**Supplementary Information**

**Evaluation of Long Fragments of Cell-free DNA in the triage of FIT+ patients enrolled in a colorectal cancer screening program: a prospective, cross-sectional study conducted in Naples, Italy**

Mauro Scimia\*, Francesco Pepe\*\*\*, Gianluca Russo\*\*\*, Umberto Malapelle\*\*\*, Simone Scimia\*, Annalaura Alfieri\*\*, Valentina Olivieri\*\*, Rachel Chuang\*, Hiromi Tanaka\*, Michael Sha\*, David Chen\*, Shuo Shen\*, Yulia Gavrilov\*\*\*\*, Stav Edelstein\*\*\*\*, Maria Antonia Bianco\*\*, Giancarlo Troncone\*\*\*.

**Affiliations:** DiaCarta Inc., CA, USA (\*); ASL NA 3 SUD, Torre del Greco (NA), Italy (\*\*); Department of Public Health, University of Naples Federico II, Naples, Italy (\*\*\*); Technostat, Israel (\*\*\*\*)

**Colonoscopy**

Eligible subjects underwent a full colonoscopy (Tables 9-10; Supplementary Information). When a subject whose blood sample was available, did not have a complete colonoscopy, he/she was excluded from the study. A completed colonoscopy procedure was defined as one reaching the caecum or the junction between the small and large intestine.

Visibility and accessibility were evaluated for the full procedure, for the left and right side and the transverse colon. Colonoscopies evaluated at least as “fair” passed this qualification phase. Caecal intubation, total-elapsed time and microscopic alteration of the bowel were also evaluated.

Macroscopic findings were assessed as to status, polyp number, location, morphology, suspected cancer, size, resection, no-lift sign, retrieval, and non-diagnostic information such as diverticula or haemorrhoids (Tables 11-12-14; Supplementary Information).

All positive colonoscopies (see later section on “Endpoints”) were referred to histopathological examination.

**Histopathology**

The following categories were reported (Tables 16-17-18-19):

* 1. Histology results: CRC, AA, NAA, and other findings
	2. CRC: identification (ID), status, lesion number, histology, pathological classification (pT, pN, pM), grade, location of metastasis, cancer location
	3. AA: ID, status, lesion number, polyp histology, dysplasia type, location, size (mm)
	4. NAA: ID, status, lesion number, location and size (mm)
	5. Final diagnosis: Negative, negative with NAA, AA, CRC

**Assay methodology**

**Long cfDNA quantification**

QuantiDNATM test (DiaCarta Inc.) is based on branched DNA (bDNA) Technology to directly detect cell-free DNA (cfDNA) in plasma without cfDNA extraction and purification. Probe designs are specific and preferable for longer cfDNA fragment measurement. 10 μL of plasma was diluted using 90 μL of 1xPBS and boiled at 95°C for 10 minutes to denature double stranded cfDNA. These diluted plasma samples were then cooled on ice for at least 5 minutes, quick vortex, and spun at 10,000xg for 1 minute at 4°C. 20 μL of clear supernatant was then loaded onto the capture plate with 80 μL of working probe mixture with Probe Set B per well. After two hours of hybridization at 55°C for 2 hours using a Thermal Incubator (DiaCarta Inc.), the reaction mixture was removed, then the wells were washed three times. Signal amplification was then applied with the sequential hybridization steps with Pre-Amplifier Probes, Amplifier Probes, then Alkaline Phosphatase conjugated Label Probes. Three wash steps were included in between. After the final incubation step with alkaline phosphatase substrate, related light units (RLUs) were then measured using the Benchtop Luminometer (DiaCarta, Inc.). The fragmented DNA Standards (25, 12.5, 6.25, 3.125, 1.5625, and 0.78125 ng/mL) provided from the test kit were performed side by side with the patient samples on the same capture plate. RLUs data were analyzed automatically using the Luminometer’s CLIA software (DiaCarta, Inc.) and point-to-point analysis was performed with the default equation. The cutoff of negative and positive was set at 6.27 ng/mL as described in the kit manual.

