Original paper

Why are Polish women diagnosed with invasive cervical cancer after negative cytology in the organized screening programme – a pilot reevaluation of negative Pap smears preceding diagnoses of interval cancers

Katarzyna Komerska[1], Anna Macios[1], Patrycja Glińska[2], Włodzimierz Olszewski[3], Joanna Didkowska[4], Urszula Wojciechowska[4], Michał F. Kamiński[1], Andrzej Nowakowski[1]

[1]Department of Cancer Prevention, Maria Sklodowska-Curie National Research Institute of Oncology, Warsaw, Poland
[2]Warsaw PhD School in Natural and BioMedical Sciences, Maria Sklodowska-Curie National Research Institute of Oncology in Warsaw, Poland
[3]Department of Pathology and Laboratory Diagnostics, Maria Sklodowska-Curie National Research Institute of Oncology in Warsaw, Poland
[4]Polish National Cancer Registry, Maria Sklodowska-Curie National Research Institute of Oncology in Warsaw, Poland

We have aimed to study reasons for reporting false-negative cytology results preceding diagnosis of interval cervical cancers (CC) in Poland.
Data on all Pap smears collected in the organised screening in 2010-2015 were retrieved from the electronic database and linked with Polish National Cancer Registry (PNCR) data. False-negative results were defined as those sampled and assessed normal up to 3.5 years before diagnosis of invasive CC. False-negative slides were then seeded among twice as many randomly selected slides from the same lab and reviewed independently by three expert cytomorphologists. New diagnosis was established when experts agreed on a result.
Of 48 selected false-negative slides, 1 case was diagnosed as a low-grade abnormality, 22 cases as a high-grade abnormalities, 3 cases as unsatisfactory for evaluation and 5 as no intraepithelial lesion of malignancy (NILM) by all three experts. There was no agreement in 17 cases. Percentages of agreement between experts was 64.6. Interobserver agreement rate was moderate with Fleiss' κ values.
Our pilot study indicates evaluation errors as the main reason of false-negative cytology preceding interval CC in the organized screening programme in Poland. True lack of abnormal cells on the slide is the next reason.

Key words: cytology, interval cancer, cervical cancer, cervical cancer screening.

Introduction

In 2018 cervical cancer (CC) was the fourth most frequently diagnosed malignancy and the fourth leading cause of cancer death worldwide among females [1]. In 2017 in Poland there were 2502 new cases and 1609 deaths due to CC which ranks it 8th as the most frequent female cancer and the 9th cancer-related reason of death in females [2]. Invasive CC are preceded by many years by preinvasive intraepithelial lesions which may be identified in screening and effectively treated. This may lead to major decline in both cancer incidence and mortality. Well-organized screening programmes can

contribute to 80% decline in CC incidence [3]. However this may be achieved only in countries with very high quality at every step of the screening programme which limits incidence of interval cancers – namely those which are diagnosed after false-negative screening tests before the next or after the last round of screening. Audit of interval cancers and deep insight into the reasons of false-negative screening tests results are a very important and integral part of quality assurance in screening programmes according to European guidelines for quality assurance in cervical cancer screening [4].

Organised Cervical Cancer Screening Programme (OCCSP) was implemented in Poland in 2006 [5]. Under the Programme, each insured Polish woman aged 25 to 59 is eligible for free of charge Pap smear test every 3 years, which is in accordance with the European guidelines [4]. Triage algorithm for abnormal smears is based on Recommendations of the Polish Gynecological Society and include repeated cytology in 6 months and colposcopy/colpscopcially directed biopsy depending on the severity of cytological abnormalities [6]. If screening cytology is negative, re-screening is scheduled after 3 years. Medical records of procedures performed in the OCCSP are stored in the central electronic database called SIMP (Polish: *System Informatyczny Monitorowania Profilaktyki*). It collects information and enables analysis of results of screening tests. The SIMP is partially linked with the treatment databases of the National Health Fund and allows selection of patients who are eligible for screening. Every diagnoses of invasive cancer in Poland is required by law to be reported to the Polish National Cancer Registry (PNCR).

