# The EC project "Setting standards for process indicators in cervical cancer screening": preliminary results

**G.Ronco*, M.van Ballegoijen+, M.Zappa°, E.Lynge§, N.Segnan* and the GISCi$ working group on organisation and evaluation**

* Centre for Cancer Epidemiology and Prevention (CPO), Torino, Italy
+ Erasmus University, Rotterdam, the Netherlands
° Centre for the Study and Prevention of Cancer (CSPO), Firenze, Italy
§ Copenhagen University, Denmark
$ Italian Group for Cervical Cancer Screening

*This study was partially supported by EC contract SC 98 200263 05F02*

## Introduction

Screening with Pap smear remains the best currently evaluated method of reducing the incidence and mortality from invasive cervical cancer. Organised programmes have shown to be the most effective, in particular because of a more rational distribution of smears in the population.
A high quality is needed in order to maximise the impact of screening at a population level, in terms of reducing incidence of and mortality from cervical cancer. For this purpose monitoring the screening process by measuring process indicators that reflect, at a population level, those aspects that more directly determine effectiveness (and human and economic cost) is essential.
Setting reference values for many indicators in cervical cancer screening is not obvious on the basis of theoretical considerations only. Studying their empirical distribution and sources of variability is a needed starting point.
Purpose of the study was investigating the range of variability of a number of possible process indicators and the sources of such variability, in order to evaluate their practical applicability and to set reference values. For this purpose data were collected from Italian and Dutch organised screening programmes and from the cohort of the Maribø County, Denmark.
Preliminary data from Italy and the Netherlands are presented.

## Cervical cancer screening in Italy and in the Netherlands - General features

So far in Italy no nation-wide programme has been implemented, but in the past years several organised programmes have been set-up on a local basis and, more recently, on a regional level. A survey conducted in 1997 identified 33 organised programmes that covered 13.5% of the Italian population 25 to 64 years old. Most of these programmes followed Italian and European recommendations, the age range is 25 to 64 years, screening interval is three years. Many organised programmes are in a phase of implementation, as a result of the guidelines delivered from the National Oncologic Commission in 1996. Within 1999, 44.3% of the female Italian population 25-64 years old is expected to be included in the target population of organised programmes.
All programmes refer women with LSIL cytology for colposcopy and most refer women with ASCUS cytology.

In the Netherlands screening started in the seventies. Around 1980 an almost nationwide coverage was reached. A newly organised programme was nationwide implemented in 1996. The age range extended from 35-53 to 30-60 years and the interval was from 3 to 5 years.
Only women with high-grade lesions are immediately referred for colposcopy. These range between 0.32% and 0.76% (weighted mean 0.56%). Women with ASCUS or LSIL are referred for colposcopy only after two cytologies, if they are not both negative.

## Materials and methods

A standardised questionnaire was prepared to collect aggregated data in a standard format on:

- Target population (by 5-year age class)
- Women invited in 1997 and among them those who performed the test within April 1998 (by 5-year age class)
- Achieved coverage (i.e. % of the target population having a smear within the last 3 (or 5) years, either in the organised programme or spontaneously
- Population screened (in some cases only women screened after invitation could be included, in others all screened women; by 5-year age class)
- Distribution of their smear results (including unsatisfactory)
- Women in the screened population referred to colposcopy and among them those who actually underwent it (by 5-year age class and index cytology)
- Their hystological diagnoses (by 5-year age class) and correlation with previous cytology.

In addition variables possibly useful in order to interpret differences were collected.
In Italy each programme, on the basis of his local information system, filled the questionnaire. Not all centres were able to provide the requested information, due mainly to insufficiencies in the computerisation of data and in some cases to lack of referral centres for colposcopy.
In the Netherlands data were mainly obtained by the Dutch Network and National Database for Pathology (PALGA), a database including all pathology results at a national level. Colposcopies are not registered.
The Positive Predictive Value (PPV) was computed on three levels of cytological result (HSIL, LSIL, and ASCUS) and in respect of two different histological results (CIN2+ AND CIN1+). In Italy PPV was calculated taking into account only women who actually underwent colposcopy while in the Netherlands the denominator included all women with the relevant cytological result. We modelled it by a logistic model, including the logit of detection rate as an offset, in order to control for different disease prevalence.
Determinants of the Detection Rate were studied by Poisson regression.

## Results

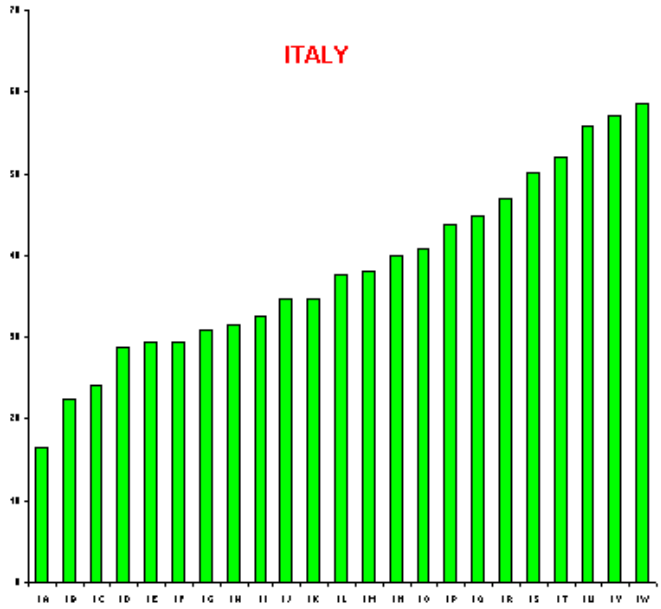
In Italy questionnaires were mailed to all 29 programmes identified as active in 1997 according to the first systematic survey of organised cervical cancer screening programmes in Italy. Among them 24 (73%) filled the questionnaires of the present survey. They were mainly located in Northern and Central Italy. Overall in 1997 these programmes covered 1,870,000 women, with a wide range of target populations (5,200 to 278,000; median 56,600). Most programmes started very recently (13 after October 1995). The overall screened population was 351,567 women, with a wide variability among centres (range 2,000 to 40,000).
In the Netherlands data were obtained for 13 different areas, covering the entire country. Overall the target population was 957,000 women. The overall screened population was 401,754 women with a

