FDA Executive Summary

Prepared for the
March 27, 2014 meeting of the
Molecular and Clinical Advisory Committee Panel

P130017
Cologuard™
Exact Sciences Corporation

Introduction

This is the FDA Executive Summary for the Exact Sciences Corporation *Cologuard*. The device is a stool DNA-based colorectal cancer screening test and is intended for use as an adjunctive screening test for the 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 or pre-malignant colorectal neoplasia. *Cologuard* is not intended as a replacement for diagnostic colonoscopy. *Cologuard* is intended to be used in conjunction with colonoscopy and other test methods in accordance with recognized screening guidelines. A positive result in *Cologuard*, as with any screening test, should be followed by colonoscopy. *Cologuard* is intended for patients who are typical candidates for colorectal cancer screening: adults of either sex, 50 years or older, who are at average risk for colorectal cancer. Exact Sciences has submitted a Premarket Approval Application (PMA) requesting marketing approval of the device under P130017. This submission has been reviewed by the Division of Immunology and Hematology Devices (DIHD), Office of In Vitro Diagnostics and Radiological Health (OIR), within the Center for Devices and Radiological Health (CDRH) of the Food and Drug Administration (FDA).

This memorandum will summarize FDA's review of the PMA, highlighting the areas for which we are seeking your expertise and input. These topics will include the device performance and clinical experience to date. At the conclusion of your review and discussion of the data presented, FDA will ask for your recommendation for whether FDA should approve this PMA, based on your interpretation of the benefit-risk assessment.

TABLE OF CONTENTS

1 PROPOSED INDICATIONS FOR USE

The sponsor has proposed the following Indications for Use statement:

Cologuard is intended for use as an adjunctive screening test for the 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 or pre-malignant colorectal neoplasia. *Cologuard* is not intended as a replacement for diagnostic colonoscopy. *Cologuard* is intended to be used in conjunction with colonoscopy and other test methods in accordance with recognized screening guidelines. A positive result in *Cologuard*, as with any screening test, should be followed by colonoscopy. *Cologuard* is intended for patients who are typical candidates for colorectal cancer screening: adults of either sex, 50 years or older, who are at average risk for colorectal cancer.

1.1 Contraindications

The sponsor has proposed the following Contraindications:

Cologuard is not suitable for everyone. This test is indicated for men and women, age 50 years or older, who are at average risk for development of colorectal cancer. Patients should inform their doctor if they:

- Have a history of colorectal cancer, adenomas, or other related cancers.
- Have had a positive result from another colorectal cancer screening method within the last 6 months.
- Have been diagnosed with a high-risk condition for colorectal cancer. High risk conditions include but are not limited to inflammatory bowel disease (IBD), chronic ulcerative colitis (CUC), Crohn's disease, Familial adenomatous polyposis (FAP), or a family history of colorectal cancer.
- Have been diagnosed with a relevant hereditary cancer syndrome. Examples include Hereditary non-polyposis colorectal cancer syndrome ("HNPCC" or "Lynch Syndrome"), or others including but not limited to 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 and Familial Hyperplastic Polyposis.

2 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 independent families of markers that are thought to be associated with CRC and AA. The first DNA family targets epigenetic changes in the form of gene promoter region methylation, specifically *NDRG4* promoter region hypermethylation and *BMP3* promoter region hypermethylation. The second DNA family targets seven specific point mutations in *KRAS*. The third family of markers is non-DNA based and detects occult hemoglobin. Beta-actin ("*ACTB*") is a reference gene measured for quantitative estimation of the total amount of human DNA present in each sample.

2.1 Device Components

Cologuard uses the following reagent components:

DNA Capture Reagents
Includes Capture Beads

DNA Preparation Reagents
Includes Denaturation Solution; Bisulfite Conversion Solution; Neutralization Solution; Desulphonation Solution (Concentrate); Binding Beads; and DNA and *QuARTS* Supplementary Lot Information Card

QuARTS Assay Reagents
Includes Carrier Solution; Elution Buffer; Oligo Mix A, Methylation; Oligo Mix B, Mutation; Enzyme Mix; DNA Calibrator 1, High Methylation; DNA Calibrator 2, Low Methylation; DNA Calibrator 3, High Mutation; and DNA Calibrator 4, Low Mutation

Hemoglobin Assay Reagents
Includes Hemoglobin Assay Plate; Sample Buffer; Antibody Conjugate
SUBS, Substrate; Stop Solution; Hemoglobin Assay Calibrator; and 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; DNA Control 2, Low with specific copy numbers of relevant methylated and non-methylated DNA; and DNA Control 3, Negative with a specific copy number of non-methylated DNA

(3) *Cologuard* Hemoglobin Control Kit containing Lyophilized Hemoglobin Control 1, High; Hemoglobin Control 2, Low derived from human whole blood and plasma containing specific concentrations of human hemoglobin; and Lyophilized

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

(4) Ancillary Materials and Bulk Assay Reagents including Stool Buffer; Inhibitor Removal Tablet; Spin Filter; Barcoded Mixing Tubes; Capture Bead Pre-wash; Capture Solution; Capture Wash; Binding Solution; Conversion Wash Concentrate; and 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.).

2.2 Principles of Operation

Cologuard uses 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 naturally 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. The test also incorporates detection of fecal occult hemoglobin.

Stool samples are collected using the *Cologuard* Collection Kit, which includes patient instructions, a protein sample tube with stool collection stick and buffer (for the fecal occult hemoglobin portion of the assay), a stool collection container (for the molecular portion of the assay), a foldable plastic bracket, a liquid preservative, and a mailing container. The mailing container is used to send the collected sample to Exact Sciences' laboratory for processing. In general, one stool sample from one stool collection is adequate for testing.

Once received in the laboratory, in the pre-processing step, 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 thawed stool homogenates; (2) bisulfite conversion of methylated DNA; and (3) DNA purification coupled with Quantitative Allele-Specific Real-time Target and Signal (*QuARTS*™) 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 Enzyme-Linked Immunosorbent Assay (ELISA) technique is used to analyze the level of hemoglobin present in the stool sample.

Cologuard provides a final qualitative result of "negative" or "positive" based on the *Cologuard* composite score, which has a range of 0-1000. The final result is "positive" if the composite score is greater than or equal to 183 and is "negative" if the composite score is less than 183. The composite score is obtained by applying the logistic function to a sum of scores and multiplying by 1000. The sum of scores is the sum of a logistic score weighting individual methylation, mutation, and hemoglobin assay results and four other scores for NDRG4, BMP3, KRAS1, and KRAS2 that could be either zero or some non-zero value. The actual composite score will not be provided in the device report.

3 BACKGROUND INFORMATION AND REGULATORY HISTORY

3.1 Colorectal Cancer

The American Cancer Society's estimates for the number of colorectal cancer cases in the United States are 96,830 new cases of colon cancer and 40,000 new cases of rectal cancer per year. Colorectal cancer is the third leading cause of cancer-related deaths in the United States when men and women are considered separately, and the second leading cause when both sexes are combined. It is expected to cause about 50,310 deaths in 2014. Overall, the lifetime risk of developing colorectal cancer is about 5%. This risk is slightly lower in women than in men (www.cancer.org). The disease begins as a benign adenomatous polyp, which develops into an advanced adenoma with high-grade dysplasia and then progresses to an invasive cancer. Invasive cancers that are confined within the wall of the colon (tumor-node-metastasis stages I and II) are curable, but if untreated, they may spread to regional lymph nodes (stage III) and then metastasize to distant sites (stage IV). Stage I and II tumors are curable by surgical excision, and up to 73% of cases of stage III diseases are curable by surgery combined with adjuvant chemotherapy. Recent advances in chemotherapy have improved survival, but stage IV disease is usually incurable (Markowitz, 2009).

3.2 Current Colorectal Cancer Screening Guidelines

There is strong evidence to suggest that screening for CRC reduces the incidence and mortality of the disease (CDC, 2013). 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 colonoscopy. Colonoscopy is considered to be the most accurate screening tool and is the reference method. Although estimates of sensitivity and specificity for colonoscopy in the published literature vary depending on, for example, what pathology is defines as positive, the US Preventive Services Task Force (USPSTF) has performed modeling and estimates the sensitivity of colonoscopy for CRC to be 95% and specificity to be 90% (U.S. Preventive Services Task Force, 2008). Non-invasive CRC screening tools include fecal occult blood test (FOBT) using either guaiac (gFOBT) or immunochemical (iFOBT – also known as fecal immunochemical test or FIT). To use FIT, no dietary or medicinal restrictions are required and the patient is asked to collect one single sample from one stool into a sampling bottle that is returned to the physician office or laboratory for testing. Patients who have a positive test result with any of these screening methods, except colonoscopy, warrant further investigation with colonoscopy to discount or confirm the presence of CRC or polyps. Currently, the assessment and comparison of CRC screening methods in a testing program over time from a population perspective are limited to data from analytic modeling.

A number of professional societies and organizations have developed guidelines for CRC screening. Although the details of the recommendations differ, there is agreement that screening for average-risk persons should start at age 50 with repeat testing over time. Guidelines published in 2009 from the American College of Gastroenterology (ACG) recommend that colonoscopy conducted every 10 years remains the preferred CRC screening strategy (Rex DK, 2009). In settings where colonoscopy is not available due to economic limitations or eligible persons are not willing to undergo colonoscopy for screening purposes, the guidelines recommend that patients be offered an alternative screening test such as flexible sigmoidoscopy every 5-10 years, CTC every 5 years, or a cancer detection test such as Hemoccult Sensa annually or Fecal DNA testing every 3 years.

In March 2008, the American Cancer Society (ACS), the U.S. Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology jointly recommended screening for colorectal cancer by 1) high-sensitivity FOBT or FIT annually, 2) flexible sigmoidoscopy every 5 years, 3) double-contrast barium enema every 5 years, 4) CTC every 5 years, 5) colonoscopy every 10 years, or 6) fecal DNA testing at an unspecified interval (Levin B, 2008). In contrast, the USPSTF does not include in its screening recommendations the option of using fecal DNA testing or virtual colonoscopy or CTC. In October 2008, the USPSTF recommended screening for CRC using 1) FOBT annually, 2) sigmoidoscopy every 5 years with FOBT every 3 years, or 3) colonoscopy every 10 years (U.S. Preventive Services Task Force, 2008).

The guidelines also vary with respect to the approach to screening for CRC in patients 75 years of age and older. Recommendations from the ACG and ACS do not specify an age limit for CRC screening (Rex DK, 2009 and Levin B, 2008). The American Society for Gastrointestinal Endoscopy recommends that the decision to cease screening at a particular age should be considered by the patient and their physician on an individual basis, taking into account the patient's health status and prior screening history (www.asge.org). The USPSTF recommends against routine screening in adults 76-85 years of age, though there may be instances when considerations on an individual basis may support CRC screening (U.S. Preventive Services Task Force, 2008). The USPSTF further recommends that CRC screening not be conducted in patients over 85 years of age.

In summary, the USPSTF recommendations are divided into three separate grades by screening age as the following:

- Grade A: recommends CRC screening using fecal occult blood testing, sigmoidoscopy, or colonoscopy in adults, beginning at age 50 and continuing until age 75 years.
- Grade C: recommends against routine screening for CRC in adults 76 to 85 years of age.
- Grade D: recommends against screening for CRC in adults older than age 85 years.

3.3 Colorectal Cancer Screening Sensitivity, Participation, Practice

Determination of test sensitivity from one-time use in a cross sectional study does not directly translate to screening program sensitivity, which is achieved through repeated testing over time (Imperiale TF, 2010). A test applied serially can have multiple opportunities to detect a lesion to the extent that results are independent at each use; however, if some lesions cannot be detected by a particular test (e.g., a person's lesion does not and will not exhibit a particular molecular alteration), then results are not independent and cumulative sensitivity would not increase for those patients. Another factor that may affect screening program sensitivity is dwell times. For lesions growing quickly, a lower sensitivity test repeated more frequently may detect more disease compared with a higher sensitivity test performed less often (Ransohoff DF, 2013). Given that screening program sensitivity cannot rely solely on test sensitivity, there are randomized trials in progress comparing screening by FIT with colonoscopy (e.g., the Colonoscopy versus Fecal Immunochemical Test in Reducing Mortality from Colorectal Cancer [CONFIRM] trial, and the Colorectal Cancer Screening in Average-Risk Population: Immunochemical Fecal Occult Blood Testing versus Colonoscopy [COLONPREV] trial) (Levin TR, 2013).