**DiaCarta’s Luminometer**

**Non-inferiority analysis**

According to the basic rules set by EMA (European Medicine Agency) and FDA (Food and Drug Administration), the investigational product met the following requirements (6,7):

* + - 1. Show an effect > 0. This requirement is likely met by the odds ratio of 1.76, that is statistically significant, with p-value = 0.009
			2. The use of the test must not imply any irreversible event, such as a fatality or permanent disability.
			3. It has to provide some tangible benefits, like less adverse events or economic advantages. By decreasing the number of colonoscopies, the alternative approach indirectly reduces the adverse events normally associated with them (bleeding, perforations, and others). Furthermore, 33.4% less patients could avoid the troubles associated with colonoscopy preparation and healthcare facilities’ workload and waiting lists could decrease, thus making a better allocation of resources possible.

The alternative approach appears to satisfy the above conditions.

With regard to the choice of margin, we observed the range of the Adenoma Detection Rate (**ADR**) of colonoscopists. According to published studies, the range that includes NAAs and worse lesions, can vary between 20% and 70% (13). A margin of -10% may thus represent the distance between two groups of colonoscopists within this range. There is a low number of studies highlighting the detection rate ranges for AAs only (**AADR**). For instance, according to Greenspan et al. (14), the variability in the range of AADR went from 2.00% to 18.18%. According to another study by Shaukat et al. (15), the range for AADR varied from 1.3% to 14.1%. In both cases, the choice of a -10% margin falls within such range. We could not find a specification for CRC detection rate (**CRCDR**) in all of these studies, so we derived the choice of a -3.8% margin applying a percentage attributable to CRC in the ADR from some published studies. By doing so, we found that the CRCDR could vary between 2% and 6% approximately, thus including the -3.8% NIM. The actual point estimate for risk difference between SOC and alternative approach is -1.04 (95% C.I. -3.16, 1.07), so the lower bound of the C.I. and the point estimate are well within the -3.8% margin and can, theoretically, prove robust to further restrictions of the NIM down to -3.2%. Also, noteworthy is that the 95% C.I. of both relative risk (**RR**) and OR pass through the unit: RR 0.7407 (0.4196, 1.3077), OR 0.7329 (0.4069, 1.3198), which hints to a non significant difference between AAP and SOC.

Although this is not part of a non-inferiority analysis, there is merit in mentioning that the upper bound of the confidence interval of the risk difference between the two approaches for CRC, crosses to the right, hinting to a possible superiority of the alternative approach.

**Logistic regression**

We used a single predictor logistic regression to evaluate the diagnostic performance of the QuantiDNA™ assay. The fragment score was converted into a binary outcome, using the cut-off of 6.27 ng/ml, with the following formula:

* + - * + $$Logit\left(P(Reference=1)\right)=β\_{0}+β\_{1}\*I\left\{Fragment\geq 6.27\right\}$$
				+ $$Anti-logit(0.2045+1.5535\*I\left\{Fragment\geq 6.27\right\}$$

Following conversion of the QuantiDNA™ test results into a binary outcome, all accuracy statistics (sensitivity, specificity, PPV, NPV, DR, OR, Positivity Rate, NLR and PLR) were calculated along with corresponding 95% confidence intervals.

Confidence intervals for sensitivity and specificity were calculated using Clopper-Pearson method.

Confidence intervals for PPV and NPV were calculated using the logit transformation, as recommended in Mercaldo et al. (2007). Specifically,

For NPV,

$$Variance\left(logit(NPV)\right)=\left[\frac{Se}{1-Se}\right]\frac{1}{\#referencepositive}+\left[\frac{1-Sp}{Sp}\right]\frac{1}{\#referencenegative}$$

where *logit(X) = LN [X / (1-X)]*

The 95% confidence intervals,

$$\frac{exp\left\{logit\left(NPV\right)\pm z\_{0.975}\sqrt{Variance\left(logit(NPV)\right)}\right\}}{1+exp\left\{logit\left(NPV\right)\pm z\_{0.975}\sqrt{Variance\left(logit(NPV)\right)}\right\}}$$

Z-score was calculated in SAS, QUANTILE (‘Normal’, 0.975);

(If Se = 100% or Sp = 0%, the formula above is not well defined. In such cases, Clopper-Pearson should be used.)

