Pathology (June 2015) **47(4)**, pp. 308–313

ANATOMICAL PATHOLOGY

High reproducibility of histological diagnosis of human papillomavirus-related intraepithelial lesions of the anal canal

Jennifer M. Roberts[1], Fengyi Jin[2], Julia K. Thurloe[1], Clare Biro[1], Isobel M. Poynten[2], Sepehr N. Tabrizi[3,4,5], Christopher K. Fairley[6], David J. Templeton[2,7], Andrew D. Carr[8], Suzanne M. Garland[3,4,5], Richard J. Hillman[9], Alyssa M. Cornall[3,4,5], Andrew E. Grulich[2] and Annabelle Farnsworth[1]

[1]GynaePath, Douglass Hanly Moir Pathology, Sydney, [2]The Kirby Institute, UNSW Australia, Sydney, NSW, [3]Regional HPV Labnet Reference Laboratory, Department of Microbiology and Infectious Diseases, The Royal Women's Hospital, Melbourne, [4]Murdoch Childrens Research Institute, Melbourne, [5]Department of Obstetrics and Gynaecology, University of Melbourne, [6]Melbourne Sexual Health Centre and Central Clinical School, Monash University, Melbourne, Vic, [7]RPA Sexual Health, Sydney Local Health District, Sydney, [8]Centre for Applied Medical Research, St Vincent's Hospital, Sydney, and [9]Western Sydney Sexual Health Centre, University of Sydney, Sydney, NSW, Australia

Summary

In a natural history study of anal human papillomavirus (HPV) infection and HPV-related lesions, we examined the reproducibility of histological high-grade squamous intraepithelial lesion (HSIL). Three expert anogenital pathologists share the reporting of histological specimens from the Study of the Prevention of Anal Cancer (SPANC), utilising Lower Anogenital Squamous Terminology (LAST) criteria. In total, 194 previously reported biopsies were randomly chosen within diagnostic strata [50 HSIL–anal intraepithelial neoplasia (AIN) 3; 45 HSIL–AIN 2; 49 'flat' low-grade squamous intraepithelial lesion (LSIL); 50 'exophytic' LSIL; and 50 negative for squamous intraepithelial lesion] and reviewed by each of these three pathologists. Consensus was defined as agreement between at least two review diagnoses, using a binary classification of HSIL and non-HSIL, or if consensus was not obtained in this way, it was achieved through a multiheader microscope session by the three pathologists. We found very high agreement between original and consensus diagnoses (Kappa = 0.886) and between each pathologist's review and consensus (Kappas = 0.926, 0.917 and 0.905). Intra-observer agreement for the three pathologists was 0.705, 1.000 and 0.854. This high level of diagnostic reproducibility indicates that the findings of SPANC should be robust and provide reliable information about HPV-related anal canal disease.

Key words: Anus neoplasms, human papillomavirus, kappa statistics, squamous intraepithelial lesion.

Received 13 November, revised 12 December, accepted 22 December 2014

INTRODUCTION

Squamous cell carcinoma (SCC) of the anal canal shares many similarities with cervical SCC. Both have a strong causal relationship with high-risk human papillomavirus (HPV), in particular HPV16, and both are preceded by an intraepithelial precursor lesion termed high-grade squamous intraepithelial lesion (HSIL) at the transformation zone. The histological appearance of HSIL is virtually identical in these two (and other anogenital) sites. However, the degree of similarity between the natural history of anal canal and cervical HSIL is uncertain.[1] Further, while the anal canal is accessible to cytological sampling, it is not clear whether anal cytological screening for HSIL (with treatment of lesions) will reduce the incidence of cancer to the same extent that programs based on cervical cytology have reduced the incidence of cervical cancer.[2,3]

The Study of the Prevention of Anal Cancer (SPANC) is a longitudinal, natural history study, exploring the epidemiology of anal HPV infection and related epithelial lesions among a community-recruited cohort of homosexual men in Sydney, Australia. The end point for many analyses in this study involves a histological diagnosis of HSIL, made on high resolution anoscopy (HRA)-guided biopsy. However, variability in biopsy interpretation has been acknowledged in both cervix and anal canal.[4–9]

In this paper, we describe histopathological diagnosis methods, and report inter-observer and intra-observer reproducibility of histological diagnosis in SPANC.