In the United States, CRC screening participation has increased in recent years. According to the Centers for Disease Control and Prevention (CDC), CRC screening

participation for persons 50 to 75 years of age increased from 54% in 2002 to 65% in 2010 (CDC, 2013). As of 2012, the proportion of individuals eligible for screening actually undergoing CRC screening was 65.1%. Among this population, the most frequently used screening modality was colonoscopy (61.7%) followed by FOBT (10.4%) and then a combination of sigmoidoscopy and FOBT (0.7%). It was further reported that the proportion of people using either FOBT or colonoscopy increased with age and was greater among those with health insurance and those with a regular health-care provider. Communication and outreach efforts across communities could facilitate progress in increasing CRC screening. About one-third of the average risk population remains unscreened.

In clinical practice, the screening guidelines are not necessarily followed. One report found serious deviations from evidence-based recommendations in United States primary care (Nadel MR, 2010). For example, it was reported that, instead of diagnostic colonoscopy as follow up for a positive FOBT test result, physicians recommend repeating the FOBT (17.8%) or using other tests (6.6%). These departures from CRC screening guidelines suggest that efforts are needed to inform physicians of appropriate test methods.

3.4 Regulatory Considerations

FDA is reviewing *Cologuard* under the Premarket Application Approval process. To appropriately define test performance, it is important to evaluate the disease spectrum representative of the screening population. Prior to the pivotal trial of *Cologuard*, FDA suggested that a cross-sectional clinical study supporting the performance of an in vitro diagnostic device for CRC screening be designed in the context of FIT performance. FIT is a recommended screening modality across different guidelines and occult blood screening is supported by long-term longitudinal follow-up (Shaukat A, 2013). Different studies have reported a range of FIT performance (Whitlock EP, 2008). In a systematic review, pooled performance characteristics were similar for 1-, 2-, and 3-sample FITs whereas studies using colonoscopy to follow up negative FIT results had lower sensitivity compared with studies using longitudinal follow-up (Lee JK, 2014). Due to heterogeneity between studies, a direct head-to-head comparison to a FIT assay with well-documented CRC screening experience in the intended use setting is warranted to assess the performance of a new in vitro diagnostic device.

In the event that *Cologuard* is approved, the practice of medicine with regard to CRC screening may change. Due to the caveats in extrapolating programmatic colorectal cancer screening performance from cross sectional data, FDA has encouraged the sponsor to propose a post-approval study (PAS) for Panel discussion concerning the need and approach for additional longitudinal performance data to adequately ensure safety and effectiveness, as the landscape of colorectal cancer screening would be changed if the device is approved. In the absence of longitudinal performance results using a newly approved device, a screening interval would not have been established. Consideration for

follow-up screening by an independent method, reference to medical guidelines, and clarifying the appropriate scope of device claims may mitigate safety concerns for potentially suboptimal cumulative sensitivity and for the cumulative false positive rate with repeated testing.

An additional consideration is the extent to which materials provided to patients and physicians properly inform decisions on screening and follow-up testing. For example, in the absence of performance data, patients could defer additional screening after a negative result indefinitely if insufficient advice on follow-up testing is provided. In light of the trends in participation and deviations from screening recommendations in actual practice, FDA would like to ensure that materials provided to patients and physicians with in vitro diagnostic devices are appropriate within the context of current screening guidelines and the intended use. The Agency seeks feedback on whether safety and effectiveness of *Cologuard* is adequately assured based on these considerations and Panel suggestions.

3.5 Marketing History

Cologuard has not been marketed in the United States or any foreign country. If approved, *Cologuard* will be made available for sale in the United States.

4 NON-CLINICAL STUDIES

The sponsor has conducted the following non-clinical studies to evaluate the analytical performance characteristics of *Cologuard*. Summaries of these studies are provided in the appendix.

- Algorithm Development and Cut-Off Determination
- Sensitivity: Limit of Blank, Limit of Detection, Limit of Quantification and Linearity
- *Cologuard* Molecular Assay Cross-Reactivity with Wild Type KRAS
- *Cologuard* QuARTS Partial Methylation Testing
- *Cologuard* Hemoglobin Assay Cross-Reactivity and Specificity
- *Cologuard* Cross-Reactivity with Non-Colorectal Cancers and Diseases
- Precision and Reproducibility (Lab-to-Lab)
- Lot-to-Lot Reproducibility
- Robustness
- Interference
- Carry-Over and Cross-Contamination *Cologuard* Testing
- Stability
- Software documentation, including test results from complete software verification and validation testing

5 CLINICAL STUDY

5.1 Study Design

The goal of the pivotal study ("Multi-Target Colorectal Cancer Screening Test for the Detection of Colorectal Advanced Adenomatous Polyps and Cancer: DeeP-C Study") was to evaluate the safety and effectiveness of *Cologuard* as a screening test for the detection of markers associated with the presence of CRC and AA. This *Cologuard* pivotal study was a prospective, multi-center trial 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. Patients were provided with the *Cologuard* collection kit as well as a separate collection tube to collect stool samples for FIT. Patients were required to undergo colonoscopy within 90 days of study enrollment and were considered to have completed the study after undergoing a colonoscopy with acceptable bowel preparation or after the 90 day period had passed.

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

Colonoscopy findings were recorded per site-specific standard of practice. Patients 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 1.

Table 1: 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

5.1.1 Inclusion and Exclusion Criteria

Patients must have met the following criteria to be eligible for the study:

- Patient is average risk for development of colorectal cancer
- Patient is 50 to 84 years of age inclusive
- Patient has not had a colonoscopy in the previous 9 years
- Patient has signed informed consent

The clinical study included two 49 year olds: one who turned 50 by the time the colonoscopy was conducted and the other who turned 50 within a few months of the colonoscopy. These two patients were categorized as negative by colonoscopy and their inclusion in the analyses did not have a significant impact on study results. The study also included one 44 year old who was also categorized as negative by colonoscopy. All three patients were called negative by both *Cologuard* and the PolyMedco FIT test used in the study. Inclusion did not have a significant impact on study results. In addition, patient 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 patients of age 65-84; 64% of patients in the study population were of age 65-84.

Patients presenting with any of the following were not included in the study:

- Patient has any condition that in the opinion of the investigator should preclude participation in the study (e.g., patient not eligible for a diagnostic colonoscopy).

- Patient has a history of colorectal cancer or advanced adenoma.
- Patient has a history of aerodigestive tract cancer
- Patient has had a prior colorectal resection for any reason other than sigmoid diverticular disease
- Patient has had overt rectal bleeding, e.g., hematochezia or melena, within the previous 30 days. (Blood on toilet paper, after wiping, does not constitute rectal bleeding)
- Patient has a diagnosis or personal history of any of the following high-risk conditions for colorectal cancer:
    - Inflammatory bowel disease (IBD) including chronic ulcerative colitis (CUC) and Crohn's disease.
    - 2 first-degree relatives who have been diagnosed with colon cancer. (Note: first-degree relatives include parents, siblings and offspring).
    - One first-degree relative with CRC diagnosed before the age of 60.
    - Patient has a family history of:
        - Familial adenomatous polyposis (also referred to as "FAP", including attenuated FAP).
        - Hereditary non-polyposis colorectal cancer syndrome (also referred to as "HNPCC" or "Lynch Syndrome").
        - Other hereditary cancer syndromes including but are not limited to 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 and Familial Hyperplastic Polyposis.
- Participation in any "interventional" clinical study within the previous 30 days in which an experimental treatment is administered or might be administered through a randomized assignment of the patient to one or more study groups.

Two patients were found to have occult IBD on colonoscopy. They were included in all analyses. For these two patients, the histological categories were 5 and 6 and *Cologuard* and PolyMedco FIT results were both falsely positive.

5.1.2 Clinical Study Objectives

From Table 1, patients may be characterized as having CRC (category 1), advanced adenoma (AA) (category 2), or negative for CRC/AA (categories 3-6) using as the reference method colonoscopy with histopathology (when required). Based on these characterizations, different sensitivities and specificities may be defined to evaluate the classification performance of a colon screening test (Table 2). CRC sensitivity is defined as the proportion of CRC patients called positive by the test whereas CRC specificity is the proportion of non-CRC patients (categories 2-6) called negative. When advanced neoplasia (AN) is defined as CRC or AA, AN sensitivity is the proportion of CRC/AA patients (categories 1-2) called positive whereas AN specificity is the proportion of

patients without CRC/AA (categories 3-6) called negative. Finally, AA sensitivity is the proportion of AA patients called positive.

The primary performance measures of the study were *Cologuard* CRC sensitivity and *Cologuard* advanced-neoplasia (AN) specificity. The sponsor chose to exclude AA's (category 2) from the specificity calculation, considering these to be positive outcomes since they are treated during colonoscopy. **In this executive summary, the per protocol sensitivity is called CRC sensitivity and the per protocol specificity is called AN specificity (which excludes CRC and AA)**.

The primary objective for *Cologuard* CRC sensitivity was a 95% one-sided lower confidence bound exceeding 65%. The primary objective of *Cologuard* AN specificity was a 95% one-sided lower confidence bound exceeding 85%.

A secondary objective was to demonstrate that *Cologuard* CRC sensitivity was non-inferior to the CRC sensitivity of a commercially available FIT test (PolyMedco), with a non-inferiority margin of 5% (i.e., non-inferiority is demonstrated if the 95% one-sided lower confidence bound on the difference between the tests is greater than -5%). If *Cologuard* were to be declared non-inferior to FIT in CRC sensitivity, then *Cologuard* was also evaluated for superiority to FIT in CRC sensitivity by determining if the difference is statistically significantly greater than 0.

An additional secondary objective was to demonstrate superiority of *Cologuard* sensitivity to detect advanced adenoma (AA) compared with PolyMedco FIT.

An additional analysis was the comparison of *Cologuard* and PolyMedco FIT on AN specificity. A formal hypothesis test was not considered, however, because according to the sponsor the two tests are designed to have different AN specificities.

To complement the primary performance measure of CRC sensitivity, FDA requested that the sponsor calculate CRC specificity, the specificity when histological categories 2-6 are considered as disease negative cases (including AAs). CRC sensitivities and CRC specificities form a complementary pair of classification measures because they include all histological categories 1-6.

Cologuard and PolyMedco FIT were also evaluated on positive and negative likelihood ratio (PLR, NLR) and on positive and negative predictive value (PPV, NPV). Table 2 provides definitions of these measures.

Table 2: Definitions of Sensitivities, Specificities, Likelihood Ratios, and Predictive Values

CRC Sensitivity†	Proportion of patients in histological category 1 who test positive
CRC Specificity	Proportion of patients in histological categories 2-6 who test negative
AN Sensitivity	Proportion of patients in histological categories 1-2 who test positive
AN Specificity†	Proportion of patients in histological categories 3-6 who test negative
AA Sensitivity	Proportion of patients in histological category 2 who test positive
Positive likelihood ratio	Sensitivity / (1 – Specificity)
Negative likelihood ratio	(1 – Sensitivity) / Specificity
Positive predictive value for CRC, AA, and non-AN	Proportion of test positive patients in histological categories 1, 2, and 3-6, respectively
Negative predictive value for CRC, AA, and non-AN	Proportion of test negative patients in histological categories 1, 2, and 3-6, respectively

†The primary performance measures.

6 CLINICAL STUDY RESULTS

6.1 Patient Accountability

The study enrolled a total of 12,766 patients at 90 sites, including both primary care point-of-referral (POR) sites and colonoscopy centers. Patients were excluded from all analyses if their colonoscopy, histopathology, or *Cologuard* results were deemed unusable. Patients were excluded for a variety of reasons, including withdrawal of consent (n=464), failure to undergo colonoscopy (n=1,168), and an unusable colonoscopy (n=304). A colonoscopy was deemed unusable if it was performed outside the study window or prior to stool collection (in violation of the study protocol) or lacked findings due to poor bowel prep, incomplete exam, or no cecum inspection. Patients were also excluded if they did not have a usable *Cologuard* result (n=817) either because no stool was submitted (n=128), the stool was unable to be tested (n=474) (e.g. stool was overweight), the *Cologuard* result was invalid (n=213), or the stool was inadvertently not sent to the laboratory for processing (n=2).

In all, a total of 10,023 patients were included in the primary analysis population. This population included 65 patients with CRC and 760 with AA. All patients who were excluded from the analyses were assessed by the sponsor for bias.

6.2 Baseline Demographics

The baseline demographic characteristics for the Primary Effectiveness Population are presented in Table 3 below. As shown in the table, the average age of patients was 64.2

years old, and there was a slightly higher percentage of female patients (5,378/10,023, 53.7%) as compared with male patients (4,645 /10,023, 46.3%). The majority of patients were White (8,422/10,017, 84.1%), with a minority of Black or African American patients (1,071/10,017, 10.7%), or Hispanic or Latino (991/10,019, 9.9%). Average BMI was 28.83 and the majority of patients never smoked (5,531 /10,019, 55.2%). Two 49-year-old patients were included in the study, which is inconsistent with the intended user population. Both patients were true negatives on *Cologuard* and their inclusion did not significantly change performance estimates. Patients that were enrolled at POR sites were similar to those enrolled at non-POR sites and to the population as a whole.