For PPV,

$$Variance\left(logit(PPV)\right)=\left[\frac{1-Se}{Se}\right]\frac{1}{\#referencepositive}+\left[\frac{Sp}{1-Sp}\right]\frac{1}{\#referencenegative}$$

The 95% confidence intervals,

$$\frac{exp\left\{logit\left(PPV\right)\pm z\_{0.975}\sqrt{Variance\left(logit(PPV)\right)}\right\}}{1+exp\left\{logit\left(PPV\right)\pm z\_{0.975}\sqrt{Variance\left(logit(PPV)\right)}\right\}}$$

(If Se = 100% or Sp = 0%, the formula above is not well defined. In such cases, Clopper-Pearson should be used.)

**ROC curve**

**Figure 1 : Receiver Operating Characteristics (ROC) Curves for Continuous ColoScape Score By Logistic Regression Model 1 as Derived from Pilot Study - Primary Reference Definition (Fragment Efficacy Analysis Set)**

AUC and confidence limits were calculated using the logistic model.

Model 1:Logit(P(Reference=1))=β\_0+β\_1\*I{Fragment≥6.27}

Primary reference is defined as subject who has the presence of Colorectal Cancer (CRC) or Advanced Adenoma (AA)

**Tables**

**Table 1 : Descriptive Statistics of Age (Years) (Fragment Efficacy Analysis Set)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Mean | Std | Min | Median | Max | N |
| 65.5 | 6.8 | 50.0 | 67.0 | 74.0 | 671 |

**Table 2 : Frequency Distribution of Gender and Race (Fragment Efficacy Analysis Set)**

|  |  |  |  |
| --- | --- | --- | --- |
| Demographic Characteristics | N | % |  |
| Gender | Female | 348 | 51.9 |
| Male | 323 | 48.1 |
| Total | 671 | 100.0 |
| Race | Caucasian | 668 | 99.6 |
| Black or African | 0 | 0.0 |
| Middle Eastern or North African | 0 | 0.0 |
| South Asian | 0 | 0.0 |
| East Asian | 0 | 0.0 |
| Eastern European | 2 | 0.3 |
| Western European | 1 | 0.1 |
| Southeast Asian or Pacific Islander | 0 | 0.0 |
| Other | 0 | 0.0 |
| Total | 671 | 100.0 |

**Table 3 : Descriptive Statistics of Height, Weight and BMI (Fragment Efficacy Analysis Set)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Parameter | Mean | Std | Min | Median | Max | N |
| Height (cm) | 165.2 | 8.3 | 147.0 | 165.0 | 190.0 | 671 |
| Weight (kg) | 76.7 | 14.3 | 43.0 | 75.0 | 180.0 | 671 |
| BMI (kg/m2) | 28.0 | 4.9 | 2.3 | 27.4 | 70.3 | 671 |

**Table 4 : Frequency Distribution of Smoking Status (Fragment Efficacy Analysis Set)**

|  |  |  |
| --- | --- | --- |
| Smoking Status | N | % |
| Never | 493 | 73.5 |
| Current | 110 | 16.4 |
| Former | 68 | 10.1 |
| Total | 671 | 100.0 |

**Table 5 : Descriptive Statistics of Daily Tobacco Use (Cigarettes Smoked Per Day) (Fragment Efficacy Analysis Set)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Mean | Std | Min | Median | Max | N |
| 17.5 | 11.1 | 1.0 | 15.0 | 60.0 | 178 |

**Table 7 : Frequency Distribution of Subjects Who Have Experienced Any Clinically Relevant Medical Condition (Fragment Efficacy Analysis Set)**

|  |  |  |
| --- | --- | --- |
| Experienced Any Clinically Relevant Medical Conditions | N | % |
| Yes | 395 | 58.9 |
| No | 276 | 41.1 |
| Total | 671 | 100.0 |

**Table 8 : Frequency Distribution of Family History (Fragment Efficacy Analysis Set)**

|  |  |  |
| --- | --- | --- |
| Family History | N | % |
| Subject has family history of FAP - Familial Adenomatosis Polyposis Syndrome | 0 | 0.0 |
| Subject has family history of HNPCC - Hereditary Non-Polyposis Colorectal Carcinoma | 1 | 0.1 |
| Subject has one first-degree relative (parents, siblings, and offspring) with CRC diagnosed before age of 60 | 35 | 5.2 |
| Subject has greater than or equal to 2 first-degree relatives who have been diagnosed with colon cancer | 4 | 0.6 |
| Subject has family history of colon cancer | 16 | 2.4 |
| Total | 671 | 100.0 |

