

Original Research Article

Interobserver agreement for Pap smear and visual inspection tests in southern

Brazil

Abstract

Aim: To verify the agreement on ASC-US or higher grade lesions between different examiners with Pap smear and visual inspection tests (VIT) in screening for cervical cancer.

Study design: A prospective cross-sectional study.

Place and Duration of Study: Rede Feminina de Combate ao Câncer de Florianópolis, a female cancer prevention in Florianópolis, southern Brazil between June 2010 and July 2012.

Methodology: Two observers evaluated 380 Pap smear test specimen and 284 visual inspection tests (VIT) of the women screened for cervical cancer in female cancer prevention center in southern Brazil. VIT started with naked eye examination after applying acetic acid to cervix (VIA) and Lugol's iodine (VILI) and was followed by additional examination with 2x magnifying glass (VIAM, VILIM).

Results: Concordance between the first and the second observer regarding Pap test was reasonably good ($\kappa=0.67$) but was even better regarding all VIT modalities (κ range 0.76-0.83). Magnifying glass did not significantly improve the agreement.

Conclusions: VIT showed very good agreement between observers. Adding VIT to Pap smear test may improve diagnostic accuracy in cervical cancer screening.

24 **Key words:** interobserver agreement, Pap test, visual inspection tests, cervical cancer.

25

26 **1 Introduction**

27 Most of the world countries rely solely on the Pap smear test for cervical cancer
28 screening. In Brazil, the National Institute for Cancer (abbrev. INCA in Portuguese)
29 recommends its use for screening 25-64 years old women every three years after two
30 consecutive negative test results [1,2]. However, various errors can occur, such as those
31 during the specimen collection and those due to different interpretations of the specimen
32 between the examiners, as well as false negative test results [3]. This paper examines
33 the second component and describes an attempt to reduce it by adding visual inspection
34 tests (VIT).

35 The ALTS study showed reasonable agreement among pathologists regarding
36 low-grade cervical intraepithelial lesion (LSIL) versus high-grade cervical
37 intraepithelial lesion (HSIL) diagnosis but also considerable disagreement in
38 differentiation between moderate and severe HSIL that is, between cervical
39 intraepithelial lesions of grades 2 and 3 (CIN2 vs. CIN3) [4]. In addition, the cytopathic
40 effect of HPV is often confounded with low-grade cervical intraepithelial neoplasia
41 (CIN1).

42 Visual inspection techniques (VIT) have been amply piloted in Asian and
43 African countries with limited resources where the access to Pap test screening is often
44 difficult and women health prevention programs are precarious. Visual inspection with
45 acetic acid (VIA) and Lugol's iodine (VILI) have been particularly popular among VIT.
46 Their advantages include a sensitivity similar to Pap test in detecting high-grade
47 cervical lesions, low cost, simple execution, availability of test result immediately after

the test application and thus the viability of the “see and treat” protocol [5-7]. Also, a short training course of 3-14 days can be administered to medical professionals of varying background in order to capacitate them for cervical screening, which is an important shortcut in places with the lack of trained screening professionals [6]. However, intensive initial training and periodic refreshment and/or upgrade courses are necessary to ensure effectiveness of this strategy [8].

The aim of this study was to verify the interobserver agreement (IOA) in screening for cervical cancer by Pap test and VIT.

2 Materials and methods

A cross-sectional study evaluated IOA for 380 Pap smear test specimen and 284 VIT of the women screened for cervical cancer in Florianopolis, the capital of the Santa Catarina state in southern Brazil. The site was a clinic specialized in this type of screening and maintained by an non-governmental organization set to prevent cancer in women (abbrev. RFCCF in Portuguese). All women who were scheduled for regular screening and signed the consent form were eligible for the study providing they had an intact uterus. The women who had been submitted to hysterectomy or diagnosed with HPV, anal or genital condyloma lesions, were excluded from the study. Among 380 women screened with Pap test, 96 were not evaluated with VIT because they missed colposcopy scheduled shortly after the Pap test.

The Pap test specimen was collected by a nurse with over 20 years of work experience. Two specimens were collected for each woman: one by scraping with Ayre spatula from ectocervix and another by endocervical brush from endocervix. Both specimens were fixed by 70% alcohol and stained by Papanicolaou technique.