Table 3: Baseline Demographics

Parameter Statistic	**All Enrolled (N=10023)**	**Specificity Subset (2-6) (N=9958)**	**Specificity Subset (3-6) (N=9198)**	**CRC Subset (N=65)**	**AA Subset (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	29.67	27.55 (4.861)

Parameter Statistic	All Enrolled (N=10023)	Specificity Subset (2-6) (N=9958)	Specificity Subset (3-6) (N=9198)	CRC Subset (N=65)	AA Subset (N=760)	FIT Secondary Effectiveness (N=65)
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)
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	21.77 (14.732)	21.13 (14.450)	28.47	27.93	28.47
Median	20.0	20.0	20.0	29.0	30.0	29.0

6.3 Analysis of Primary Study Objectives

The primary effectiveness population consists of 10,023 patients. To facilitate analysis of primary study objectives, these patients are cross-classified in a contingency table by *Cologuard* result and histology result in Table 4.

Table 4: Cross-Classification of 10,023 Patients by *Cologuard* Result and Histological Category, n (%)

EXACT	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)

Sensitivities and Specificities:

Sensitivity and specificity for CRC (categories 1 vs. 2-6) and AN (categories 1-2 vs. 3-6) the sensitivity for AA (category 2) are provided in Table 5.

Table 5: Sensitivities and Specificities of *Cologuard*, %, 2-Sided 95% CI (Fraction) (n=10,023)

Type	Specificity	Sensitivity
CRC	84.4, 83.7-85.1 (8405/9958)	92.3, 83.0-97.5 (60/ 65)
AN	86.6, 85.9-87.3 (7967/9198)	46.3, 42.9-49.8 (382/825)
AA		42.4, 38.8-46.0 (322/760)

Primary Analysis of *Cologuard* CRC sensitivity:

Cologuard CRC sensitivity for CRC was 92.3% (60/65) with 95% one-sided lower confidence bound 84.5% (Clopper-Pearson method). Thus, the primary study objective of *Cologuard* CRC sensitivity greater than 65% and a 95% one-sided lower confidence bound exceeding 65% was met.

Primary Analysis of *Cologuard* AN specificity:

Cologuard AN specificity (categories 3-6) was 86.6% (7967/9198), with 95% one-sided lower confidence bound 86.0% (Clopper-Pearson method). Thus, the primary study objective of *Cologuard* AN specificity greater than 85% with a 95% one-sided lower confidence bound exceeding 85% was met.

6.4 Secondary and Additional Analyses Comparing *Cologuard* with FIT

The secondary effectiveness population consists of 9,989 patients with available *Cologuard*, PolyMedco FIT, and histology results. To facilitate analysis of the secondary study objectives, basic contingency tables are provided in Table 6.

Table 6: Comparison of *Cologuard* with FIT (n=9,989)

Category 1, CRC

	FIT	FIT
EXACT	Negative	Positive
Negative	4	1
Positive	13	47

Category 2, AA

	FIT	FIT
EXACT	Negative	Positive
Negative	407	29
Positive	170	151

Categories 3-6

	FIT	FIT
EXACT	Negative	Positive
Negative	7787	149
Positive	908	323

Sensitivities and Specificities:

Sensitivity and specificity for CRC (categories 1 vs. 2-6) and for AN (categories 1-2 vs. 3-6) and the sensitivity for AA are provided for *Cologuard* and FIT in the secondary analysis population of 9,989 patients as shown in Table 7.

Table 7: Sensitivities and specificities for *Cologuard* and FIT (n=9989)

Cologuard

Type	Specificity	Sensitivity
CRC	84.4, 83.6-85.1 (8372/9924)	92.3, 83.0-97.5 (60/ 65)
AN	86.6, 85.9-87.3 (7936/9167)	46.4, 42.9-49.8 (381/822)
AA		42.4, 38.8-46.0 (321/757)

FIT

Type	Specificity	Sensitivity
CRC	93.4, 92.9-93.9 (9272/9924)	73.8, 61.5-84.0 (48/ 65)
AN	94.9, 94.4-95.3 (8695/9167)	27.7, 24.7-30.9 (228/822)
AA		23.8, 20.8-27.0 (180/757)

Secondary CRC Sensitivity Comparison:

In this population, CRC sensitivity was 92.3% (60/65) for *Cologuard* and 73.8% (48/65) for PolyMedco FIT, resulting in a difference of 18.5%. The one-sided 95% lower confidence bound on the difference was 8.0%. Thus, the secondary study objective that *Cologuard* was non-inferior to Poly FIT in CRC sensitivity was met with respect to

protocol-specified non-inferiority margin 5.0% because the lower bound 8.0% is greater than –5.0%.

The two CRC sensitivities were significantly different, as indicated by a McNemar test p value p=0.002. Thus, *Cologuard* sensitivity can be declared superior to FIT sensitivity. FDA calculates that the difference 18.5% in CRC sensitivity has two-sided 95% CI [7.3%, 30.4%] (exact inference method).

Secondary AA Sensitivity Comparison:

The sensitivity to detect AA was 42.4% (321/757) for *Cologuard* and 23.8% (180/757) for PolyMedco FIT, resulting in a difference of 18.6%. The McNemar test p value was < 0.001, indicating the two AA sensitivities were significantly different. Thus, the secondary study objective that *Cologuard* was superior to Poly FIT in AA sensitivity was met. FDA calculates that the difference 18.6% in AA sensitivity has two-sided 95% CI [15.3%, 22.1%] (asymptotic inference method).

AN Specificity Comparison:

Cologuard was compared with PolyMedco FIT on AN specificity (categories 3-6). AN specificity was 86.57% (7936/9167) for *Cologuard* and 94.85% (8695/9167) for FIT. The difference –8.28% is significantly less than zero (95% CI –8.97, –7.61) (asymptotic inference method).

Per FDA request, *Cologuard* was compared with PolyMedco FIT on CRC specificity (categories 2-6). CRC specificity was 84.36% (8372/9924) for *Cologuard* and 93.43% (9272/9924) for FIT. The difference –9.07% is significantly less than zero (95% CI –9.76, –8.40) (asymptotic inference method).

Sub-categories of AA:

For adenoma with carcinoma in situ or high grade dysplasia (Category 2.1) sensitivity was 69.2% (27/39) for *Cologuard* compared with 46.2% (18/39) for FIT, representing a sensitivity advantage for *Cologuard* in this sub-category. For serrated lesions and polyps (Category 2.4), sensitivity was 42.4% (42/99) for *Cologuard* compared with 5.1% (5/99) for FIT, representing a sensitivity advantage for *Cologuard* in this sub-category as well.

Full Stratification of Data by Six Histological Categories:

For completeness, a fully stratified contingency table cross-classifying patients by *Cologuard*, FIT, and each of the six histology categories is provided Table 8.

Table 8: Study Cross-Classification of 9,989 patients by *Cologuard*, PolyMedco FIT, and Histological Category, n (%)

Cologuard	FIT	Category 1 (CRC)	Category 2 (AA)	Category 3 (1-2 >5 mm)	Category 4 (≥3 <10 mm)	Category 5 (1-2 ≤5 mm)	Category 6 (6.1 or 6.2)
Negative	Negative	4 (6.2)	407 (53.8)	584 (78.5)	287 (68.5)	1461 (84.5)	5455 (86.9)
Negative	Positive	1 (1.5)	29 (3.8)	18 (2.4)	15 (3.6)	30 (1.7)	86 (1.4)
Positive	Negative	13 (20.0)	170 (22.5)	96 (12.9)	72 (17.2)	173 (10.0)	567 (9.0)
Positive	Positive	47 (72.3)	151 (19.9)	46 (6.2)	45 (10.7)	66 (3.8)	166 (2.6)

After histological category in Table 8 was collapsed to CRC (category 1) or non-CRC (categories 2-6), FDA fit a log-linear model the data to evaluate the structure of the associations between *Cologuard*, FIT, and histological category that is revealed by the data. All of the two-way interactions between *Cologuard*, FIT, and binary histological category (non-CRC, CRC) were highly statistically significant ($p < 0.0001$). This model did not exhibit lack of fit relative to the model with the three-way interaction (p value 0.7793).

To assess if *Cologuard* informed for CRC status beyond FIT, FDA performed a logistic regression of binary CRC status on binary predictors for *Cologuard* and FIT (n=9,989). The coefficient on *Cologuard* was highly statistically significant ($p < 0.001$), indicating it was informative for CRC status after adjustment for FIT.

Marginal tables cross-classifying *Cologuard* by histology and FIT by histology are provided in Tables 9 and 10.

Table 9: *Cologuard* by Histological Category

Cologuard	Category 1 (CRC)	Category 2 (AA)	Category 3 (1-2 >5 mm)	Category 4 (≥3 <10 mm)	Category 5 (1-2 ≤5 mm)	Category 6 (6.1 or 6.2)
Negative	5 (7.7)	436 (57.6)	602 (80.9)	302 (86.2)	1491 (86.2)	5541 (88.3)
Positive	60 (92.3)	321 (42.4)	142 (19.1)	117 (13.8)	239 (13.8)	733 (11.7)

Table 10: FIT by Histological Category

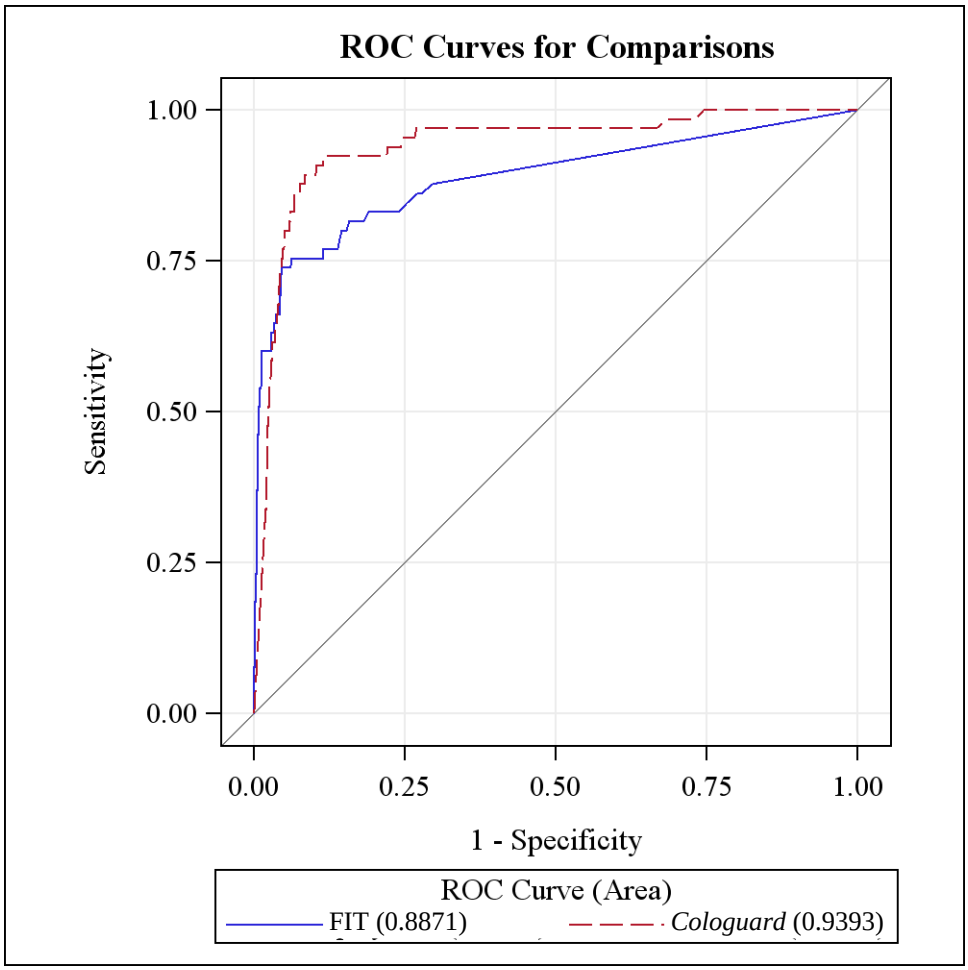
FIT	Category 1 (CRC)	Category 2 (AA)	Category 3 (1-2 >5 mm)	Category 4 (≥3 <10 mm)	Category 5 (1-2 ≤5 mm)	Category 6 (6.1 or 6.2)
Negative	17 (26.2)	577 (76.2)	680(80.9)	359 (85.7)	1634 (94.5)	6022 (96.0)
Positive	48 (73.8)	180 (23.8)	64 (9.1)	60 (14.3)	96 (5.5)	252 (4.0)

Chi-square hypothesis testing reveals that the distribution of histological category is highly significantly different between *Cologuard* negative and positive test results (p value < 0.001) and between Poly FIT negative and positive test results (p value < 0.001).