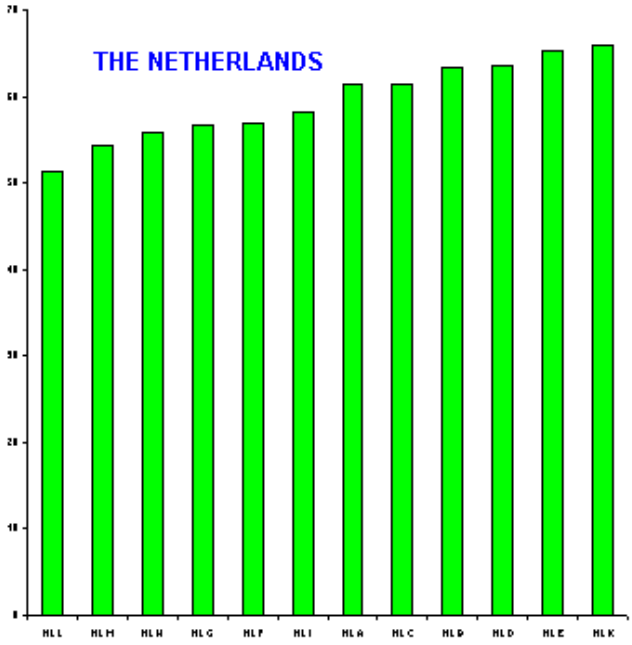
range from 26,000 to 63,000.

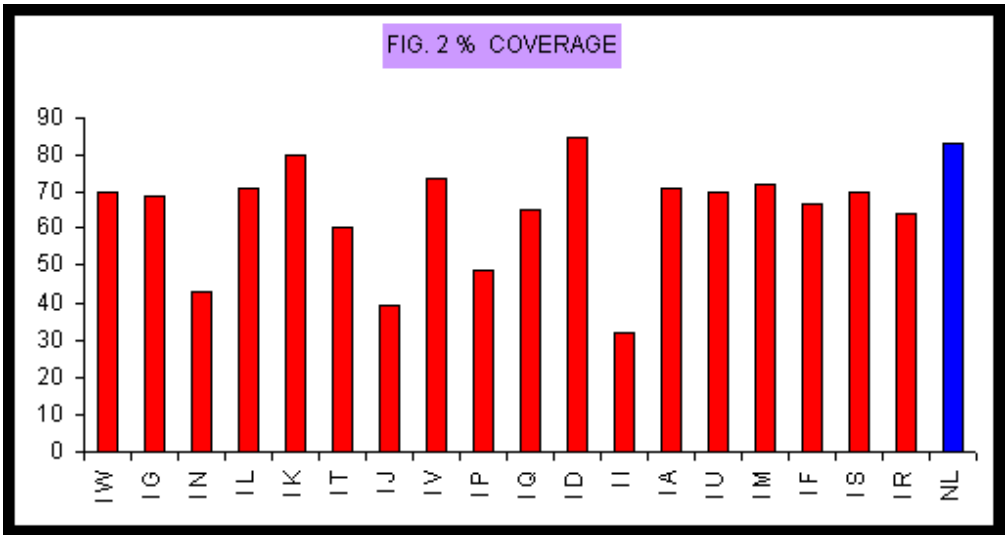
## Compliance to invitation and coverage

In Italy overall 636,638 women were invited and among them 229,928 (36.12%) accepted. There was a large variability in compliance to invitation, ranging from 17% to 58% (Fig. 1). It depended largely on the type of organisation. The lowest value (17%) was found in a programme that, each year invites all women resulting not to have performed a test in the last 3 years. In programmes inviting only women spontaneously uncovered in the last 3 years the (weighted) average compliance was 26,7 vs. 39.9 in programmes inviting all women in the target population ($p<0.001$). A higher compliance was also found in programmes inviting at a pre-fixed appointment vs. those with a free one (43.4% vs. 23.8%, $p<0.01$). However there is a high correlation between type of women invited and type of invitation.
In the Netherlands 936,956 women were invited and 540,714 (57.17%) complied. Compliance ranged from 51% to 66%.
Only part of Italian programmes (19 out of 24) provided data on coverage, failure being mainly caused by lack of information on spontaneous activity. Data ranged from 32% to 85% (Fig. 2). However in centres with the lowest values estimates are based on largely incomplete information on spontaneous activity, leading to severe underestimation.
Most programmes provided a similar estimate of about 70% and we are confident that this can be considered as a true estimate of the actual situation.
In the Netherlands coverage estimate are not available by area. A national 5-year coverage estimate is 83%.

