SUMMARY OF SAFETY AND EFFECTIVENESS DATA (SSED)

I. GENERAL INFORMATION

Device Generic Name:	Stool DNA-Based Colorectal Cancer Screening Test
Device Trade Name:	*Cologuard*™
Device Procode:	PHP
Applicant's Name and Address:	Exact Sciences Corporation 441 Charmany Drive Madison, WI 53719
Date of Panel Recommendation:	March 27, 2014
Premarket Approval Application (PMA) Number:	P130017
Date of FDA Notice of Approval:	August 11, 2014

Priority Review: Granted priority review status on June 12, 2013 because *Cologuard* is a first of a kind device that requires breakthrough technology.

II. INTENDED USE AND INDICATIONS FOR USE

Intended Use

Cologuard is intended for the qualitative detection of colorectal neoplasia associated DNA markers and for the presence of occult hemoglobin in human stool. *Cologuard* is for use with the *Cologuard* collection kit and the following instruments: BioTek ELx808 Absorbance Microplate Reader; Applied Biosystems® 7500 Fast Dx Real-Time PCR; Hamilton Microlab® STARlet; and the Exact Sciences System Software with *Cologuard* Test Definition.

Indications for Use

Cologuard is intended for the qualitative detection of colorectal neoplasia associated DNA markers and for the presence of occult hemoglobin in human stool. A positive result may indicate the presence of colorectal cancer (CRC) or advanced adenoma (AA) and should be followed by diagnostic colonoscopy. *Cologuard* is indicated to screen adults of either sex, 50 years or older, who are at typical average-risk for CRC. *Cologuard* is not a replacement for diagnostic colonoscopy or surveillance colonoscopy in high risk individuals.

III. CONTRAINDICATIONS

Cologuard is intended for use with patients, age 50 years and older, at average risk who are typical candidates for CRC screening. *Cologuard* was not clinically evaluated for the following types of patients:

- Patients with a history of colorectal cancer, adenomas, or other related cancers.
- Patients who have had a positive result from another colorectal cancer screening method within the last 6 months.
- Patients who have been diagnosed with a condition that is associated with high risk for colorectal cancer. These include but are not limited to:
    - Inflammatory Bowel Disease (IBD)
    - Chronic ulcerative colitis (CUC)
    - Crohn's disease
    - Familial adenomatous polyposis (FAP)
    - Family history of colorectal cancer
- Patients who have been diagnosed with a relevant familial (hereditary) cancer syndrome, such as Hereditary non-polyposis colorectal cancer syndrome (HNPCCC or Lynch Syndrome), Peutz-Jeghers Syndrome, MYH-Associated Polyposis (MAP), Gardner's syndrome, Turcot's (or Crail's) syndrome, Cowden's syndrome, Juvenile Polyposis, Cronkhite-Canada syndrome, Neurofibromatosis, or Familial Hyperplastic Polyposis.

IV. WARNINGS AND PRECAUTIONS

The warnings and precautions can be found in the *Cologuard* labeling.

V. DEVICE DESCRIPTION

Cologuard is an *in vitro* diagnostic device designed to analyze patients' stool for the presence of colorectal cancer (CRC) and pre-malignant colorectal neoplasia ("Advanced Adenoma" or "AA") through detection of hemoglobin, multiple DNA methylation and mutational markers, and the total amount of human DNA. Specifically, *Cologuard* is designed to detect three (3) independent categories of biomarkers that exhibit an additive association with CRC and AA. The first category of biomarkers targets epigenetic changes in the form of gene promoter region methylation. The specific methylated gene targets include N-Myc Downstream-Regulated Gene 4 (*NDRG4*) and Bone Morphogenetic Protein 3 (*BMP3*).

The second category targets seven (7) specific gene mutations in V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (*KRAS).* The third biomarker is non-DNA based and detects occult hemoglobin. Additionally, beta-actin ("*ACTB"*) is a reference gene used for confirmation and quantitative estimation of the total amount of human DNA present in each sample. Results from the methylation, mutation, and hemoglobin assays are integrated by the Exact Sciences Analysis Software to determine a Positive or Negative reportable result or invalid result. *Cologuard* cannot distinguish between methylcytosine and hydroxymethylcytosine.

Cologuard uses the following reagent components:

DNA Capture Reagents
CAP BDS, Capture Beads

DNA Preparation Reagents
DEN SLN, Denaturation Solution
BIS SLN, Bisulfite Conversion Solution
NEU SLN, Neutralization Solution
DES SLN, Desulphonation Solution (Concentrate)
BND BDS, Binding Beads
DNA and *QuARTS* Supplementary Lot Information Card

***QuARTS* Assay Reagents**
CAR SLN, Carrier Solution
ELU BFR, Elution Buffer
MIX A, Oligo Mix A, Methylation
MIX B, Oligo Mix B, Mutation
ENZ, Enzyme Mix
D CAL 1, DNA Calibrator 1, High Methylation
D CAL 2, DNA Calibrator 2, Low Methylation
D CAL 3, DNA Calibrator 3, High Mutation
D CAL 4, DNA Calibrator 4, Low Mutation

Hemoglobin Assay Reagents
Hb PLATE, Hemoglobin Assay Plate
SMP BFR, Sample Buffer
CONJ, Antibody Conjugate
SUBS, Substrate
STP SLN, Stop Solution
Hb CAL, Hemoglobin Assay Calibrator
Hemoglobin Assay Supplementary Lot Information Card

In addition, the following components are required for use of *Cologuard*:

(1) *Cologuard* Collection Kit containing the patient instructions, a protein sample tube with stool collection stick and buffer, a stool collection container, a foldable plastic bracket, a liquid preservative and a mailing container.

(2) *Cologuard* DNA Control Kit containing:

- DNA Control 1, High and DNA Control 2, Low with specific copy numbers of relevant methylated and non-methylated DNA
- DNA Control 3, Negative with a specific copy number of non-methylated DNA

(3) *Cologuard* Hemoglobin Control Kit containing:

- Lyophilized Hemoglobin Control 1, High and Hemoglobin Control 2, Low derived from human whole blood and plasma containing specific concentrations of human hemoglobin

- Lyophilized Hemoglobin Control 3, Negative derived from human whole blood and plasma with no human hemoglobin

(4) Ancillary Materials and Bulk Assay Reagents:

- STL BFR, Stool Buffer
- TABLT, Inhibitor Removal Tablet
- FILT, Spin Filter
- TUBES, Barcoded Mixing Tubes
- PRE WSH, Capture Bead Pre-wash
- CAP SLN, Capture Solution
- CAP WSH, Capture Wash
- BND SLN, Binding Solution
- CNV WSH, Conversion Wash Concentrate
- Hb WSH, Hemoglobin Assay Wash Concentrate

(5) BioTek ELx808 Absorbance Microplate Reader multichannel ELISA reader.

(6) Applied Biosystems® 7500 Fast Dx Real-Time PCR Instrument with integrated thermal cycler and fluorimeter.

(7) Capture Incubator for automation of DNA capture hybridization.

(8) Capture Aspirator for automation of DNA capture clean-up washes.

(9) Hamilton Microlab®[1] STARlet for automation of the DNA preparation and *QuARTS* assay set up process.

(10) Exact Sciences System Software with *Cologuard* Test Definition.

(11) Other general lab equipment specified (centrifuge, shaker, bottle top dispenser, mixer etc.).

Principles of Operation

Cologuard involves stool DNA-based (sDNA) testing, which detects molecular markers of altered DNA that are contained in the cells shed by CRC or AA into the lumen of the large bowel. The DNA markers are released from cells that continuously slough from the lining of the colon into the stool. Through the use of selective enrichment and amplification techniques, sDNA tests are designed to detect even very small amounts of the DNA markers to identify CRC or AA. In addition, the test incorporates detection of fecal occult hemoglobin. Hemoglobin levels are quantified using the *Cologuard* FIT test, which is an Enzyme-Linked Immunosorbent Assay (ELISA) designed to quantify Hemoglobin in stool.

Stool samples are collected using the *Cologuard* Collection Kit, which includes patient instructions, a protein sample tube with stool collection stick and buffer, a stool collection container, a foldable plastic bracket, a liquid preservative, and a mailing container. The mailing container is used to send the collected sample to a lab for processing.

[1] Microlab® is a registered trademark of Hamilton Company.

Once received, the stool sample is weighed, diluted, homogenized, and aliquots of the homogenates are taken and frozen. After pre-processing the *Cologuard* test begins with: (1) target specific capture to isolate DNA from frozen stool homogenates; (2) the aliquot for the methylation assay is treated with bisulfite conversion reagents; and (3) DNA purification coupled with Quantitative Allele-Specific Real-time Target and Signal (*QuARTS*™)[2] amplification. The *QuARTS*™ amplification technology combines the routinely used molecular biology techniques of real-time PCR and invasive cleavage chemistry to perform allele-specific amplification and detection of methylated target DNA (*NDRG4, BMP3),* specific DNA point mutations (*KRAS*) and total human DNA (*ACTB).*

In a parallel workflow, a quantitative ELISA technique is used to analyze the level of hemoglobin present in the stool sample. The final *Cologuard* result is determined utilizing a composite score based on a patient's individual methylation, mutation, and hemoglobin assay results. The score is calculated by multiplying a patient's individual methylation, mutation, and hemoglobin assay results by a constant marker specific weighting factor. The aggregate of these individually weighted marker results determines the composite score (range of 0-1000), which is then compared to a cut-off (183) to determine a positive or negative result. The final result is positive if the composite score is greater than or equal to 183. The final result is negative if the composite score is less than 183. The actual composite score will not be provided in the device report.

VI. ALTERNATE PRACTICES AND PROCEDURES

Conventional screening for CRC includes both invasive and non-invasive options. Invasive tools include flexible sigmoidoscopy, double contrast barium enema, computed tomography colonography (CTC) and conventional colonoscopy. Colonoscopy is considered to be the most accurate screening tool.

Other than stool DNA-based testing, non-invasive CRC screening tools include guaiac-based fecal occult blood testing (gFOBT) and immunochemical-based fecal occult blood testing (FIT).

Patients who have a positive test on an invasive or non-invasive screening, with the exception of colonoscopy itself, warrant further investigation through conventional colonoscopy to rule out and/or remove the presence of CRC or AA.

Cologuard is a first of a kind device that requires breakthrough technology.

VII. MARKETING HISTORY

Cologuard has not been marketed in the United States or any foreign country.

VIII. POTENTIAL ADVERSE EFFECTS OF DEVICE ON HEALTH

Due to the nature of the noninvasive stool collection process, potential adverse events (AEs) caused by or related to testing with *Cologuard* are unlikely. The primary risk associated with the

[2] *QuARTS*™ is a trademarked brand name that the sponsor uses with the product.

Cologuard test is a false assay result (i.e., a false positive or a false negative result). All positive test results should lead to a colonoscopy. In the instance of a false negative result on *Cologuard*, there is a possibility that a case of CRC or AA could go undetected.

IX. SUMMARY OF PRECLINICAL STUDIES

Nonclinical studies were conducted by Exact Sciences to evaluate the analytical performance characteristics of *Cologuard*. The studies are described below.

A. Algorithm Development and Cut-Off Determination

The cut-offs and the algorithm for the *Cologuard* sDNA-based colorectal cancer screening test were established based on an evaluation of a panel of donor samples that were categorized by colonoscopy. Variable selection for the *Cologuard* model was performed as a stepwise selection with the main variables assessed one at a time based on their respective statistical significance. The total sample size of the dataset for algorithm development included 953 samples, including 794 normal samples, 73 advanced adenomas and 86 cancers.

The derived *Cologuard* algorithm sensitivity and specificity compared to colonoscopy outcome was assessed based on a data set of 1003 samples that included the original 953 samples used to build the algorithm, plus 50 samples tested with the hemoglobin component of *Cologuard*, but collected with a different protein collection tube. The achieved sensitivity of approximately 98% for cancer and approximately 57% for advanced adenoma met the pre-defined acceptance criteria.

B. Sensitivity: Limit of Blank (LoB), Limit of Detection (LoD), Limit of Quantification (LoQ) and Linearity

LoB, LoD, and LoQ studies were performed for both the methylation and mutation component (i.e., molecular assay) and the hemoglobin assay component of *Cologuard* based on guidance from the CLSI Standard: EP17-A (*Protocols for Determination of Limits of Detection and Limits of Quantitation; Approved Guideline*). For molecular assays, such as the *QuARTS* component of *Cologuard*, the signal from the blank wells is absent. Therefore, the LoD and LoQ were established through means independent of a Limit of Blank (LoB) measurement. For *NDRG4*, *BMP3*, *KRAS* 38A, *KRAS* 35T, and *ACTB*, a minimum of 60 replicates per marker near the LoD concentration were tested across 6 samples at the expected LoD concentration within a dilution series in order to use probit analysis to predict LoD. For LoQ, a minimum of 60 replicates per marker near the anticipated LoQ concentration was tested across 6 samples and the lowest concentration with total error less than that of the total error goal of 20% CV on log strands was the determined LoQ.