**Table 9 : Frequency Distribution of Colonoscopy Evaluation of Visibility and Accessibility (Fragment Efficacy Analysis Set)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Location | Excellent | Good | Fair | Poor | Total |
| N | % | N | % | N | % | N | % | N | % |
| Full Colonoscopy | 0 | 0.0 | 653 | 97.3 | 13 | 1.9 | 5 | 0.7 | 671 | 100.0 |
| Left Side of the Colon | 0 | 0.0 | 657 | 97.9 | 9 | 1.3 | 5 | 0.7 | 671 | 100.0 |
| Right Side of the Colon | 0 | 0.0 | 658 | 98.1 | 8 | 1.2 | 5 | 0.7 | 671 | 100.0 |
| Transverse Colon | 0 | 0.0 | 658 | 98.1 | 8 | 1.2 | 5 | 0.7 | 671 | 100.0 |

**Table 10 : Descriptive Statistics of Total Elapsed Time (Minutes) (Fragment Efficacy Analysis Set)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Mean | Std | Min | Median | Max | N |
| 15.1 | 1.1 | 5.0 | 15.0 | 30.0 | 671 |

**Table 11 : Frequency Distribution of Colonoscopy Findings (Fragment Efficacy Analysis Set)**

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | Yes | No | Total |
| N | % | N | % | N | % |
| 1 or 2 colonoscopies | 303 | 45.2 | 368 | 54.8 | 671 | 100.0 |
| At least one polyp | 303 | 100.0 | 0 | 0.0 | 303 | 100.0 |
| At least one suspected cancer polyp | 28 | 9.2 | 275 | 90.8 | 303 | 100.0 |
| At least one resection was done | 282 | 93.1 | 21 | 6.9 | 303 | 100.0 |
| At least one no-lift sign | 30 | 9.9 | 273 | 90.1 | 303 | 100.0 |
| At least one retrieval | 281 | 92.7 | 22 | 7.3 | 303 | 100.0 |

From the second parameter down, percentages were calculated out of total number of subjects who had at least one colonoscopy and were included in the analysis set.

**Table 12 : Descriptive Statistics for the Number of Polyps Found (Fragment Efficacy Analysis Set)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Mean | Std | Min | Median | Max | N |
| 1.5 | 0.9 | 1.0 | 1.0 | 5.0 | 303 |

**Table 14 : Frequency Distribution of the Largest Polyp Size (Fragment Efficacy Analysis Set)**

|  |  |  |
| --- | --- | --- |
| Polyp Size | N | % |
| 0-9 mm | 141 | 46.5 |
| 10-19 mm | 106 | 35.0 |
| 20-29 mm | 22 | 7.3 |
| > 29 mm | 34 | 11.2 |
| Total | 303 | 100.0 |

Each patient counted once in the size subgroup based on the size of the largest polyp.

**Table 15 : Frequency Distribution of Subjects for Whom Histopathology was Performed (Fragment Efficacy Analysis Set)**

|  |  |  |
| --- | --- | --- |
| Histopathology was Performed | N | % |
| Yes | 302 | 45.0 |
| No | 369 | 55.0 |
| Total | 671 | 100.0 |

**Table 16 : Frequency Distribution of Histopathology Results (Fragment Efficacy Analysis Set)**

|  |  |  |
| --- | --- | --- |
| Histopathology Results | N | % |
| Colorectal Cancer, stages I-IV | 27 | 8.9 |
| Advanced Adenoma | 119 | 39.4 |
| Non-advanced Adenoma | 198 | 65.6 |
| Negative, no other findings | 6 | 2.0 |
| None of the Above | 1 | 0.3 |
| Total | 302 | 100.0 |

Percentages were calculated out of all subjects who had histopathology in the analysis set. Multiple selections per subject were allowed

**Table 17 : Frequency Distribution of Final Diagnosis (Fragment Efficacy Analysis Set)**

|  |  |  |
| --- | --- | --- |
| Final Diagnosis | N | % |
| Colorectal Cancer | 27 | 8.9 |
| Advanced Adenoma | 114 | 37.7 |
| Negative, with Non Advanced Adenoma | 155 | 51.3 |
| Negative, no other findings | 6 | 2.0 |
| Total | 302 | 100.0 |

