Chapter 8

Summary, discussion, and guidelines

Before cost-effective implementation of nation-wide hrHPV testing in cervical screening can be envisaged an optimal balance between clinical sensitivity and specificity should be reached to minimize redundant follow-up pwwrocedures. In order to translate such clinical requirements into test performances we performed the studies described in this thesis. Our findings are discussed below and, inferred from these, directions to guidelines for hrHPV DNA test requirements for use in primary screening are given.

In the first part of this thesis **(Chapter 2-4)** we aimed to find out which of the commonly used hrHPV DNA detection assays has the most optimal balance between clinical sensitivity and specificity for high-grade CIN and cervical cancer. To that end several two-way comparisons were performed.

It appeared that, compared to hc2, a commercially available DNA in situ hybridization (ISH) assay lacked sufficient sensitivity for prevalent high-grade CIN or cervical cancer **(Chapter 2)**. Since in situ hybridization positivity was only found in smears displaying relatively high viral loads, as deduced from the corresponding hc2 RLU/CO values, it can be concluded that the ISH assay used suffered form an insufficient analytical sensitivity to reach an optimal clinical sensitivity

Comparison of the automated version of the hc2 assay and GP5+/6+-PCR in a cross-sectional study of women participating in a population-based cervical screening trial, revealed that both assays had nearly similar sensitivities for high-grade CIN or cervical cancer **(Chapter 3)**. However, an higher overall positivity rate of the hc2 resulted in a significantly lower clinical specificity compared

to that of GP5+/6+-PCR. An important observation was that we could significantly improve the clinical specificity of hc2 for high-grade CIN or cervical cancer by increasing the cut-off for positivity (from 1 to 2 RLU/CO) without a loss in clinical sensitivity. Differences in clinical specificity and sensitivity between the hc2 and GP5+/6+-PCR disappeared at a cut-off of 3 RLU/CO. Therefore, both assays are considered similarly accurate in the detection of high-grade CIN or cancer, provided that the hc2 cut-off is slightly adapted.

In **Chapter 4** we compared in a case-control format the GP5+/6+-PCR-EIA with the ultra-sensitive SPF_{10} assay to detect ≥CIN2 in women with normal cytology. The application of SPF_{10} did not lead to an increase in clinical sensitivity for high-grade CIN or cervical cancer, but instead a significantly decrease in clinical specificity compared to that of the GP5+/6+-PCR. The extra positivity scored by SPF_{10} mainly involved infections characterized by a very low viral load that do not result in high-grade CIN or cervical cancer.

Taken together, the data from **Chapters 2 to 4** indicate that the analytical sensitivity of the ISH assay is too low, and that of SPF_{10} too high for use in a screening setting. The clinical performances of GP5+/6+-PCR-EIA and hc2 were nearly at the same level, in line with data of large screening trials (1-3). However, the clinical performance of these assays remains subjected to improvement. In order to address this issue we performed additional viral load analysis for the most common high-risk HPV types on women that are GP5+/6+-PCR positive for the respective types. To that purpose we firstly developed and validated real-time PCR assays to quantify the amount of HPV16, -18,

-31, and –33 load, as well as that of the β-globin housekeeping gene **(Chapter 5)**. Using these assays, we were able to calculate the amount of HPV copies per cell.

In a subsequent cross-sectional study we applied these assays to define HPV16, -18, -31, and 33 type-specific viral load thresholds above which all women with abnormal cytology and prevalent ≥CIN3 lesions were discerned **(Chapter 6)**. These thresholds were based on the 33rd percentile of the lowest viral load values found in women with normal cytology participating in a population-based screening program. Hence, prevalent ≥CIN3 lesions may be excluded in at maximum 33% of the women positive for the respective HPV types when using additive viral load analysis. It furthermore became clear that at this threshold, absolute viral load values differed considerably between the different types.

Subsequently, in **Chapter 7** we evaluated various viral load thresholds, defined in **Chapter 6**, in a longitudinal population-based screening trial with 18 months of follow-up. Despite the fact that for most of the examined types viral load levels were predictive for high-grade CIN or cervical cancer, none of the thresholds could exclude 18 month cumulative CIN2, CIN3, or cervical cancer. This was mainly due to the wide overlap between viral load levels in women with and without high-grade lesions. In addition, viral load assessment had no additive value over cytology when applied to women with a positive hrHPV GP5+/6+-PCR-EIA test in this screening population.