Central Coordination Centre (Polish: *COK – Centralny Ośrodek Koordynujący*) within the Department of Cancer Prevention, Maria Sklodowska-Curie National Research Institute of Oncology, Warsaw, Poland, based on a contract with the Ministry of Health has, among others, the task to monitor quality of services in screening programmes run nationally in the country. These activities were limited mainly to on-site visits since the roll-out of the OCCSP but were gradually developed after 2017 into a more complete set of quality assurance activities. An audit of interval CC have never been performed and false-negative slides have never been reviewed in Poland before.

Therefore the aim of this pilot study was to get insight into the reasons for obtaining false-negative results of Pap tests preceding diagnosis of interval CC as an integral part of quality assurance process in the Polish OCCSP.

Material and methods

In 2018 COK retrieved data from SIMP on all Pap smears collected in the OCCSP in 2010-2015. There were 1 599 434 Pap smears collected in 2010 (795 992) and 2011 (803 442) within the OCCSP. Each of 1 575 662 women participating in the OCCSP, identified by her personal identification number, was attributed to her last slide. Slides were evaluated according to Bethesda 2001 terminology.

Data from SIMP concerning smears taken in 2010-2011 were then linked by personal identification number with data obtained from the PNCR on invasive cancers and their histology based on International Classification of Diseases for Oncology (ICD-O-3) coding collected by the PNCR.

False-negative Pap test results were defined as those assessed as No Intraepithelial Lesion of Malignancy (NILM) according to the Bethesda system in women diagnosed with invasive CC and recorded in the PNCR within 42 months since Pap test sampling.

For each false-negative Pap test random two slides collected in 2010-2015 were selected from the same laboratory. Blinded review of false-negative slides seeded among those randomly selected (in accordance with European guidelines for quality assurance in cervical cancer screening [4] stating that slides should be re-screened alongside negative and/or positive controls and the labels concealed) was performed independently by three cytomorphologists, each one with more than 20 years of experience in Pap test evaluation – among them two pathomorphologists, from one laboratory. Cytomorphologists were aware that false-negative slides were mixed among other slides, but had no knowledge about the number of them.

Due to lack of data we cannot state whether these interval cancers were detected by a subsequent screening test (in opportunistic screening before the end of interval) or were they symptomatic.

Statistical analysis

Diagnosis made by each expert was categorized into four categories based on the Bethesda system (TBS):

- NILM,
- unsatisfactory for evaluation,
- low-grade abnormalities: atypical squamous cells of undetermined significance (ASC-US), low-grade squamous intraepithelial lesion (LSIL),
- high-grade abnormalities: atypical squamous cells, cannot exclude HSIL (ASC-H), high-grade squamous intraepithelial lesion (HSIL), atypical glandular cells (AGC)/adenocarcinoma (ADC), squamous cell carcinoma (SCC).

If there was no agreement between all three experts slide was put in a category "No agreement".

We have selected this very strict way of reaching a final diagnosis and coding of cases for statistical analysis to reassure that only slides assessed by all three experts as satisfactory for evaluation and NILM

Table I. Results of review of 48 false negative smears: number of slides evaluated coherently by 3 experts with histopathological results of cervical cancer cases (unsatisfactory for evaluation vs. NILM vs. low-grade abnormalities vs. high-grade abnormalities).

Diagnosis made by 3 experts	Number of slides by histological diagnosis			False-negative slides evaluated coherently by 3 experts, N (%)
	SCC	ADC	SARC	
Unsatisfactory for evaluation	2	1	0	3 (6.3)
NILM	3	2	0	5 (10.4)
Low-grade abnormalities	0	1	0	1 (2.1)
High-grade abnormalities	17	4	1	22 (45.8)
No agreement	13	4	0	17 (35.4)
Total	35	12	1	48 (100.0)

NILM – no intraepithelial lesion of malignancy; SCC – squamous cell carcinoma; ADC – adenocarcinoma; SARC – carcinosarcoma

should trigger retesting at a standard interval or end of screening (for women aged 59 years).