**Table 18 : Descriptive Statistics of Polyp Size (mm) (Fragment Efficacy Analysis Set)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Mean | Std | Min | Median | Max | N | N Missing |
| 15.1 | 6.9 | 6.0 | 12.0 | 50.0 | 142 | 1 |

**Non-Advanced Adenoma**

**Table 19: Frequency Distribution of Polyp Location in Bowel (Fragment Efficacy Analysis Set)**

|  |  |  |  |
| --- | --- | --- | --- |
| Location | NP | NS | %S |
| Ileum | 0 | 0 | 0.0 |
| Ileocecal value | 0 | 0 | 0.0 |
| Cecum | 21 | 17 | 8.4 |
| Ascending colon (right) | 61 | 57 | 28.2 |
| Hepatic flexure | 0 | 0 | 0.0 |
| Transverse colon | 47 | 41 | 20.3 |
| Splenic flexure | 0 | 0 | 0.0 |
| Descending colon (left) | 33 | 31 | 15.3 |
| Sigmoid colon | 73 | 67 | 33.2 |
| Rectum | 50 | 40 | 19.8 |
| Missing | 1 | 1 | 0.5 |
| Total | 286 | 202 | 100.0 |

Np = number of polyps that were found in the location. Ns = number of subjects with at least one polyp found in the location.

Percentages were calculated out of total number of subjects with at least one non-advanced adenoma polyp.

**Table 20 : Accuracy Parameters by Logistic Regression Model as Derived from Pilot Study – Primary Reference Definition**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Parameter | Total N | True Outcome | Estimate (%) | Lower 95% CL | Upper 95% CL | P-Value |
| Sensitivity | 141 | 107 | 75.9 | 68.0 | 82.7 |  |
| Specificity | 530 | 190 | 35.8 | 31.8 | 40.1 |  |
| PPV | 447 | 107 | 23.9 | 21.9 | 26.0 |  |
| NPV | 224 | 190 | 84.8 | 80.3 | 88.4 |  |
| DR | 671 | 107 | 15.9 | 13.3 | 18.9 |  |
| PR | 671 | 447 | 66.6 | 62.9 | 70.2 |  |
| NLR |  |  | 0.67 | 0.49 | 0.92 |  |
| PLR |  |  | 1.18 | 1.06 | 1.32 |  |
| OR |  |  | 1.76 | 1.15 | 2.69 | 0.009 |
| Youden's J statistic |  |  | 0.117 |  |  |  |

Confidence intervals for sensitivity and specificity were calculated using Clopper-Pearson method

Confidence intervals for PPV and NPV were calculated using the logit transformation, as recommended in Mercaldo et al. (2007).

Confidence intervals for PLR and NLR were calculated using the delta method, as described in Zhou, Obuchowski, and McClish, 2nd ed, 2011 (p. 113).

Model: Anti-logit(0.2045+1.5535\*I{Fragment≥6.27})

Primary reference is defined as subject who has the presence of Colorectal Cancer (CRC) or Advanced Adenoma (AA)

The analysis was performed on the Efficacy analysis set.

**Table 21 : Detection Rate (DR) by Lesion Type and Group (Fragment Efficacy Analysis Set)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Cancer Type | Group | Total N | True Outcome | Estimate (%) | Lower 95% CL | Upper 95% CL |
| CN | QuantiDNA | 671 | 107 | 15.9 | 13.3 | 18.9 |
| SOC | 671 | 141 | 21.0 | 18.0 | 24.3 |
| AA | QuantiDNA | 671 | 87 | 13.0 | 10.5 | 15.7 |
| SOC | 671 | 114 | 17.0 | 14.2 | 20.0 |
| CRC | QuantiDNA | 671 | 20 | 3.0 | 1.8 | 4.6 |
| SOC | 671 | 27 | 4.0 | 2.7 | 5.8 |

DR is calculated as the number of true positive cases/total cases.