Taken together, our data indicate that additional viral load assessment cannot enhance the clinical specificity of the GP5+/6+-PCR assay. On the other hand, an increase in analytical sensitivity relative to GP5+/6+-PCR

would result in a marked decrease in clinical specificity, without having a major impact on the clinical sensitivity. Therefore, it can be concluded that, amongst the assays analyzed, the GP5+/6+-PCR and the compatible hc2 suit best for screening purposes. This is in line with the finding that these two assays have proven to be of clinical value in large screening trials (1-3) and therefore can be considered as clinically validated for cervical screening.

Since both median and clinically relevant viral load levels in women of a screening population differ markedly between the various hrHPV types (chapter 6) clinical test requirements cannot easily be translated into analytical test requirements in terms of setting universal assay cut-off points. Instead, clinical criteria should form the basis of formulating directions to guidelines for hrHPV test requirements for primary cervical screening. A proposal for such guidelines and clinical validation strategy necessary for fulfillment of these guidelines in the European setting are presented below. In addition, indications for constant quality assurance of hrHPV testing by the laboratories are proposed.

Requirements of HPV tests in primary cervical screening

In a primary cervical screening setting a HPV detection assay should fulfill the following requirements:

1. The candidate test should have a clinical sensitivity for ≥CIN3 of at least 95% ± 3%. This percentage is chosen to assure a very high negative predictive value of the HPV detection assay and is based on data obtained from large clinical trials performed

by hc2 and GP5+/6+-PCR. A meta analysis showed the pooled sensitivity for hc2 to be 97.9% (95%CI: 95.9-99.9) in primary screening in Europe and North-America (4). The GP5+/6+-PCR revealed a sensitivity of 94.1% (95%CI: 91.7-95.9) (5).

2. The candidate test should have a clinical specificity for ≥CIN3 of at least 95% ± 2% in women over 40 years of age. This percentage is based on data obtained from the POBASCAM indicating that the GP5+/6+-PCR has a clinical specificity for women >40 years of age of 97.3% (95%CI: 97.1-97.5) (personal communication J. Berkhof and C. Meijer). In European trials, the hc2 displays a clinical specificity for women >35 years of age of at least 93.3% (95%CI: 92.9-93.6) (6). Based on the compatibility between both tests, hc2 is assumed to fit the criterion of clinical specificity in women >40 years of age. The reason for selecting this age category is that in women >40 years of age HPV prevalence determined by hc2 and GP5+/6+-PCR is constant and comparable between the different European countries (6).This clinical specificity percentage is required in order to assure a most adequate positive predictive value of the HPV detection assays that result in a cost-effective cervical screening.

3. The candidate test should display an intra- and inter-laboratory reproducibility of at least 90% ± 3%. The hc2 and GP5+/6+-PCR had an intra-laboratory reproducibility of 92.2% (95%CI: 90.4-93.7) and 92% (95%CI: 84.5-99.5), respectively (7;8). This is to ensure a robust and highly reliable performance of the test in clinical practice.

Since these test requirements were deduced from data obtained with hrHPV hc2 and GP5+/6+-PCR, it is obvious that these assays meet these requirements. Hc2 and GP5+/6+-PCR detect 13 and 14 hrHPV types, respectively (i.e. both assays detect HPV16, -18, -31, -33, -35, -39, -45, -51, -52, -56, -58, -59, and –68; GP5+/6+-PCR additionally detects HPV66). It is complicated to conclude the minimal number of hrHPV types that a HPV detection assay should be able to detect as there is still debate about the carcinogenicity of certain HPV types that have been rarely detected in carcinomas (9-11). Nevertheless, as long as the HPV detection assay complies with the abovementioned criteria, it is of minor importance to what extent uncommon hrHPV types are actually targeted.

Validation guidelines for candidate HPV assays

To realize the abovementioned test requirements for clinical application in cervical screening, a candidate HPV detection assay has to be validated on samples that originate from a population-based screening cohort. In order to ensure a good representation of this population the candidate test has to be compared with a clinically validated test. The following validation strategy is advised:

1. The clinical sensitivity should be determined by testing at least 200 smears of randomly selected women from the population-based screening cohort with cytologically >BMD having histologically confirmed ≥CIN3. Based on the requirements mentioned before the HPV negativity rate in these >BMD smears should not exceed 8%.

2. The clinical specificity should be determined by testing at least 1,500 smears of women >40 years of age with normal cytology that

are randomly selected from the subset of women from the population-based cohort without histologically confirmed ≥CIN2. The HPV positivity rate in these smears should not exceed 7% in a European population.