**FIG. 1 % COMPLIANCE TO INVITATION**





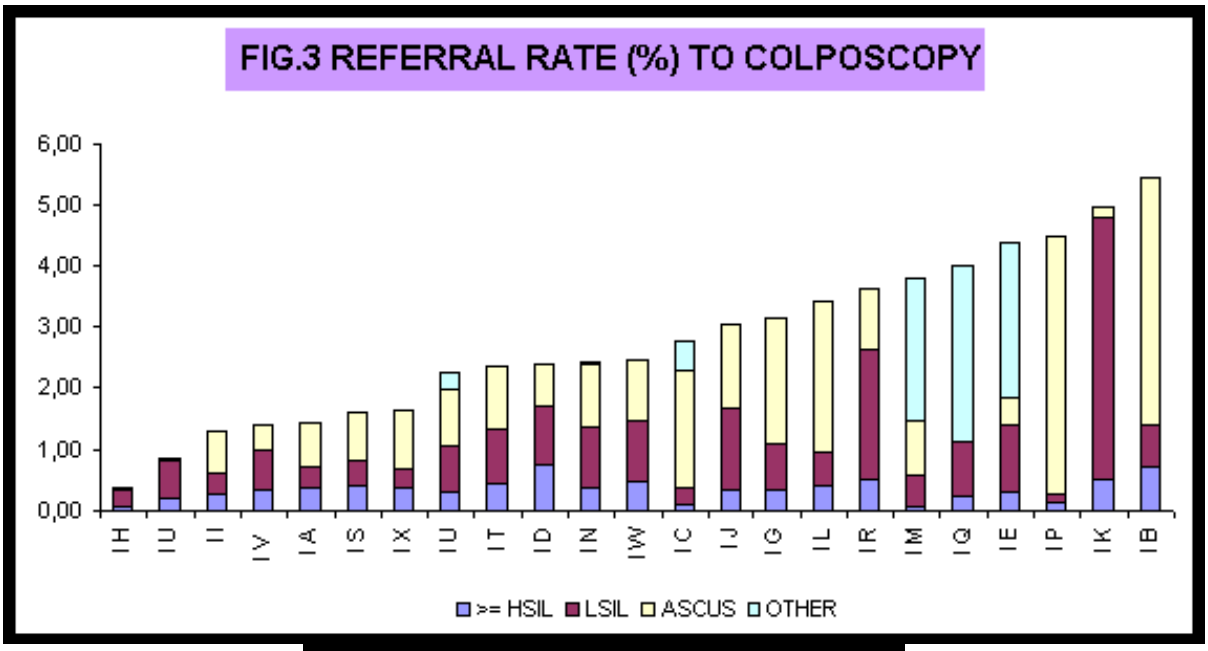



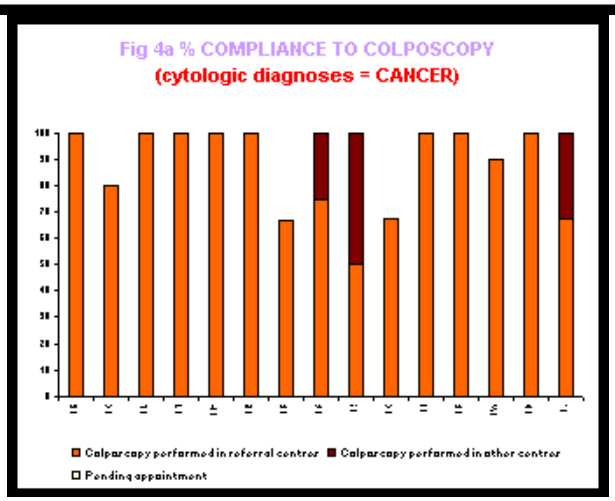

Coverage is within 3 years for Italian programmes and within 5 years for The Netherlands.

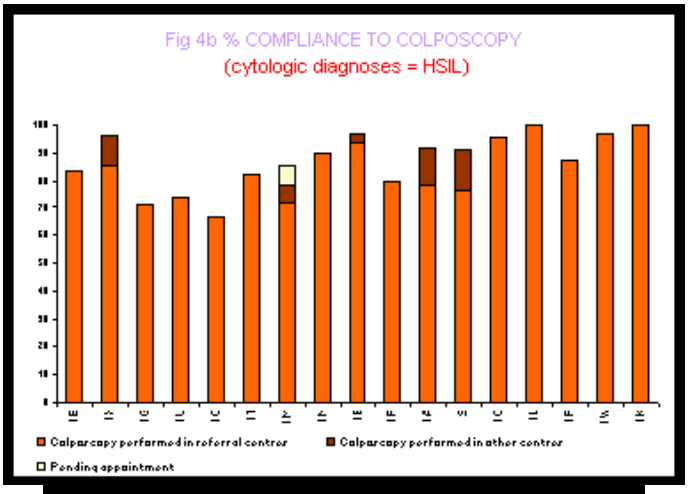
## Referral rate and compliance to colposcopy

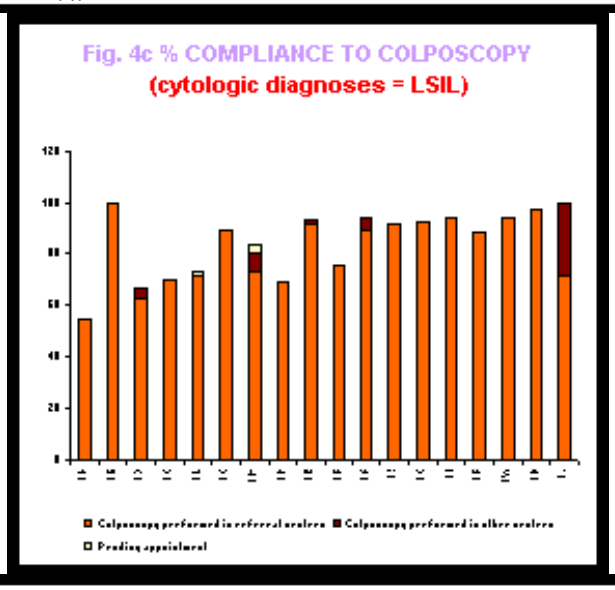
Figure 3 reports the percentage of women referred for colposcopy by each Italian centre, according to the cytological report. Variability was large, ranging from 0.4% to 5.4%. These figures did not change after standardisation for 5-year age class (range: 0.4%-5.1%). In most programmes (14 out of 24) referral rate was below 3%.
Differences partly reflect variation in the Positive Predictive Value of cytology but also the variability in the detection rate of histologically confirmed lesions (see below).
They were also partly due to differences in protocol: indeed the highest values were observed in centres where women with ASCUS cytology are immediately referred for colposcopy. In one programme, at that time, women were invited to colposcopy at the second consecutive smear to be repeated.