Confidence interval was calculated using Clopper-Pearson method

Model 1: Anti-logit(0.2045+1.5535\*I{Fragment≥6.27})

**Table 22 : Non-Inferiority Analysis on Detection Rate (DR) by Lesion Type (Fragment Efficacy Analysis Set)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Cancer Type | Non-inferiority Margin (%) | Risk Difference (%) | Lower 95% CL | Upper 95% CL | P-Value |
| CN | -10 | -5.07 | -9.23 | -0.90 | 0.010 |
| AA | -10 | -4.02 | -7.89 | -0.16 | 0.001 |
| CRC | -3.8 | -1.04 | -3.16 | 1.07 | 0.005 |

The non-inferiority hypothesis was tested using the Farrington-Manning method with one-sided alpha=0.025.



Point estimates and 95% C.I. lower bound for AA, CRC and CN entirely within margin (confirms non-inferiority)

**Table 23 : Accuracy Parameters by Logistic Regression Model as Derived from Pilot Study for AA Cases – Primary Reference Definition**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Parameter | Total N | True Outcome | Estimate (%) | Lower 95% CL | Upper 95% CL | P-Value |
| Sensitivity | 114 | 87 | 76.3 | 67.4 | 83.8 |  |
| Specificity | 557 | 197 | 35.4 | 31.4 | 39.5 |  |
| PPV | 447 | 87 | 19.5 | 17.7 | 21.4 |  |
| NPV | 224 | 197 | 87.9 | 83.7 | 91.2 |  |
| DR | 671 | 87 | 13.0 | 10.5 | 15.7 |  |
| PR | 671 | 447 | 66.6 | 62.9 | 70.2 |  |
| NLR |  |  | 0.67 | 0.47 | 0.95 |  |
| PLR |  |  | 1.18 | 1.05 | 1.33 |  |
| OR |  |  | 1.76 | 1.11 | 2.81 | 0.017 |
| Youden's J statistic |  |  | 0.117 |  |  |  |

Confidence intervals for sensitivity and specificity were calculated using Clopper-Pearson method

Confidence intervals for PPV and NPV were calculated using the logit transformation, as recommended in Mercaldo et al. (2007).

Confidence intervals for PLR and NLR were calculated using the delta method, as described in Zhou, Obuchowski, and McClish, 2nd ed, 2011 (p. 113).

Model: Anti-logit(0.2045+1.5535\*I{Fragment≥6.27})

Primary reference is defined as subject who has the presence of Colorectal Cancer (CRC) or Advanced Adenoma (AA)

The analysis was performed on the Fragments analysis set.

**Table 24 : Accuracy Parameters by Logistic Regression Model as Derived from Pilot Study for CRC Cases – Primary Reference Definition**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Parameter | Total N | True Outcome | Estimate (%) | Lower 95% CL | Upper 95% CL | P-Value |
| Sensitivity | 27 | 20 | 74.1 | 53.7 | 88.9 |  |
| Specificity | 644 | 217 | 33.7 | 30.0 | 37.5 |  |
| PPV | 447 | 20 | 4.5 | 3.6 | 5.6 |  |
| NPV | 224 | 217 | 96.9 | 94.2 | 98.3 |  |
| DR | 671 | 20 | 3.0 | 1.8 | 4.6 |  |
| PR | 671 | 447 | 66.6 | 62.9 | 70.2 |  |
| NLR |  |  | 0.77 | 0.40 | 1.47 |  |
| PLR |  |  | 1.12 | 0.89 | 1.41 |  |
| OR |  |  | 1.45 | 0.60 | 3.49 | 0.404 |
| Youden's J statistic |  |  | 0.078 |  |  |  |

Confidence intervals for sensitivity and specificity were calculated using Clopper-Pearson method

Confidence intervals for PPV and NPV were calculated using the logit transformation, as recommended in Mercaldo et al. (2007).

Confidence intervals for PLR and NLR were calculated using the delta method, as described in Zhou, Obuchowski, and McClish, 2nd ed, 2011 (p. 113).

Model: Anti-logit(0.2045+1.5535\*I{Fragment≥6.27})

Primary reference is defined as subject who has the presence of Colorectal Cancer (CRC) or Advanced Adenoma (AA)

The analysis was performed on the Fragments analysis set.