MATERIALS AND METHODS

Ethics approval for SPANC was granted by the Human Research Ethics Committee at St Vincent's Hospital, Darlinghurst, NSW.

A detailed description of the design of SPANC has been published elsewhere.[10] In brief, the study aims to recruit 600 homosexual men aged ≥35 years, both HIV-positive and HIV-negative, from Sydney community-based settings. Signed informed consent was obtained from all participants. Each participant has five clinic visits over a 3-year period and recruitment will be complete in mid-2015. At each visit, men undergo a digital anorectal exam, an anal Papanicolaou (Pap) ThinPrep (Hologic, USA) test for anal cytology and HPV DNA detection and genotyping, followed by HRA during which lesions suspected of being HPV-related are biopsied.

Histopathological processing and diagnoses

All cytological and histological specimens are referred to Gynaepath, a specialist anogenital unit within Douglass Hanly Moir Pathology, a large private general pathology laboratory in Sydney. Three anatomical pathologists, each with approximately 20 years of experience in diagnosing anogenital HPV-related pathology, are involved in the study, and share the reporting of all specimens.

Print ISSN 0031-3025/Online ISSN 1465-3931
DOI: 10.1097/PAT.0000000000000246

Biopsies are received in 10% formalin and are processed in a routine fashion. Seven levels of each biopsy are prepared. Levels 1–3 and 5–7 are routinely stained with haematoxylin and eosin (H&E). Level 4 is prepared on a coated slide and left unstained for potential use for histochemical staining or immunostaining. Each case is viewed and reported by one of the three study pathologists. All reporting is performed blinded to clinical factors (other than age) and previous results.

Reporting of the biopsies is in accordance with criteria, terminology and recommendations of the Lower Anogenital Squamous Terminology (LAST) Project.[11,12] The following details are particularly relevant to this study. Firstly, with respect to non-invasive HPV-related disease, the following terms are used: negative for squamous intraepithelial lesion (SIL), low-grade squamous intraepithelial lesion (LSIL) and high-grade squamous intraepithelial lesion (HSIL). Secondly, within the category of HSIL, further subdivision is performed, into anal intraepithelial neoplasia (AIN) grade 2 and AIN grade 3. Thirdly, immunostaining for p16 INK4A (p16) is performed on the unstained spare slide in the following two circumstances: to differentiate HSIL from a benign mimic, in particular, atypical immature metaplasia, but also inflammatory/reactive changes, tangential sectioning and partial epithelial denudation; and when a diagnosis of AIN 2 is proposed, in order to confirm positive p16 reactivity. If the result is negative in the latter circumstance, the lesion is then classified as LSIL or negative for SIL. Fourthly, immunostaining for p16 is not used in formulating diagnoses of straightforward LSIL or straightforward HSIL–AIN 3. And finally, immunostaining for p16 is reported as positive if there is 'continuous strong nuclear or nuclear plus cytoplasmic staining of the basal cell layer with extension upwards involving at least one third of the epithelial thickness'.[11,12] Any other pattern of staining is regarded as negative.

In addition to the LAST categories above, a decision was made to subcategorise LSIL into 'exophytic' and 'flat', based on presence or absence respectively of architectural features of condylomata acuminata. This decision reflected the possibility that these two variations of LSIL may be associated with different HPV types.[13,14]

Sub-study of reliability and repeatability of histological diagnoses

In order to assess inter- and intra-observer variability, we aimed to select a stratified random sample of 200 previously reported baseline visit biopsies (from a total of 502 reported biopsies), including 50 biopsies from each of the major diagnostic categories (HSIL–AIN 3, HSIL–AIN 2, LSIL and Negative). However, only 45 AIN 2 biopsies had been reported at the time of the study and only 49 LSIL biopsies were included, resulting in a total of 194 biopsies. Each biopsy had its two H&E slides de-identified. The three pathologists in turn each reviewed these H&E slides. If the pathologist required a p16 immunostain for diagnostic purposes, this was either retrieved from file by clerical staff (if it had been performed at the time of original reporting) or it was performed at that time on the stored spare unstained slide. The pathologist was blinded to when p16 staining was performed.