Linearity and Linear Range studies using concentrations above and below the anticipated linear range were tested in the molecular assay and hemoglobin assay components of *Cologuard*. Linearity studies were performed based on guidance from CLSI Standard: EP6-A (*Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach; Approved Guideline).* All markers were individually assessed at 9 levels spanning 5 logs including concentrations 30% above and below the anticipated upper LoQ and LoD, respectively. The

markers were tested at 8 replicates total per level per marker, 2 replicates per plate, 2 instruments with 1 operator per instrument. The 9 concentrations were (in log strands per reaction): 5.59, 5.48, 5.30, 4.30, 3.30, 2.30, 1.30, 1.00, and 0.85

In summary, analytical sensitivity characteristics for *Cologuard* were observed as follows (**Table 1**): The methylation markers NDRG4, BMP3, and BT-ACT have a LoD at 0.702-0.738 log strands. KRAS was assigned a LoD value to that of KRAS 35T, the KRAS mutant with the highest LoD. The KRAS LoD is 1.058 log strands with a CI range that encompasses the lower cutoff used in the *Cologuard* software. The established LoD meets criteria of the ability to detect one percent mutation or methylation when 3.000 log strands of ACTB are present and the LoD is less than or equal to 1.300 log strands. The molecular assay LoQ is 1.176 log strands per reaction. This exceeds the input requirement that the LoQ must be less than 2.000 log strands. The molecular assays demonstrate good linearity over at least 5 logs and that R^2 is $= \geq 0.996$ for all targets.

Table 1: Analytical Sensitivity Characteristics Summary

Performance Characteristic	**Molecular Assay**	**Hemoglobin Assay**
Limit of Blank	Not Applicable	0.4 ng/mL
Limit of Detection	Methylation Markers: *NDRG4*, *BMP3* and *ACTB* 0.702 to 0.738 log strands Mutation Markers: *KRAS* 1.058 log strands	1.3 ng/mL
Limit of Quantification	LoQ ≤ 1.176 log strands	4.8 ng/mL
Assay linearity	$R^2 = \geq 0.996$ Linear range = 1.1760 to 5.591 log strands	Linear range = 4.8 ng/mL to 500 ng/mL No hook effect observed for concentrations up to 100 µg/mL

C. *Cologuard* Molecular Assay Cross-Reactivity with Wild Type KRAS

The potential for cross-reactivity with wild type *KRAS* was evaluated by testing two levels of *KRAS* wild type DNA in the *Cologuard QuARTS*™ methylation and mutation assays. *KRAS* wild type DNA was assessed at levels of 20,000 copies of wild type *KRAS*, which is greater than the average expected to be seen in normal human stool samples, and 200,000 copies of wild type *KRAS*, 10 times higher. Average strand recovery and standard deviations for *NDRG4*, *BMP3*, *KRAS1*, and *KRAS2* were calculated. The percentage of cross-reactivity of the two levels of wild type *KRAS* for the *QuARTS*™ Mutation and methylation assays was determined, and cross-

reactivity percentages for each of the test levels and no target control ("NTC") were calculated after subtracting the background NTC.

Results from this study indicated that cross-reactivity for wild type *KRAS* at 200,000 copies was 0% for the methylation assay and 0.01% for the mutation assay.

D. *Cologuard QuARTS*™ Partial Methylation Testing

Many genes have elevated methylation in their promoter region in CRC, whereas the same genes have low levels of methylation in normal colon epithelial cells. It is believed that highly methylated promoter region sequences in BMP3 and NDRG4 correlates to CRC and AA and low level methylation correlates to normal tissue with the *QuARTS*™ technology. The DNA oligonucleotides used in the *Cologuard* methylation assay are designed to be a perfect match to fully methylated DNA in *NDGR4* and *BMP3*.

The analytical specificity of the DNA methylation assay component of *Cologuard* was tested against partially methylated *BMP3* and *NDRG4* DNA targets using the *QuARTS*™ assay. The testing utilized synthetic DNA targets that contained all possible permutations of partial methylations in the *QuARTS*™ assay footprint region of *BMP3* and *NDRG4*.

The study results demonstrated that *Cologuard* is specific for highly methylated DNA, specifically highly methylated *NDRG4* and *BMP3*. At least five of eight potential methylation sites for *BMP3* and five of nine potential methylation sites for *NDRG4* have to be methylated for any reactivity in *Cologuard*. With respect to *NDRG4*, the percent cross-reactivity was 2.5%, indicating that the analytical specificity for total methylations in *NDRG4* is 97.5%. With respect to *BMP3*, the percent cross-reactivity was 1.8%, indicating that the analytical specificity for total methylations in *BMP3* is 98.2%, above the 95% specificity outlined in the pre-defined acceptance criteria.

E. *Cologuard* Hemoglobin Assay Cross-Reactivity and Specificity

The ability of the Hemoglobin Assay to detect hemoglobin in specimens heterozygous for Hemoglobin S (HbS) and Hemoglobin C (HbC) was evaluated. Samples used for testing Hb variants consisted of a stool sample background spiked with normal, HbS heterozygous, or HbC heterozygous whole blood. The Hemoglobin Assay detected both HbS and HbC variants, when comparing equivalent volumes of blood from normal and heterozygous variant specimens.

Additionally, cross-reactivity of *Cologuard* Hemoglobin Assay with animal hemoglobin and myoglobin was evaluated. Samples used for testing animal blood cross-reactivity consisted of a stool sample spiked with animal whole blood. Samples used for testing myoglobin cross-reactivity consisted of a stool sample spiked with prepared meat extracts or purified myoglobin. Mean HbC concentrations for all animal hemoglobin and myoglobin samples were less than the limit of detection (LoD) of the assay (1.3 ng/mL) after the mean concentration of the Hb Negative Stool Sample was subtracted, indicating that no cross-reactivity was detected.

F. *Cologuard* Cross-Reactivity with Non-Colorectal Cancers and Diseases

The potential for cross-reactivity with non-colorectal cancers was evaluated by testing 151 specimens from subjects with other cancers, including diseases other than CRC that have a potential association with the GI tract, or inflammatory conditions that could affect the screening population for *Cologuard*. The diseases and cancers tested are listed in **Table 2** below. Samples were tested with both the molecular and hemoglobin assay components of *Cologuard*. Overall *Cologuard* Scores were then generated to assess whether reactivity was found with any of these non-CRC samples.

Cancers in organs connected to the digestive tract (i.e., pancreas and liver) may shed markers that could be detected by *Cologuard.* As such, it is expected that a certain level of reactivity will be observed in cases of these cancers. The results are highlighted in **Table 2** below.

Table 2: Incident Rates and Contribution to *Cologuard* Positivity for Non-CRC Diseases and Cancers

Disease or Cancer*	Number of specimens tested	Incident rate per 10,000**	% Positivity of *Cologuard*	Number positive *Cologuard* calls in 10,000 subjects
Bladder Cancer	17	2.3	17.6%***	0.4
Breast Cancer	14	12.4	0.0%***	0.0
Esophagus Cancer	11	0.5	18.2%***	0.1
Gynecologic Cancer	11	2.0	36.4%	0.7
Hepatic Cancer	6	0.8	50%	0.4
IBD	18	1.0	38.9%	0.4
Lung Cancer	10	6.5	20.0%***	1.3
Lupus	17	0.2-0.8	11.8%***	0.1
Pancreas Cancer	12	1.2	41.6%	0.5
Prostate Cancer	12	15.5	8.3%***	1.3
Rheumatoid Arthritis	15	4.1	26.7%***	1.1
Stomach Cancer	8	0.8	25.0%***	0.2
Total per 10,000 subject		NA	NA	6.5

*Listed value for gynecologic cancer is the sum of ovarian and cervix uteri cancers.
**For cancers, figures were obtained from the National Cancer Institute (http://seer.cancer.gov/statfacts/index.html). For other diseases, figures were obtained from the Centers for Disease Control and Prevention (http://www.cdc.gov).
***Not significantly greater than what would be expected in a "normal" population.

Based on the results of this study, considering the non-CRC diseases and cancers where the percent positivity was slightly higher than would be expected in a normal population, the expected positivity for the tested diseases would result in only a minimal (0.02%) decrease in specificity for *Cologuard* (or two positive calls per 10,000 screening patients tested).

G. Precision and Reproducibility (Lab-to-Lab)

A laboratory-to-laboratory precision and reproducibility study was performed to assess variation of the *Cologuard* assay measurement system based on guidance from the CLSI Standard: EP15-A2 (*User Verification of Performance for Precision and Trueness; Approved Guideline*). As part of the study, a variance component analysis was performed by sample type for the *Cologuard* system to estimate the components of precision for each source of variation (operator, run, site, and replicate) as well as total variation for each individual marker and the overall *Cologuard* Score.

The study was performed at three sites with a minimum of two operators at each site. A total of 22 *Cologuard* runs were performed at each site, 11 per operator. Each run involved 42 samples, including six replicates of each of the following: four stool pool samples (negative, high negative, low positive and high positive) and three control samples (negative, low positive and high positive).

For the molecular assay component of *Cologuard*, the stool sample types were prepared by combining characterized residual stool samples. The samples were characterized as positive or negative for CRC based on colonoscopy results. Subsequently, these residual clinical stool specimens were tested with the *Cologuard* assay to establish the planned DNA content of samples for use in this study. Spiked synthetic DNA was used to create the contrived control samples.

For the hemoglobin assay component of *Cologuard*, the clinical stool pools were prepared by adding fresh whole blood to normal patient stool pools. Specifically, whole blood was spiked into stool samples and diluted to the appropriate concentration. Control samples (including negative, low, and high controls) were provided to each testing site in lyophilized form for reconstitution prior to testing.

Percent agreement between sites was evaluated by generating two-by-two (2 x 2) contingency tables for negative and positive results for all site pairs, calculating the average positive agreement (APA) and average negative agreement (ANA), and calculating the exact two-sided lower 95% confidence interval by the Clopper-Pearson method. The resulting lower confidence limit was then compared to the target agreement rate of 0.95. The lower confidence interval for percent agreement of all site pairs was ≥ 0.95. Inter-site agreement shows minimal variation.

Descriptive statistics were separately calculated for all marker/sample combinations. %CV was calculated only for samples with an expected positive result. Inter-site descriptive statistics are provided below (**Table 3**).

Table 3: Inter-Site Descriptive Statistics for the *Cologuard* Score

Sample	N	Mean	Lower 95% CL for Mean	Upper 95% CL for Mean	Std Dev	Total %CV
Negative Stool Pool	387	9.98	9.65	10.31	3.31	NA
High Negative Stool Pool	394	62.92	60.24	65.61	27.14	NA
Low Positive Stool Pool	393	391.11	383.66	398.36	74.13	18.96
High Positive Stool Pool	394	978.34	977.44	979.24	9.13	0.93
Negative Control	392	6.35	6.26	6.44	0.90	NA
Low Positive Control	393	626.24	621.39	631.09	48.91	7.81
High Positive Control	393	963.38	962.30	964.46	10.89	1.13

Overall the assay was reproducible with inter-site agreement values of the lower confidence interval of >95% and all of the positive *Cologuard* Scores had inter-site CVs of less than 20% (**Table 3**).

An additional multi-operator prospective study was conducted to evaluate the intermediate precision and repeatability of the processes developed for use with the *Cologuard* assay with high negative and low positive stool samples containing levels of DNA or hemoglobin that together, give a *Cologuard* Score near the cut-off of the *Cologuard* assay. The study was performed at one site with two operator teams. A total of 22 *Cologuard* runs were performed during the study, in which each operator team performed 11 complete runs, with each run requiring 2 shifts to complete. Each run involved 12 samples, including six replicates of each of the high negative and low positive stool samples. A single lot of *Cologuard* reagents and controls was used throughout the study.

Percent agreement between operators was evaluated by generating two-by-two (2 x 2) contingency tables for negative and positive results, calculating the weighted average negative agreement (ANA) and average positive agreement (APA), and calculating the exact two-sided lower 95% confidence interval. The lower confidence interval for total agreement of all pairs was >0.95% agreement of all site pairs. Descriptive statistics were calculated and are shown in **Table 4** below.