Fleiss' κ was calculated as an index of all three observers agreement. Significance level of 0.05 was established. Values of κ coefficient was interpreted according to Landis and Koch [7] as: <0 – poor; 0.0-0.2 – slight; 0.2-0.4 – fair; 0.4-0.6 – moderate; 0.6-0.8 – substantial and 0.8-1.0 – almost perfect agreement.

Results

We obtained 379 records of interval CC in 2010-2011, including 374 with histopathological confirmation by ICD-O-3 codes. Altogether for our pilot study we managed to collect 142 Pap smears taken in 2010-2015 from three laboratories based in Silesian Voivodship (141 with negative result, 1 with positive result), including 48 cases with histopathological confirmation of CC within 42 months after the negative result. Some of the selected smears were utilised or damaged and excluded from the study, but we managed to obtain the final proportion of false-negative smears to those randomly selected close to 1:2.

Review of 48 false-negative slides

Among 48 women with false-negative Pap tests results obtained in 2010-2011 selected for this pilot analysis, 35 were diagnosed with invasive squamous cell carcinomas, 12 had confirmed adenocarcinomas and 1 was confirmed as carcinosarcoma. Every Pap sample was a conventional cytology slide. The unweighted Cohen's κ value for pathomorphologists (0.508) was lower than between any other pair of experts (0.713 and 0.557), so we concluded that in this study the position of the reviewer was not relevant to the level of agreement.

Diagnosis was established when three experts agreed on a result. Results of slide review by three experts with histopathological results are presented in Table I. There were 31 diagnoses altogether (64.6% of all slides). Of 48 smears all three experts diagnosed 1 case as a low-grade abnormality (2.1%), 22 (45.8%) cases as high-grade abnormalities, 3 (6.3%) cases as unsatisfactory for evaluation and 5 (10.4%) as NILM (Table I). There was no agreement in 17 (35.4%) cases.

Of 12 smears taken before adenocarcinoma confirmation, 4 was re-classified by experts as high-grade abnormalities, 1 as low-grade abnormalities, 2 as a NILM and 1 as unsatisfactory for evaluation. There was no agreement between experts on 4 cases.

All three experts diagnosed unanimously 26 (54.2%) cases as abnormal.

The interobserver agreement by percentages was 64.6. Fleiss' κ coefficient was 0.593 for three experts.

Review of control slides

Among 94 control smears that we selected for this pilot analysis, 93 were previously diagnosed as NILM and 1 as LSIL.

There were 49 diagnoses established (three experts' agreement, 52.1% of all control smears). There were 5 diagnoses of low-grade abnormalities (5.3%), 8 (8.5%) cases diagnosed as a high-grade abnormalities, 7 (7.4 %) cases as unsatisfactory for evaluation and 29 (30.9%) as NILM (Table II). There was no agreement in 45 (47.9%) cases. Without the distinction between low-grade and high-grade abnormalities, the experts classified 22 (23.4%) smears as abnormal.

The interobserver agreement by percentages was 52.3. Fleiss' κ coefficient was 0.500 for three experts.

Discussion

Exfoliative cytology, although enabled implementation of mass screening for cervical neoplasia in developed countries as a fairly simple, inexpensive and non-invasive method, is not a perfect screening test. As a subjective method relying on visual, microscopic interpretation of cells by human eye, it has

Table II. Results of review of 94 control smears: number of slides evaluated coherently by 3 experts (unsatisfactory for evaluation vs. NILM vs. low-grade abnormalities vs high-grade abnormalities)

Diagnosis	Control slides evaluated coherently by 3 experts, N (%)
Unsatisfactory for evaluation	7 (7.4)
NILM	29 (30.9)
Low-grade abnormalities	5 (5.3)
High-grade abnormalities	8 (8.5)
No agreement	45 (47.9)
Total	94 (100.0)

NILM – no intraepithelial lesion of malignancy

limited sensitivity and moderate reproducibility and accuracy [8, 9, 10]. According to the literature, interpretation of cytological specimens is dependent on the experience, skills and time spent on examination of the smear by the screener and a consequence of this is inter- and intra-observer variation [4]. Interobserver agreement of three experts in our study was moderate. Previous studies showed similar results in cervical cytology [8, 11, 12].

