High sensitivity of five colorectal screening programmes with faecal immunochemical test in the Veneto Region, Italy

Manuel Zorzi,1 Chiara Fedato,1 Grazia Grazzini,2 Fiorella Carmen Stocco,1 Flavio Banovich,3 Antonio Bortoli,4 Luigi Cazzola,5 Adriana Montaguti,6 Tina Moretto,7 Marco Zappa,2 Marcello Vettorazzi1

ABSTRACT
Objective Although guaiac-based faecal occult blood test screening has been shown to be effective in reducing colorectal cancer (CRC) mortality, it has been criticised mostly for its low sensitivity. Italian CRC screening programmes are based on immunochemical tests (iFOBT). We collected and analysed the interval cancers (ICs) found by five screening programmes to estimate their sensitivity.

Methods ICs were identified in subjects who had a negative result in a screening examination from 2002 to 2007 (N=267 789); data were linked with 2002–2008 hospital discharge records. Analysis was based on the follow up of 468 306 person-years. The proportional incidence-based sensitivity was estimated overall and by sex, age class, time since last negative iFOBT result, anatomical site, and history of screening (first or subsequent test).

Results Overall, 126 ICs were identified, compared to 572 expected cancers. The proportional incidences were 15.3% and 31.0% in the first and the second interval-years, respectively, with an overall episode sensitivity of 78.0% (95% CI: 73.8 to 81.6). Sensitivity was higher for males than females (80.1% vs 74.8%); no differences were observed by age, anatomical site or between programmes. The test sensitivity of iFOBT was 82.1% (95% CI 78.1% to 85.3%).

Conclusions iFOBT-based screening programmes showed a high performance in terms of sensitivity as estimated through the IC rates. The screening schedule utilised in our programmes (single iFOBT, positivity threshold of 100 ng Hb/ml of sample solution, inter-screening interval of 2 years) shows low rates of missed cancers that are diagnosed during the interval. HDR are a convenient and reliable source of data for IC studies.

INTRODUCTION
Cancer screening is aimed primarily at reducing deaths from specific cancers. If such reduction exists, it occurs many years after the screening implementation. For this reason, it is highly desirable to have early indirect indicators of screening efficacy for monitoring screening programme quality and its potential impact on cancer mortality.

Sensitivity estimates of screening programmes using interval cancer (cancers that are detected after a negative screening episode and before the next invitation for screening) rates is one of the most important of such indicators. Different definitions of sensitivity in cancer screening have been proposed. Traditionally, sensitivity is specified as to the screening test and it measures the ability of the test itself to identify disease in the detectable preclinical phase in those screened. According to Hakama, we also estimated the episode sensitivity, which measures how much of the disease the screening test and diagnostic confirmation combined are able to identify in those screened, since some of the interval cancers are identified in persons who originally tested positive but were confirmed negative.1

The estimate of screening sensitivity requires the follow-up of individuals who have had a negative screening test (with or without further assessment) in order to identify interval cancers (ICs) before
a subsequent invitation or within a screening interval period. ICs include a spectrum of tumours, from those that either did not exist or were undetectable at the previous screening round to those that were detectable but missed.

Colorectal cancer (CRC) screening based on the guaiac faecal occult blood test (gFOBT) has been shown to be effective in reducing CRC mortality in four randomised studies. Nevertheless, the sensitivity of CRC screening programmes has not been estimated routinely within a public health policy. In fact, few studies reported sensitivity estimates evaluating interval cancers exclusively for gFOBT screening programmes. Many studies have already evaluated immunochemical FOB (iFOBT) sensitivity using colonoscopy as the ‘gold standard’. Such a method is biased by lead time and over-diagnosis. Only two studies investigated iFOBT sensitivity by identifying ICs with cancer registries: Zappa obtained a sensitivity of 82% (95% CI 67% to 92%) while Castiglione reported a sensitivity of 74% (95% CI 57% to 85%).

