

Title: Interobserver Reproducibility of cervical histological diagnoses obtained by pathologists under routine conditions of a pragmatic randomized trial (ASCUS-COL trial)

Authors: MC Agudelo¹, M Riveros², A Baena¹, D Suescun³, C Lopez⁴, LJ Gomez⁵, J Castaño^{4,5}, M Roldan^{4,5}, P Sasieni⁶, M Almonte⁷, R Herrero⁷, PE Castle⁸, M Stoler⁹, GI Sanchez¹

¹Grupo Infección y Cáncer, Universidad de Antioquia, Medellín, Colombia.. ² Hospital Pablo Tobón Uribe, Medellín, Colombia ³Laboratorio de Patología y Citología Suescun, Medellín, Colombia. ⁴Departamento de Patología Hospital Universitario San Vicente Fundación, Medellín, Colombia. ⁵Dinamica IPS, Medellín, Colombia. ⁶Centre for Cancer Prevention, Queen Mary University of London, London, UK. ⁷Prevention and Implementation Group, International Agency for Research on Cancer. Lyon, France. ⁸Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA, ⁹Department of Pathology and Laboratory Medicine, University of Virginia, Charlottesville, VA, USA.

Background. Adequate interpretation and reproducibility of cervical histopathological diagnoses is fundamental to decide whether to treat women with abnormal screening tests. We evaluated the interobserver reproducibility of histopathological diagnoses of biopsies and endocervical curettages (ECC) obtained under routine conditions in a pragmatic trial that compared 3 strategies (IC=Immediate colposcopy, RC=repeat cytology and HPV=high-risk HPV test) to manage women with ASC-US cytology. **Methods.** 2.332 slides corresponding to 646 diagnostic reports of all CIN1+ cases and a subset of 393 negative were recovered from pathology labs that provided services to the 2.661 women included in the trial. Slides were re-read blindly and retrospectively by two expert pathologists who issued a single diagnosis. The percentage agreement and unweighted and weighted Kappa values were obtained for histological diagnoses with unique specimens: 652 biopsies and 153 ECCs for 5 (Negative, CIN1, CIN2, CIN3, SCC and Adenocarcinoma) or 3 (Neg, CIN1 and \geq CIN2) categories. **Results.** Agreement and kappa value, respectively, were 51.5% (95%CI:47.7-55.3%) and 0.32 (95%CI:0.28-0.37%) for biopsies and 66.7% (95%CI:58.9-73.6%) and 0.38 (95%CI: 0.27-0.48%) for ECCs. Agreement and kappa values did not differ when using the 3 categories. Among 91 lesions diagnosed as CIN3 by expert pathologists, 15 (16.5%) have been classified as CIN3, 39 (42.9%) as CIN2, 36 (39.6%) as CIN1, and 1 (1.1%) as Negative under routine conditions. Fourteen and 7% of undercalled diagnoses were observed in IC and HPV arms, respectively. There was no undercalling in the RC arm. **Conclusions.** Reproducibility was significantly higher in women with \geq ASC-US in RC arm (square-weighted kappa 0.90) and hrHPV+ in HPV arm (square-weighted kappa 0.70) than in IC arm (square-weighted kappa 0.40). If experts have the correct diagnosis, about 40% of women were not properly referred to treatment. Correct adjudication of discordant diagnoses is required to estimate the degree of accuracy of histopathology diagnoses.