ROC Analysis:

Receiver operating characteristic curves ("ROC curve") were generated plotting the CRC sensitivity (category 1) vs. AN false positive fraction, i.e., 1 – AN specificity (categories 3-6) for *Cologuard* and for PolyMedco FIT by varying the cut-off for each as shown in Figure 1. This ROC analysis corresponds to the primary analysis in which AA patients (category 2) are excluded. The area under the ROC curve (AUC) is the probability that a randomly selected CRC patient has a test value greater than a randomly selected non-AN patient (categories 3-6). For this analysis, AUC was 93.9% for *Cologuard* and 88.7% for FIT. The two sided p-value for the difference was statistically significant (p=0.0372).

Figure 1: ROC Curves for CRC Sensitivity (Category 1) vs. 1 – AN Specificity (Categories 3-6): *Cologuard* vs. PolyMedco FIT (n = 9,176)



For the 9989 patient dataset, FDA generated the ROC curves of CRC sensitivity vs. CRC false positive fraction (1 – CRC specificity) for *Cologuard*, for PolyMedco FIT, and for the FIT component of *Cologuard*, referred to in this executive summary as EXACT FIT. The ROC curves are shown in Figure 2 and the analysis is summarized in Table 11. The AUC for this ROC curve is the probability that a randomly selected CRC patient (category 1) has a test value greater than a randomly selected non-CRC patient (categories 2-6). The AUC was 92.98% for *Cologuard*, 91.94% for EXACT FIT, and 88.02% for PolyMedco FIT. AUC was significantly greater for *Cologuard* than for PolyMedco FIT (p = 0.0496), significantly greater for EXACT FIT than PolyMedco FIT (p = 0.0292), but not significantly greater for *Cologuard* than EXACT FIT (p=0.5507). The Figure also superimposes on the ROC curves the CRC sensitivity / CRC false positive fraction pair at the operating points or cut-off for each test: 183 for *Cologuard* score and 101 ng/mL for PolyMedco FIT. To permit this analysis, a cut-off of 204 ng/mL was used for EXACT FIT, however, the EXACT FIT component of *Cologuard*

does not use a “cut-off” within the context of the test algorithm. That operating point for EXACT FIT was chosen to match PolyMedco FIT on specificity to facilitate comparison of the two FIT tests, separate from the overall *Cologuard* score.

Figure 2: ROC Curves for CRC sensitivity (category 1) vs. 1 – CRC specificity (categories 2-6) for *Cologuard* composite score, PolyMedco FIT (FIT), and *Cologuard* FIT component of *Cologuard* (EXACT FIT), n=9989. The pair (CRC sensitivity, 1 – CRC specificity) is also plotted at cut-offs 183 for *Cologuard*, 204 ng/mL for EXACT FIT, and 101 ng/mL for PolyMedco FIT.

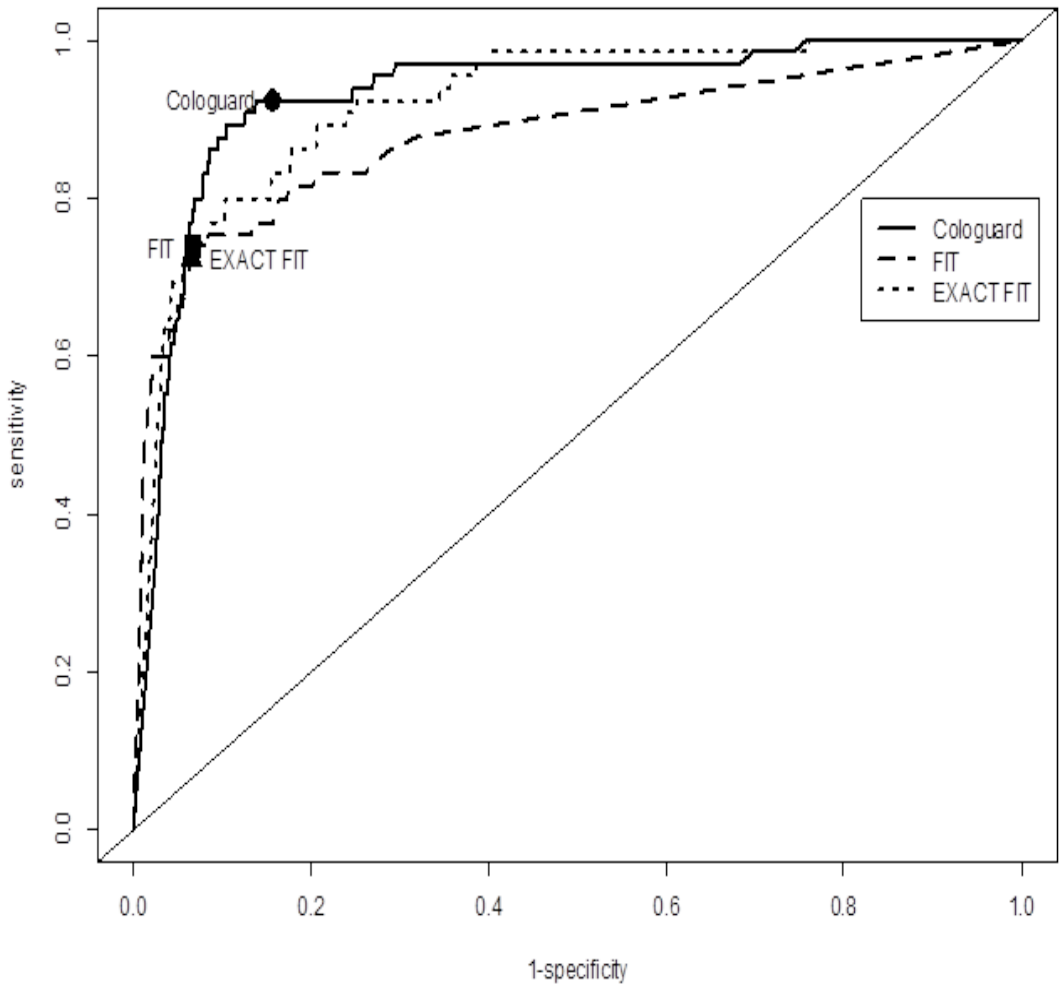


Table 11: Comparison of AUC (%) between *Cologuard* composite score, FIT component of *Cologuard* (EXACT FIT), and PolyMedco FIT (PolyMedco FIT) (n=9,989)

Comparison	AUC1	AUC2	Difference	Wald 95% CI†	p-value†
Cologuard – Poly FIT	93.0	88.0	5.0	0.03, 9.9	0.0485
EXACT FIT – Poly FIT	91.9	88.0	3.9	0.4, 7.5	0.0289
Cologuard – EXACT FIT	93.0	91.9	1.0	-2.4, 4.5	0.5507

†Based on standard error and covariance of the Mann-Whitney U-statistics.

In Figure 2, *Cologuard*, PolyMedco FIT, and EXACT FIT are compared on the AUC of the CRC ROC curve (CRC sensitivity vs. 1 – CRC specificity). Alternatively, the three tests were compared on the AUC of the AN ROC curve (AN sensitivity vs. 1 – AN specificity), as depicted in Figure 3. The AN AUC was significantly larger for *Cologuard* (73.3%) than PolyMedco FIT (66.7%) and EXACT FIT (69.3%), with p-values < 0.0001 and 0.0002, respectively (Table 12).

Figure 3: AN ROC curves (AN sensitivity vs. 1 - AN specificity) for *Cologuard* composite score, PolyMedco FIT (FIT), and *Cologuard* FIT component of *Cologuard* (EXACT FIT), n=9989

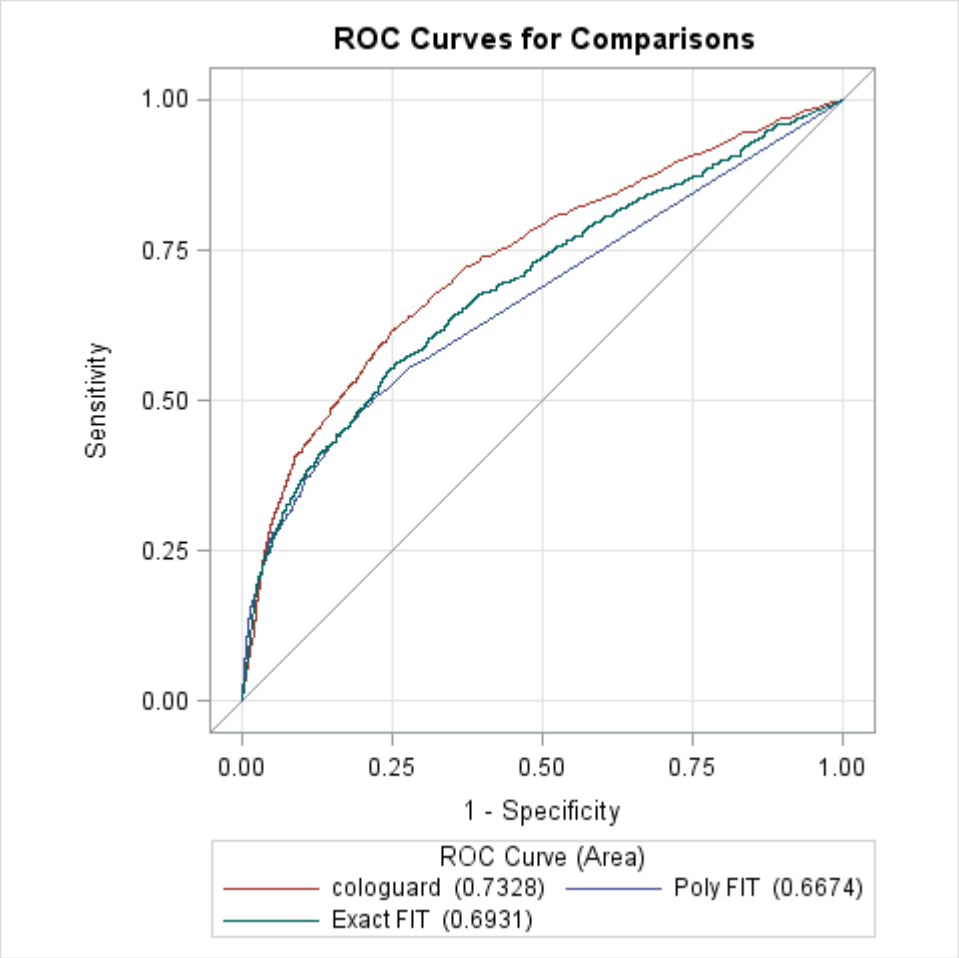


Table 12: Comparison of AUC (%) between *Cologuard* composite score, FIT component of *Cologuard* (EXACT FIT), and PolyMedco FIT (PolyMedco FIT), n=9989

Comparison	AUC1	AUC2	Difference	Wald 95% CI†	p-value†
Cologuard – Poly FIT	73.3	66.7	6.5	4.4, 8.7	<0.0001
EXACT FIT – Poly FIT	69.3	66.7	2.6	0.9, 4.2	0.0020
Cologuard – EXACT FIT	73.3	69.3	4.0	1.9, 6.0	0.0002

†Based on standard error and covariance of the Mann-Whitney U-statistics.

In another analysis, FDA employed logistic regression to consider if the binary results of the tests based on their cut-offs are significant predictors of binary CRC status (CRC vs. non-CRC). When all three tests are included in the model as binary predictors, EXACT FIT was not a significant predictor ($p = 0.2104$) but *Cologuard* ($p < 0.0001$) and PolyMedco FIT ($p < 0.0001$) were.

Benefit-Risk Analysis:

FDA performed a benefit-risk analysis comparing the diagnostic yield of *Cologuard* with PolyMedco FIT in a screening population of 100,000 patients. Diagnostic yield was projected based on the prevalence of CRC, AA, and non-AN (not CRC or AA), estimated from the 10,840 patients with a valid histological category and based on the fractions of these three patient groups who tested positive by the two tests using the 9,989 patients on whom valid results were available for both tests (Table 13).