Table 4: Inter-Operator Descriptive Statistics for the *Cologuard* Score

Sample	N	Median	Mean	Lower 95% CL for Mean	Upper 95% CL for Mean	Std Dev	%CV
High Negative Stool Pool	132	141.9	142.0	139.1	145.0	17.4	12.2
Low Positive Stool Pool	132	238.5	236.5	232.7	240.2	21.8	9.2

Result from the additional testing demonstrated that the assay was reproducible with inter-operator agreement values of the lower confidence interval of >95% and all of the positive *Cologuard* Scores had inter-operator CVs of less than 20%.

H. Lot-to-Lot Reproducibility

Lot-to-Lot reproducibility was evaluated for *Cologuard* based on guidance from the CLSI Standards: EP5-A2 (*Evaluation of Precision Performance of Quantitative Measurement Methods; Approved Guideline*); EP15-A2 (*User Verification of Performance for Precision and Trueness; Approved Guideline*); EP12-A2 (*User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline*); and I/LA28-A2 (*Quality Assurance for Design Control and Implementation of Immunohistochemistry Assays; Approved Guideline*).

Lot-to-Lot reproducibility was assessed by testing a sample panel comprised of seven samples containing various levels of DNA and hemoglobin, using three lots of *Cologuard* reagents and controls.

For the molecular assay component of *Cologuard*, the stool sample types were prepared by combining characterized residual stool samples available to Exact Sciences. The samples were characterized as positive or negative for CRC based on colonoscopy results. Subsequently, these residual clinical stool specimens were tested with the *Cologuard* assay to establish the planned DNA content of samples for use in this study. Spiked synthetic DNA was used to create the contrived control samples.

For each sample in the panel, there were 24 sample results per lot and 72 sample results for the entire study. Across the seven samples in the panel, there were 168 results per lot, and 504 results for the entire study.

The mean, SD, %CV, N, minimum value and maximum value were calculated for each marker or each lot and test sample. Additionally, *Cologuard* Scores were determined. Percent positive results for the *Cologuard* Score were analyzed across lots and for lot to lot. Variance component analyses were also conducted.

Descriptive statistics were calculated for all marker/sample combinations, including median, mean, mean upper and lower 95% confidence intervals, standard deviation, and coefficient of variation values (**Table 5**). %CV was calculated only for controls with expected result of positive. Descriptive statistics were calculated both within and across lots. Descriptive statistics for this study are shown below. The *Cologuard* Score %CV values for positive samples were within the pre-specified acceptance criteria, ranging between 0% and 16.8%.

Table 5: Descriptive Statistics for Lot-to-Lot *Cologuard* Score

Sample Name	N	Median	Mean	Lower 95% CL for Mean	Upper 95% CL for Mean	Std Dev	CV
Negative Stool Pool	72	9.47	11.39	10.19	12.58	5.07	NA
High Negative Stool Pool	72	64.46	57.74	51.12	64.36	28.18	NA
Low Positive Stool Pool	71	380.75	373.93	359.03	388.84	62.98	16.84
High Positive Stool Pool	71	973.92	972.88	970.36	975.40	10.64	1.09
Negative Control	70	6.33	6.40	6.21	6.59	0.79	NA
Low Positive Control	71	584.09	579.52	570.09	588.95	39.85	6.88
High Positive Control	71	1000	1000	1000	1000	0	0

Percent agreement between lots was evaluated by generating 2 x 2 tables for negative and positive results for all lot pairs, calculating the average positive agreement (APA) and average negative agreement (ANA). Testing of samples with various levels of hemoglobin and DNA markers demonstrated a percent agreement for positive and negative samples across multiple lots between 98.6% and 100%, with a lower confidence limit above 95%.

The study demonstrated that *Cologuard* results are reproducible across multiple reagent lots.

I. Robustness

The *Cologuard* performance was assessed in response to defined variable factors at specific steps in the test procedure, using both the molecular assay and hemoglobin assay components of *Cologuard*. The processing steps analyzed in this study are the steps at which operator variability or error are most likely to occur. Three total instrument and operator sets were used for the study.

Cologuard Molecular Assay Robustness

The Molecular Assay portion of the *Cologuard* assay involves 1) sample preparation, 2) DNA capture, 3) DNA preparation, and 4) the QuARTS assay. Each of these segments is further divided into individual steps that in the final workflow are either semi- or fully-automated. These steps were categorized into the following categories: 1) Processing step optimized in development or separate study; 2) Process step is controlled by instrumentation/software or is not variable unless Instructions for Use (IFU) is not followed (i.e., operator error); or 3) Process step in which operator can introduce variability. The manual steps that are in category 1 were optimized in separate studies. Category 2 items have mitigations built into the process, primarily

in the instrumentation/software or in the IFU. All category 3 items were tested in the robustness study and described here.

Results when these various factors were introduced into the processing steps were compared to the expected results for a positive stool sample, a control sample with high levels of mutation and methylation markers, and a control sample with moderate levels of mutation and methylation markers. Fourteen replicates of each sample type were used. Analysis of these samples assumed a hemoglobin value of zero, when calculating overall *Cologuard* score. Factors tested included the following:

- Factors related to DNA capture, including wait times between processing steps, amount of reagents added, and duration of storage at the appropriate temperatures;
- Factors related to the amount of time various instruments are paused during the automated DNA preparation and *QuARTS* assay steps of the *Cologuard* process; and
- Factors related to the amount of time between plate assembly and processing during the *QuARTS* assay step.

The results for the molecular assay component of *Cologuard* showed that time between plate assembly and processing during the *QuARTS*™ assay step and the number of days the captured DNA was stored at the appropriate temperatures could have a detectable effect on assay response. Testing demonstrated that the prepared *QuARTS*™ plate should be processed within 30 minutes and captured DNA could be tested for up to four days.

Cologuard Hemoglobin Assay Robustness

Results when these factors were introduced into the processing steps were compared to the expected results for a stool sample with a known level of endogenous hemoglobin and a high and low control sample with high and low levels of hemoglobin. The study tested 16 replicates of each sample type. Analysis of these results involved comparing the resulting hemoglobin concentration with the expected hemoglobin concentration. Factors tested include the following:

- Time between steps during plate preparation;
- Incubation times for antibodies and substrates; and
- Time between steps during plate reading phase.

Results for the hemoglobin assay component of *Cologuard* showed that substrate incubation time had a detectable effect on assay performance. Testing demonstrated that a substrate incubation time of 15 ± 1.5 minutes would result in acceptable assay performance.

J. Interfering Substances

Cologuard Molecular Assay Interference Testing

Interference with the molecular assay component of *Cologuard* was evaluated using 55 common substances that potentially could be present in stool materials. Testing was performed using 16 replicates of positive and negative stool homogenate samples, with and without interfering

substances. All samples were processed through the entire molecular test component of *Cologuard*, evaluating the methylation and mutation markers for *Cologuard* score calculations to assess whether interference was observed.

Cologuard molecular assay was evaluated with potential interfering substances in the following categories:

- Common lotions, creams, and feminine over-the-counter products;
- Stool softeners, anti-diarrhea, and laxative products;
- Anti-acids and upset stomach relief products;
- Animal genomic DNA of commonly edible animals (both high and low levels);
- Urine and alcohol;
- A mixture of common vegetables and fruits; and
- Fecal Fats (fatty acids and cholesterol).

For samples known to be positive, no differences were observed in the overall *Cologuard* results for spiked samples versus unspiked samples. Comparisons of the mean *Cologuard* score for each interferent group with the mean score for the unspiked control revealed no statistically significant differences. No interference with the molecular assay component of *Cologuard* was observed for any of the tested substances.

Cologuard Hemoglobin Assay Interference Testing

Interference with the hemoglobin assay component of *Cologuard* was evaluated using 46 common substances that potentially could be present in stool materials. Testing was performed using 16 replicates of positive and negative stool homogenate samples, with and without interfering substances. All samples were processed through the hemoglobin assay component of *Cologuard.* Samples were evaluated for inhibition or enhancement of hemoglobin concentrations in spiked and un-spiked samples to assess whether interference was observed.

Cologuard hemoglobin assay was evaluated with potential interfering substances in the following categories:

- Common lotions, creams, and feminine over-the-counter products;
- Urine;
- Stool softeners, anti-diarrhea, and laxative products;
- Anti-acids and upset stomach relief products;
- Antibiotics, anti-inflammatories, anti-fungal drugs, pain relievers, and decongestants;
- A mixture of common vegetables and fruits;
- Fats and lipids;
- Alcohol; and
- Iron sulfate (as found in oral supplements);
- Vitamin C; and
- DNA Stabilization Buffer (preservative solution provided in the *Cologuard* Collection Kit for the whole stool sample used in the molecular assay).

A comparison of the mean hemoglobin concentration results indicated there were no statistical differences between the mean hemoglobin concentrations in test and control samples in both the 'positive' and 'normal' stool pools. None of the substances tested interfered with the *Cologuard* hemoglobin assay.

K. Carry-over and Cross-contamination *Cologuard* Testing

Carry-over Evaluation

Sequential runs of high positive and negative samples were used to evaluate carry-over contamination for each assay component of *Cologuard*. Testing of the molecular assay and hemoglobin assay components was conducted in two separate studies.

For the molecular assay (methylation/mutation assay), the testing involved two consecutive runs of high positive DNA samples, composed of 10x high level run controls diluted in Tris, EDTA and non-human DNA, followed by a run of negative samples composed of Tris, EDTA and non-human DNA. A total of 43 high positive samples and 3 run controls were used in each high positive run. A total of 43 negative samples and 3 run controls were used for the negative run.

For the hemoglobin assay, the testing involved two consecutive runs of high positive hemoglobin samples, composed of 100,000 ng/mL hemoglobin, followed by a run of negative samples composed solely of the protein preservative solution from the hemoglobin sample collection tube. The high positive samples consisted of a hemoglobin level that is much higher than the quantitative range of the assay, which identifies all samples >500 ng/mL as greater than the maximum range of the assay. For the high positive runs, a total of 86 high positive hemoglobin samples were used. For the negative run, 86 negative samples were used. In each run, the signal obtained on the controls was utilized to ensure the validity of the run.

Results from the molecular assay and hemoglobin assay carry-over analyses demonstrated that there was no carry-over in the *Cologuard* assay.

Cross-contamination Evaluation

Cross-contamination testing of *Cologuard* was based on a checkerboard study design, alternating high positive and negative samples, to evaluate the potential for contamination from the positive to the negative samples within a run. Testing of the molecular assay and hemoglobin assay components was conducted in two separate studies.

For the molecular assay, 22 high positive samples, 21 negative samples, and three run control samples were used. The high positive samples for this study were composed of 10x high level run controls diluted in Tris, EDTA and non-human DNA, and the negative samples were composed of Tris, EDTA and non-human DNA. One run was performed and samples were processed using the *Cologuard* molecular process from the semi-automated front end sample processing through the automated processing.

For the hemoglobin assay, a total of 43 high hemoglobin and 43 negative hemoglobin samples were used. As in the carry-over study, the high positive samples contained 100,000 ng/mL hemoglobin, while the negative samples consisted solely of the protein preservative solution

from the hemoglobin sample collection tube. Three runs were performed and samples were processed using the *Cologuard* hemoglobin process.

Results from the cross-contamination analysis for the molecular assay demonstrated that the molecular assay component of *Cologuard* and the associated instruments needed to run the assay performed as intended and met the study acceptance criteria. Specifically, one well experienced some cross-contamination (52 strands of ACTB), however, this was within the pre-specified acceptance criteria, which dictated that no more than three wells could exhibit 10-100 strands of ACTB and no single well could exhibit more than 100 strands.

The high hemoglobin samples utilized in this study contained hemoglobin levels that are approximately 50 times higher than the median positive hemoglobin values observed in colorectal cancer subjects (Levi et. al, 2007). The high hemoglobin concentrations tested in this study are much higher than would be expected in use of *Cologuard*.

Testing of the Hemoglobin Assay cross-contamination showed a low level of contamination (~0.01%). Signal was observed in 4 out of 43 negative samples with an average detectable hemoglobin level of 11 ng/mL (0.011%). This calculates to a 0.011% contamination level in those four samples. As the hemoglobin assay involves several manual steps (e.g., manual washing and reagent addition), repeat testing was conducted, in which no cross contamination was observed.

Under normal use conditions, low level contamination observed in this study would be negligible. However, this study provides evidence that cross contamination is possible due to the manual steps in the assay processing.

L. Stability Studies

In-Use Stability: Molecular Assay Stability Under Standard Operating Conditions

The stability of reagents used in the molecular assay component of *Cologuard* was evaluated following guidance from CLSI standard: EP25-A (*Evaluation of Stability of In Vitro Diagnostic Reagents; Approved Guideline*). The purpose of this testing was to determine reagent stability after opening the containers and using them under potential user operating conditions. All reagents required for the molecular assay were tested.

