*Health Statistical Yearbook*. Beijing: Peking Union Medical College Press, 2011. (in Chinese)

中华人民共和国卫生部. *中国卫生统计年鉴*. 北京: 中国协和医科大学出版社, 2011.

[5] Qiao YL, Zhang WH, Li L, et al. A cross-sectional comparative trial of multiple techniques to detect cervical neoplasia. *Acta Academiae Medicinae Sinicae*, 2002, 24(1): 50-53. (in Chinese)  
乔友林,章文华,李凌,等. 宫颈癌筛查方法的横断面比较研究. *中国医学科学院学报*, 2002, 24(1): 50-53.

[6] Society AC. What are the key statistics about cervical cancer? [http://web.archive.org/web/20071030194439/http://www.cancer.org/docroot/CRI/content/CRI\\_2\\_1X\\_What\\_are\\_the\\_key\\_4\\_statistics\\_for\\_cervical\\_cancer\\_8.asp?sitearea=](http://web.archive.org/web/20071030194439/http://www.cancer.org/docroot/CRI/content/CRI_2_1X_What_are_the_key_4_statistics_for_cervical_cancer_8.asp?sitearea=). 2006.

[7] Zhou XH, Obuchowski NA, McClish DK. *Statistical Methods in Diagnostic Medicine(诊断医学统计学)*. 宇传华,译. 北京: 人民卫生出版社, 2005.

[8] Leisenring W, Alonzo T, Pepe MS. Comparisons of predictive values of binary medical diagnostic tests for paired designs. *Biometrics*, 2000, 56(2): 345-351.

[9] Wang W, Davis CS, Soong SJ. Comparison of predictive values of two diagnostic tests from the same sample of subjects using weighted least squares. *Stat Med*, 2006, 25(13): 2215-2229.

[10] Nofuentes JAR, Castillo JD, Alonso MAM. Global hypothesis test to simultaneously compare the predictive values of two binary diagnostic tests. *Comput Stat Data Anal*, 2012, 56(5): 1161-1173.

[11] CCTS Work Group, Wang T, Yi D. Statistical consideration for multiplicity in clinical trial. *Chin J Health Stat*, 2012, 29(3): 445-450. (in Chinese)  
CCTS 工作组,王彤,易东. 临床试验中多重性问题的统计学考虑. *中国卫生统计*, 2012, 29(3): 445-450.

[12] Grizzle JE, Starmer CF, Koch GG. Analysis of categorical data by linear models. *Biometrics*, 1969, 25(3): 489-504.

[13] Belinson J, Qiao YL, Pretorius R, et al. Shanxi Province Cervical Cancer Screening Study: a cross-sectional comparative trial of multiple techniques to detect cervical neoplasia. *Gynecol Oncol*, 2001, 83(2): 439-444.

[14] Chen PY. Some questions about methods for diagnostic trial. *Journal of Peking University: Health Science*, 2010, 42(6): 764-766. (in Chinese)  
陈平雁. 关于诊断试验方法的若干问题. *北京大学学报: 医学版*, 2010, 42(6): 764-766.

[15] Qiao YL, Sellors JW, Eder PS, et al. A new HPV-DNA test for cervical-cancer screening in developing regions: a cross-sectional study of clinical accuracy in rural China. *Lancet Oncol*, 2008, 9(10): 929-936.

(收稿日期: 2012-08-31)

(本文编辑: 张林东)