Most common reasons why false-negative reports may be generated are: inadequate sampling, handling or staining of the sample, misinterpretation of a positive smear containing abnormal cells as a negative slide [13, 14, 15, 16]. A false-negative "normal" cytology imposes no indication for colposcopy and return for screening after an interval or exit from screening for oldest screening cohorts, so diagnosis and treatment are delayed and prognosis may worsen because invasive cancers may be diagnosed at symptomatic, advanced stages.

To maximally limit the false-negative reports, organised cytology-based screening programmes require extensive quality assurance measures. Audit of interval cancers is an important part of these activities [4].

In our pilot study we have for the first time identified major reasons for obtaining a negative result of screening cytology preceding diagnosis of interval CC in the Polish OCCSP as a part of quality assurance process. Over 60% of Pap smears initially classified as NILM preceding diagnosis of invasive CC were re-classified as abnormal/unsatisfactory for evaluation by expert review (including 45.8% of them evaluated as high-grade abnormalities), 10.4% initially negative slides in routing screening practice were assessed as satisfactory for evaluation and graded normal (NILM) by expert review. Moderate agreement (Fleiss' K 0.593) was noted between 3 independent expert cytologists on blinded reassessment of false-negative Pap smears. In 35.4% of slides the expert review was discordant. On the grounds of our pilot results we may point at interpretation/assessment errors as the main causes of false-negative cytology results in the Polish OCCSP. The results of review of control slides seems to uphold that conclusion – three experts independently agreed on finding high-grade abnormalities in 8.5% and low-grade abnormalities in 5.3% of control slides.

Our experts' review results that re-classify 54.2% smears as positives, are consistent with those of Kenter *et al*. reporting 53% of negative smears sampled 3.5 years before the diagnosis of squamous cell carcinoma re-classified as positive [17]. In audit performed in Denmark 9.8% of the 112 women diagnosed with a cervical malignancy were found to have false-negative cytological samples (defined as "normal" cytological sample prior to cancer diagnosis with post-audit diagnosis of HSIL or worse) [18]. DeMay reports in his study 51.9% of slides reclassified from "normal" to "abnormal" (of 655 reported smears) [19].

Many solutions to limit the rate of assessment errors have been proposed and implemented worldwide such as lab and personnel certification/recertification, continuous educations and training programs, rapid or two-stage reassessment of all or randomly selected slides [4]. Initiatives have been undertaken in Poland to implement new or intensify some of these measures which are already in place. Also rescreening is mandatory in the OCCSP in Poland and rapid rescreening of 10% of slides is dominating.

Quality of the smear is an important component of assessment of the Pap sample according to TBS (2001 or revisioned 2014) [20]. There are various determinants of "unsatisfactory for evaluation" quality of conventional Pap smears such as inadequate cellularity, obscuring blood, obscuring inflammation [21, 22]. Some of them may be attributed to sampling errors [19, 23]. Sample assessed as unsatisfactory for evaluation should be retaken promptly. In our pilot study 6.3% of false-negative smears initially assessed as satisfactory for evaluation were reassessed as unsatisfactory for evaluation by all three experts. In 2008, of all smears taken within the OCCSP 0.9% were unsatisfactory [24], in 2018 and 2019 according to SIMP there were respectively 0.6% and 0.7% unsatisfactory smears. In studies conducted in the United States, Netherlands or Italy the percentage of unsatisfactory smears was between 1.11 and 1.4 [25, 26]. The rate of unsatisfactory conventional cervical cytology is considerably lower in the Polish OCCSP. Our results indicate that more attention should be paid to the quality of the smear by Polish cytomorphologists and the rate of "unsatisfactory for evaluation" should be higher which should limit the unsatisfactory

quality of the smear as the reason for obtaining false negative Paps.

In our pilot study minority of false-negative Pap smears (10.4%) was related to true lack of abnormal cells on the slides and were reassessed normal by three experts blinded to the original diagnosis. There may be two causes of this phenomenon: 1) sampling errors or 2) the nature of the lesion which does not allow for proper collection of abnormal cells such as its submucosal location deep in the glands covered by normal epithelium or endocervical localisation. The first is potentially addressed by further training of gynaecologists and midwives to improve sampling quality. The second may be potentially limited by implementation of more sensitive screening tools such as molecular HPV testing.