Table 13: Positive Fraction of *Cologuard* and PolyMedco FIT by Histological Type (n=9989)

Histological Type	Prevalence (n=10840)	*Cologuard* positive fraction (n=9989)	Poly FIT positive fraction (n=9989)
CRC	0.70% (76/10840)	92.31% (60/65)	73.85% (48/65)
AA	7.58% (822/10840)	42.40% (321/ 757)	23.78% (180/ 757)
Not CRC or AA	91.72% (9942/10840)	13.43% (1231/9167)	5.15% (472/9167)

Among 100,000 patients in a screening population, 700, 7580, and 91,720 are expected to be CRC, AA, and not AN (neither CRC nor AA). *Cologuard* is expected to test positive for 129 more CRC patients than PolyMedco FIT as shown in Table 14. These patients may be referred to colonoscopy as a next step. *Cologuard* is also expected to test positive for 1,412 more AA patients than PolyMedco FIT. However, *Cologuard* is expected to falsely test positive for 7,594 more non-AN patients than PolyMedco FIT. From these numbers, compared with PolyMedco FIT, *Cologuard* is expected to detect 1 more CRC patient and 11 more AA patients for every 59 more false positives on non-AN patients (rounded to nearest integer).

Table 14: Comparison of for *Cologuard* and PolyMedco FIT on Expected Diagnostic Yield in a Screening Population (hypothetical n=100,000 screening population)

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

Since a cross sectional study addresses test sensitivity but not screening program sensitivity, FDA would like the panel to discuss issues of interpreting performance from a cross sectional study in the context of a screening program with patients using the test for the first time as well as follow-up implications. FDA seeks advisory committee perspective on the acceptability of this tradeoff of sensitivity and specificity for initial testing without longitudinal repeat performance data.

PANEL DISCUSSION QUESTION #1

The Deep-C study met the primary objectives with respect to both required sensitivity and specificity of *Cologuard* compared to colonoscopy, with 92.3% sensitivity for CRC, 42.4% sensitivity for AA, and 86.6% specificity for AN.

With respect to the secondary objectives, *Cologuard* sensitivity was higher than FIT for both CRC and AA (92.3% vs. 73.8% and 42.4% vs. 23.8%, respectively). Although not a secondary objective, *Cologuard* AN specificity was lower than FIT (86.6% vs. 94.9%).

a. Do these conclusions adequately demonstrate effectiveness of *Cologuard* within the contexts of the proposed intended use and current recommendations for CRC screening?

b. Based on the results of the pivotal clinical study, do the data provided allow for adequate assessment of the benefits and risks of *Cologuard*?

6.5 Additional Effectiveness Analyses

6.5.1 Likelihood Ratios and Predictive Values

Likelihood Ratios:

In addition to sensitivity and specificity for CRC and AN, the positive and negative likelihood ratios for *Cologuard* were calculated from the study data. The positive likelihood ratio (PLR) is sensitivity / (1 – specificity), the ratio of the true positive fraction to the false positive fraction. The negative likelihood ratio (NLR) is (1 – sensitivity) / specificity, the ratio of the false negative fraction to the true negative fraction. Larger values of PLR and smaller values of NLR indicate better classification of disease status. Additionally, the positive predictive value (next subsection) is monotone increasing as PLR increases and the negative predictive value is monotone decreasing as NLR increases.

For the 10,023 patient dataset, positive and negative likelihood ratios for *Cologuard* were generated using CRC sensitivity (category 1) and AN specificity (categories 3-6) (Table 15). This analysis corresponds to the primary analysis in which AA cases (category 2) are excluded. Based on *Cologuard* CRC sensitivity 92.3% (60/65) and AN specificity 86.6% (7967/9198), the positive likelihood ratio is 6.9 (95% CI 6.3, 7.5), indicating that a person with CRC would be 6.9 times more likely to have a positive *Cologuard* result than someone without CRC or AA. The negative likelihood ratio is 0.089 (95% CI 0.38, 0.21), indicating that someone without CRC or AA is approximately 11 times (1/0.089) more likely to test negative on *Cologuard* than someone with CRC.

Likelihood ratios may also be evaluated on complementary pairs of sensitivity and specificity that include all histological categories 1-6 as shown in Table 15. Using *Cologuard* CRC sensitivity 92.3% (60/65) and CRC specificity 84.4% (8405/9958), the positive likelihood ratio is 5.9 (95% CI 5.4, 6.4) and the negative likelihood ratio is 0.0911 (0.04, 0.21). Using *Cologuard* AN sensitivity 46.3% (382/825) and AN specificity 86.6% (7967/9198), the positive likelihood ratio is 3.5 (95% CI 3.2, 3.8) and the negative likelihood ratio is 0.62 (95% CI 0.58, 0.66).

For the 9,989 patient dataset, *Cologuard* and PolyMedco FIT may be compared on positive and negative likelihood ratios using complementary pairs of sensitivity and specificity (Table 16). On this dataset, CRC sensitivity and CRC specificity are 92.3% (60/65) and 84.4% (8372/9924) for *Cologuard* and 73.8% (48/65) and 93.4% (9272/9924) for FIT, leading to positive and negative likelihood ratios of 5.9 (95% CI 5.4, 6.4) and 0.091 (95% 0.04, 0.21) for *Cologuard* and 11.2 (95% CI 9.6, 13.1) and 0.28 (95% CI 0.19, 0.42) for FIT. The greater PLR for FIT than *Cologuard* implies that *Cologuard* has a smaller CRC positive predictive value than FIT. However, the greater

NLR for FIT than *Cologuard* implies that *Cologuard* has a greater CRC negative predictive value than FIT.

Table 15: *Cologuard* Positive and Negative Likelihood Ratios (PLR, NLR) for complementary pairs of CRC and AN sensitivity and specificity (n=10,023)

Type	Specificity	Sensitivity	PLR %, 95% CI	NLR %, 95% CI
CRC	84.4% (8405/9958)	92.3 (60/ 65)	5.9 (5.4, 6.4)	0.0911 (0.04, 0.21)
AN	86.6% (7967/9198)	46.3% (382/825)	3.5 (3.2, 3.8)	0.62 (0.58, 0.66)

Table 16: Positive and Negative Likelihood Ratios (PLR, NLR) for complementary pairs of CRC and AN sensitivity and specificity, *Cologuard* vs. PolyMedco FIT (n=9,989)

Cologuard

Type	Specificity	Sensitivity	PLR %, 95% CI	NLR %, 95% CI
CRC	84.4 (8372/9924)	92.3 (60/ 65)	5.9 (5.4, 6.4)	0.091 (0.04, 0.21)
AN	86.6 (7936/9167)	46.4 (381/822)	3.5 (3.2, 3.8)	0.62 (0.58, 0.66)

FIT

Type	Specificity	Sensitivity	PLR %, 95% CI	NLR %, 95% CI
CRC	93.4 (9272/9924)	73.8 (48/ 65)	11.2 (9.6, 13.1)	0.28 (0.19, 0.42)
AN	94.9 (8695/9167)	27.7 (228/822)	5.4 (4.7, 6.2)	0.76 (0.73, 0.80)

Predictive Values:

FDA calculated the predictive values of a *Cologuard* positive result for CRC, AA, and categories 3-6 and the data is presented in Table 17. The positive predictive values are the fractions of these patients that test positive by *Cologuard*. The negative predictive values are the fractions that test negative. The positive predictive value for CRC is 3.72% (60/1613). Considering that CRC prevalence in the study is 0.65% (65/10023), a test positive patient is estimated to be 5.7 times more likely to have CRC than the general population (3.72/0.65). Likewise, considering that the PPV for AA is 19.96% and that AA prevalence is 7.58% (760/10023), a test positive patient is estimated to be 2.6 times more likely to have AA than the general population (19.96/7.58).

Table 17: Predictive Values of *Cologuard*, %, 2-sided 95% CI (Fraction), n=10023

Cologuard	CRC Category 1	AA Category 2	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)

Adjustment for Missing Cologuard Results:

The sponsor provided an accounting of patients with missing histology, *Cologuard* results, and FIT results. A number of patients did not have a *Cologuard* result for various

reasons (Section 6.1). FDA compared the distribution of histological category between those with and without *Cologuard* results (Table 19).

Table 19: Distribution of Histological Category for Patients With/Without *Cologuard* Results

Status of *Cologuard* result	N	Category 1 (CRC)	Category 2 (AA)	Category 3 (1-2 >5 mm)	Category 4 (≥3 <10 mm)	Category 5 (1-2 ≤5 mm)	Category 6 (6.1 or 6.2)
Missing	817	11	58	66	26	156	500
no stool sample	128	3	11	17	6	23	68
stool sample unusable	474	5	35	33	13	98	290
Invalid	213	3	12	15	7	35	141
unanalyzed	2	0	0	1	0	0	1
Valid	10023	65	760	749	419	1735	6295
Negative	8410	5	438	607	302	1496	5562
Positive	1613	60	322	142	117	239	733

FDA adjusted the estimates of *Cologuard* performance (sensitivity, specificity, predictive value) for missing *Cologuard* data by assuming that within each histologic category, the distribution of *Cologuard* missing results, when available, would be expected to have the same distribution as *Cologuard* non-missing results. The data are then said to be *missing at random*. To obtain estimates of performance that include patients without a *Cologuard* result (n=817) as well as patients with a *Cologuard* result (n=10023), patients with missing results were distributed (fractionally) within histological category across *Cologuard* negative and positive categories in the same proportion as was observed for patients with *Cologuard* results. This imputation of the missing results is shown in Table 20.

Table 20: FDA Adjusted Estimates of *Cologuard* Performance Including Missing *Cologuard* Data

Status of *Cologuard* result	N	Category 1 (CRC)	Category 2 (AA)	Category 3 (1-2 >5 mm)	Category 4 (≥3 <10 mm)	Category 5 (1-2 ≤5 mm)	Category 6 (6.1 or 6.2)
Missing	817	11	58	66	26	156	500
Imputed Positive	685.52	0.85	33.43	53.49	18.74	134.51	441.78
Imputed Negative	131.48	10.15	24.57	12.51	7.26	21.49	58.22
Valid	10023	65	760	749	419	1735	6295
Negative	8410	5	438	607	302	1496	5562
Positive	1613	60	322	142	117	239	733
Combined	10840	76	818	815	445	1891	6795
Negative	9095.52	5.85	471.43	660.49	320.74	1630.51	6003.78
Positive	1744.48	70.15	346.57	154.51	124.26	260.49	791.22

Adjusted for missing data, *Cologuard* CRC sensitivity is unchanged, 92.3% (70.15/76) with 95% one-sided lower confidence bound 84.5%. Thus, the adjusted *Cologuard* CRC sensitivity is greater than 65% with statistical significance, a primary study objective. The unchanged 97.5% one-sided lower confidence bound is 83.0%, which also exceeds 65%.

Adjusted for missing data, *Cologuard* AN specificity (categories 3-6) was 86.6%, with 95% one-sided lower confidence bound 86.0% (FDA-computed). Thus, the adjusted *Cologuard* AN specificity is greater than 85% with statistical significance, a primary study objective. The 97.5% one-sided lower confidence bound is 85.9% (FDA-computed), which also exceeds 85%.

Adjusted for missing data, several types of sensitivity and specificity were estimated and shown in Table 21.

Table 21: Sensitivity and Specificity of *Cologuard*, Adjusted for Missing *Cologuard* Results, % (2-sided 95% CI†, n=10840)

	Cologuard	*Cologuard*
Type	Specificity	Sensitivity
CRC	84.4% (83.7, 85.1)	92.3% (84.1, 96.9)
AN	86.6% (85.9, 87.3)	46.6% (43.2,50.0)
AA		42.4% (38.9, 45.9)

† 95% CIs were obtained using a Bayesian imputation model

Adjusted for missing data, predictive values of a *Cologuard* positive result for CRC, AA, and categories 3-6 were estimated and shown in Table 22.

Table 22: Predictive Values (%) of *Cologuard* (95% CI), Adjusted for Missing *Cologuard* Results, % (95% CI†) (n=10840)

Cologuard	CRC, Category 1	AA, Category 2	Categories 3-6
Negative	0.07 (0.02,0.14)	5.18 (4.73, 5.66)	94.75 (94.27, 95.21)
Positive	3.99 (3.12, 4.95)	19.83 (17.97, 21.70)	76.18 (74.18, 78.15)

† 95% CIs were obtained using a Bayesian imputation model.

6.5.2 Sub-Group Analyses

The DeeP-C study results were also analyzed according to various demographic characteristics, as well as lesion size and location.

Results by Gender:

Sensitivity of *Cologuard* was higher for males than for females, for both CRC and AA. As shown in Table 23 below, *Cologuard* sensitivity for CRC was 100.0% (34/34) for males, and 83.9% (26/31) for females, a statistically significant difference (exact p value = 0.02057 using Fisher-Freeman-Halton method). Sensitivity for AA was 44.7% (201/450) for males, and 39.0% (121/310) for females, an insignificant difference (p = 0.1353).

Table 23: *Cologuard* Positive Fractions (%) by Gender, n=10,023.