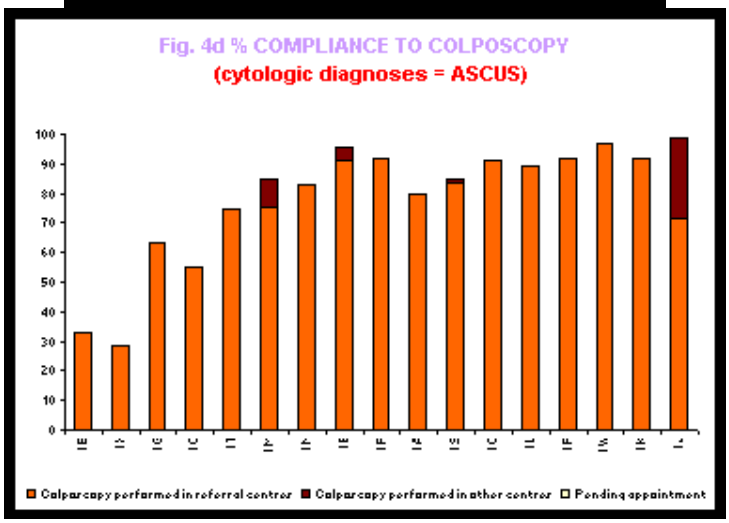
In the Netherlands data on recommended action are not available.
In Italy compliance to colposcopy ranged from 67% to 100% when the index cytology was invasive cancer or HSIL, from 55% to 100% when it was LSIL and from 28% to 71% when the index cytology was ASCUS. Overall 86.3% of women with LSIL or more severe cytology attended (82.5% considering only women having colposcopy in referral centres). This is less would be advisable. However it has to be taken into account that in many programmes data about colposcopies performed outside referral centres are lacking or incomplete and that in some cases these are plausibly a substantial proportion.
The Dutch system does not register colposcopy data. Therefore compliance to colposcopy is not available.






## Positive Predictive Value (PPV)

Figure 5 shows, for each programme, the Positive Predictive Value of having a histologically confirmed CIN2+ or CIN1+ by smear result.
In most programmes the PPV of cytology of carcinoma cells for CIN2+ was near 100%; anyway such cytological reports are quite rare. As far as HSIL smear is concerned, in Italy the PPV for CIN2+ ranged from 34% to 100%, but in most centres it was higher than 60%. In the Netherlands it

was below 60% in one programme only. In the case of LSIL the PPV showed a high variability in Italy, while it was more stable in the Netherlands, with the exception of area A. For ASCUS smears the PPV for CIN1+ ranged from 2% to 80% in Italy (being very low in some programmes) and was very low in the Netherlands. In some Italian programmes ASCUS cases are referred for colposcopy only after smear repeat.
Variability in PPV partly depends on different criteria of interpretation of cytology but also, as well known, on prevalence of disease. In addition random variation can be remarkable in some case, given small numbers. In order to control for these factors we computed the Likelihood Ratios of a positive test (LR+ = post-test odd of disease/ pre-test odd). LR+ does not depend on disease prevalence. As an example, Table 1 reports, for each programme, the PPV of a cytology ASCUS or more severe in predicting histologically confirmed CIN2+, its expected value in the hypothesis of constant LR+, and the corresponding LR+. The, weighted, average LR+ value was 51 in Italy (i.e. the frequency of histologically proved CIN2+ is roughly 50 times greater in women with ASCUS+ cytology than in the average of screened women) and 25 in the Netherlands (RR=0.50 p<0.0001). Some programmes have high LR+ values despite a low PPV, showing that the latter was the result of low disease prevalence. However, variability remains high: in 5 Italian and 2 Dutch programmes the observed LR+ was significantly lower than that predicted. This confirms that remarkable differences in interpretation of cytology exist.
The LR+ for cytology LSIL+ in predicting CIN2+ was 88 in Italy and 132 in the Netherlands (RR=1.49 p<0.0001).

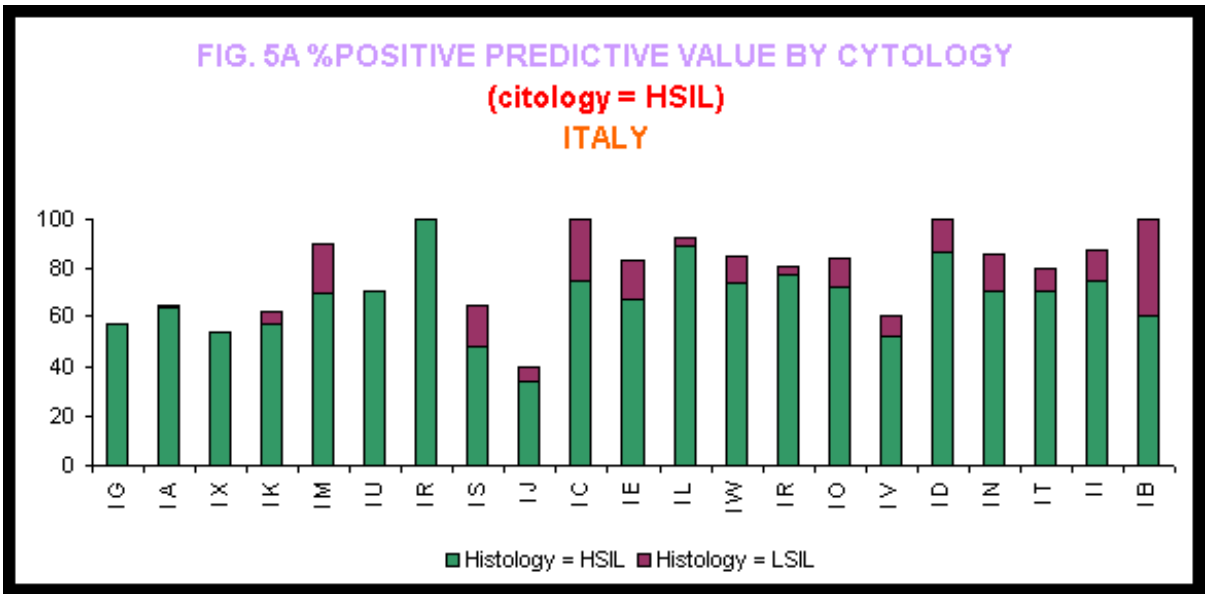


FIG. 5B % POSITIVE PREDICTIVE VALUE BY CYTOLOGY

(citology = LSIL)