In our study close to 20% of interval CC cancers were adenocarcinomas and this rate is over twice as high as the proportion of this histotype among all diagnosed CCs in Poland. These findings are consistent with many other studies indicating lower sensitivity of cytology in detecting glandular neoplasia than most common squamous cell lesions [27, 28]. In extensively screened populations this may be responsible for a relative increase in adenocarcinoma incidence among all cervical cancer in recent years [29, 30].

We are aware of the limitations of our pilot analysis. A relatively small number of slides was reassessed and majority of them (64.6%) was retrieved from just one laboratory for logistic reasons and accessibility. However, we have already initiated revision of all false-negative slides from all laboratories in Poland beginning from 2010 up to most recent times to obtain a full picture of reasons for false-negative reports of cytology results preceding interval cancers in the Polish OCCSP. We have also already used the most difficult and controversial slides as the educational material during organized courses for cytomorphologists. Also, our 3 experts came from one laboratory which could have influence review. At present the review of all accessible false-negative slides is performed by 3 most experienced experts in gynaecological cytology in Poland with more than 30 years of experience, the title of Professor of medicine and extensive background as researchers and trainers. These actions in the framework of quality assessment and assurance should result in more effective CC screening in the country, even after possible changes within the OCCSP i.e. transition to primary HPV-based screening with reflex Liquid Based Cytology incorporated into triage protocols.

Conclusions

In our pilot analysis interpretation errors followed by true absence of abnormal cells on the slides are responsible for false-negative Pap smear results preceding diagnosis of interval CC in the Polish OCCSP. Our pilot results initiated full, comprehensive review of all false-negative slides and further development of quality assurance in the screening program in our country.

The authors declare no conflict of interest.

References

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68: 394-424.
2. Didkowska J, Wojciechowska U, Czaderny K, et al. Cancer in Poland in 2017. Polish National Cancer Registry, Department of Epidemiology and Cancer Prevention 2019.
3. Antilla A, Aoki D, Arbyn M, et al. IARC Handbooks of Cancer Prevention. Vol. 10: Cervix Cancer Screening 2005.
4. Arbyn M, Anttila A, Jordan J, et al. European guidelines for quality assurance in cervical cancer screening, second edition, IARC 2008.
5. Stępień M, Epidemiologia i profilaktyka nowotworów szyjki macicy. Kliniczne czynniki ryzyka zachorowań. In: Nowotwory szyjki macicy. Centrum Medyczne Kształcenia Podyplomowego Warszawa 2011.
6. Management of abnormal PAP smear – consensus guidelines of the National Cervical Cancer Screening Programme in Poland Coordinating Centre, the Polish Gynaecologic Society, the Polish Society of Pathologists and the Polish Society of Colposcopy and Uterine Cervix Pathology. Ginekol Pol 2009; 80: 129-133.
7. StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LLC.
8. Cai B, Ronnett BM, Stoler M, et al. Longitudinal Evaluation of Interobserver and Intraobserver Agreement of Cervical Intraepithelial Neoplasia Diagnosis Among an Experienced Panel of Gynecologic Pathologists. Am J SurgPathol 2007; 31: 1854-1860.
9. Stoler MH, Schiffman M. Atypical Squamous Cells of Undetermined Significance-Low-grade Squamous Intraepithelial Lesion Triage Study G. Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study. JAMA 2001; 285: 1500-1505.
10. Baena A, Guevara E, Almonte M, et al. Factors related to inter-observer reproducibility of conventional Pap smear cytology: a multilevel analysis of smear and laboratory characteristics. Cytopathology 2017; 28: 192-202.
11. Stoler M, Schiffman M. Interobserver reproducibility of cervical cytologic and histologic interpretations realistic estimates from the ASCUS-LSIL triage study. JAMA 2001; 285: 1500-1505.
12. Sørbye SW, Suhrke P, Revå BW, et al Accuracy of cervical cytology: comparison of diagnoses of 100 Pap smears read by four pathologists at three hospitals in Norway. BMC Clin Pathol 2017; 17: 18.
13. Koss LG, Sherman ME, Cohen MB, et al. Significant reduction in the rate of false-negative cervical smears with neural network-based technology (PAPNET Testing System). Hum Pathol 1997; 28: 1196-1203.
14. Barton SE, Jenkins D, Hollingworth A, et al. An explanation for the problem of false-negative cervical smears. Br J Obstet Gynaecol 1989; 96: 482-485.