71 The Pap test results were classified according to the Bethesda System of 2001
72 [9]. The first observer to make the diagnosis was an experienced laboratory technician
73 specialized in cytological tests. The second observer was a pathologist specialized in
74 cytopathology and employed by the University Hospital of the Federal University of
75 Santa Catarina. Both observers had to classify specimen as altered (thus requiring
76 further investigation) or otherwise, where the alteration of interest was defined as ASC-
77 US or higher grade. In the case the two observers disagreed, a third observer was called
78 to decide. The latter was a pathologist specialized in cytopathology and employed by
79 the same laboratory as the second observer. The first observer has access to the medical
80 records but did not know whether the specimen under evaluation was part of the study
81 or not. The second and third observers did not have access to the medical records nor
82 the knowledge of the other observer's diagnosis.

83 VIT were performed by the same nurse who collected the Pap test specimen. She
84 took a 130 hours course on VIT, based on IARC manual [10] and administered by an
85 experienced gynecologist specialized in colposcopy and cervical cancer screening. The
86 VIT started with the application of a 5% acetic acid to the cervix for one minute and a
87 naked eye examination to make the VIA diagnosis. Next, a 2% Lugol's iodine solution
88 was applied in the presence of a 100 W cold light in order to make the VILI diagnosis.
89 Both VIA and VILI naked eye examinations were followed by additional examination
90 with 2x magnifying glass in order to make VIAM and VILIM diagnoses. The second
91 VIT observer was a gynecologist who re-examined all study participants but was blind
92 to the findings of the first one. In addition, colposcopy was performed for all of them.
93 All VIT results were classified as positive or negative according to the IARC manual
94 [10].

95 Both VIA and VIAM diagnoses were classified as negative under following
 96 conditions after the application of acetic acid: in the absence of acetowhite areas or if
 97 these were pale or ill-defined, whitish with undiscernible margins blending with the
 98 epithelium, areas with salient white line underlying the squamo-columnar junction
 99 (SCJ), angular acetowhite areas, dispersed white areas far away from the SCJ or in the
 100 external orifice of the cervix (EOC) if the SCJ was not visualized, acetowhiteness of the
 101 mucus or Nabothian cysts and normal acetowhiteness in the columnar epithelium.
 102 Positive diagnosis was defined when the visible presence of well-defined, opaque
 103 acetowhite areas abutting a fully visible SCJ or the transformation zone or around EOC
 104 if SCJ was not visible, as well as acetowhiteness of the entire cervix or its growth [10].

105 Negative VILI and VILIM diagnoses included the following findings: normal
 106 cervix where squamous epithelium turned black or brown and the columnar epithelium
 107 remained pale, colorless or partially brown in the transformation zone, or thin yellow,
 108 ill-defined non-iodine uptake areas on the cervix located far away from the SCJ, with or
 109 without extension to the vagina. Positive diagnosis was defined as well defined, thick,
 110 mustard or saffron yellow area touching the SCJ or occupying a large portion of the
 111 cervix [10].

112 Biopsy was performed for all VIT-positive cases, including white acetic
 113 epithelium, mosaic or point-like forms, atypical vascular presentations, enlarged
 114 glandular orifices and iodine-negative areas. Differently from the IARC manual, the
 115 cases considered difficult to classify as either positive or negative on Pap screening
 116 were treated as positive and investigated further.

Individual clinical and laboratory data were digitalized and analyzed by Stata software version 12.1. Kappa statistic was used to measure agreement between observers. The type I error was fixed at 5% ($P < 0.05$).

3 Results

Mean age of the women who participated in the study was 32.9 years (range 14.0-54.0). The age of the first menarche was reported at 12.8 years (range 10.0-16.0) and that of the first sexual intercourse was at 16.8 years (range 13.0-34).

Concordance of 0.67 between the first and the second observer regarding Pap test was reasonably good (Table 1). However, concordance indices between these two observers and the third one were poor (kappa values of 0.14 and 0.23).

Table 1– Interobserver agreement for Pap test: kappa statistic and its 95% confidence intervals (CI)

Pap Test	Observer 2		Kappa CI 95%	Observer 3		Kappa CI 95%
	(+)	(-)		(+)	(-)	
Observer 1			0,77			0,14
(+)	33	17	(0.65-0.88)	12	6	(0.0-0.45)
(-)	10	293		10	9	
Total	70	310		22	15	
Observer 2						0,23
(+)	-	-		18	9	(0.0-0.57)
(-)	-	-		4	6	
Total				22	15	

Pap test positive (+) if lesion grade $\geq$ ASC-US, otherwise negative (-)

140

141 Concordance between the first and the second observer regarding VIT was pretty
142 good for all modalities considered separately as well as pulled together, with kappa
143 values ranging between 0.76 and 0.83 (Table 2).