Subgroup	CRC	AA	Categories 3-6
Male	100.0 (34/34)	44.7 (201/450)	14.2 (592/4161)
Female	83.9 (26/31)	39.0 (121/310)	12.7 (639/5037)

Meanwhile, AN specificity (categories 3-6) for *Cologuard* (100% – the positive fraction in Table 23) was 85.8% (3,569/4,161) for males and 87.3% (4,398/5,037) for females. While similar, the difference was still statistically significant (p = 0.0313).

Results by Race and Ethnicity:

With respect to race, *Cologuard* sensitivity for CRC was higher among White patients (53/55, 96.4%), than among Black/African-American patients (5/8, 62.5%), with a statistically significant difference of p = 0.01239 (Table 24). The single Asian CRC case was detected (1/1, 100.0%) as was the Other case (1/1, 100.0%). Taking all the data together, variation in CRC sensitivity by race group was statistically significant (p = 0.02618). Nonetheless, the results observed in Black or African- American patients may have been driven by the low overall number of cancer cases in that subpopulation.

CRC sensitivity was higher for Hispanic or Latino patients (8/9, 88.9%) than for Black/African-Americans, but lower than for Whites. When CRC sensitivity for Hispanic or Latino patients is compared with CRC sensitivity for Not Hispanic or Latino patients (92.9%, 52/56), the difference is insignificant (p = 0.5375). As shown in Table 19 below, sensitivity for AA was similar for White (271/641 42.3%) and Black/African-American (36/85, 42.4%) patients. Variation in AA sensitivity did not significantly vary by race group (p = 0.6964). Sensitivity was also similar among Hispanic/Latino patients (23/59, 39.0%). *Cologuard* sensitivity for AA was lower among Asian patients (4/13, 30.8%) and high for American Indian or Alaskan Natives (3/4, 75.0%), compared with other groups.

Table 24: *Cologuard* Positives Fraction (%) by Race (n=10,023)

Subgroup	CRC	AA	Categories 3-6
White	96.4 (53/55)	42.3 (271/641)	14.1 (1087/7726)
Black or African American	62.5 (5/8)	42.4 (36/85)	10.1 (99/978)
Asian	100.0 (1/1)	30.8 (4/13)	6.5 (16/245)
American Indian or Alaska Native	NA (0/0)	75.0 (3/4)	25.0 (8/32)
Native Hawaiian or other Pacific Islander	NA (0/0)	NA (0/0)	8.7 (2/23)
Other	100.0 (1/1)	43.8 (7/16)	9.5 (18/189)

Cologuard specificity for CRC was > 85% for all racial and ethnic groups other than American Indian/Alaska Native. Specificity was 93.5% (229/245) for Asian patients, 91.3% (21/23) for Native Hawaiian/Pacific Islander patients, and 90.7% (837/923)

among Hispanic or Latino patients. Specificity was similar for White (6,639/7,726, 85.9%) and Black/African-American (879/978, 89.9%) patients in this analysis, and lowest for American Indian/Alaskan Native patients (24/32, 75.0%). However, variation in AN specificity varied significantly by race group ($p < 0.001$), a result due more to large sample size than large differences in AN specificity.

Results by Age:

Cologuard sensitivity for CRC was consistent across all age groups as shown in Table 25. Sensitivity for patients 65 years of age and older ranged from 88.9% to 100.0%. Although sensitivity was 75% for patients age 60-64, the number of CRC cases was small in this age group (n = 4), and one CRC case was not detected by *Cologuard*. Variation in CRC sensitivity did not significantly vary by age group ($p = 0.5972$). With respect to AA, sensitivity was similar across all age groups, with sensitivity as high as 46.8% for patients between the ages of 70 and 79. The variation was not statistically significant ($p = 0.6556$).

Cologuard specificity for AN (CRC and AA) was highest for younger patients and lower for older patients. Specificity was greater than 80% for all age groups other than patients > 75 years old. Variation in AN specificity by age group was highly statistically significant ($p < 0.001$).

Table 25: *Cologuard* Positive Fraction (%) by Age Group (n=10,023)

Subgroup	CRC	AA	Categories 3-6
<60 years	100.0 (7/7)	38.0 (65/171)	7.8 (212/2703)
60-64 years	75.0 (3/4)	42.1 (24/57)	11 (84/765)
65-69 years	95.0 (19/20)	41.5 (125/301)	14.3 (481/3352)
70-74 years	88.9 (16/18)	46.8 (72/154)	17.5 (274/1566)
75-79 years	100.0 (6/6)	46.8 (29/62)	22.2 (137/617)
80-84 years	90.0 (9/10)	46.7 (7/15)	22.1 (43/195)

The study was designed with a goal to enroll at least 75% of patients age 65-84 to increase the point prevalence of CRC. Because the study was enriched for older age patients, *Cologuard* sensitivity and specificity for CRC may be adjusted to the distribution of age categories from the US census population (Table 26).

Age adjusted CRC sensitivity is 90.9% (FDA 95% CI 79.3-97.6) for *Cologuard*, compared with unadjusted CRC sensitivity 92.3% (95% CI 83.0-97.5). Age adjusted CRC specificity (categories 2-6) is 85.8% (FDA 95% CI 85.0-86.6), compared with unadjusted CRC specificity 84.4% (95% CI 83.7-85.1).

Table 26: *Cologuard* CRC Sensitivity (SE) and CRC Specificity (SP) Adjusted to Age Distribution of US Census Population (n=10,023)

		Non-CRC		CRC					
Age	US %†	Test –	Test +	Test –	Test +	n	%	SE	SP
50-59	44.56	2597	277	0	7	2881	28.7	100.0	90.4
60-64	18.71	714	108	1	3	826	8.2	75.0	86.9
65-69	13.05	3047	606	1	19	3673	36.7	95.0	83.4
70-74	9.93	1374	346	2	16	1738	17.3	88.9	80.0
75-79	7.61	513	166	0	6	685	6.8	100.0	75.5
80-84	6.14	160	50	1	9	220	2.2	90.0	76.2
Observed		8405	1553	5	60	10023	100.0%	92.3	84.4
Weighted to US		8537.8	1413.2	6.6	65.4	10023		90.9††	85.8††

*Categories 2-6; †Based on 2011 US Census Data. ††Misreported to sponsor as 93.8 and 87.3 at time of the printing of their Executive Summary.

Similarly, *Cologuard* sensitivity and specificity for AN may be adjusted to the distribution of age categories from the US census population (not shown).

Age-adjusted positive and negative likelihood ratios may also be computed for *Cologuard* using age-adjusted sensitivity and specificity for either CRC or AN (not shown).

Results by Lesion Size and Cancer Stage:
Exact Sciences evaluated *Cologuard* results by lesion size, as well as cancer stage. Sensitivity of *Cologuard* decreased with lesion or lesion size. The amount of DNA shed from cancerous or pre-cancerous tissue in the colon is generally thought to increase with increased mass or lesion size.

CRC sensitivity was > 90% for most lesion sizes. Sensitivity for CRC was highest for patients with CRCs ≥ 30 mm (32/34, 94.1%) and lowest for patients with CRCs 5-9 mm in size (4/5, 80.0%). Sensitivity by cancer stage was the highest for patients with Stage II cancers (21/21, 100.0%) and Stage III cancers (9/10, 90%). Sensitivity of *Cologuard* for AA was also higher among patients with AAs of larger sizes.

Specificity of *Cologuard* for CRC was 86.2% (1,847/2,142), for patients with CRCs < 5 mm in size, and 79.7% (1,523/1,912) for patients with CRCs 5-9 mm in size. These results are summarized in Tables 27 and 28.

Table 27: *Cologuard* Positive Fraction (%) by Lesion Size

Largest lesion size	CRC	AA	Categories 3-6
<5 mm	NA (0/0)	20.0 (2/10)	13.8 (295/2142)
5-9 mm	80.0 (4/5)	32.1 (18/56)	20.3 (389/1912)
10-19 mm	92.9 (13/14)	39.0 (225/577)	0/0
20-29 mm	91.7 (11/12)	64.6 (51/79)	0/0
≥ 30 mm	94.1 (32/34)	68.4 (26/38)	0/0

Table 28: *Cologuard* Positive Fraction (%) by Stage

Stage	CRC
I	89.7 (26/29)
II	100.0 (21/21)
III	90.0 (9/10)
IV	75.0 (3/4)

Results by Lesion Location:
Cologuard results also were assessed by lesion location. As shown in Table 29, sensitivity of *Cologuard* for CRC was 90% or greater, regardless of lesion location. Sensitivity of *Cologuard* for AA was greatest among patients with distal AAs (133/238, 55.9%).

Table 29: *Cologuard* Positive Fraction (%) by Lesion Location

Lesion Location	CRC	AA	Categories 3-6
Proximal	90.0 (27/30)	33.0 (143/433)	16.6 (343/2066)
Distal	91.7 (22/24)	55.9 (133/238)	17.9 (246/1377)
Rectal	100.0 (11/11)	51.1 (45/88)	15.5 (95/612)

Specificity of *Cologuard* for CRC was similar for different lesion locations: 83.4% for patients with proximal CRCs, 82.1% for patients with distal CRCs, and 84.5% for patients with rectal CRCs.

6.6 Labeling Considerations

The study design was focused on patients of average risk who agreed to participate in screening by colonoscopy. It was not designed to yield performance information when used as a substitute for colonoscopy in settings of heightened clinical concern including high risk patients (e.g. predisposition due to genetics or gastrointestinal disease), diagnostic colonoscopy (e.g. patients with signs or symptoms), or surveillance colonoscopy (e.g. patients with personal history of colon cancer or polyps). Performance of the device in patients who refused screening colonoscopy cannot be determined from this study.

The DeeP-C study was designed to include patients ages 50-84; performance of the device in patients ages 85 and above cannot be determined from the study. In addition, the studies presented were not designed to evaluate subgroups and subgroup analysis should be interpreted with that in mind. FDA requests panel feedback regarding device labeling to emphasize device performance across groups.

PANEL DISCUSSION QUESTION #2:
Are there patient subgroups, such as age (e.g., ages 75-79, 80-84, 85 and above), gender, and race/ethnicity where considerations for device performance merit additional labeling?

Since the sponsor has not provided longitudinal performance for *Cologuard*, there is uncertainty regarding optimal follow-up of a negative result. Note that performance evaluation of *Cologuard* in the studies presented here was in the context of annual FIT. The lack of data regarding device performance in patients previously testing negative may prompt a preference for additional testing including other methods with consideration of medical guidelines.

PANEL DISCUSSION QUESTIONS #3:
The Deep-C study conducted by Exact Sciences was not designed to provide follow-up data on patients that tested negative with *Cologuard*.

a. What is appropriate labeling to assure safety and effectiveness for follow-up evaluation of patients testing negative with *Cologuard?* The FDA would like feedback on follow-up test interval and modality, use of guidelines, and other possible follow-up approaches.

7 POST-APPROVAL STUDY

The inclusion of a Post-Approval Study section in this summary should not be interpreted to mean that FDA has made a decision or is making a recommendation on the approvability of this PMA device. The presence of a post-approval study plan or commitment does not in any way alter the requirements for premarket approval and a recommendation from the Panel on whether the risks outweigh the benefits. The premarket data must reach the threshold for providing reasonable assurance of safety and effectiveness before the device can be found approvable and any post-approval study could be considered. The issues noted below are FDA's comments regarding potential post-approval studies, for the Panel to include in the deliberations, should FDA find the device approvable based upon the clinical premarket data.

The FDA review team has discussed with the sponsor that if *Cologuard* is approved, a post-approval study (PAS) may be required as a condition of approval for this device. Through review of the Premarket Data, FDA has identified the following postmarket concerns and recommends that a PAS be conducted to address the programmatic performance in relationship to screening interval including:

- the negative to positive conversion rate
- the diagnostic yield
- and predictive values in postmarket setting

FDA and the sponsor have begun to work together to design this study. However, there are still outstanding issues that require panel input. An overview of the proposed PAS protocol is provided below.

7.1 Overview of Proposed Post- Approval Study

The DeeP-C trial was a point-in-time (cross-sectional) study that established *Cologuard's* performance characteristics in the intended use population. However, no evidence was generated to understand how *Cologuard* performs over the course of a screening program in an average risk population. The purpose of the study is to collect longitudinal data on patients for whom *Cologuard* has recently been prescribed.

The study objective is to collect longitudinal data annually on patients prescribed *Cologuard* over the course of 3 years and to assess the risk of CRC/AA among those with a positive *Cologuard* test at the third year of follow-up (T3) compared to baseline (T0).

The primary endpoint for this study is to assess the risk of CRC/AA among those with a positive *Cologuard* test at the third year of follow-up (T3) compared to baseline (T0).