ITALY

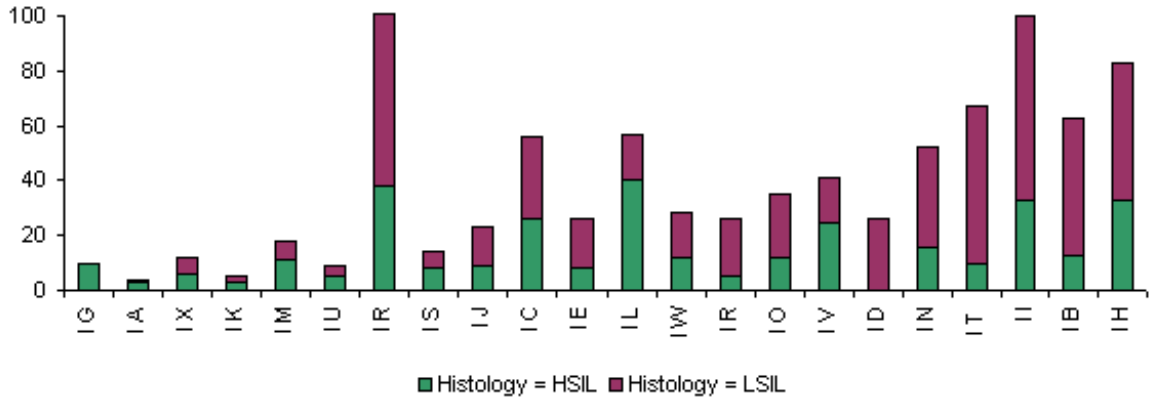



FIG. 5C % POSITIVE PREDICTIVE VALUE BY CYTOLOGY

(citology = ASCUS)

ITALY

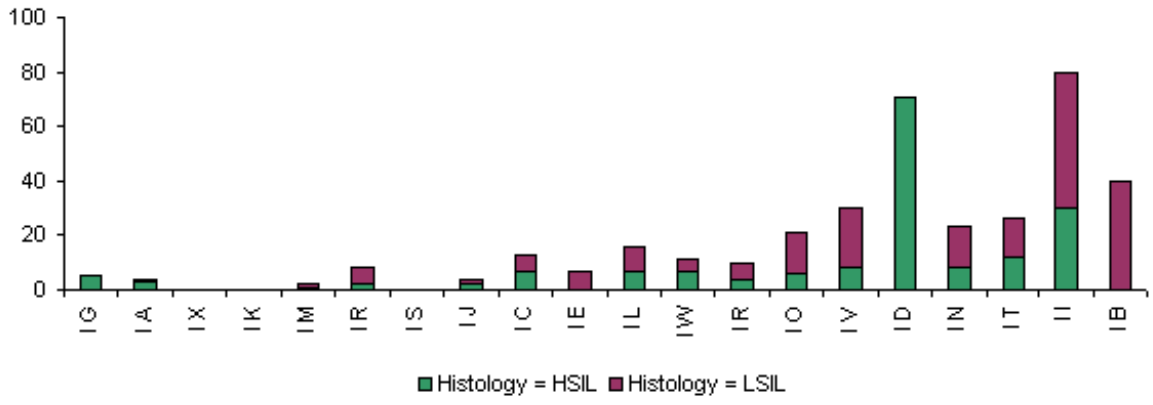


FIG. 5A % POSITIVE PREDICTIVE VALUE BY CYTOLOGY

(citology = HSIL)

THE NETHERLANDS

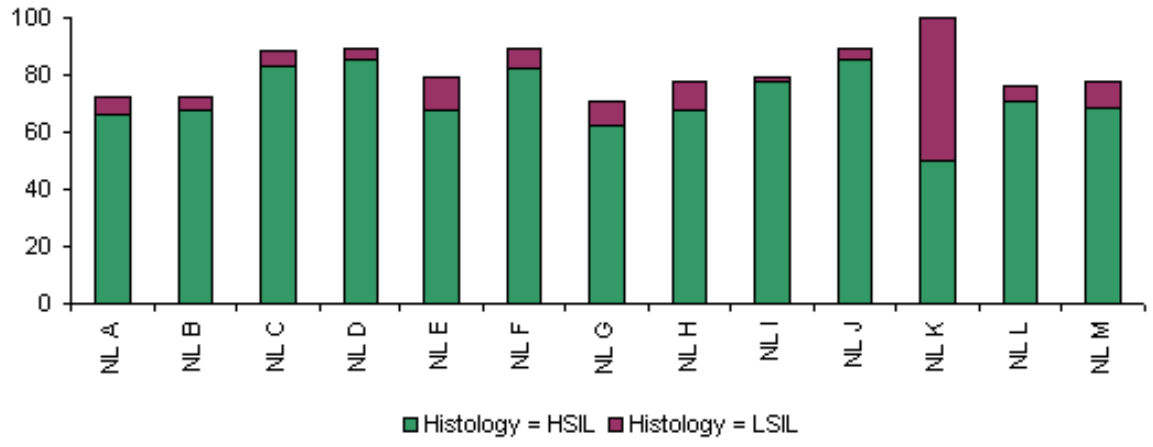
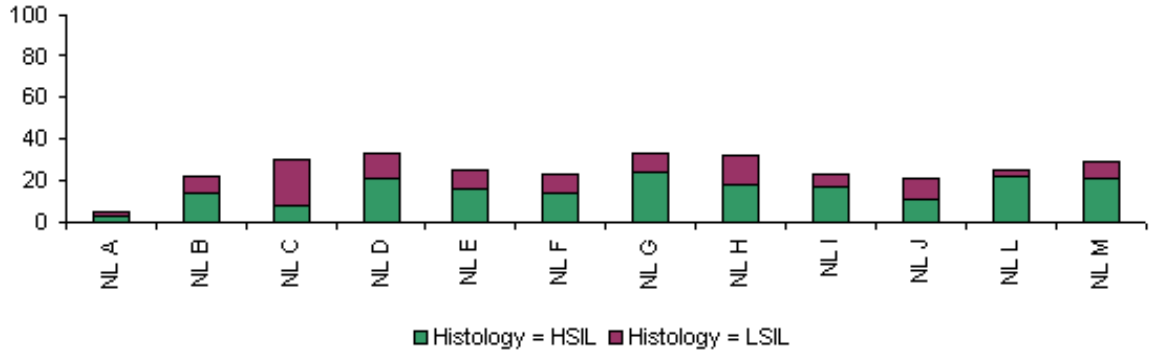