144 Table 2– Interobserver agreement for visual inspection tests: kappa statistic and its 95%
145 confidence intervals (CI)

Test	Observador 2		<i>Kappa</i>	Test	Observer 2		<i>Kappa</i>
	(+)	(-)	CI 95%		(+)	(-)	CI 95%
VIA				VIAM			
Observer 1			0.77	Observer 1			0.77
(+)	66	23	(0.65-0.88)	(+)	67	22	(0.65-0.88)
(-)	4	191		(-)	5	190	
Total	70	214		Total	72	212	
VILI				VILIM			
Observer 1			0.81	Observer 1			0.83
(+)	71	22	(0.70-0.93)	(+)	69	20	(0.71-0.94)
(-)	0	191		(-)	0	195	
Total	71	213		Total	69	215	
VIT							
Observer 1			0.76				
(+)	155	4	(0.64-0.87)				
(-)	30	95					
Total	185	99					

146

147 VIA = visual inspection with acetic acid;
148 VIAM = visual inspection with acetic acid magnified;
149 VILI = visual inspection with lugol's iodine;
150 VILIM = visual inspection with lugol's iodine magnified;
151 VIT = visual inspection tests;
152 VIT (+) = any positive results on VIA, VIAM, VILI or VILIM;
153 VIT (-) = VIA, VIAM, VILI and VILIM all negative.

154

155 4 Discussion

156 Interobserver agreement is an important component of diagnostic accuracy of
157 both Pap and VIT tests. Adequate training of health professionals for cervical cancer

158 screening plays a key role in achieving good accuracy of these tests, particularly in
159 reducing the false negative rate, as reported in another Brazilian study [11].

160 A reasonable agreement between the first and the second observer regarding Pap
161 test result, and their poor agreement with the third observer, may be explained by
162 inadequate staining and/or fixation of the specimen as reported occasionally by the
163 former. Low quality of the staining material, high workload for the screening
164 professionals and different classification standards for morphological alteration used by
165 the observers may have contributed to disagreement on Pap test diagnoses in some
166 cases. However, it should be kept in mind that the third observer re-examined only the
167 specimen with discrepant diagnosis between the first and the second observer, which
168 were likely more difficult to classify and therefore to agree upon.

169 Despite above obstacles, the concordance between the first two observers in the
170 present study ($\kappa=0.77$) was better than the one reported in another study based on
171 hundred specimen which were initially evaluated by an experienced laboratory
172 technician and revised by three pathologists, achieving mean kappa value of 0.56 [12].
173 Yet another Brazilian study also found good agreement between the first two observers
174 ($\kappa=0.67$) but their poor agreement with the third one ($\kappa=0.02$) [13]. The
175 concordance between the first observer and an expert panel was pretty good
176 ($\kappa=0.75$), thus indicating a good diagnostic accuracy of the first observer. A lack of
177 precise guidelines for diagnostic interpretation of Pap test contributed to only modest
178 agreement ($\kappa=0.46$) in an important multicentric study [4] which concluded that
179 better training and more rigorous quality control may improve the agreement level. The
180 present study range of kappa values (0.76-0.83) confirm this conclusion regarding VIT.
181 Extensive training of the observers with particular attention to the difficulties in

182 diagnostic classification has been emphasized in the literature [5,7,14-16] but its routine
183 implementation among screening professionals remains a challenge.

184 Small magnification (2x) did not improve the IOA of either VIA or VILI in the
185 present study. However, the studies with larger magnification (4x) showed its beneficial
186 effect on diagnostic accuracy, especially in improving test specificity [17].

187 VILI and VILIM had the highest IOA rates, probably due to the fact that the
188 diagnoses were made immediately after the Lugol's iodine application as opposed to a
189 one minute wait for VIA. In addition, the whitening of acetic acid in VIA may be
190 difficult to localize when the area of lesion is small, within transformation zone, next to
191 a large area of immature metaplasia or glandular tissue. Also, acetic acid may
192 occasionally cause bleeding of the cervix which interferes with accurate VIT diagnosis
193 [18].