**Table 25 : Accuracy Parameters by Logistic Regression Model as Derived from Pilot Study – Secondary Reference Definition**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Parameter | Total N | True Outcome | Estimate (%) | Lower 95% CL | Upper 95% CL | P-Value |
| Sensitivity | 297 | 201 | 67.7 | 62.0 | 73.0 |  |
| Specificity | 374 | 128 | 34.2 | 29.4 | 39.3 |  |
| PPV | 447 | 201 | 45.0 | 42.3 | 47.6 |  |
| NPV | 224 | 128 | 57.1 | 51.8 | 62.3 |  |
| DR | 671 | 201 | 30.0 | 26.5 | 33.6 |  |
| PR | 671 | 447 | 66.6 | 62.9 | 70.2 |  |
| NLR |  |  | 0.94 | 0.76 | 1.17 |  |
| PLR |  |  | 1.03 | 0.92 | 1.15 |  |
| OR |  |  | 1.09 | 0.79 | 1.51 | 0.604 |
| Youden's J statistic |  |  | 0.019 |  |  |  |

Confidence intervals for sensitivity and specificity were calculated using Clopper-Pearson method

Confidence intervals for PPV and NPV were calculated using the logit transformation, as recommended in Mercaldo et al. (2007).

Confidence intervals for PLR and NLR were calculated using the delta method, as described in Zhou, Obuchowski, and McClish, 2nd ed, 2011 (p. 113).

Model: Anti-logit(0.2045+1.5535\*I{Fragment≥6.27})

Secondary reference is defined as subject who has the presence of Colorectal Cancer (CRC), Advanced Adenoma (AA) or Non Advanced Adenoma (NAA)

The analysis was performed on the Efficacy analysis set.

**Table 26 : Accuracy Parameters by Logistic Regression Model and AA Polyp Histology as Derived from Pilot Study – Primary Reference Definition**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Value | Parameter | Model 1 | Model 2 | Model 3 |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | Total N | n | Estimate | Lower 95% CL | Upper 95% CL | Total N | n | Estimate | Lower 95% CL | Upper 95% CL | Total N | n | Estimate | Lower 95% CL | Upper 95% CL |
| Adenoma-villous growth pattern (>=25%) | Sensitivity | 47 | 35 | 74.5 | 59.7 | 86.1 | 49 | 8 | 16.3 | 7.3 | 29.7 | 47 | 38 | 80.9 | 66.7 | 90.9 |
| PPV | 35 | 35 | 100.0 | 90.0 | 100.0 | 8 | 8 | 100.0 | 63.1 | 100.0 | 38 | 38 | 100.0 | 90.7 | 100.0 |
| Serrated lesion >/= 10mm | Sensitivity | 26 | 21 | 80.8 | 60.6 | 93.4 | 27 | 2 | 7.4 | 0.9 | 24.3 | 26 | 21 | 80.8 | 60.6 | 93.4 |
| PPV | 21 | 21 | 100.0 | 83.9 | 100.0 | 2 | 2 | 100.0 | 15.8 | 100.0 | 21 | 21 | 100.0 | 83.9 | 100.0 |
| Tubular adenoma >/= 10mm | Sensitivity | 42 | 32 | 76.2 | 60.5 | 87.9 | 45 | 5 | 11.1 | 3.7 | 24.1 | 42 | 34 | 81.0 | 65.9 | 91.4 |
| PPV | 32 | 32 | 100.0 | 89.1 | 100.0 | 5 | 5 | 100.0 | 47.8 | 100.0 | 34 | 34 | 100.0 | 89.7 | 100.0 |

n = True Outcome

Sensitivity and PPV are displayed as percentage.

Confidence interval for sensitivity was calculated using Clopper-Pearson method

Confidence interval for PPV was calculated using the logit transformation, as recommended in Mercaldo et al. (2007). For estimated values of (0,1) Clopper-Pearson method was used.

Model 1: Anti-logit(0.2045+1.5535\*I{Fragment≥6.27})

Model 2: Anti-logit(0.2441+1.4817\*Mutation)

Model 3: Anti-logit(0.1844+1.5024\*I{Fragment≥6.27}+1.4824\*Mutation)

For model 1, the analysis was performed on the Fragment Efficacy Analysis Set. For model 2, the analysis was performed on the Mutation Analysis Set. For model 3, the analysis was performed on the Efficacy Analysis Set.

Polyp-level covariates: For high grade dysplasia, a subject was counted in the "high grade" group if they had at least one polyp marked as "Polyp Histology = Adenoma with carcinoma in situ or high grade dysplasia" or "Dysplasia Type = High".