The aim of the present study was to estimate the test and episode sensitivity of some CRC screening programmes taking place in the Veneto Region, through the identification of ICs obtained by the linkage of screening archives and hospital discharge records (HDRs). HDRs are produced essentially for administrative purposes; however, a great deal of information such as full individual identification and clinical data can be obtained from them. HDRs are computerised and available for consultation with a shorter delay compared to that of Cancer Registry data. For these reasons, IC identification through HDRs could be a feasible and easily applied method, reproducible almost everywhere, even in the absence of a cancer registry. Such a method has already been employed successfully in evaluating mammography screening sensitivity.

**POPULATION AND METHODS**

**Setting and population**

This study was carried out in five cancer screening programmes (Local Health Units of Alto Vicentino, Bussolelango, Dolo Mirano, Feltre and Pieve di Soligo) of the Veneto Region, Italy. Local Health Units are the public agencies that organise and administer the health services of groups of municipalities, including screening programmes.

The main features and protocol of Italian bowel screening programmes have been reported in detail. In the Veneto Region, screening programmes for CRC started in 2002 and are still under implementation. The screening protocol is addressed to all residents aged 50–69, who are invited via mail every 6 months. Those who are non-responders to the first invitation or within a screening interval period. ICs, identified by linkage, was checked to confirm the diagnosis by consulting pathology databases or the hospital charts and the date sequence (date of diagnosis neither preceding nor later than 2 years compared to the date of screening). Only invasive cases were included in the study. We defined as ‘cancerised adenoma’ the colorectal adenoma containing carcinoma invading the submucosa but not the muscular layer.

**Identification of interval cancers**

To achieve a complete identification of ICs, the traditional method is usually based on a Cancer Registry database. In Italy, a National Cancer Registry does not exist, whereas in the Veneto Region the Regional Cancer Registry does not cover all the Local Health Units. In addition, a Cancer Registry provides information with a significant delay (3–5 years) and for this reason it was not feasible for the purposes of this study. Consequently, we used regional HDR concerning the period between 2002 up to 31 December 2008 and HDR relative to extra-regional hospitalisations that were available until 31 December 2006.

We included in the study subjects screened from the beginning of programmes up to 31 December 2007. Subjects screened until 2006 were followed up for 2 years after each screening episode, those screened during 2007 only for 1 year and were therefore included only in the first-interval year analysis. Each programme provided a database of screened subjects with full identity data and the date of screening test. These databases were linked with the HDR databases in order to obtain a list of screened people who had been discharged by the hospital with a diagnosis of (or possibly correlated with) colorectal cancer within the 2 years following the screening episode (the codes utilised to highlight a HDR are shown in table 1).

This list included screen-detected cancers and a list of possible IC cases. A large proportion of ICs were already known by the screening programme as they had been diagnosed at the screening centre. Each case in the additional group of possible ICs, identified by linkage, was checked to confirm the diagnosis by consulting pathology databases or the hospital charts and the date sequence (date of diagnosis neither preceding nor later than 2 years compared to the date of screening). Only invasive cases were included in the study. We defined as ‘cancerised adenoma’ the colorectal adenoma containing carcinoma invading the submucosa but not the muscular layer.

<table>
<thead>
<tr>
<th>Table 1 Codes and description used to identify hospital discharge records with a diagnosis of (or possibly correlated with) colorectal cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Code DRG/ICD9</strong></td>
</tr>
<tr>
<td>DRG 146</td>
</tr>
<tr>
<td>DRG 147</td>
</tr>
<tr>
<td>DRG 148</td>
</tr>
<tr>
<td>DRG 149</td>
</tr>
<tr>
<td>DRG 172</td>
</tr>
<tr>
<td>DRG 173</td>
</tr>
<tr>
<td>ICD9 153</td>
</tr>
<tr>
<td>ICD9 154</td>
</tr>
<tr>
<td>ICD9 211</td>
</tr>
<tr>
<td>ICD9 2113</td>
</tr>
<tr>
<td>ICD9 2114</td>
</tr>
<tr>
<td>ICD9 2119</td>
</tr>
<tr>
<td>ICD9 230</td>
</tr>
<tr>
<td>ICD9 2303</td>
</tr>
<tr>
<td>ICD9 2304</td>
</tr>
<tr>
<td>ICD9 2309</td>
</tr>
<tr>
<td>ICD9 235</td>
</tr>
<tr>
<td>ICD9 2352</td>
</tr>
</tbody>
</table>
**Statistical analysis**

