Reviewer's report

Title: Genotype distribution of human papillomavirus (HPV) in cervical cytologic specimens in a region of southeast Spain.

Version: 1 Date: 3 March 2009

Reviewer: Jean-Luc Pretet

Reviewer's report:

In their study, Conesa-Zamora et al. collected 432 cervical samples with different cytological diagnoses. They were obtained from 358 women living in the Murcia region of Spain. The authors analyzed the distribution of HPV genotypes using two different kits based on DNA chips. Overall, they identified HPV16 as being the most prevalent type (22.5%) and HPV58 as being the second most prevalent type (6.4%) without taking into account the lesion type. Then they indicate that the prevalence of HPV18 is rather low and propose that the region of Murcia may be an interesting place to assess the effect of cross protection against HPV58 induced by HPV vaccination.

However, some methodological points limit the robustness of the data. Moreover the presentation of the result is sometimes unclear. For these reasons the reviewer cannot accept the publication of this manuscript in its present form.

Major Compulsory Revisions:

1- The description of HPV prevalence does not take into account the cytological diagnosis. But it is known that HPV distribution varies according to the lesion type. HPV16 surely represents the most frequently detected type whatever the cytological diagnosis is (1st position in ASC-US, N/B, LSIL, HSIL, cf. Figure 1). In contrast, the prevalence of HPV58 varies from 17% (ASC-US, 3rd position) to 2% in LSIL (8th position). While HPV58 is the second genotype in term of frequency in this series of samples, it is hazardous to conclude that this genotype is the second most prevalent type in the region of Murcia.

Is the sample size large enough to detect rare genotypes (less than 5% for example) with a good precision (2 or 3 % for example)? This is important to be able to draw sound conclusions regarding type distribution.

2- The authors have used two different kits to perform their study. They indicated that both tests are DNA chip based test. The reviewer would like to point out that sensitivity of each assay may be different for the detection of specific type of HPV. For example it is documented that Linear array from Roche Diagnostics and InnoLipa from Innogenetics (both based on reverse blot hybridization) are not equivalent in term of sensitivity to identify some HPV types. Thus the authors should not mix results obtained with 2 different kits. Either they do not include results obtained with the HPV GenoArray or present separate results.

3- The prevalence of HPV infection according to cytology diagnosis is sometimes
questionable regarding the literature published so far. For example, HPV positivity > 50% in N/B samples appears really high. The authors nevertheless indicate in their discussion that this rate of infection is within the range of HPV positivity described in two other studies performed in Spain (ref 7 and ref 26). This argument is not acceptable because

(i) Spain is probably one the country with the lowest incidence of HPV infection around the world

(ii) the comparison with the two Spanish studies is not reliable. Indeed, in reference 7, Gonzales-Bosquet et al. analyzed only abnormal smears (almost 80% of HSIL and LSIL, only 23% in the submitted manuscript) but no normal smears. Interestingly, Gonzales-Bosquet et al. report HPV prevalence consistent with published meta-analyses. In reference 26, de Antonio et al. do not refer to any cytological diagnosis, rendering the comparison with the Conesa-Zamora’s study invalid.

In contrast, HPV16 is detected in only 30% of HSIL which appears low. In figure 1, HPV type distribution should be presented in order of decreasing frequency for clarity. The relevance of type distribution in HSIL is questionable due to the low number of samples analyzed.

- 116 specimens correspond to samples collected during the follow-up of 42 women. This strategy of sample collection have surely biased the distribution of some HPV types as the probability to find the same HPV from 2 or 3 consecutive samples is likely to be higher than by chance.
- The result presentation lacks clarity:

For example, the total number of HPV positive samples (n=251) appears in table 2. This precludes a rapid understanding of table 1 in which the authors indicates that the percentage was calculated from 251 HPV positive samples.

**Level of interest:** An article of limited interest

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I have received fees from sanofi Pasteur MSD