The secondary objectives are to:

- To evaluate the distribution of colorectal epithelial lesions (by Category) among positive *Cologuard* patients at T0 and at T3
- To evaluate the predictive values of a positive *Cologuard* at T0 and at T3.

7.1.1 Sample Collection

A *Cologuard* Collection Kit will be distributed to each enrolled patient per investigator prescription. The Collection Kit is sent to the patient's home with instructions for stool collection and sample return. Patients should collect the stool sample by following the instructions provided.

Stool collection must be completed within 30 days of enrollment at T0 and within 90 days of T3 date (up to 3 years $\pm$90 days from T0).

All in-coming samples will be assessed for acceptability by the processing laboratory as defined by the Sponsor's Instructions for Use (IFU). If a sample does not meet the defined acceptance criteria, the laboratory will notify the ordering physician. Repeat stool collection must occur within the protocol defined window period. If a second sample cannot be collected within the defined protocol window, the patient may be replaced through additional enrollment.

7.1.2 Colonoscopy and Histopathology Procedure

Positive *Cologuard* results are to be followed by diagnostic colonoscopy. The time between a positive *Cologuard* result (T0) and colonoscopy may not exceed 90 days. The time between the *Cologuard* result (T3) and colonoscopy may not exceed 90 days. Bowel preparation procedures and colonoscopy will be performed following established standard practice at each clinical site.

Sites must acquire and maintain histopathologic reports of the interpretation of endoscopic biopsies, polypectomy specimens, and excisional surgical pathology specimens for all patients with tissue excised at colonoscopy.

Several measures will be taken to ensure high quality diagnostic colonoscopies in this study. A completed colonoscopy procedure for patients with no findings will be defined as reaching the cecum, unless otherwise noted. Cecal intubation will be documented with photographic evidence and/or documentation of cecal intubation.

Any tissue biopsied must be sent for histopathological review. Histopathological results will be categorized as defined in Appendix 4.

The sponsor may request access to histopathology slides and completed reports for review of clinical features. A sponsor review of the slides will not replace the site diagnosis and all analysis related to study objectives will be based on results of the histopathology examination conducted at the clinical sites.

7.1.3 Patient Selection

The study will enroll approximately 1,830 patients at approximately 20 sites. The following eligibility criteria are designed to select patients who are at average risk for colorectal cancer screening. All relevant medical and non-medical conditions should be taken into consideration when deciding whether this protocol is suitable for a particular patient. No waivers to these criteria will be permitted.

7.1.4 Inclusion and Exclusion Criteria

Patients must meet the following criteria at the time of enrollment to be eligible for the study:

- Patient is average risk for development of colorectal cancer
- Patient is 50 to 84 years of age inclusive
- Patient has not a had a colonoscopy in the previous 9 years
- Patient signs informed consent

Patients presenting with any of the following at the time of enrollment will not be included in the study:

- Patient has any condition that in the opinion of the investigator should preclude participation in the study (e.g., patient not eligible for a diagnostic colonoscopy).
- Patient has a history of colorectal cancer or advanced adenoma.
- Patient has a history of aerodigestive tract cancer
- Patient has had a prior colorectal resection for any reason other than sigmoid diverticular disease
- Patient has had overt rectal bleeding, e.g., hematochezia or melena, within the previous 30 days. (Blood on toilet paper, after wiping, does not constitute rectal bleeding)
- Patient has a diagnosis or personal history of any of the following high-risk conditions for colorectal cancer:
    - Inflammatory bowel disease (IBD) including chronic ulcerative colitis (CUC) and Crohn's disease.
    - 2 first-degree relatives who have been diagnosed with colon cancer. (Note: first-degree relatives include parents, siblings and offspring).
    - One first-degree relative with CRC diagnosed before the age of 60.
    - Patient has a family history of:
        - Familial adenomatous polyposis (also referred to as "FAP", including attenuated FAP).
        - Hereditary non-polyposis colorectal cancer syndrome (also referred to as "HNPCC" or "Lynch Syndrome").
        - Other hereditary cancer syndromes including but are not limited to 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 and Familial Hyperplastic Polyposis.
- Participation in any "interventional" clinical study within the previous 30 days in which an experimental treatment is administered or might be administered through a randomized assignment of the patient to one or more study groups.

7.1.5 Sample Size Determination

The sample size for the post approval study is determined based on the projected risk of a CRC/AA and a positive *Cologuard* test at year three (T3) when baseline (T0) is negative for both CRC/AA and *Cologuard* result, using the results observed in the Deep-C Study. DeeP-C *Cologuard* results per each age-gender stratum by Category ("Category" is the classification used in the Deep-C study based on colonoscopy findings) were used to project the *Cologuard* percent positives and Category distribution among *Cologuard* positives at each annual follow-up visit (Table 30). Under the assumption that the results of *Cologuard* will be independent between the T0 and T3 tests, the overall percents positive can be projected to drop 1.8% between the first study visit (baseline; T0) and Year 3 (T3) from 15.8% to 14.0% (Table 30 top row). Repeated surveillance is projected to decrease the percent of *Cologuard* positive patients with CRC from 3.8% to 0.3% with a corresponding decrease in the percent of *Cologuard* positive patients with AA from

20.4% to 14.4%. The combined CRC/AA percents among *Cologuard* positives are projected to drop from 24.2% for the initial *Cologuard* evaluation among positives to 14.7% at T3.

Table 30: Projected Percents *Cologuard* Positive

	Year 0 (T0)	**Year 3 (T3)**
%*Cologuard* Positives	15.8%	14.0%
Category Distribution among Positives		
CRC	0.038	0.003
AA	0.204	0.144
Adenoma	0.087	0.092
Category 4	0.074	0.066
Category 5	0.148	0.168
Category 6	0.450	0.527
Category Totals	100% of 15.8%	100% of 14.0%

A total of 1,830 patients will be enrolled to have a minimum of 946 patients at the year 3 visit (T3) assuming 15.8% and 14.0% positivity rates at T0 and T3 respectively (Table 31), with 15% annualized lost to follow up (LTFU). In addition, it is assumed that another 15% of patients will be lost after the T3 *Cologuard* due to colonoscopy refusal. Table 2 displays the projected numbers of *Cologuard* positives and negatives at baseline including those lost to follow-up (LFU).

Table 31: Projected Numbers of *Cologuard* Positives and Negatives Over Time

	T0	T1	T2	T3	T3 with Colonoscopy
Cologuard Negative	1540			814	692
Cologuard Positive	290			132	112
Totals	1830	1309	1113	946	804
		Assuming 15% annualized lost to follow-up			Assuming 15% Colonoscopy Refusal

A meaningful benefit would be a 45% relative decrease in the combined CRC/AA colonoscopy results from 3.82% (15.8% x 0.242) at T0 to 2.06% (14.0% x 0.147) at T3, based *Cologuard* positive results. As shown in Table 31, a total of 946 patients followed for three years would have at least 80% power to confirm that the percentage of patients with CRC/AA at year 3 (T3) is statistically significantly less than at baseline; this number includes a provision that 85% of the *Cologuard* positives at T3 undergo a colonoscopy at T3 (Table 32) so the effective number to be evaluated at T3 is 804 patients. Under these collective assumptions, it is expected that the actual percentage of patients with CRC/AA at T3 is reduced by 45% compared to T0 (from 3.82% (70 of 1,830 patients) to 2.06% (17 of 804 patients) (Table 32).

	Parameters
Test significance level, alpha	0.025
1 or 2 sided test?	1
Null hypothesis proportion, p_0	0.038
Alternative proportion, p_A	0.0206
Power (%)	84
N	804

Table 32: Exact One-sided Binomial Test for *Cologuard* Proportion Positive

It should be noted that the null hypothesis proportion (P_0) at baseline (T0) used in the sample size calculations is based on the observed results in the DeeP-C Study; however, depending on the age-gender distribution and the prevalence of CRC/AA in the post approval study population, this estimate will be assessed following completion of T0 data collection to verify the sample size assumptions and the sample size may be increased accordingly, if necessary.

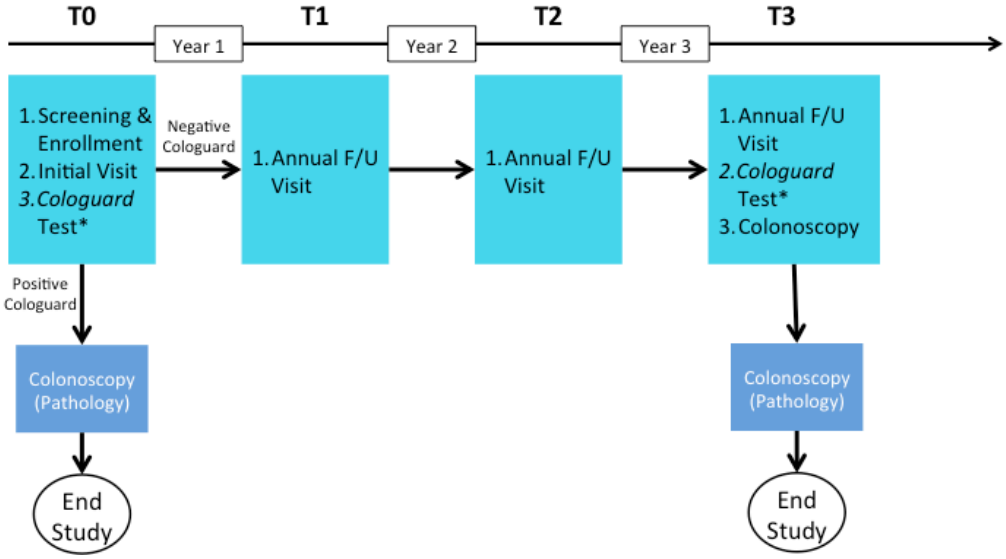
The proportion of patients with a positive *Cologuard* and CRC/AA findings at T0 and T3 will be reported with point estimates and two-sided exact 95% confidence intervals.

The distribution of colorectal epithelial lesions among positive *Cologuard* patients at T0 and T3 will be reported with counts and proportions.

The positive predictive value of *Cologuard* at T0 and T3 will be presented with point estimates and two-sided exact 95% confidence intervals.

7.1.6 Study Schema

Figure 4: Study Schema



Current CRC screening guidelines recommend that patients undergo routine screening with repeat testing over time. Within the context of these guidelines, the performance of *Cologuard* has not been evaluated upon repeat testing. The premarket data was solely based on cross-sectional studies in which specimens were tested at one time point. Thus, a post-approval study intended to evaluate the longitudinal performance of the *Cologuard* in the postmarket setting may support long term safety and effectiveness.

FDA DISCUSSION QUESTION #4:
The proposed device claim does not rule out repeating testing as part of a colorectal cancer screening program. Cross-sectional performance at one time point may not translate to longitudinal performance over time. Data was not provided to support repeat testing with *Cologuard*.

a. *Cologuard* claims do not specify a testing interval. Please discuss whether a longitudinal study should be required to address long-term safety and effectiveness.

In the event that a post-approval study is recommended as a condition of PMA approval, there are further considerations to be addressed regarding the study design. First, the appropriate study population criteria may depend on the intended use of the device. Second, not all study outcomes, such as proportion of positive patients and positive predictive value, are associated with study hypotheses. Third, the appropriateness of the proposed study hypothesis to support repeat testing with *Cologuard* is unclear (e.g. in comparison to FIT as an annual screening method). It is also uncertain the extent that satisfactory performance from subsequent testing would be supported.

For discussing the proposed longitudinal study, the Agency would appreciate advisory committee feedback concerning key issues such as whether the study population is appropriate given the intended use of the device, whether the proposed investigation demonstrates statistically and clinically meaningful performance for repeat testing at an appropriate interval over an adequate follow-up period, whether allowing study participants to forgo annual FIT testing is appropriate based on current evidence, and whether comparison to annual FIT performance should factor into study design and/or conduct.

FDA DISCUSSION QUESTION #5:
Assuming that a longitudinal study is needed to evaluate performance with repeat *Cologuard* testing, is the proposed longitudinal study adequate to address the following issues? If not, how should they be modified?

- Appropriate study population
- Length of screening interval and patient follow-up
- Performance (e.g. number of test negative to positive conversions, diagnostic yield of significant findings, predictive values, adherence to screening and diagnostic follow-up)
- Comparison for statistically and clinically meaningful performance (e.g. annual FIT)
- Suitability of patients forgoing the option of annual FIT screening during the study duration
- Study size, feasibility
- Other panel suggestions

8 QUESTIONS FOR PANEL DISCUSSION (RESTATED FROM ABOVE)

1. The Deep-C study met the primary objectives with respect to both required sensitivity and specificity of *Cologuard* compared to colonoscopy, with 92.3% sensitivity for CRC and 86.6% AN specificity.