Sensitivity estimates were based on the proportional incidence methodology.\(^{10}\) This approach compares interval cancers within a given interval following a negative screening or a positive test followed by a negative assessment, with the expected incident cancers in the absence of screening according to the formula:

\[
\text{Sensitivity} = 1 - \frac{\text{OI}(t)}{\text{EI}}
\]

where OI = observed interval cancers during time t and EI = expected incident cancers.

The expected number of cancers is calculated by applying the underlying incidence (in the absence of screening) to the person-years at risk for the screened subjects.

Incidence rates were obtained from the Veneto Tumour Registry, which covers about 2,300,000 subjects, 48.8\% of the entire regional resident population. We calculated the 5-year age and gender specific incidence rates and applied them to the person-years distribution of screened subjects. Incidence rates in the interval 1999–2002 were used, being the period immediately preceding the start of the screening activity. Person-years at risk were calculated from the date of the first negative screening episode.

The episode sensitivity of screening programmes was estimated using ICs following a negative iFOBT plus those occurring after a second level assessment, either positive or negative, or in subjects who refused any assessment. The inclusion of the latter, that were considered a failure in terms of the programmes, represents a different application of the method suggested by Hakama.\(^1\)

For this estimate person-years at risk were calculated for all screened subjects, both with a negative or a positive iFOBT, except those who received a diagnosis of cancer. Sensitivity was estimated both overall and by gender, age at iFOBT, time since last negative iFOBT (first or second year of the interval), cancer location (rectum and recto-sigmoid junction: ICDO=154; colon: ICDO=158), colon sub-site (proximal - up to and including the transverse, vs distal) and history of screening (subjects at first or subsequent screening test).

The Veneto Tumour Registry did not provide the sub-site specific incidence rates. In order to calculate the colon sub-site specific sensitivity we estimated these rates by applying to the Veneto incidence rates the age and gender specific sub-site distribution of cases reported in an archive of 54,500 CRCs by the Italian Association of Cancer Registries.\(^{19}\) 41\% of that case-series were proximal and 59\% distal.

95\% CIs were calculated on the basis of the normal distribution and the exact Poisson distribution, as appropriate.\(^{20}\)

To calculate the test sensitivity of iFOBT we considered only the ICs occurring after a negative test.

Our sensitivity method may be subject to selection bias, because it compares the incidence of interval CRC in attenders with the underlying incidence expected in the general population, assuming similar CRC incidence in attenders and non-attenders.\(^{21}\) We therefore carried out an unbiased estimate of sensitivity, by including in the formula the incidence rates among the non-attenders.\(^1\) The latter were calculated by record linkage of the archive of non-attenders and the CRCs recorded by the Veneto Tumour Registry.

**RESULTS**

Table 2 shows the principal data about the five programmes that took part in the study. Overall, 267,769 screening episodes took place between 2002 and 2007 (males 47.6\%, 173,859 of which were a first screening test and 93,010 a subsequent one. Overall, the attendance to subsequent screening of subjects who had already been screened was 90.3\%. iFOBT was positive in 13,388 subjects (5\%), of which 12,095 (90.3\%) accepted to undergo recommended assessments whereas 1295 refused any further diagnostic workup (9.7\%). Complete colonoscopy was carried out in 94.2\% of cases. A total of 748 cancers were detected at screening with a detection rate of 2.8 × 1000 screened. The TNM stage of screen-detected CRCs was I in 56\%, II in 15.7\% and III or IV in 19.2\% subjects respectively (in 9.1\% cases the stage was not registered). Advanced adenomas (larger than 9 mm, or with high-grade dysplasia, or villous component >20\%) were detected in 5213 subjects (detection rate 12.0 × 1000 screened).