3. The intra- and inter-laboratory reproducibility should be determined by evaluation of at least 500 smears, 30% of which tested positive in a reference laboratory using a clinically validated assay. Comparative analysis with a clinically validated test (i.e. hc2 and GP5+/6+-PCR) performed on the same smears should result in a percentage of agreement of at least 92% (kappa value of at least 0.5).

Laboratory guidelines for HPV testing

Laboratories performing the HPV test should comply to quality assurance (QA), including internal quality control (IQC), external quality assessment (EQA) and quality improvement (QI). To realize QA, at least the following items should be fulfilled:

1. The laboratory should have an infrastructure for molecular testing, including separate laboratories for sample identification/ preparation, DNA extraction, DNA amplification, if applicable, and detection.

2. The laboratory should have accreditation for clinical molecular testing and should comply with standard operation procedures (SOP) and good laboratory practice (GLP) guidelines.

3. The laboratory should process approximately 15,000 smears per year to assure consistency, reliability, and accuracy of results reported.

4. The HPV test performance of the laboratory should be monitored by proficiency testing

including regular intra- and inter-laboratory evaluation of smear specimens and/or cell line dilution series sent by a (inter)national reference laboratory. An example of such an independent organization in the Netherlands is the 'Quality-control Clinical Pathology Association' (i.e. Stichting Kwaliteitstoetsing Klinische Pathologie, SKKP).

In conclusion, the data collected in this thesis indicate that within a cervical screening setting hrHPV tests should exhibit specific requirements to assure high clinical sensitivity and at the same time high clinical specificity to limit unnecessary follow-up. The studies described in this thesis resulted in more insight into the clinical test characteristics of various hrHPV DNA detection assays. The assays mainly differed in clinical specificity, which for the largest part can be attributed to differences in the detection rate of transient HPV infections characterized by low viral loads. Such infections do not cause malignancies, are therefore clinically irrelevant, and potentially harmful in a screening setting. In order to avoid such drawbacks specific HPV test guidelines were proposed. It can be expected that future implementation of hrHPV testing is accelerated once (inter)national consensus is obtained about such guidelines.

Reference List

1. Bulkmans N, Berkhof J, Rozendaal L, van Kemenade F, Boeke A, Bulk S et al. Human papillomavirus DNA testing for the detection of cervical intraepithelial neoplasia grade 3 and cancer: 5-year follow-up of a randomised controlled implementation trial. Lancet 2007.

2. Mayrand MH, Duarte-Franco E, Rodrigues I, Walter SD, Hanley J, Ferenczy A et al. Human papillomavirus DNA versus Papanicolaou screening tests for cervical cancer. N.Engl. J.Med. 2007;357(16):1579-88.

3. Naucler P, Ryd W, Tornberg S, Strand A, Wadell G, Elfgren K et al. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. N.Engl.J.Med. 2007;357(16):1589-97.

4. Arbyn M, Sasieni P, Meijer CJ, Clavel C, Koliopoulos G, Dillner J. Chapter 9: Clinical applications of HPV testing: A summary of meta-analyses. Vaccine 2006;24 Suppl 3:S78-S89.

5. Bulk S, Bulkmans NW, Berkhof J, Rozendaal L, AJ PB, Verheijen RH et al. Risk of high-grade cervical intra-epithelial neoplasia based on cytology and high-risk HPV testing at baseline and at 6-months. Int.J.Cancer 2007;121(2):361-7.

6. Cuzick J, Clavel C, Petry KU, Meijer CJ, Hoyer H, Ratnam S et al. Overview of the European and North American studies on HPV testing in primary cervical cancer screening. Int.J.Cancer 2006;119(5):1095-101.

7. Castle PE, Wheeler CM, Solomon D, Schiffman M, Peyton CL. Interlaboratory reliability of Hybrid Capture 2. Am.J.Clin.Pathol. 2004;122(2):238-45.

8. Jacobs MV, Snijders PJ, Voorhorst FJ, Dillner J, Forslund O, Johansson B et al. Reliable high risk HPV DNA testing by polymerase chain reaction: an intermethod and intramethod comparison. J.Clin.Pathol. 1999;52(7):498-503.

9. Munoz N, Bosch FX, de Sanjose S, Herrero R, Castellsague X, Shah KV et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer. N.Engl.J.Med. 2003;348(6):518-27.

10. Cogliano V, Baan R, Straif K, Grosse Y, Secretan B, El Ghissassi F. Carcinogenicity of human papillomaviruses. Lancet Oncol. 2005;6(4):204.

11. Munoz N, Castellsague X, de Gonzalez AB, Gissmann L. Chapter 1: HPV in the etiology of human cancer. Vaccine 2006;24S3:S1-S10.