15. Rokita W. [The diagnostic value of cytology and colposcopy in women with cervical intraepithelial neoplasia]. Ginekol Pol 2011; 82: 607-611.
16. Kenter GG, Schoonderwald EM, Koelma IA, et al. The cytological screening history of 469 patients with squamous cell carcinoma of the cervix uteri; does interval carcinoma exist? Acta Obstet Gynecol Scand 1996; 75: 400-403.
17. Renshaw AA, Analysis of error in calculating the false-negative rate in the interpretation of cervicovaginal smears. Cancer 1997; 81: 264-271.
18. Kirschner B, Poll S, Rygaard C, et al. Screening history in women with cervical cancer in a Danish population-based screening program, Gynecologic Oncol 2011; 120: 68-72.
19. DeMay RM. Cytopathology of false negatives preceding cervical carcinoma. Am J Obstet Gynecol 1996; 175 (4 Pt 2): 1110-1113.
20. Solomon D, Davey D, Kurman R, et al. The 2001 Bethesda System: terminology for reporting results of cervical cytology. JAMA 2002; 287: 2114-2119.
21. Moriarty AT, Clayton AC, Zaleski S, et al. Unsatisfactory reporting rates: 2006 practices of participants in the college of american pathologists interlaboratory comparison program in gynecologic cytology. Arch Pathol Lab Med 2009; 133: 1912-1916.
22. Gupta S, Sodhani P, Sardana S, et al. Clinical determinants and smear characteristics of unsatisfactory conventional cervicovaginal smears. Eur J Obstet Gynecol Reprod Biol 2013; 168: 214-217.
23. Ruba S, Schoolland M, Allpress S, et al. Adenocarcinoma in situ of the uterine cervix: screening and diagnostic errors in Papanicolaou smears. Cancer 2004; 102: 280-287.
24. Spaczynski M, Michalska M, Januszek-Michalecka L. [Central Coordinating Center Summary of the National Cervical Cancer Screening Program in Poland in the period from January 1st till December 31st 2008]. Ginekol Pol 2009; 80: 220-226.
25. Castle PE, Bulten J, Confortini M, et al. Age-specific patterns of unsatisfactory results for conventional Pap smears and liquid-based cytology: data from two randomised clinical trials. BJOG 2010; 117: 1067-1073.
26. Amies AM, Miller L, Lee SK, et al. The effect of vaginal speculum lubrication on the rate of unsatisfactory cervical cytology diagnosis. Obstet Gynecol 2002; 100 (5 Pt 1): 889-892.
27. Baldur-Felskov B, Munk C, Nielsen TS, et al. Trends in the incidence of cervical cancer and severe precancerous lesions in Denmark, 1997-2012. Cancer Causes Control 2015; 26: 1105-1116.
28. Bray F, Carstensen B, Møller H, et al. Incidence trends of adenocarcinoma of the cervix in 13 European countries. Cancer Epidemiol Biomarkers Pre 2005; 14: 2191-2199.
29. Wang J, Elfström KM, Andrae B, et al. Cervical cancer case-control audit: Results from routine evaluation of a nationwide cervical screening program. Int J Cancer 2020; 146: 1230-1240.
30. Smith HO, Tiffany MF, Qualls CR, et al. The rising incidence of adenocarcinoma relative to squamous cell carcinoma of the uterine cervix in the United States – a 24-year population-based study. Gynecol Oncol 2000; 78: 97-10.

Address for correspondence

Katarzyna Komerska
Department of Cancer Prevention
Maria Sklodowska-Curie National Research Institute of Oncology
Roentgena 5
02-781 Warsaw, Poland
e-mail: katarzyna.komerska@pib-nio.pl