During 1999–2002, in the Veneto Region, the crude incidence rates for colorectal cancer were equal to 84.8 per 100,000 residents in males and 65.0 per 100,000 for females and increased with age (males: from 62.5 per 100,000 in subjects 50–54 years old to 281.4 in those 65–69; females: from 56.5 to 136.4, respectively).\(^2\)

The estimates of sensitivity were calculated on 267,021 person-years of follow-up in the first interval-year after a negative screening episode and 201,285 in the second (table 3). Overall, 126 ICs were identified, 50 in the first and 76 in the second interval-year after a screening episode, compared to 327 and 245 expected cancers. Therefore the proportional incidence was respectively 15.5\% and 31.0\% in the first and the second interval-year with an overall sensitivity of 78.0\% (95% CI 75.8\% to 81.6\%). The episode sensitivity corrected for selection bias was 74.9\% (95% CI 70.1\% to 79.1\%).

The sensitivity of the single programmes ranged from 67.0\% (Feltre) to 80.0\% (Dolo-Mirano). The proportional incidence in the first and second interval-year was 14\% and 38\% for Alto Vicentino, 16\% and 29\% for Bussolengo, 15\% and 26\% for Dolo-Mirano, 23\% and 45\% for Feltre, 14\% and 28\% for Pieve di Soligo.

Overall, the sensitivity for subjects at the first screening test was 77.1\% (proportional incidence at the first and second

---

**Table 2** Main data of the bowel screening programmes involved in the study

<table>
<thead>
<tr>
<th>Programme</th>
<th>Date of start</th>
<th>Attendance to invitation (%)</th>
<th>Screened subjects (n)*</th>
<th>iFOBT positivity (%)</th>
<th>Attendance to colonoscopy (%)</th>
<th>Screen detected cancers (n)</th>
<th>Detection rate for cancer (× 1000 screened)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alto Vicentino</td>
<td>May 2004</td>
<td>79.5</td>
<td>45983</td>
<td>4.4</td>
<td>95.1</td>
<td>158</td>
<td>3.4</td>
</tr>
<tr>
<td>Bussolengo</td>
<td>May 2004</td>
<td>66.0</td>
<td>46296</td>
<td>5.1</td>
<td>93.2</td>
<td>135</td>
<td>2.8</td>
</tr>
<tr>
<td>Dolo Mirano</td>
<td>May 2002</td>
<td>60.7</td>
<td>65469</td>
<td>4.9</td>
<td>87.4</td>
<td>147</td>
<td>2.2</td>
</tr>
<tr>
<td>Feltre</td>
<td>Jan 2004</td>
<td>64.0</td>
<td>20912</td>
<td>4.4</td>
<td>93.5</td>
<td>106</td>
<td>5.1</td>
</tr>
<tr>
<td>Pieve di Soligo</td>
<td>Oct 2002</td>
<td>67.0</td>
<td>87109</td>
<td>5.4</td>
<td>89.5</td>
<td>202</td>
<td>2.3</td>
</tr>
<tr>
<td><strong>OVERALL</strong></td>
<td></td>
<td>66.5</td>
<td>267769</td>
<td>5.0</td>
<td>90.3</td>
<td>748</td>
<td>2.8</td>
</tr>
</tbody>
</table>

*From the date of start up to 31 December 2007 for all programmes.

iFOBT: faecal immunochemical test.
In fact, the current schedule distribution by stage was worse for proximal cases (up to the TNM stage I, 22.3% at stage II and 56.5% at stage III to 74.9% to 83.1%). The test sensitivity corrected for selection bias was 79.3% (95% CI 78.1% to 85.3%), with a proportional incidence at the caecum, three ascending and one transverse). At the time of colonoscopy (12.6%) and 8 (6.3%) had had a positive iFOBT followed iFOBT positive subjects who had refused to undergo colonoscopy. Of the 85 cases with known stage at diagnosis, 21.2% were at stage I, 70.6% of which were at stage III–IV, than for those distal (46.8%).