FIG. 5B % POSITIVE PREDICTIVE VALUE BY CYTOLOGY

(citology = LSIL)

THE NETHERLANDS



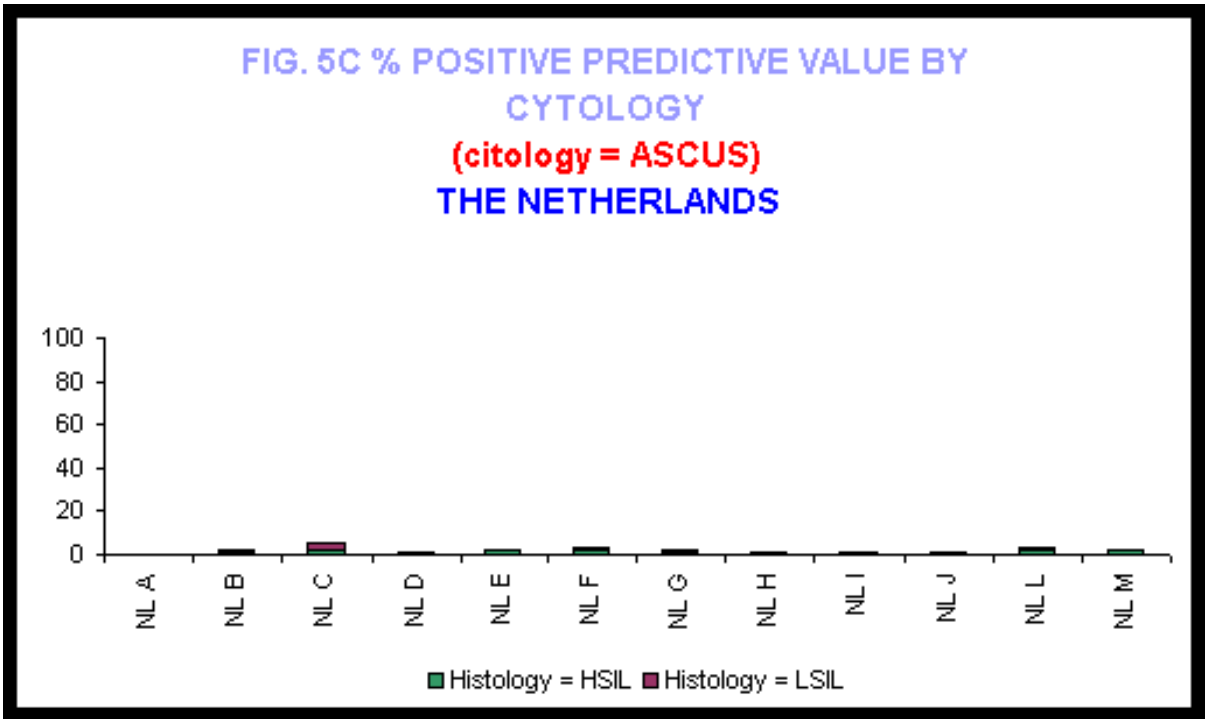



**TABLE 1**
**PPV AND LR+ OF ALL CASES ASCUS OR MORE SEVERE IN PREDICTING A HISTOLOGY OF CIN2 OR MORE SEVERE**

| PROGRAMME | % OBSERVED PPV | OBSERVED/ PREDICTED* PPV (1) | LR+ |
|---|---|---|---|
| NL A | 3.9 | 0.35 ** | 8.00 |
| I K | 8.8 | 0.48** | 21.99 |
| I B | 20.0 | 0.56* | 22.67 |
| I R | 14.74 | 0.69** | 32.12 |
| NL H | 8.68 | 0.70** | 16.88 |
| I G | 12.48 | 0.75** | 36.09 |
| I L | 23.61 | 0.81* | 38.00 |
| I J | 9.09 | 0.99 | 50.03 |
| I E | 16.48 | 1.00 | 50.30 |
| I W | 21.53 | 1.01 | 51.33 |
| NL F | 10.25 | 1.04 | 26.12 |
| I N | 22.79 | 1.05 | 53.74 |