In this manner, three 'review' diagnoses were obtained. Agreement of at least two review diagnoses was defined as 'consensus' diagnosis. If consensus was not achieved in this way, the three pathologists were required to discuss the case at the multiheader microscope to achieve a consensus diagnosis.

Therefore for each biopsy, there was an original diagnosis, three review diagnoses (one by each of the reviewing pathologists) and a consensus diagnosis.

The primary hypothesis was that there would be a high level of agreement between the original diagnosis and the consensus diagnosis. We also anticipated high agreement between each reviewer's diagnosis and the consensus diagnosis and between the three reviewers. Each of these comparisons constitutes inter-observer repeatability. Finally, we predicted that there would be very high agreement between the original and review diagnoses of each pathologist (intra-observer repeatability). Comparisons were based on a binary classification, namely HSIL and non-HSIL, as in the context of both a natural history study and a screening program leading to treatment of HSIL, this is the distinction that is thought to have clinical significance.[15]

Statistical analyses were performed using Stata 12.1 (Stata Corporation, USA). We calculated Kappa statistics and 95% confidence intervals (CIs) to estimate the inter- and intra-observer agreements between the three pathologists. The unweighted kappa statistic was used to analyse binary data. Levels of agreement as reported by Landis and Koch[16] were used as follows: <0.0 = 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.

RESULTS

By December 2012, a total of 283 participants were enrolled into SPANC and this histological sub-study included 194 biopsies from 141 men. The median age of these men was 49 years (range 35–76 years).

Inter-observer repeatability

Table 1 is a summary of the prevalence of the different grades of anal squamous intraepithelial lesion in the 194 biopsies as reported originally, at consensus and for each of the three reviews.

Comparing original and consensus diagnoses, the agreement for diagnosis of HSIL versus other diagnoses was in the 'almost perfect' range ($\kappa = 0.886$, 95% CI 0.802–0.936).

Within the HSIL category, there was 'movement' between AIN 2 and AIN 3, with 11 original AIN 3 results becoming AIN 2 at consensus. The converse occurred for 10 original AIN 2 results. In addition, four original AIN 3 cases and seven original AIN 2 cases were downgraded to Negative or LSIL at consensus. Within the Negative/LSIL category, there was also 'movement', predominantly within the category [i.e., between Negative, exophtic LSIL (E-LSIL) and flat LSIL (F-LSIL)]. Negative and E-LSIL numbers increased at consensus, predominantly at the expense of F-LSIL.

There was excellent agreement between each pathologist's review and consensus $\kappa = 0.926$ (95% CI 0.852–0.964), 0.917 (95% CI 0.840–0.958) and 0.905 (95% CI 0.823–0.949) for pathologists A, B and C, respectively.

There was 'almost perfect' agreement when the results of all three pathologists were compared ($\kappa = 0.839$, 95% CI 0.766–0.891).

Table 1 A summary of the prevalence of the different grades of anal squamous intraepithelial lesion in the 194 biopsies as reported originally, at consensus and for each of the three reviews

	Original		Consensus		Reviewer A		Reviewer B		Reviewer C	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%
Negative/LSIL	99	51.0	110	56.7	113	58.3	102	52.6	117	60.3
Neg	50	25.8	61	31.4	28	14.4	64	33.0	66	34.0
Exophytic LSIL	25	12.9	32	16.5	34	17.5	27	13.9	32	16.5
Flat LSIL	24	12.4	17	8.8	51	26.3	11	5.7	19	9.8
HSIL	95	49.0	84	43.3	81	41.8	92	47.4	77	39.7
AIN 2	45	23.2	39	20.1	37	19.1	37	19.1	39	20.1
AIN 3	50	25.8	45	23.2	44	22.7	55	28.4	38	19.6

AIN, anal intraepithelial neoplasia; HSIL, high-grade squamous intraepithelial lesion; LSIL, low-grade squamous intraepithelial lesion.