To estimate the sensitivity of our HDR-based approach, we calculated the proportion of cases registered in the years 2002–2005 by the Veneto Tumour Registry that would have been lost, had HDR been the only source of data. Out of 1405 registered cancers, 1386 would have been identified through a HDR, with a sensitivity of 98.6%, while 19 cases would have been lost (1.4%).

**DISCUSSION**

This is one of the first evaluations of the test and episode sensitivity of iFOBT-based bowel screening programmes. The analysis was based on the follow-up of 468,306 person-years relative to five different screening programmes. We recorded an episode sensitivity of 78.0% and an overall test sensitivity of 82.1%. The values obtained by the single programmes were homogeneous. The lack of standards limits the possibility of a final judgment on these results. Nevertheless, these data seem very satisfying when compared, for instance, to the standard of breast cancer screening (proportional incidence <30% in the first interval-year and <30% in the second).24

According to the national surveys of the Italian screening programmes, iFOBT has proven to be a very good first-level test as regards process indicators (acceptability, positivity rate, positive predictive value at colonoscopy) and early indicators of impact (detection rates of carcinoma and advanced adenoma).15 24 25 Our results add a positive evaluation of iFOBT also in terms of sensitivity, estimated through the IC rates. Moreover, the schedule utilised in our programmes (single iFOBT, positivity threshold of 100 ng Hb/ml of sample solution, inter-screening interval of 2 years), that is currently used by all Italian programmes, seems to adequately control the burden of missed cancers that are diagnosed during the interval.

One unexpected outcome of this study was the identification of subgroups of subjects (identified by age, gender, first or subsequent screening test) at a higher IC rate, who could benefit from a more intensive screening protocol such as lowering the positivity cut-off and doubling the number of sampled bowel movements, alone or combined.26 In fact, the current schedule showed a homogeneously high sensitivity whichever subgroup we took into consideration and we were unable to highlight any situation where a more sensitive approach could be indicated.

Sensitivity registered in the second-interval year was suboptimal and the adoption of a shorter inter-screening interval would probably increase sensitivity by converting to ‘screen-detected’ some of the interval cancers occurring in the second year. Of course this strategy should be evaluated also in terms of costs and organisational sustainability.

Our results suggest that the adoption of a more specific strategy could be taken into consideration, for instance utilising a higher positivity threshold. In many Italian programmes, affected by long waiting times for endoscopic assessment, this could be the case of subjects at subsequent screening episodes that showed high positivity rates (3.9% in 2007) but that have differed from that observed by most statistics in the absence of organised screening programmes.27–32

As expected, the distribution by stage at diagnosis of the interval cancers is much worse than that of the screen-detected cases (more than 50% of which are at stage I, according to the Italian national surveys)14 24 25 but not different from that observed by most statistics in the absence of organised screening programmes.27–32 A deeper insight into this aspect could be obtained from the survival analysis of interval cancers, which is still ongoing.
Regarding anatomical site, the sensitivity for carcinomas of the colon overall was good (77%), while it was lower for the proximal tract (65%). We observed a high proportion of proximal cases, which were affected by a worse stage distribution. These data seem to indicate a poorer performance of iFOBT for proximal lesions, as already described by some authors but not by others. Moreover, five out of the eight ICs that were diagnosed after a negative colonoscopy were proximal. Improving adenoma detection at colonoscopy could reduce the rate of ICs, particularly of those proximal.

The comparison with the literature is difficult, since most studies evaluated the sensitivity of guaiac test-based programmes, whose performance has been shown to be lower than that of iFOBT. The published results of sensitivity range from 50% to 96%, with higher performances recorded by rehydrated Hemoccult and when a higher number of samples are taken.