| | | | |
|---|---|---|---|
| I T | 23.14 | 1.05 | 53.90 |
| I O | 19.07 | 1.19 | 62.19 |
| NL J | 17.00 | 1.21 | 31.39 |
| NL I | 14.15 | 1.23 | 31.78 |
| I C | 13.51 | 1.24 | 64.50 |
| NL M | 19.20 | 1.34 | 35.51 |
| NL B | 11.07 | 1.37 | 35.42 |
| I S | 16.54 | 1.39 | 74.16 |
| I X | 22.58 | 1.40 | 76.50 |
| I M | 8.02 | 1.47 | 76.36 |
| NL L | 13.14 | 1.51 | 39.85 |
| NL D | 25.43 | 1.54 | 43.16 |
| I A | 19.00 | 1.64 | 90.29 |
| NL G | 14.76 | 1.66 | 44.40 |
| NL C | 25.60 | 1.66 | 47.48 |
| NL K | 20.00 | 1.89 | 53.00 |
| NL E | 14.51 | 1.95 | 53.06 |
| I P | 19.40 | 2.00 | 113.14 |
| I U | 21.92 | 2.61 | 155.08 |

1. Predicted values were computed by a logistic model having the logit of PPV for CIN2+ as dependent variable, just an intercept as explanatory variable and the logit of the detection rate of CIN2+ as an offset variable. Therefore it assumes a constant LR+.
   * standardised deviance residual <= -1.65
   ** standardised deviance residual <= - 2.31.

### Detection rate (DR)

The weighted average row DR of histologically confirmed CIN2+ lesions was 3.10 per thousand screened women in Italian programmes and 5.41 in the Netherlands. When considering women aged 30 to 64, in a Poisson model the DR (adjusted for age and recent start) was significantly higher in the Netherlands than in Italy (rate ratio 1.85 $p<0.0001$). In Italy programmes that started from less than 3 years (therefore all screened women are at the first round within the organised programme) had a significantly higher detection rate (rate ratio 1.82 $p<0.0001$). This could be expected "a priori" but these programmes largely correspond to a specific region (Emilia Romagna). Therefore some confounding with local factors is possible.

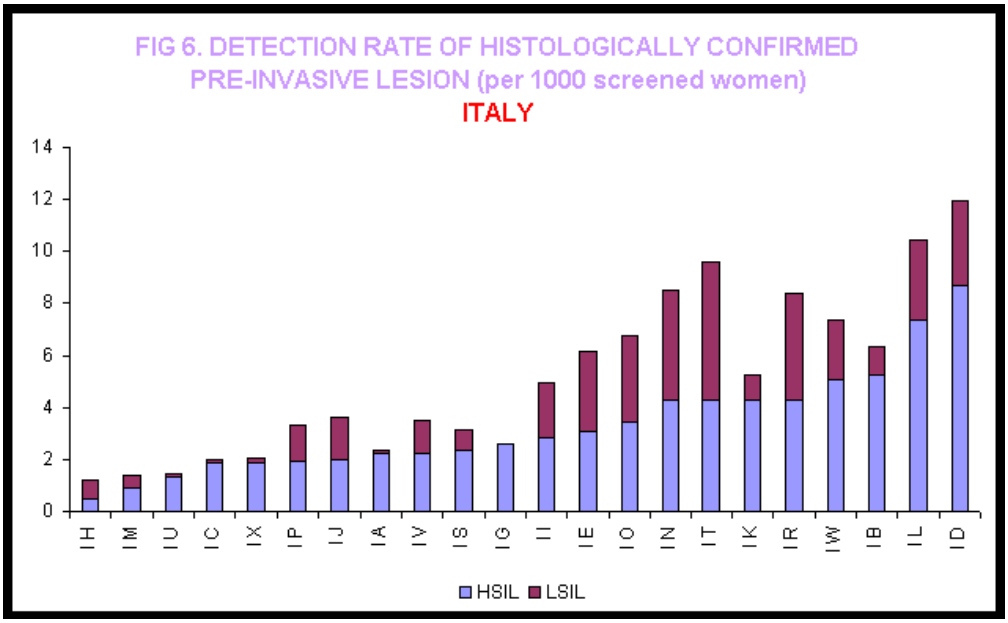
Table 2 reports the observed number of histologically confirmed CIN2 or more severe lesions and that predicted by a model assuming a common DR for all programmes in the Netherlands and, in Italy, one for programmes started from at least 3 years and another one for those started more recently.
Part of variability can be explained by random fluctuations of rates due to the small number of screened women in some centres. Nevertheless out of 28 programmes with valid data 10 had a value significantly different from that predicted.
Differences seem to be too large to reflect real differences between the considered geographic areas within each country. In principle differences in detection rate could reflect differences in sensitivity but the observed variation is too large to admit this as the only explanation. In addition a very low incidence of interval cancers was observed in some of the programmes showing the lowest detection rates, like in programme I M.