194 A large multicentric study in India and Africa achieved moderate IOA for VIA,
195 with mean kappa value of 0.38 (range 0.15-0.65) across 11 sites [19]. Considerably
196 higher kappa of 0.77 for VIA in the present study may be due to longer training of the
197 observers and lesser specimen variation in a single center study. The same type of study
198 in Karnataka, India, found low-to-moderate agreement (kappa=0.36) between a nurse
199 and a physician after a 2-week training to diagnose NIC-2 or higher grade lesion by
200 VIA, using IARC diagnostic criteria [20]. In the view of these results, it seems worth
201 providing a more extensive training for screening professionals in order to improve IOA
202 with VIT.

203 Average kappa of 0.57 was obtained for IOA among three VIA experts who
204 evaluated 144 photographs of the cervix after application of acetic acid, without any aid
205 from the clinical data [21]. The women who participated in the study were from India,

206 South Africa and Peru. The results indicate potential of VIA for controlling the quality
207 of screening programs at an affordable cost, even in the countries with limited
208 resources. Where there is a substantial lack of trained screening professionals, nurses
209 can be trained to perform the screening. However, the effectiveness of the training
210 critically depends on its periodic reinforcement, mainly in order to reduce false positive
211 test results and consequently the screening costs [20]. Internet may provide a low-cost
212 alternative to sending digitalized photographs of the cervix to a quality control center.
213 Monitoring whether local screening site has achieved reasonable IOA (kappa 0.40 or
214 higher) can be used among criteria of a good quality cervical cancer prevention [6]

215 **5 Conclusions**

216 IOA for VIT was very good in the present study as indicated by the kappa range
217 between 0.77 and 0.83 for VIA, VIAM, VILI and VILIM. Magnifying (2x) did not
218 improve the IOA.

219 IOA was better for VIT than for Pap test, although the difference did not reach
220 statistical significance.

221 Better standardization of VIT diagnostic criteria, extensive training of screening
222 professionals followed by periodic refreshment courses and quality control monitoring
223 of screening sites hold promise of significantly improving cervical cancer screening.

224 **ETHICAL APPROVAL**

225 The Federal University of Santa Catarina ethics committee approved the study protocol
226 under number 681/10 on February 26, 2010.

227

228

References

230

231 1. Brasil. Ministério da Saúde. Instituto Nacional do Câncer. Estimativas 2014. Rio
232 de Janeiro: INCA; 2014: 39-40. Accessed May 5 2014. Available:
233 <http://www.inca.gov.br/estimativa/2014/estimativa-24042014.pdf> (Portuguese).

234

235 2. Brasil. Ministério da Saúde. Instituto Nacional do Câncer. Diretrizes brasileiras
236 para o rastreamento do câncer do colo do útero / Instituto Nacional de Câncer.
237 Coordenação geral de ações estratégicas. Divisão de apoio à rede de atenção
238 oncológica. *In*: Capítulo 1 – Rastreio de lesões precursoras do câncer do colo do
239 útero. Rio de Janeiro: INCA, pp 29-34, 2011. Accessed December 10 2013.
240 Available:
241 [http://www1.inca.gov.br/inca/Arquivos/Diretrizes_rastreamento_cancer_colo_ut](http://www1.inca.gov.br/inca/Arquivos/Diretrizes_rastreamento_cancer_colo_utero.pdf)
242 [ero.pdf](http://www1.inca.gov.br/inca/Arquivos/Diretrizes_rastreamento_cancer_colo_utero.pdf). (Portuguese).

243

244 3. World Health Organization-WHO. Cytology Screening in Middle-income
245 Countries. *In*: Cervical cancer screening in developing countries: report of a
246 WHO consultation. Geneva, Switzerland, pp 13–24, 2002. Accessed July 18
247 2014. Available: <http://whqlibdoc.who.int/publications/2002/9241545720.pdf>.

248

249 4. Stoler MH, Schiffman M. Interobserver reproducibility of cervical cytologic and
250 histologic interpretations: realistic estimates from the ASCUS-LSIL Triage
251 Study. *JAMA*. 2001;285(11):1500–5.

252

253 5. Denny L, Quinn M, Sankaranarayanan R. Chapter 8: Screening for cervical
254 cancer in developing countries. *Vaccine*. 2006;24 Suppl 3:71–7.

255

256 6. Sankaranarayanan R, Nessa A, Esmey PO, Dangou J-M. Visual inspection
257 methods for cervical cancer prevention. *Best Pract and Res Clin Obstet Gynecol*.
258 2012;26(2):221–32.