Although the methods underlying these results are not totally comparable with ours, it seems that the sensitivity of our programmes is higher than that obtained from most trials and from service screening programmes. Our results are comparable to those of the two other studies that analysed the proportional incidence of a iFOBT-based programme.

In the Minnesota Colon Cancer Control Study, sensitivity ranged from 89% to 96% according to the method used, but in that study screening was annual and a set of six guaiac slides (rehydrated Hemoccult) was used at each screening round.

Other studies have assessed FOBT sensitivity using colonoscopy as the gold standard: sensitivity estimates ranged from 66% to 100% and were directly associated with the number of day sampling (3 vs 1), stage (advanced vs early CRC) and right vs left colon. However, studies comparing FOBT results with colonoscopy for assessing FOBT sensitivity are not comparable with our study due to lead time or overdiagnosis biases.

The gathering of the necessary data to calculate the proportional incidence is difficult, since the underlying incidence, in the absence of screening, and a complete list of ICs, are needed. This may explain why this indicator is generally neglected by service screening and its monitoring is rarely carried out. Therefore, the use of HDR is particularly useful and appealing even for screening programmes of areas not covered by Cancer Registries. It is a convenient procedure as the data from HDR are available long before that of Cancer Registries and are the most reliable source of data where Cancer Registries are not available. Thus, screening programmes should currently utilise them.

The main limitation of our HDR-based approach is the possible loss of cases. In fact, some cancerised adenomas may be treated exclusively during colonoscopy in an outpatient setting and therefore not picked up through HDR. It is plausible that this is a limited phenomenon: since most cancerised adenomas are asymptomatic, it is unlikely that, after a negative iFOBT, a subject may undergo an endoscopic examination only for preventive purposes. In fact, our estimates on the CRC registered during 2002–2003 by the Veneto Tumour Registry showed a very small proportion of missed cases (1.4%).

HDR relative to extra-regional hospitalisations after 31 December 2006 were not available. Since none of the ICs diagnosed between 2002 and 2006 was identified exclusively as a result of extra-regional hospitalisations, the loss of cases, if any, is likely to be small.

The studies by Zappa and Castiglione are based on a Cancer Registry as the only source of data for ICs. Since their results are very similar to ours, this may represent a further indirect proof of the good quality of our HDR-based approach.

The sensitivity analysis we carried out to address selection bias showed a small decrease in sensitivity compared to the non-corrected estimates. Higher attendance rates among non-smokers and in subjects adopting healthy protective behaviours, such as regular physical activity, have been reported. Thus the risk of CRC among the compliers could be lower than average and using the incidence of the general population would overestimate the expected incidence cancers and hence sensitivity.

The programmes used two different tests (OC-Sensor and FOB Gold). The study was not designed to compare different screening tests; however, we did not observe any difference in sensitivity between them and both of them were included in the overall analysis.

To estimate the expected number of cases, we utilised the CRC incidence rates relative to 1999–2002 but we did not account for the underlying incidence trend (annual increase: males +1.6%, females +0.8%). It is plausible, however, that accounting for such trend would determine a small increase of the expected cases and hence a marginal increase of sensitivity.

CONCLUSIONS

Although gFOBT screening has been shown to be effective in reducing CRC mortality, it has been criticised, mostly for its low sensitivity. The recent spread of service screening programmes will hopefully convey the publication of other studies on interval cancers rates and sensitivity estimates. Studies on iFOBTs are still very few; however, our data seem to represent a further stepping stone towards a very positive judgment of these tests for the first level of colorectal cancer screening.

Acknowledgements Thanks to Denise Kilmartin for the English revision of the draft.

Funding The evaluation and analysis was done at the Veneto Tumour Registry, IOV IRCCS.

Competing interests None.

Contributors MZo and MV designed the study; MZo, GG and MZa drafted the manuscript. CF and CFS assembled the data and did the statistical analysis. FB, AB, LC, AM and TM ran the screening centres and collected the data. All authors critically reviewed the manuscript and approved the final version for publication. MZo is guarantor.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES