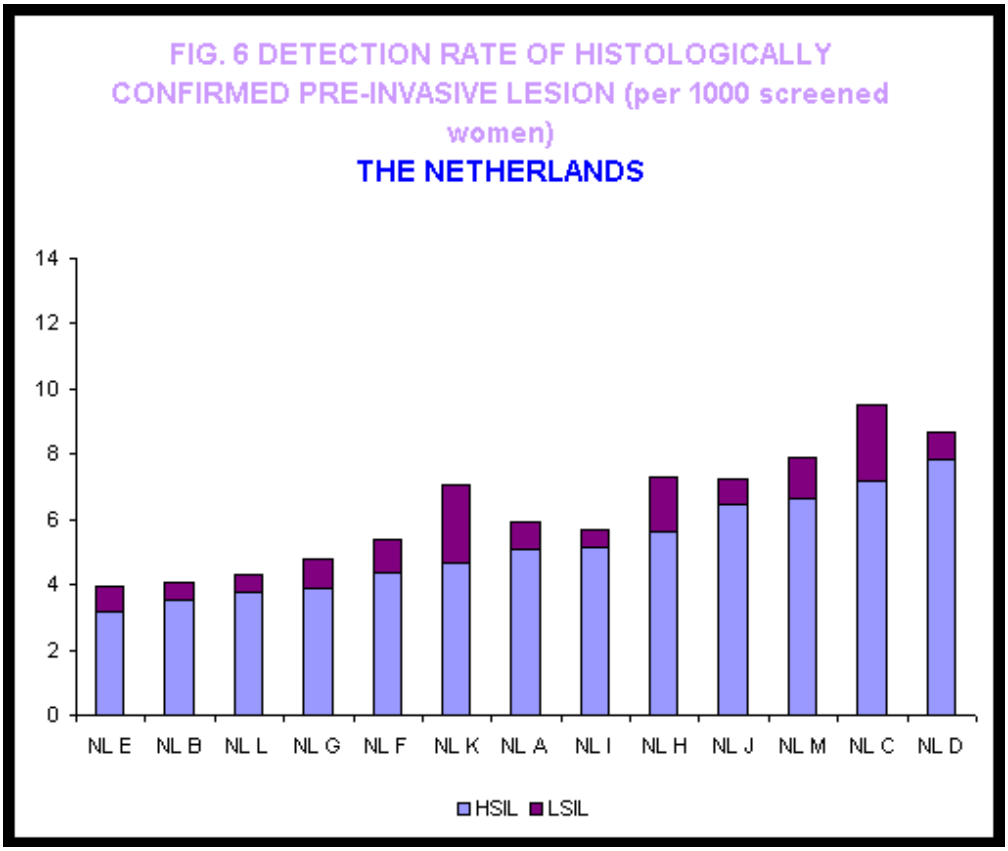
**TABLE 2**
**OBSERVED AND EXPECTED CIN2+ CASES**

| PROGRAMME | OBSERVED CIN2+ | PREDICTED CIN2+ (1) | OBSERVED/ PREDICTED |
|---|---|---|---|
| I H | 2 | 6.58 | 0.30 |
| I M | 19 | 33.21 | 0.57* |
| NL E | 49 | 81.54 | 0.60** |
| I U | 13 | 19.12 | 0.68 |
| NL G | 87 | 122.97 | 0.71** |
| NL L | 193 | 268.94 | 0.73** |
| NL B | 62 | 82.84 | 0.75* |
| NL K | 2 | 2.59 | 0.77 |
| I E | 26 | 33.07 | 0.79 |
| I O | 73 | 91.14 | 0.80 |
| NL F | 55 | 62.92 | 0.87 |
| I K | 30 | 33.58 | 0.89 |
| I G | 69 | 76.13 | 0.91 |
| I S | 21 | 23.04 | 0.91 |
| NL A | 140 | 147.62 | 0.95 |
| NL I | 187 | 195.84 | 0.95 |
| I A | 84 | 81.49 | 1.03 |
| NL H | 406 | 392.37 | 1.03 |

| | | | |
|---|---|---|---|
| I N | 117 | 107.65 | 1.09 |
| I T | 102 | 90.63 | 1.13 |
| NL D | 140 | 122.98 | 1.14 |
| I B | 3 | 2.63 | 1.14 |
| NL M | 569 | 479.99 | 1.19** |
| NL J | 110 | 79.95 | 1.38** |
| I X | 12 | 8.69 | 1.38 |
| NL C | 142 | 101.46 | 1.40** |
| I R | 32 | 17.50 | 1.83** |
| I L | 46 | 24.56 | 1.87** |

1. From a Poisson model having the number of screened women as an offset variable and intercept, country and a term for recent (< 3 years before survey) start as explanatory variables.
   * p<0.05
   ** p<0.01
   In I B, I L, I K and I O only uncovered women are invited; only women screened after invitation included in study population. In I M all women are invited; only women screened after invitation included in study population. In all other cases all screened women were included in the study population






In our judgement, in Italy, the largest determinant of such variability is in selection processes for the study populations, resulting in different histories of prior Pap-testing. Such different histories are only partially approximated by a recent start of the organised programme, since spontaneous activity was already present everywhere. Selection plausibly depends on different approaches to integration with spontaneous activity (in some programme only women not spontaneously covered are invited while others invite all women, independently of previous screening history) and in criteria of inclusion in the study population (some programme could include only women screened in the organised programme while others included all). However modelling these factors is difficult given the small number of observations. Unfortunately in Italy a reliable measure of previous screening history of the study population was not available.
In addition selection for risk factors of the study population can be present, especially in programmes in a starting phase. Finally local differences in histology interpretation can play a role.

## Discussion

Variability was actually found to be high for many indicators, especially in Italy. In this country part of variability is likely to be an artefact, due to difficulties of some programmes in collecting data. Furthermore many of the most extreme values were based on a small number of screened women (either because the programme was in a starting phase or because it was very small).
In some programme data showed atypical situations that need local in-depth investigation and correction, if needed.
In general data show relevant differences in cytological reporting criteria. These differences resulted in a variation of predictive values, even within countries adopting different protocols for managing women with the same cytological result.
It is difficult to assess if these differences resulted in different sensitivity.
Unfortunately, in cervical screening evaluation, Detection Rate showed not to be usable alone as an

indicator of sensitivity. It should be related to incidence of invasive cancer in the absence of screening, but this is not possible, for cervical cancer, given that at least "spontaneous" activity is widespread from many years. Data on interval cancers are being collected in the Netherlands and in some Italian programmes. They are also are expected to be largely influenced by "baseline" risk. However we hope that their combination with Detection Rate data allows a more thorough interpretation of sources of variability.

Data show the need that good information systems are implemented in organised screening programmes, in order to produce data in a standardised way.

The value of a number of possible indicators varies according to the local frequency of pre-invasive lesions and can be influenced by organisational features. An effort to identify indicators stable to these factors will be done. This is requested in order to allow homogeneous evaluation and quality assurance methods in EC countries and comparison between such countries.