Samples were processed with the molecular assay component of *Cologuard*, using these reagents, to determine the in-use stability of the reagents and the effect of the various factors on *Cologuard* results. The samples used in the in-use stability study for the various *Cologuard* reagent groups included DNA calibrators; High Positive and Low Positive control samples consisting of synthetic targets in stool collection buffer; a Negative DNA control sample; DNA positive and negative run controls; and a positive stool sample.

The study demonstrated that *Cologuard* reagents are stable when opened or stored for variable times before use under standard operating conditions. Specifically:

- Multiple-use reagents stored at room temperature are stable for up to six weeks from the open date.

- Capture Beads that have been pre-washed and stored at 2-8°C are stable for up to 13 days.
- Pre-washed Capture Beads are stable for up to six hours at room temperature prior to use.

Single-use reagents that are used on the automated system are stable on the Hamilton Microlab® STARlet deck for up to 4 hours prior to the start of the run.

Freeze-Thaw Stability

A freeze-thaw stability study was conducted to evaluate the stability of the *QuARTS*™ assay reagents when subjected to repeated freeze/thaw events. The *QuARTS*™ assay reagents tested included only those assay components normally stored frozen (-25 to -15°C):

1) Oligo Mix A, Methylation;
2) Oligo Mix B, Mutation;
3) Enzyme Mix;
4) DNA Calibrator 1 High Methylation;
6) DNA Calibrator 2 Low Methylation;
7) DNA Calibrator 3 High Mutation; and
8) DNA Calibrator 4, Low Mutation.

Materials from one lot of each assay component were subjected to 0, 2, 4, and 6 freeze-thaw cycles. Each component was then tested in the *Cologuard* molecular assay component using the *Cologuard* DNA Controls (i.e., DNA Control 1, High Positive and DNA Control 2, Low Positive), which did not undergo freeze-thaw cycling. The study tested 16 replicates for each component and each freeze-thaw cycle. Calibrators used during testing to assess assay validity and to generate curves for sample concentration assessment were not subjected to freeze-thaw cycling. Log strands for each marker were compared to those for samples where the reagents did not undergo freeze thaw cycling.

All log strand results for all samples were statistically equivalent to those that did not undergo freeze thaw cycling, thereby demonstrating that the *Cologuard QuARTS*™ assay reagents are stable for six freeze thaw events.

Real-Time Stability

Real-time stability testing of *Cologuard* was conducted by evaluating the functional performance of three reagent lots over a period of 41 weeks. Each lot was comprised of unique batches of reagents, which were tested at various time points over 41 weeks.

Samples that were used to evaluate hemoglobin assay reagent stability consisted of negative stool matrix spiked with whole blood to create samples with a low and high hemoglobin concentration. Samples for evaluation of molecular assay reagent stability consisted of negative stool matrix spiked with oligonucleotides that contain the marker sequences. Oligonucleotides for *NDRG4, BMP3, BTACT, KRAS1, KRAS2*, and *ACT* were spiked into the negative stool samples to create samples with a low and high level of sDNA samples. At each time point, seven replicates of samples and controls were tested.

The results of the real time stability studies demonstrated that overall the components of the *Cologuard* assay gave similar results through the 41 week study. These data supported the 6 month shelf life currently assigned to the *Cologuard* assay reagents.

Collection Kit Testing

The following studies were conducted for the collection kit:

- Leak testing to ensure that the collection kit can be used and shipped in accordance with the directions, without sample leakage.
- Stability testing to demonstrate that the collection kit can preserve samples of varying levels of concentration for various time periods.
- Shipping stress testing to demonstrate that the collection kit can withstand the typical stresses of shipment to and from the user and the laboratory.
- Biocompatibility testing and additional shelf life and stability testing.
- Usability and human factors issues study demonstrating that patients can successfully utilize the collection kit in an at-home environment

X. SUMMARY OF PRIMARY CLINICAL STUDY

The pivotal study ("Multi-Target Colorectal Cancer Screening Test for the Detection of Colorectal Advanced Adenomatous Polyps and Cancer: DeeP-C Study") was conducted to generate data to support the safety and effectiveness of *Cologuard* as a screening test for the detection of markers associated with the presence of colorectal cancer (CRC) and advanced adenoma (AA). To evaluate the performance of *Cologuard*, the *Cologuard* test result (negative or positive) was compared with the histopathological result from optical colonoscopic examination and histopathological diagnosis of all significant lesions discovered during the colonoscopy and either biopsied or removed. Based on this comparison, *Cologuard* sensitivity (true positive fraction) was 92.3% (60/65) for subjects with a histopathological diagnosis of CRC and 42.4% (322/760) for subjects with a diagnosis of AA. For subjects without a diagnosis of CRC or AA, *Cologuard* specificity (true negative fraction) was 86.6% (7967/9198). Furthermore, among subjects having a valid *Cologuard* test result and also a valid test result from a commercially available FIT (OC FIT-CHEK, Polymedco, Inc.) ("FIT"), both sensitivity for CRC and sensitivity for AA were higher for *Cologuard* (92.3%, 42.4%) than for FIT (73.8%, 23.8%), and both differences (18.5%, 18.6%) were significantly different from zero (p=0.002, 0.001). However, for subjects without CRC or AA, specificity was lower for *Cologuard* (86.6%) than for FIT (94.9%), and the difference (-8.3%) was significantly different from zero ($p < 0.0001$).

An overview of the study design and results is provided below.

A. Study Design

The *Cologuard* pivotal study was a prospective, cross-sectional, multi-center study (DeeP-C study) that began enrollment of study participants on June 30, 2011. A total of 12,776 patients were enrolled from 90 sites in the U.S. and Canada, including both colonoscopy centers and primary care sites, with study participation concluding on February 4, 2013. Subjects were provided with a collection kit, which they used to collect stool samples for *Cologuard* and FIT testing. Subjects subsequently underwent colonoscopy within 90 days of study enrollment.

The stool samples for analysis with *Cologuard* were sent to a central biorepository for batch testing at one of three laboratories while the stool samples for the FIT were sent to a single laboratory for testing. Samples tested with *Cologuard* were assayed by laboratory technicians blinded to the results of colonoscopy and the FIT results. Results from *Cologuard* and the FIT test were compared to the results of an optical colonoscopic examination, and histopathological diagnosis of all significant lesions discovered during the colonoscopy and either biopsied or removed.

Colonoscopy findings were recorded per site specific standard of practice. Subjects with no findings were categorized as negative by colonoscopy. Histopathological results from biopsied tissue or excised lesions were categorized based on the most clinically significant lesion present (i.e., the index lesion) by a central pathologist according to the pre-specified standards outlined in **Table 6**.

Table 6: Histopathological category definitions

Category	Findings
1	CRC, all stages (I-IV)
2	Advance adenoma, including the following subcategories: 2.1 – Adenoma with carcinoma *in situ*/high grade dysplasia, any size 2.2 – Adenoma, villous growth pattern (≥25%), any size 2.3 – Adenoma ≥ 1.0 cm in size, or 2.4 – Serrated lesion, ≥ 1.0 cm in size
3	1 or 2 adenoma (s), >5 mm in size, or < 10 mm size, non-advanced
4	≥ 3 adenomas, <10mm, non-advanced
5	1 or 2 adenoma(s), ≤5 mm in size, non-advanced
6	Negative – No neoplastic findings 6.1 – negative upon histopathological review 6.2 – no findings on colonoscopy, no histopathological review

B. Inclusion and Exclusion Criteria

Subjects eligible for enrollment in the study were of both genders between the ages of 50 and 84 years (inclusive), who were at average risk for development of colorectal cancer and asymptomatic for gastrointestinal symptoms warranting diagnostic colonoscopy. In addition, subject enrollment was age-weighted toward a slightly older population to increase the point prevalence of colorectal cancer in this study. An effort was made to enroll the majority of subjects of age 65-84; 64% of subjects in the actual study population were of age 65-84.

C. Clinical Performance Measures

The performance of *Cologuard* was evaluated based on comparison of the test result with the histopathological category (**Table 6** above). Two co-primary performance measures were pre-specified: *Cologuard* sensitivity for subjects diagnosed with CRC (histopathological category 1), and *Cologuard* specificity for subjects without a diagnosis of CRC or AA (categories 3-6). For subjects with CRC, *Cologuard* sensitivity is the fraction of CRC subjects called positive by the *Cologuard* test (true positive fraction). Defining advanced neoplasia (AN) as CRC or AA, *Cologuard* specificity for AN is the fraction of non-AN subjects called negative by the *Cologuard* test (true negative fraction). For the study to be successful, the co-primary analysis required that the lower bound of the 95% one-sided confidence interval was greater than 65% for *Cologuard* sensitivity for CRC and greater than 85% for *Cologuard* specificity for AN. It should be noted that sensitivity for CRC and specificity for AN are not complimentary in that advanced adenoma (AA, histopathological category 2) is left out of the definition of both measures.

Two secondary performance evaluations were pre-specified: *Cologuard* was evaluated for non-inferiority to FIT in sensitivity for CRC and for superiority to FIT in sensitivity for AA (fraction of AA subjects testing positive). Per the pre-specified protocol, *Cologuard* would be declared non-inferior to FIT in sensitivity for CRC if the one-sided 95% confidence interval lower bound on the *Cologuard* – FIT difference exceeded -5%. If *Cologuard* were to be declared non-inferior to FIT in CRC sensitivity, then evaluation for superiority to FIT in CRC sensitivity was permitted and declared if the difference was positive and its one-sided p-value (based on exact McNemar test) was less than 0.025. Likewise, per protocol, *Cologuard* would be declared superior to FIT in AA sensitivity if the *Cologuard* – FIT difference was positive and the one-sided McNemar p-value was less than 0.025.

D. Accountability of PMA Cohort

The study enrolled a total of 12,766 subjects at 90 sites, including both primary care point-of-referral (POR) sites and colonoscopy centers. A total of 2,753 subjects were excluded from the primary analysis population due to unusable data (e.g., no colonoscopy). A total of 10,023 subjects were included in the primary analysis population. This population included 65 subjects with CRC. Analysis was conducted to rule out bias associated with the subjects excluded from the analysis population.

E. Study Population Demographics and Baseline Parameters

The baseline demographic characteristics for the Primary Effectiveness Population are presented in **Table 7** below. As shown in the table, the average age of subjects was 64.2 years old, and there was a slightly higher percentage of female subjects (5,378/10,023, 53.7%) as compared with male subjects (4,645/10,023, 46.3%). The majority of subjects were White (8,422/10,017, 84.1%), although 10.7% of the population were Black or African American subjects (1,071/10,017). Nearly 10% of subjects were Hispanic or Latino (991/10,019, 9.9%). Average BMI was 28.83 and the majority of subjects never smoked (5,531 /10,019, 55.2%). It should be noted that two 49-year-old subjects and one 44-year-old subject were included in the study, which is inconsistent with the intended use population. Each of these subjects was a true negative and their inclusion did not notably impact data analyses.

Subjects that were enrolled at POR sites were similar to those enrolled at non-POR sites and to the population as a whole.