259

260 7. World Health Organization-WHO. Visual Inspection with Acetic Acid
261 Application (VIA) as an Alternative Approach to Cytology Screening in Low-
262 Income Countries. *In*: Cervical cancer screening in developing countries: report
263 of a WHO consultation. Geneva, Switzerland, pp 25-36, 2002. Accessed Aug 10
264 2014. Available: <http://whqlibdoc.who.int/publications/2002/9241545720.pdf>.

265

266 8. Sherigar B, Dalal A, Durdi G, Pujar Y, Dhumale H. Cervical cancer screening
267 by visual inspection with acetic acid-interobserver variability between nurse and
268 physician. *Asian Pacific J Cancer Prev*. 2010;11(3):619–22.

269

- 270 9. Solomon D, Davey D, Kurman R, Moriarty A, O'Connor D, Prey M, et al. The
271 2001 Bethesda System: terminology for reporting results of cervical cytology.
272 JAMA. 2002;287(16):2114–19.
273
- 274 10. Sankaranarayanan R, Wesley RS. A practical manual on visual screening for
275 cervical neoplasia. Lyon: IARCPress, pp 1-35, 2003. Accessed June 25 2012.
276 Available: <http://screening.iarc.fr/doc/viavilimanual.pdf>.
277
- 278 11. Russo E, Kupek E. Parallel screening with Papanicolaou smear and visual
279 Inspection tests significantly reduced false negative screening results in southern
280 Brazil. Br J Med Med Res. 2014;4(14):2783-95.
281
- 282 12. Alli PM, Ollayos CW, Thompson LD, Kapadia I, Butler DR, Williams BH, et al.
283 Telecytology: intraobserver and interobserver reproducibility in the diagnosis of
284 cervical- vaginal smears. Human Pathol. 2001;32(12):1318–22.
285
- 286 13. Santos ALF, Derchain S, Calvert E, Martins M, Dufloth R, Martinez E.
287 Desempenho do exame colpocitológico com revisão por diferentes observadores
288 e da captura híbrida II no diagnóstico da neoplasia intra-epitelial cervical graus 2
289 e 3. Cad Saúde Pública. 2003;19(4):1029–37. (Portuguese).
290
- 291 14. Sankaranarayanan R, Shastri SS, Basu P, Mahé C, Mandal R, Amin G et al. The
292 role of low-level magnification in visual inspection with acetic acid for the early
293 detection of cervical neoplasia. Cancer Detect and Prev. 2004;28:345–51.
294
- 295 15. Muwonge R, Manuel MDG, Filipe AP, Dumas JB, Frank MR,
296 Sankaranarayanan R. Visual screening for early detection of cervical neoplasia
297 in Angola. Int J Gynecol Obstet. 2010;111(1):68–72.
298
- 299 16. Ngoma T, Muwonge R, Mwaiselage J, Kawegere J, Bukori P, Sankaranarayanan
300 R. Evaluation of cervical visual inspection screening in Dar es Salaam,
301 Tanzania. Int J Gynecol Obstet. 2010;109(2):100–4.
302
- 303 17. Sritipsukho P, Thaweekul Y. Accuracy of visual inspection with acetic acid
304 (VIA) for cervical cancer screening: a systematic review. J Med Assoc Thai.
305 2010;93 Suppl 7:254–61.
306
- 307 18. Bomfim-Hyppólito S, Franco ES, Franco RGDMM, De Albuquerque CM,
308 Nunes GC. Cervicography as an adjunctive test to visual inspection with acetic
309 acid in cervical cancer detection screening. Int J Gynecol Obstet.2006;92(1):58–
310 63.
311

- 312 19. Sankaranarayanan R, Basu P, Wesley R, Maché C, et al. Accuracy of visual
313 screening for cervical neoplasia: results from IARC multicentric study in India
314 and Africa. *Int J Cancer*. 2004;110:907-13.
315
- 316 20. Sherigar B, Dalal A, Durdi G, Pujar Y, Dhumale H. Cervical cancer screening
317 by visual inspection with acetic acid-interobserver variability between nurse and
318 physician. *Asian Pacific J Cancer Prev*. 2010;11 (3):619–22.
319
- 320 21. Sellors JW, Jeronimo J, Sankaranarayanan R, Wright TC, Howard M,
321 Blumenthal PD. Assessment of the cervix after acetic acid wash: inter-rater
322 agreement using photographs. *Obstet Gynecol*. 2002;99 (4): 635–40.