Table7: Baseline Demographics – Primary Effectiveness Subjects

Parameter Statistic	All Enrolled (N=10023)	Specificity Subset (Cat. 2-6) (N=9958)	Specificity Subset (Cat. 3-6) (N=9198)	CRC Subset (Cat. 1) (N=65)	AA Subset (Cat. 2) (N=760)	FIT Secondary Effectiveness (N=65)
Age (years) at Screening						
n	10023	9958	9198	65	760	65
Mean (SD)	64.2 (8.42)	64.1 (8.41)	64.0 (8.44)	70.2 (7.92)	65.4 (7.93)	70.2 (7.92)
Median	66	66	66	70	66	70
Min, Max	44, 84	44, 84	44, 84	50, 84	50, 84	50, 84
Gender, n (%)						
Male	4645 (46.3)	4611 (46.3)	4161 (45.2)	34 (52.3)	450 (59.2)	34 (52.3)
Female	5378 (53.7)	5347 (53.7)	5037 (54.8)	31 (47.7)	310 (40.8)	31 (47.7)
Race, n (%)						
White	8422 (84.1)	8367 (84.1)	7726 (84.0)	55 (84.6)	641 (84.5)	55 (84.6)
Black or African American	1071 (10.7)	1063 (10.7)	978 (10.6)	8 (12.3)	85 (11.2)	8 (12.3)
Asian	259 (2.6)	258 (2.6)	245 (2.7)	1 (1.5)	13 (1.7)	1 (1.5)
American Indian or Alaska Native	36 (0.4)	36 (0.4)	32 (0.3)	0 (0.0)	4 (0.5)	0 (0.0)
Native Hawaiian or Other Pacific Islander	23 (0.2)	23 (0.2)	23 (0.3)	0 (0.0)	0 (0.0)	0 (0.0)
Other	206 (2.1)	205 (2.1)	189 (2.1)	1 (1.5)	16 (2.1)	1 (1.5)
Missing	6	6	5	0	1	0
Ethnicity, n (%)						
Hispanic or Latino	991 (9.9)	982 (9.9)	923 (10.0)	9 (13.8)	59 (7.8)	9 (13.8)
Not Hispanic or Latino	9028 (90.1)	8972 (90.1)	8272 (90.0)	56 (86.2)	700 (92.2)	56 (86.2)
Missing	4	4	3	0	1	0
BMI (kg/m2) at Baseline						
n	10015	9950	9190	65	760	65
Mean (SD)	28.83 (5.836)	28.84 (5.841)	28.77 (5.817)	27.55 (4.861)	29.67 (6.068)	27.55 (4.861)
Median	28.0	28.0	27.9	26.8	29.0	26.8
Min, Max	13.3, 68.2	13.3, 68.2	13.3, 68.2	19.3, 42.4	16.3, 59.9	19.3, 42.4
Smoking History, n (%)						
Never Smoked	5531 (55.2)	5498 (55.2)	5157 (56.1)	33 (50.8)	341 (44.9)	33 (50.8)

Parameter Statistic	All Enrolled (N=10023)	Specificity Subset (Cat. 2-6) (N=9958)	Specificity Subset (Cat. 3-6) (N=9198)	CRC Subset (Cat. 1) (N=65)	AA Subset (Cat. 2) (N=760)	FIT Secondary Effectiveness (N=65)
Former Smoker	3589 (35.8)	3564 (35.8)	3279 (35.6)	25 (38.5)	285 (37.5)	25 (38.5)
Current Smoker	903 (9.0)	896 (9.0)	762 (8.3)	7 (10.8)	134 (17.6)	7 (10.8)
If Former or Current Smoker, Daily Use, n (%)						
<1/2 Pack Per Day	2162 (48.3)	2154 (48.4)	1970 (48.9)	8 (25.0)	184 (44.0)	8 (25.0)
1 Pack Per Day	1585 (35.4)	1569 (35.3)	1418 (35.2)	16 (50.0)	151 (36.1)	16 (50.0)
>1 Pack Per Day	732 (16.3)	724 (16.3)	641 (15.9)	8 (25.0)	83 (19.9)	8 (25.0)
Missing	13	13	12	0	1	0
If Former or Current Smoker, # Years Smoking						
n	4480	4448	4029	32	419	32
Mean (SD)	21.82 (14.733)	21.77 (14.732)	21.13 (14.450)	28.47 (13.488)	27.93 (15.959)	28.47 (13.488)
Median	20.0	20.0	20.0	29.0	30.0	29.0
Min, Max	0.0, 70.0	0.0, 70.0	0.0, 70.0	1.0, 60.0	1.0, 65.0	1.0, 60.0

F. Safety and Effectiveness Results

1. Safety Results

Adverse effects that occurred in the PMA clinical study:

With respect to safety, due to the design of the study and the nature of the stool collection process, AEs caused by or related to the stool collection procedure were not expected. As a result, events associated with potential errors in use of the collection kit and any product complaints were captured in the safety analyses. There were no cases in which the study investigator believed the product contributed to a serious adverse event, and only 4 adverse events were reported. Events included a broken fingernail, cut finger, leg pain related to a fall during stool collection and sprained hand. None of the AEs experienced in the study were deemed "serious", all were categorized as "mild" events. None of the events led to the subject discontinuing the study.

Additionally, one subject died of unrelated causes prior to undergoing colonoscopy. The subject met all eligibility criteria and successfully collected a stool sample, but did not present for the subsequent colonoscopy.

2. Effectiveness Results

Primary Effectiveness Evaluations (Sensitivity for CRC/Specificity for AN)

The primary effectiveness population consisted of 10,023 subjects with a valid histopathological diagnosis and a valid *Cologuard* result. The basic data table for primary effectiveness evaluation is provided (**Table 8**).

Table 8: Distribution of *Cologuard* Result by Histological Category (%), n = 10,023

***Cologuard* Result**	**CRC (Category 1)**	**AA (Category 2)**	**Categories 3-6**
Negative	5 (7.7)	438 (57.6)	7967 (86.6)
Positive	60 (92.3)	322 (42.4)	1231 (13.4)

The primary objectives of the DeeP-C study – demonstration of greater than 65% *Cologuard* sensitivity for CRC and greater than 85% *Cologuard* specificity for AN – were successfully met. Specifically, *Cologuard* sensitivity for CRC was 92.3% (60/65) with a one-sided 95% confidence interval lower bound of 84.5% (**Table 9**). *Cologuard* specificity for AN was 86.6%, with a one-sided 95% confidence interval lower bound of 86.0% (**Table 10**). Further, the two-sided 95% confidence interval was 83.0-97.5% for *Cologuard* CRC sensitivity and 85.9-87.3% for *Cologuard* AN specificity.

Table 9: Sensitivity for CRC – Primary Effectiveness Subjects with Valid *Cologuard* Positive Result (N=65)

Case Category	n/N (%)
1: CRC Stages 1-4	60/65 (92.3%)
Sensitivity Based on Category 1: Primary (one-sided 95% CI lower bound)	92.3% (>84.5%)
Sensitivity Based on Category 1: Supportive (one-sided 97.5% CI lower bound)	92.3% (>83.0%)

[1] Percentages based on valid test results within a category.
[2] Lower bounds calculated using an exact one-sided binomial test.

Table 10: Specificity for AN – Primary Effectiveness Subjects with Valid *Cologuard* Negative Result (N=9198)

Case Category	n/N (%)
3: 1-2 Adenomas 5 – <10 mm	607/749 (81.0%)
4: ≥3 Adenomas <10 mm, Non-advanced	302/419 (72.1%)
5: 1-2 Adenomas <5 mm, Non-advanced	1496/1735 (86.2%)
6.1: Negative upon histopathological review	1543/1821 (84.7%)
6.2: No findings on colonoscopy, no histopathological review	4019/4474 (89.8%)
Specificity Based on Categories 3-6: Primary (one-sided 95% lower bound)	86.6% (>86.0%)
Specificity Based on Categories 3-6: Supportive (one-sided 97. 5% lower bound)	86.6% (>85.9%)

[1] Percentages based on valid test results within a category.
[2] Lower bounds calculates using an exact one-sided binomial test.
[3] As noted above, one 44-year-old and two 49-year-old true negative subjects were included in the analysis population, although they would not be included in the intended user population.

Secondary Effectiveness Evaluations

The secondary effectiveness population consisted of 9,989 subjects with a valid histopathological diagnosis, a valid *Cologuard* result, and a valid FIT result. The basic data table for secondary effectiveness evaluation is provided (**Table 11**).

Table 11: Distribution of *Cologuard* and FIT Results by Histological Category, n = 9,989 CRC (Category 1)

	FIT	
Cologuard	**Negative**	**Positive**
Negative	4	1
Positive	13	47

AA (Category 2)

	FIT	
Cologuard	Negative	Positive
Negative	407	29
Positive	170	151

Categories 3-6

	FIT	
Cologuard	Negative	Positive
Negative	7787	149
Positive	908	323

Cologuard was compared with FIT for non-inferiority in sensitivity for CRC with respect to margin 5% and for superiority in sensitivity for advanced adenoma (AA). Secondary performance goals comparing *Cologuard* with FIT were evaluated in subjects having valid results from both tests.

Sensitivity for CRC was greater for *Cologuard* (92.3%, 60/65) than for FIT (73.8%, 48/65) (**Table 12 and Figure 1**), for a difference of 18.5%. *Cologuard* identified 13 CRCs that were missed by FIT. FIT identified one CRC that was missed by *Cologuard*. The one-sided 95% confidence interval lower bound on the *Cologuard* – FIT difference was 8.0%, which exceeds -5%, indicating that *Cologuard* is non-inferior to FIT in sensitivity for CRC with respect to the pre-defined non-inferiority margin 5%. Because *Cologuard* was declared non-inferior to FIT in sensitivity for CRC, it is statistically justifiable and was permissible per protocol to evaluate it for superiority to FIT as well. Sensitivity for CRC was significantly greater for *Cologuard* than for FIT (two-sided McNemar exact p value 0.0018), indicating that *Cologuard* is superior to FIT in sensitivity for CRC. Finally, for the *Cologuard* – FIT difference of 18.5% in CRC sensitivity, the two-sided 95% exact confidence interval was 7.2-30.4% (**Table 13**).

Table 12: Overall Sensitivity: CRC Subset (Category 1) - Secondary Effectiveness Subjects with Valid Results from Both *Cologuard* and FIT Tests (N=65)

	Cologuard	**FIT**
1: CRC Stages 1-4 **(n/N (%))**	60/65 (92.3%)	48/65 (73.8%)

	Cologuard	FIT
Sensitivity Based on Category 1: Primary (one-sided 95% lower bound)	92.3% (>84.5%)	73.8% (>63.4%)
Sensitivity Based on Categories 1: Supportive (one-sided 97. 5% lower bound)	92.3% (>83.0%)	73.8% (>61.5%)

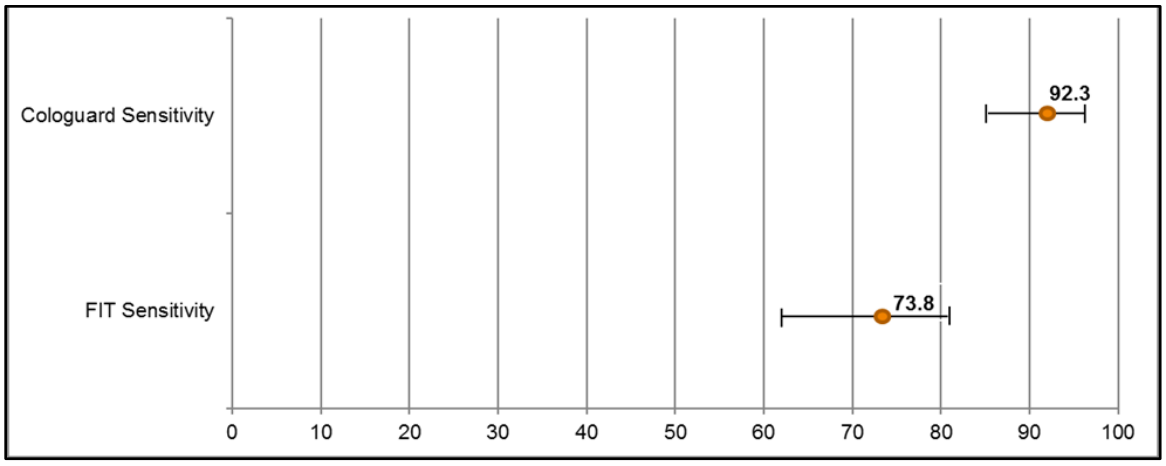
[1] Percentages based on valid test results within a category.
[2] Lower bounds calculated using an exact one-sided binomial test.

Table 13: Sensitivity Non-Inferiority and Superiority Test – CRC Subset (Category 1)

	FIT Outcome			McNemar test p-value
***Cologuard* Outcome**	**Negative**	**Positive**	**Totals**	
Negative, n (%)	4 (80.0)	1 (20.0)	5	0.0018
Positive, n (%)	13 (21.7)	47 (78.3)	60	
Totals	17	48	65	

[1] p-value is from a McNemar paired comparison test of the discordant pairs.

Figure 1: CRC Sensitivity



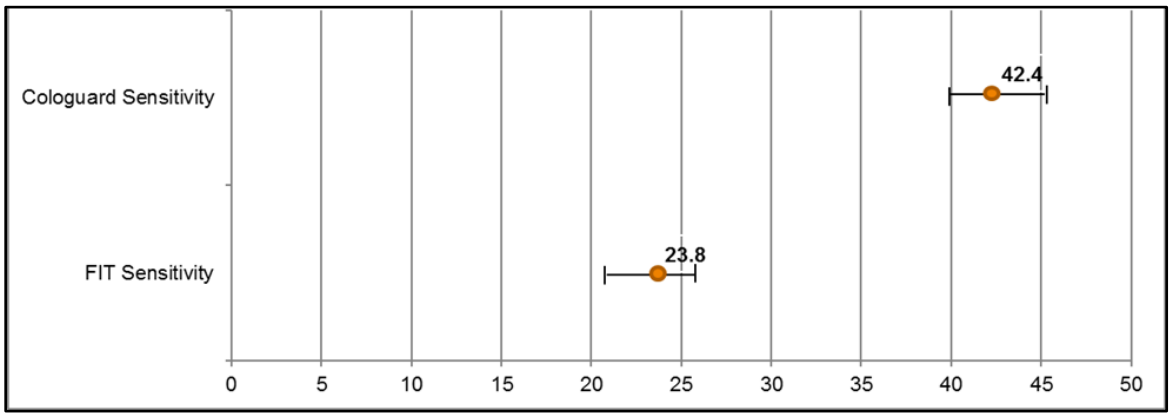
For histopathologically-confirmed AAs, sensitivity was greater for *Cologuard* (42.4%, 321/757) than for FIT (23.8%, 180/757) (**Table 14**). The difference of 18.6% was significantly different from zero (two-sided exact McNemar p value < 0.001), indicating that *Cologuard* is superior to FIT in sensitivity for AA. *Cologuard* identified 170 AA cases that were not called positive by the FIT test, while FIT identified 29 AA cases that were not called positive by *Cologuard* test. Finally, the two-sided 95% CI for the *Cologuard* – FIT difference of 18.6% was 15.3-22.1% (**Figure 2**).

Table 14: Sensitivity Superiority Test – AA Subset (Category 2)

Cologuard Outcome	FIT Outcome			McNemar test p-value
	Negative	Positive	Totals	
Negative, n (%)	407 (93.3)	29 (6.7)	436	<0.0001
Positive, n (%)	170 (53.0)	151 (47.0)	321	
Totals	577	180	757	

[1] p-value is from a McNemar paired comparison test of the discordant pairs.

Figure 2: AA Sensitivity



Additional Effectiveness Analyses

The combined sensitivity for CRC and AA subjects was also analyzed *post hoc*. The sensitivity for CRC/AA was 46.3% (381/822) for *Cologuard* and 27.7% (228/822) for FIT, for a difference of 18.6% (**Table 15**).

Table 15: Sensitivity for Advanced Neoplasia (CRC + AA)

Category	*Cologuard* (N=822)	FIT (N=822)
Category 1 Only	92.3% (60/65)	73.8% (48/65)
Categories 1-2	46.4% (381/822)	27.7% (228/822)

For subjects without CRC or AA, the specificity (fraction of subjects called negative) was smaller for *Cologuard* (86.6%, 7936/9167) than for FIT (94.9%, 8695/9167) (**Table 16**). The difference in specificity (–8.3%) was significantly different from zero ($p < 0.0001$). The two-sided 95% confidence interval on the difference was (–9.0%, –7.6%).

For subjects without CRC or AA (categories 3-6), a positive test result is considered a false positive. The false positive fraction is 1 – specificity and was significantly higher for *Cologuard*

(13.4%) than for FIT (5.1%) ($p < 0.0001$). On the other hand, for subjects with CRC or AA, the true positive fraction was higher for *Cologuard* (46.3%) than for FIT (27.7%) (**Table 16**).

For subjects without CRC (categories 2-6), the specificity (fraction of subjects called negative) was smaller for *Cologuard* (84.4%, 8372/9924) than for FIT (93.4%, 9272/9924). The difference was –9.1% with two-sided 95% confidence interval (–9.8%, –8.4%). The *Cologuard* specificity for CRC (84.4%) together with its sensitivity for CRC (92.3%) form a complimentary pair spanning the entire study population. By comparison, the FIT specificity for CRC was higher (93.4%) while its sensitivity for CRC was lower (73.8%) than for *Cologuard*.

Table 16: Specificity – Specificity Subset (Categories 3-6)

		FIT Outcome		
	***Cologuard* Outcome**	**Negative**	**Positive**	**Totals**
Categories 3-6	**Negative, n (%)**	7787 (98.1%)	149 (1.9%)	7936
	Positive, n (%)	908 (73.8%)	323 (26.2%)	1231
	Totals	8695	472	9167

Cologuard was also compared with FIT on the Receiver Operating Characteristic curve ("ROC curve"). For tests yielding (but not necessarily reporting) a continuous or ordinal value (measurement or score), a threshold or cut-off may be applied to define test positive and test negative results. The ROC curve is a plot of all possible pairs of sensitivity and 1 – specificity (true and false positive fractions) generated by varying the cut-off over the entire range of observed values.

For CRC, the ROC curves are displayed for *Cologuard* and FIT (**Figure 3**). In the figure, the false positive and true positive fractions associated with cut-offs used by the tests are superimposed. Also displayed in the figure is the area under the ROC curve (AUC) for each test. For *Cologuard* the AUC was 93.0%, indicating that a probability is 93.0% that a randomly chosen CRC subject has a higher underlying *Cologuard* composite score than a randomly chosen non-CRC subject. For FIT this probability was 88.0%.

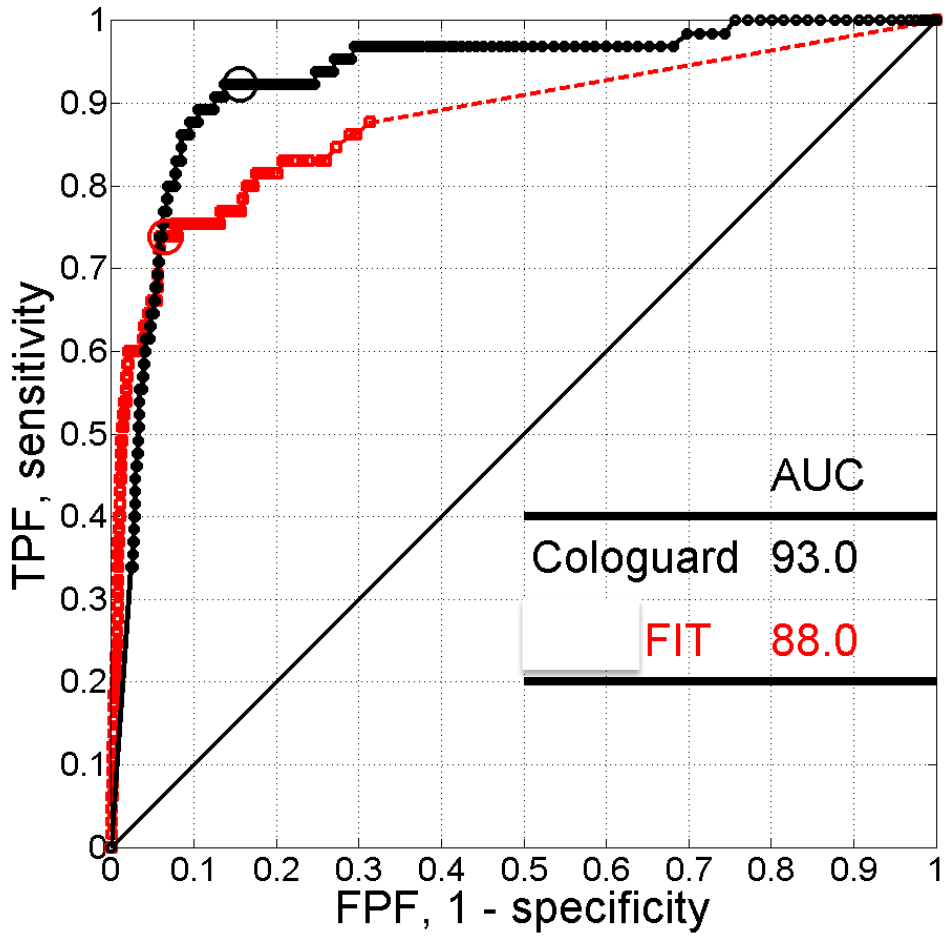


Figure 3: ROC curves for CRC: Cologuard and FIT. The ROC curve plots sensitivity for CRC (Category 1) vs. 1 – specificity for non-CRC (Categories 2-6).

For AN, the ROC curves are displayed for *Cologuard* and FIT (**Figure 4**). For *Cologuard*, the AUC for AN was 73.3%, indicating that a probability is 73.3% that a randomly chosen subject with CRC or AA has a higher underlying *Cologuard* composite score than a randomly chosen subject without CRC or AA. For FIT this probability was 66.7%.

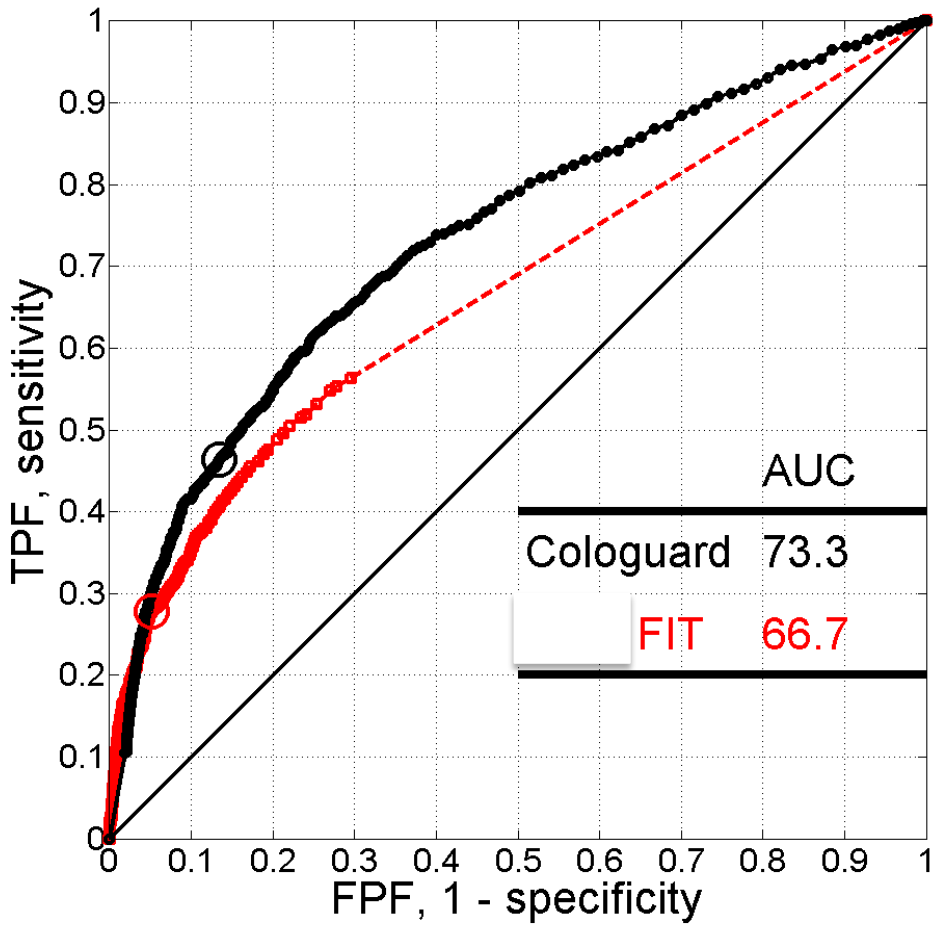


Figure 4*: ROC curves for AN: Cologuard and FIT. The ROC curve plots sensitivity for advanced neoplasia (AN, categories 1-2) vs. 1 – specificity for non-AN (Categories 3-6).*

In addition to the sensitivity and specificity for CRC and AA, the positive and negative likelihood ratios for *Cologuard* were calculated from the study data. Results demonstrated a positive likelihood ratio of 5.9 for CRC, indicating that a person with CRC would be 5.9 times more likely to have a positive *Cologuard* result than someone without CRC. The negative likelihood ratio for CRC was 0.09, indicating that someone without CRC is approximately 11 times (1/0.09) more likely to test negative on *Cologuard* compared to someone with CRC.

Results also demonstrated a positive likelihood ratio of 3.2 for AA (**Table 17**), indicating that a person with AA would be 3.2 times more likely to have a positive *Cologuard* results than someone without AA or CRC. The negative likelihood ratio for AA was 0.67, indicating that someone without AA or CRC is approximately 1.5 times (1/0.67) more likely to test negative on *Cologuard* compared to someone with AA.

Table 17: Positive and Negative Likelihood Ratios

	Category 1 (CRC) vs Categories 2-6	Category 2 (AA) vs Categories 3-6
Positive Likelihood Ratio (PLR)		
Sensitivity, %	92.3 (60/65)	42.4 (322/760)
1-Specificity, %	15.6 (1553/9958)	13.4 (1231/9198)
PLR	5.9	3.2
95% Confidence Interval	(5.4, 6.4)	(2.9, 3.5)
Negative Likelihood Ratio (NLR)		
1-Sensitivity, %	7.7 (5/65)	57.6 (438/760)
Specificity, %	84.4 (8405/9958)	86.6 (7967/9198)
NLR	0.09	0.67
95% Confidence Interval	(0.04, 0.21)	(0.63, 0.71)

Analysis was also performed to calculate the positive and negative predictive values ("PPV" and "NPV") for *Cologuard* (**Table 18**). As with any CRC screening test, the PPV is impacted by the very low prevalence of CRC in the general population. The PPV was calculated to be 3.72% (60/1613) for CRC and 19.86% (322/1613) for AA. Meanwhile, the NPV was 94.73%.

Table 18: Positive Predictive Value – Primary Effectiveness Subjects

Cologuard	Category 1 (CRC)	Category 2 (AA)	Categories 3-6
Negative	0.06, 0.02-0.14 (5/8410)	5.2, 4.7- 5.7 (438/8410)	94.7, 94.2-95.2 (7967/8410)
Positive	3.72, 2.85-4.76 (60/1613)	20.0, 18.0-22.0 (322/1613)	76.3, 74.2-78.4 (1231/1613)

*2-Sided 95% CIs

3. Sub-Group Analyses

The DeeP-C study results were also analyzed according to various demographic characteristics, as well as lesion size and location. Performance goals were not pre-specified for subgroup analysis. No attempt was made to adjust p values for multiple subgroup analyses.

Results by Gender

Sensitivity of *Cologuard* was higher for males than for females, both for CRC and AA. *Cologuard* sensitivity for CRC was 100.0% for males, compared with 83.9% for females (**Table 19)**. *Cologuard* sensitivity for AA was 44.7% for males, compared with 39.0% for females.

Table 19: *Cologuard* Sensitivity by Gender (Categories 1 and 2)

Gender, n/N (%)	**Category 1 (CRC)**	**Category 2 (AA)**
Male	34/34 (100.0)	201/450 (44.7)
Female	26/31 (83.9)	121/310 (39.0)

[1] Sensitivity calculated as number of positives (CRC or AA) divided by subjects with CRC or AA, respectively.

Meanwhile, specificity of *Cologuard* for subjects with neither CRC nor AA (AN) was very similar for females as compared with males. As shown in **Table 20** below, Specificity for non-AN was 87.3% (4,398/5,037) for females compared with 85.8% (3,569/4,161) for males.

Table 20: *Cologuard* Specificity by Gender

Gender, n/N (%)	**Categories 3-6[1]**
Male	3569/4161 (85.8)
Female	4398/5037 (87.3)

[1] Specificity calculated as number of negatives among subjects without CRC or AA.

Results by Race and Ethnicity

With respect to race, *Cologuard* sensitivity for CRC was higher among White subjects (53/55, 96.4%) than among Black or African-American subjects (5/8, 62.5%). There was only one Asian CRC case in the study (1/1, 100.0%) (**Table 21)**. However, the results observed in Black or African-American and Asian subjects may well have been driven by the low overall number of cancer cases in that subpopulation. Sensitivity among Hispanic or Latino subjects (8/9, 88.9%) was also high, although again the sample size was small. Sensitivity for AA was similar for White (271/641 42.3%) and Black/African-American (36/85, 42.4%) subjects. Sensitivity was also similar among Hispanic/Latino subjects (23/59, 39.0%). *Cologuard* sensitivity for AA was lower among Asian subjects (4/13, 30.8%) and higher for American Indian or Alaskan Natives (3/4, 75.0%), compared with other groups. When subgroups for race, gender and age are considered together as predictors of the log odds of a *Cologuard* positive result, no statistical evidence was found for significant variation by race group in either *Cologuard* sensitivity for CRC or *Cologuard* sensitivity for AA (p = 0.0872, 0.7447).

Table 21: *Cologuard* Sensitivity by Race and Ethnicity, CRC and AA Subsets (Categories 1 and 2)

Subgroup	Category 1 (CRC)	Category 2 (AA)
Race, n/N (%)		
White	53/55 (96.4)	271/641 (42.3)
Black or African American	5/8 (62.5)	36/85 (42.4)
Asian	1/1 (100.0)	4/13 (30.8)
American Indian or Alaska Native	0/0	3/4 (75.0)
Native Hawaiian or Other Pacific Islander	0/0	0/0
Other	1/1 (100.0)	7/16 (43.8)
Ethnicity, n/N (%)		
Hispanic or Latino	8/9 (88.9)	23/59 (39.0)
Not Hispanic or Latino	52/56 (92.9)	298/700 (42.6)

[1] Sensitivity calculated as number of positives (CRC or AA) divided by subjects with CRC or AA.

Cologuard specificity for subjects without CRC or AA (categories 3-6) varied among race groups (p-value = 0.001) (**Table 22**). Specificity was highest for Asian (93.5%, 229/245) and Native Hawaiian/Pacific Islander subjects (91.3%, 21/23) and lowest for American Indian/Alaska Native subjects (75.0%, 24/32). Statistical significance remained after controlling for gender and age in a multiple regression model (p-value = 0.0002). Specificity was also high (90.7% (837/923)) among Hispanic or Latino subjects.

Table 22: *Cologuard* Specificity by Race and Ethnicity – Primary Effectiveness Subjects

Subgroup	Categories 3-6
Race, n/N (%)	
White	6639/7726 (85.9)
Black or African American	879/978 (89.9)
Asian	229/245 (93.5)
American Indian or Alaska Native	24/32 (75.0)
Native Hawaiian or Other Pacific Islander	21/23 (91.3)
Other	171/189 (90.5)
Ethnicity, n/N (%)	
Hispanic or Latino	837/923 (90.7)
Not Hispanic or Latino	7127/8272 (86.2)

[1] Specificity calculated as number of negatives among subjects without CRC or AA.

Results by Age

Cologuard sensitivity for CRC was consistently high across all age groups (**Table 23**), ranging from 88.9-100.0% for age groups with at least six subjects. Although sensitivity was 75% for subjects of age 60-64, the number of CRC cases was particularly small in this age group (n = 4), and only one CRC case was not detected by *Cologuard*.

Cologuard sensitivity for AA increased from 38.0% for subjects of age < 60 to 46.8% for subjects between age 70-79 (**Table 23**). In a multiple logistic regression analysis controlling for gender, race, lesion size and lesion location, variation in *Cologuard* sensitivity for AA was significant (p-value=0.019) and from the model fit (not shown) sensitivity is estimated to increase from a baseline value of 41.5% to 44.4% with a 5 year age increase.

Table 23: *Cologuard* Sensitivity by Age

Age, n/N (%)	**Category 1 (CRC)**	**Category 2 (AA)**
<60 years	7/7 (100.0)	65/171 (38.0)
60-64 years	3/4 (75.0)	24/57 (42.1)
65-69 years	19/20 (95.0)	125/301 (41.5)
70-74 years	16/18 (88.9)	72/154 (46.8)
75-79 years	6/6 (100.0)	29/62 (46.8)
80-84 years	9/10 (90.0)	7/15 (46.7)

[1] Sensitivity calculated as number of positives (CRC or AA) divided by subjects with CRC or AA.

[2] Two 49-year-old subjects and one 44-year-old subject were included in the analysis population, although they would not be included in the intended use population.

Cologuard specificity for subjects without CRC or AA (categories 3-6) was highest for younger subjects and lowest for older subjects, ranging from 77.8-92.2% (**Table 24**). In a multiple logistic regression analysis controlling for gender and race, variation in *Cologuard* specificity for non-AA/CRC was significant and from the model fit (not shown) specificity is estimated to decrease from a baseline value of 85.7% to 82.8% with a 5 year age increase

Table 24: *Cologuard* Specificity by Age

Age, n/N (%)	**Categories 3-6**
<60 years	2491/2703 (92.2)
60-64 years	681/765 (89.0)
65-69 years	2871/3352 (85.7)
70-74 years	1292/1566 (82.5)
75-79 years	480/617 (77.8)
80-84 years	152/195 (77.9)

[1] Specificity calculated as number of negatives among subjects without CRC or AA.

[2] Two 49-year-old subjects and one 44-year-old subject were included in the analysis population, although they would not be included in the intended use population.

Results by Lesion Size and Cancer Stage

Cologuard results were evaluated by lesion size, as well as cancer stage (**Table 25**). Sensitivity of *Cologuard* increased with lesion size, as would be expected for a stool-based DNA test of this type. The amount of DNA shed from cancerous or pre-cancerous tissue in the colon is generally expected to increase with increased mass or lesion size.

As shown in the table below, sensitivity was > 90% for most lesion sizes. Sensitivity for CRC was highest for subjects with CRCs ≥ 30 mm (32/34, 94.1%) and lowest for subjects with CRCs 5-9 mm in size (4/5, 80.0%). Sensitivity by cancer stage was generally high and was the highest for subjects with Stage II cancers (21/21, 100.0%) and Stage III cancers (9/10, 90%). Sensitivity of *Cologuard* for AA was higher among subjects with AAs of larger lesion sizes, and this trend remained after controlling for gender, race and age in a multiple logistic regression analysis (p-value < 0.0001).

Table 25: *Cologuard* Sensitivity within Lesion Subgroups

Subgroup	Category 1 (CRC)	Category 2 (AA)
Largest Lesion Size, n/N (%)		
<5 mm	0/0	2/10 (20.0)
5-9 mm	4/5 (80.0)	18/56 (32.1)
10-19 mm	13/14 (92.9)	225/577 (39.0)
20-29 mm	11/12 (91.7)	51/79 (64.6)
>=30 mm	32/34 (94.1)	26/38 (68.4)
Stage, n/N (%)		
I	26/29 (89.7)	N/A
II	21/21 (100.0)	N/A
III	9/10 (90.0)	N/A
IV	3/4 (75.0)	N/A
Unknown*	1/1 (100.0)	N/A

[1] Sensitivity calculated as number of positives (CRC or AA) divided by subjects with CRC or AA.

Specificity of *Cologuard* for subjects without AA or CRC was stratified by lesion size (**Table 26)**. Specificity of *Cologuard* for CRC was 86.2% (1,847/2,142), for subjects with CRCs < 5 mm in size, and 79.7% (1,523/1,912) for subjects with CRCs 5-9 mm in size.

Table 26: *Cologuard* Specificity by Lesion Size – Primary Effectiveness Subjects

Largest Lesion Size, n/N (%)	Categories 3-6
<5 mm	1847/2142 (86.2)
5-9 mm	1523/1912 (79.7)
10-19 mm	0/0
20-29 mm	0/0
>=30 mm	0/0

[1] Specificity calculated as number of negatives among subjects without CRC or AA.

Results by Lesion Location

Cologuard sensitivity was also assessed by lesion location (**Table 27**). Sensitivity of *Cologuard* for CRC was 90% or greater, regardless of lesion location. Sensitivity of *Cologuard* for AA was higher among subjects with distal AAs (133/238, 55.9%) and lower among subjects with proximal AAs (143/433, 33.0%). The variation in AA sensitivity by lesion location was statistically significant after controlling for gender, race, age and lesion size in a multiple logistic regression analysis ($p < 0.0001$).

Table 27: *Cologuard* Sensitivity by Lesion Location

Lesion Location, n/N (%)	Category 1 (CRC)	Category 2 (AA)
Proximal	27/30 (90.0)	143/433 (33.0)
Distal	22/24 (91.7)	133/238 (55.9)
Rectal	11/11 (100.0)	45/88 (51.1)

[1] Sensitivity calculated as number of positives (CRC or AA) divided by subjects with CRC or AA.

Specificity of *Cologuard* for subjects without CRC or AA (categories 3-6) was high, regardless of lesion location. Specificity of *Cologuard* was 83.4% for subjects with proximal CRCs, 82.1% for subjects with distal CRCs, and 84.5% for subjects with rectal CRCs (**Table 28**).

Table 28: *Cologuard* Specificity by Lesion Location – Primary Effectiveness Subjects

Lesion Location, n/N (%)	Categories 3-6
Proximal	1723/2066 (83.4)
Distal	1131/1377 (82.1)
Rectal	517/612 (84.5)

[1] Specificity calculated as number of negatives among subjects without CRC or AA.

Results by AA Subcategories

Numerically greater sensitivity for *Cologuard* compared to FIT was observed across all sub-categories of AA. For example, sensitivity for adenoma with carcinoma in situ or high grade dysplasia (Category 2.1) was 69.2% for *Cologuard*, compared to 46.2% for FIT. *Cologuard* identified 43.0% of serrated lesions, whereas FIT sensitivity for these lesions was 5.1%.

Diagnostic Yield

In a hypothetical screening population of 100,000 subjects, the benefit risk for *Cologuard* was evaluated relative to FIT in detection of CRC (**Table 29**). In this analysis, the prevalences of CRC, AA, and non-AN were assumed to be 0.70%, 7.58%, and 91.72%, the same as those observed in the DeeP-C study among 10,840 subjects with a histopathological result. The fraction of CRC, AA, and non-AN subjects testing positive by *Cologuard* and by FIT were assumed to be the same as those observed in the DeeP-C secondary analysis population (n=9,989). Among 100,000 hypothetical subjects, 700 are expected to have CRC. Among the 700 CRC subjects, the number testing positive is expected to be 647 for *Cologuard* and 518 for FIT. These are numbers of true positive (TP) test results for CRC. Among the remaining 99,300 non-CRC subjects, the number testing positive is expected to be 15,529 for *Cologuard* and 6,524 for FIT. These are numbers of false positive (FP) test results for non-CRC. The ratio of the number of FPs to TPs in the screening population is therefore 24.0 for CG and 12.6 for FIT.

Taking the difference between the numbers of true positive results, *Cologuard* is expected to detect 129 more CRC subjects than FIT. However, *Cologuard* is also expected to yield 9005 more false positive test results than FIT on non-CRC subjects. As the last column of the first panel indicates, *Cologuard* is expected to detect one more CRC subject than FIT at the expense of 70 more false positive results on non-CRC subjects.

Further, assuming the risk of an adverse event due to colonoscopy is 0.68%, *Cologuard* is expected to yield 61 more adverse events on colonoscopy than FIT among non-CRC subjects referred to colonoscopy due to an FP result. Thus, as the last column of the second panel indicates, *Cologuard* is expected to detect one more CRC subject than FIT at the expense of 0.50 extra non-CRC subjects being referred to colonoscopy based on a FP result and experiencing an adverse event from that procedure.

Table 29: Diagnostic Yield of *Cologuard* (*CG*) and FIT in a Screening Population of 100,000 Subjects, CRC.

Effectiveness Evaluation: Expected Number of True and False Positive Results for CRC

Histological Type	E(N)	*CG*+	FIT+	Difference	Difference/129
CRC	**700**	**647**	**518**	**+129**	**+1**
Non-CRC	**99300**	**15529**	**6524**	**+9005**	**+70**
FPs per TP		**24.0**	**12.6**		

Safety Evaluation: Expected Number of Adverse Events on Colonoscopy as a Result of a False Positive Referral

Histological Type	E(N)	*CG*+	FIT+	Difference	Difference/129
Non-CRC	**99300**	**105.6**	**44.4**	**+61.2**	**+0.5**

†Assumes risk is 0.68% of an adverse event during colonoscopy. [3]

Similarly, the benefit-risk of *Cologuard* relative to FIT was evaluated for detection of AN (CRC or AA) (**Table 30**). In a population of 100,000 subjects, 8280 are expected to have AN. *Cologuard* is expected to detect 1542 more subjects with AN than FIT at the expense of 7,594 FP results on non-AN subjects. The ratio of the expected number of FPs for non-AN to TPs for AN is 3.2 for *Cologuard* and 2.0 for FIT. As the last column in the first panel indicates, *Cologuard* is expected to detect one more AN subject than FIT at the expense of 5 more FP results on non-AN subjects. As the last column in the second panel indicates, *Cologuard* is expected to detect one more AN subject than FIT at the expense of 0.03 extra non-AN subjects being referred to colonoscopy based on a FP result and experiencing an adverse event from that procedure.

[3] Rutter, Johnson, Miglioretti *et al*. (2012) Adverse events after screening and follow-up colonoscopy. *Cancer Causes Control* 23(2):289-96.

Table 30: Diagnostic Yield of *Cologuard* (*CG*) and FIT in a Screening Population of 100,000 Subjects, AN (CRC or AA).

Effectiveness Evaluation: Expected Number of True and False Positive Results for AN

Histological Type	E(N)	*CG*+	FIT+	Difference	Difference/129
AN	**8280**	**3863**	**2321**	**+1542**	**+1**
Non-AN	**91720**	**12316**	**4722**	**+7594**	**+5**
FPs per TP		**3.2**	**2.0**		

Safety Evaluation: Expected Number of Adverse Events on Colonoscopy as a Result of a False Positive Referral†

Histological Type	E(N)	*CG*+	FIT+	Difference	Difference/129
Non-AN	**91720**	**83.7**	**32.1**	**+51.6**	**+0.03**

†Assumes risk is 0.68% of an adverse event during colonoscopy. [4]

The primary risk associated with the *Cologuard* test is a false assay result (i.e., a false positive or a false negative result). The *Cologuard* false negative rate was estimated to be 7.7%. The total positivity rate for *Cologuard* is estimated to be 16%. Considering all positive results, approximately 1 of 4 positive results is predicted to result in a diagnosis of CRC or AA. Thus, in a hypothetical population of 100,000 users, *Cologuard* would identify 129 more CRC cases and 1412 more AA cases compared to use of FIT and for every 59 additional false positive results compared to FIT, use of *Cologuard* would be associated with identifying 12 more true positives (1 CRC and 11 AA) (**Table 31**).

[4] Rutter, Johnson, Miglioretti *et al*. (2012) Adverse events after screening and follow-up colonoscopy. *Cancer Causes Control* 23(2):289-96.

Table 31: Expected Diagnostic Yield in a Screening Population (Hypothetical Screening Population of 100,000)

Histological Type	N	Expected Number of Positives		Difference	Difference ÷ 129
		Cologuard	FIT		
CRC	700	647	518	+129	+1
AA	7580	3216	1803	+1412	+11
Not CRC or AA	91720	12316	4722	+7594	+59

G. Financial Disclosure

The Financial Disclosure by Clinical Investigators regulation (21 CFR 54) requires applicants who submit a marketing application to include certain information concerning the compensation to, and financial interests and arrangement of, any clinical investigator conducting clinical studies covered by the regulation. The pivotal clinical study included 401investigators, including both primary and sub-investigators of which none were full-time or part-time employees of the sponsor and one had disclosable financial interests/arrangements as defined in 21 CFR 54.2(a), (b), (c) and (f) and described below:

- Compensation to the investigator for conducting the study where the value could be influenced by the outcome of the study: None
- Significant payment of other sorts: None
- Proprietary interest in the product tested held by the investigator: One
- Significant equity interest held by investigator in sponsor of covered study: One

The applicant has adequately disclosed the financial interest/arrangements with clinical investigators. Statistical analyses were conducted by FDA to determine whether the financial interests/arrangements had any impact on the clinical study outcome. The information provided does not raise any questions about the reliability of the data.

XI. PANEL MEETING RECOMMENDATIONS AND FDA'S POST-PANEL ACTION

A. Panel Meeting Recommendation

At an advisory meeting held on March 27, 2014, the Molecular and Clinical Genetics Panel of the Medical Devices Advisory Committee voted 10-0 that there is reasonable assurance the device is safe, 10-0 that there is reasonable assurance that the device is effective, and 10-0 that the benefits of the device do outweigh the risks in patients who meet the criteria specified in the proposed indication.
http://www.fda.gov/AdvisoryCommittees/Calendar/ucm384559.htm

B. FDA's Post-Panel Action

All the panel recommendations are being followed.

The sponsor will conduct a PAS. The detail of the proposed study are as follows:

1. Study Objective: The study objective is to collect longitudinal data on subjects prescribed *Cologuard* over the course of 3 years.

2. Study Design: Prospective, longitudinal, multi-center study.

3. Study length: 5 years, 3 years of longitudinal subject follow up. Subjects will be required to complete the *Cologuard* test at baseline (T0) and at year 3 (T3). Subjects with a positive *Cologuard* test at T0 will undergo diagnostic colonoscopy and then will be discontinued from the study. Subjects with negative *Cologuard* test results at T0 will remain in the study, repeat the *Cologuard* test and undergo a colonoscopy at T3. Subjects will undergo annual follow-up at T1 and T2 to evaluate for changes in medical history. Final analysis will be conducted after the last subject completes follow-up at T3 or discontinues the study.

4. Primary Endpoint: The difference between the positive predictive value (PPV) at T3 (PPV3) and 1 minus the negative predictive value (NPV) at T3 (NPV3).

5. Sample Size: 1,830 Men and women between the ages of 50 and 84 inclusive, who are at average risk of developing colorectal cancer.

XII. CONCLUSIONS DRAWN FROM NONCLINICAL AND CLINICAL STUDIES

A. Effectiveness Conclusions

Data from the analytical studies demonstrated acceptable analytical sensitivity, analytical specificity and precision and reproducibility of *Cologuard.*

The pivotal clinical study established *Cologuard* sensitivity for CRC of 92.3% and a specificity of 86.6%. The lower bounds of the one-sided 95% confidence intervals for these results exceeded the thresholds set in the study protocol. As such, *Cologuard* satisfied the primary performance measure for the study.

In addition, the study successfully demonstrated superiority of *Cologuard* to FIT for detection of CRC ($p=0.0018$) and AA ($p<0.0001$). *Cologuard* demonstrated benefit over FIT, identifying 13 CRC cases that were not identified by FIT. Meanwhile, in only 1 case did FIT identify a CRC case that was not identified by *Cologuard*. Overall, *Cologuard* yielded a 20.0% incremental benefit over FIT for CRC detection. Similarly, for AA detection, *Cologuard* successfully identified 178 AA cases that were not identified by FIT. Meanwhile, FIT only identified 29 AA cases that were not identified by *Cologuard.* Overall, *Cologuard* had a 22.5% incremental benefit for AA detection. Finally, *Cologuard* sensitivity for CRC was demonstrated across a variety of age groups, racial/ethnic groups, and in both men and women.

In conclusion, the pivotal study demonstrated that *Cologuard* met and exceeded the primary performance measure of the study. Additionally, *Cologuard* met and exceeded the secondary performance measures of the study, demonstrating non-inferiority and superiority to FIT. *Cologuard* was sensitive and specific for CRC and provides incremental value over currently available non-invasive screening tests for CRC.

B. Safety Conclusions

Risks associated with the collection of the stool sample necessary for the *Cologuard* test were very minimal. During the pivotal clinical study of 12,776 patients, only 4 mild adverse events were reported and were associated with the sample collection kit.

With respect to the *Cologuard* test itself, the primary risk relates to a false assay result (i.e., a false positive or a false negative result). All positive test results should lead to a colonoscopy. Adverse events commonly associated with colonoscopy include abdominal discomfort and bowel irregularity post-procedure. Rare adverse events associated with colonoscopy include bleeding, intestinal perforation, and adverse reaction to the sedation resulting in respiratory and/or cardiac events, stroke and death. In the instance of a false negative result on *Cologuard*, there is a possibility that a case of CRC or AA could go undetected.

C. Benefit-Risk Conclusions

The probable benefits of the device are based on data collected in the clinical study ("DeeP-C") conducted to support PMA approval as described above. The clinical benefit of *Cologuard* was demonstrated in an analysis of efficacy and safety data obtained from patients who are typical candidates for CRC screening, adults of either sex, 50 years or older, who were at average risk for CRC (DeeP-C Study). Based on these data, *Cologuard* provides a safe and effective additional screening tool for detection of CRC and AA.

When used for screening, a positive result should be followed by colonoscopy for diagnosis. The risks associated with the device are similar to other *in vitro* diagnostic assays and are associated with risks resulting from false results. A false positive result could result in an additional invasive screening procedure, such as colonoscopy, and thus expose patients to the attendant risks associated with such a procedure. A false negative result with *Cologuard* could potentially delay colonoscopy and a potentially delayed diagnosis of disease. The clinical data in this application demonstrated that the *Cologuard* was sensitive and specific for CRC and provides incremental value over currently available non-invasive screening tests for CRC. Furthermore, data from the analytical studies demonstrated acceptable analytical performance of the test. In conclusion, given the available information above, the data support that for the qualitative detection of colorectal neoplasia associated DNA markers and for the presence of occult hemoglobin in human stool, the probable benefits outweigh the probable risks.

D. Overall Conclusions

The data in this application support the reasonable assurance of safety and effectiveness of this device when used in accordance with the indications for use. Data from the DeeP-C study

support the utility of *Cologuard* to screen for the presence of CRC or AA in adults of either sex, 50 years or older, who are average-risk for CRC.

XIII. CDRH DECISION

CDRH issued an approval order on August 11, 2014. The final conditions of approval can be found in the approval order.

The applicant's manufacturing facilities were inspected and found to be in compliance with the device Quality System (QS) regulation (21 CFR 820).

XIV. APPROVAL SPECIFICATIONS

Directions for use: See device labeling

Hazards to Health from Use of the Device: See Indications, Contraindications, Warnings, Precautions, and Adverse Events in the device labeling.

Post-approval Requirements and Restrictions: See approval order.

XV. REFERENCE

Rutter CM, Johnson E, Miglioretti DL, et al. (2012). Adverse events after screening and follow-up colonoscopy. *Cancer Causes Control* 23:289-96.