   With respect to the secondary objectives, *Cologuard* sensitivity is higher than FIT for both CRC and AA (92.3 vs. 73.8 and 42.4 vs. 23.8, respectively). Although not a secondary objective, *Cologuard* AN specificity is lower than FIT (86.6 vs. 94.9).

   a. Do these conclusions adequately demonstrate effectiveness of *Cologuard* within the contexts of the proposed intended use and current recommendations for CRC screening?

   b. Based on the results of the pivotal clinical study, do the data provided allow for adequate assessment of the benefits and risks of *Cologuard*?

2. Are there patient subgroups, such as age (e.g., ages 75-79, 80-84, 85 and above), gender, and race/ethnicity where considerations for device performance merit additional labeling?

3. The Deep-C study conducted by Exact Sciences was not designed to provide follow-up data on patients that tested negative with *Cologuard*.

   a. What is appropriate labeling to assure safety and effectiveness for follow-up evaluation of patients testing negative with *Cologuard?* The FDA would like feedback on follow-up test interval and modality, use of guidelines, and other possible follow-up approaches.

4. The proposed device claim does not rule out repeating testing as part of a colorectal cancer screening program. Cross-sectional performance at one time point may not translate to longitudinal performance over time. Data was not provided to support repeat testing with *Cologuard*.

   a. *Cologuard* claims do not specify a testing interval. Please discuss whether a longitudinal study should be required to address long-term safety and effectiveness.

5. Assuming that a longitudinal study is needed to evaluate performance with repeat *Cologuard* testing, is the proposed longitudinal study adequate to address the following issues? If not, how should they be modified?
   - Appropriate study population

- Length of screening interval and patient follow-up
- Performance (e.g. number of test negative to positive conversions, diagnostic yield of significant findings, predictive values, adherence to screening and diagnostic follow-up)
- Comparison for statistically and clinically meaningful performance (e.g. annual FIT)
- Suitability of patients forgoing the option of annual FIT screening during the study duration
- Study size, feasibility
- Other panel suggestions

9 QUESTIONS FOR BALLOT VOTE

The following questions relate to the approvability of *Cologuard*.

1. Is there reasonable assurance that *Cologuard* is safe for use in patients who meet the criteria specified in the proposed intended use?

2. Is there reasonable assurance that *Cologuard* is effective for use in the patients who meet the criteria specified in the proposed intended use?

3. In patients who meet the criteria specified in the proposed intended use, do the benefits outweigh the risks for use of *Cologuard*?

FDA looks forward to a productive Panel discussion regarding these issues.

10 APPENDIX

The following non-clinical studies have been reviewed by the FDA:

Algorithm Development and Cut-Off Determination

The objective of this study was to establish cut-offs and establish the algorithm for the *Cologuard* CRC screening test system by evaluating a panel of donor samples that have been categorized by colonoscopy. Selection of variables 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 was 953, including 794 normal pathology 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 from 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.

After the initial cut-off was determined for *Cologuard*, the company verified the robustness of the logistic regression-based predictive algorithm and refined the risk score cut-off using a combination of computer simulations and statistical cross-validation techniques such as Leave-One-Out cross-validation ("LOOCV") and 10-fold cross-validation analyses. Furthermore, various simulations were also performed on the *Cologuard* cut-off study data (n=953) to determine the best estimate of *Cologuard* precision.

Sensitivity: Limit of Blank, Limit of Detection, Limit of Quantification and Linearity

LoB is the highest value expected in a series of test results on a sample that contains no analyte. LoD is a critical performance characteristic of a laboratory method and is defined as the lowest concentration of the analyte that can be confidently detected by the method. 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.

Linearity is the ability (within a working range) to provide results that are directly proportional to the concentration of the analyte in the test sample. 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). These results are summarized in Table 33.

Table 33: Analytical sensitivity characteristics for *Cologuard*

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

<u>*Cologuard* Molecular Assay Cross-Reactivity with Wild Type KRAS</u>

The objective of this study was to test the effect of the presence of wild-type KRAS on the QuARTS reaction. Exact Sciences evaluated the potential for cross-reactivity with wild type *KRAS* 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*. 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. These results are highlighted in Table 34 below.

Table 34: *Cologuard* Cross-Reactivity with Wild-Type KRAS

	Methylation Assay*		
	NDRG4	*BMP3*	*BTACT1***
Wild Type *KRAS* Level	Mean Strands	Mean Strands	Mean Strands
400,000 Strands (200,000 Copies)	0%	0%	0%
40,000 Strands (20,000 Copies)	0%	0%	0%
	Mutation Assay*		
	KRAS1	*KRAS2*	*ACT2****
Wild Type *KRAS* Level	Mean Strands	Mean Strands	Mean Strands
400,000 Strands (200,000 Copies)	0.01%	0.01%	0%
40,000 Strands (20,000 Copies)	0%	0%	0%

*When strand levels derived from the cross-reactivity reactions were below the LOD of the respective reaction, a cross-reactivity level of 0% was assigned.
***BTACT* refers to how the *Cologuard* software characterizes the *ACTB* in the methylation assay.
****ACT* refers to how the *Cologuard* software characterizes the *ACTB* in the mutation assay.

<u>*Cologuard* QuARTS Partial Methylation Testing</u>

In CRC lesions, many genes have elevated methylation in their promoter region, whereas the same genes have low levels of methylation in normal colon epithelial cells. Exact Sciences previously demonstrated that highly methylated promoter region sequences in BMP3 and NDRG4 correlates to presence of CRC and AA and low level methylation correlates to presence of 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 objective of this study was to determine the analytical specificity of the DNA Methylation Assay in the presence of partially methylated BMP3 and NDRG4 DNA.

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 sites of eight for *BMP3* and five sites of nine for *NDRG4* have to be methylated to generate 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 acceptance criteria.

<u>*Cologuard* Hemoglobin Assay Cross-Reactivity and Specificity</u>

The objective of this study was to assess the ability of the Hemoglobin Assay to detect hemoglobin in specimens heterozygous for common variants Hemoglobin S (HbS) and Hemoglobin C (HbC). 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 background spiked with animal whole blood. Samples used for testing myoglobin cross-reactivity consisted of a stool sample background spiked with prepared meat extracts or purified myoglobin. Thirteen replicates of each sample type were tested with the *Cologuard* Hemoglobin Assay.

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 is expected.

<u>*Cologuard* Cross-Reactivity with Non-Colorectal Cancers and Diseases</u>

The objective of this study is to evaluate specimens collected from patients with cancers and diseases other than CRC for reactivity in the *Cologuard* assay. Exact Sciences evaluated the potential for reactivity with non-colorectal cancers by testing 151 specimens from patients 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 x 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 was projected that a certain level of reactivity would be observed in cases of these cancers. The results are highlighted in Table 35 below.

Table 35: Incident Rates and Contribution to *Cologuard* Positivity for Non-CRC Diseases and Cancer

Disease or Cancer*	Number of specimens tested	Incident rate per 10,000**	% Positivity of *Cologuard*	Number additional positive *Cologuard* call in 10,000 subjects
Bladder Cancer	17	2.3	- -	- -
Breast Cancer	14	12.4	- -	- -
Esophagus Cancer	11	0.5	- -	- -
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	- -	- -
Lupus	17	0.2-0.8	- -	- -
Pancreas Cancer	12	1.2	41.6%	0.5
Prostate Cancer	12	15.5	- -	- -
Rheumatoid Arthritis	15	4.1	- -	- -
Stomach Cancer	8	0.8	- -	- -
Total per 10,000 subjects		NA	NA	2.0

*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).

Based on the results of this study, the expected positivity for the tested diseases would result in a 0.02% decrease in specificity for *Cologuard* (or two positive calls per 10,000 screening patients tested).

Precision and Reproducibility (Lab-to-Lab)

The objective of this study was to determine the overall precision and reproducibility of the *Cologuard* assay results by testing a sample panel consisting of constructed samples containing various levels of DNA targets and hemoglobin. The primary objective is to determine how much variation in the data is due to the *Cologuard* assay measurement system. A laboratory-to-laboratory precision and reproducibility study was performed to assess variation of the *Cologuard* assay measurement system with a design similar to those recommended in CLSI Standard: EP05-A2 (*Evaluation of Precision Performance*

of Quantitative Methods). 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 (100, 200, 300), with a minimum of two operators at each site. A total of 22 *Cologuard* runs were performed at each site with 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), supplied by Exact Sciences.

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.

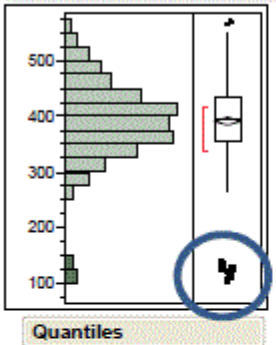
For the hemoglobin assay component of *Cologuard*, the clinical stool pools were prepared by adding fresh whole blood to normal patient stool pools. Specifically, whole blood was spiked into stool samples and diluted to the appropriate concentration. Control samples (including negative, low, and high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

For the low positive sample, the distribution of the *Cologuard* score and the methylation, mutation, and hemoglobin assay results in replicate testing are displayed in Figure 5.

Figure 5: Distribution of *Cologuard* score and individual measurements in replicate testing

Distributions

Cologuard Composite Score



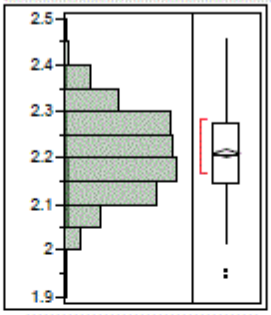
Quantiles

100.0%	maximum	571.05
99.5%		568.98
97.5%		524.64
90.0%		480.76
75.0%	quartile	434.09
50.0%	median	394.44
25.0%	quartile	354.92
10.0%		314.46
2.5%		134.25
0.5%		107.86
0.0%	minimum	100.78

Moments

Mean	390.57863
Std Dev	73.290896
Std Err Mean	3.7159959
Upper 95% Mean	397.88464
Lower 95% Mean	383.27262
N	389

LT1_NDRG4



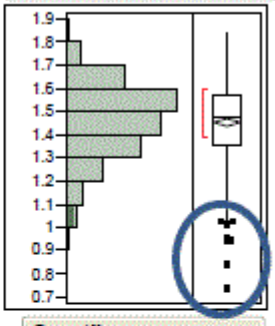
Quantiles

100.0%	maximum	2.4563
99.5%		2.4346
97.5%		2.3724
90.0%		2.3169
75.0%	quartile	2.2734
50.0%	median	2.2074
25.0%	quartile	2.1477
10.0%		2.1010
2.5%		2.0482
0.5%		1.9523
0.0%	minimum	1.9414

Moments

Mean	2.2090269
Std Dev	0.0852243
Std Err Mean	0.004321
Upper 95% Mean	2.2175225
Lower 95% Mean	2.2005313
N	389

LT1_BMP3

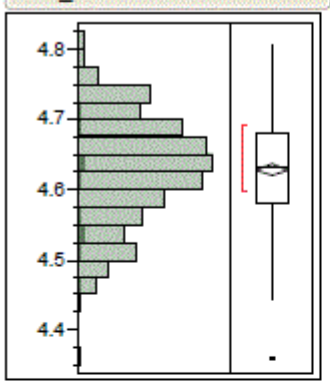


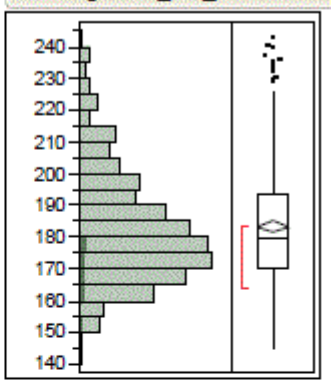
Quantiles

100.0%	maximum	1.8386
99.5%		1.8126
97.5%		1.7473
90.0%		1.6596
75.0%	quartile	1.5771
50.0%	median	1.4768
25.0%	quartile	1.3550
10.0%		1.2352
2.5%		1.0357
0.5%		0.8262
0.0%	minimum	0.7308

Moments

Mean	1.4556238
Std Dev	0.1705964
Std Err Mean	0.0086496
Upper 95% Mean	1.4726297
Lower 95% Mean	1.4386179
N	389





Quantiles

100.0%	maximum	4.8075
99.5%		4.8038
97.5%		4.7551
90.0%		4.7277
75.0%	quartile	4.6813
50.0%	median	4.6332
25.0%	quartile	4.5797
10.0%		4.5214
2.5%		4.4755
0.5%		4.4390
0.0%	minimum	4.3575

Moments

Mean	4.6286702
Std Dev	0.0744352
Std Err Mean	0.003774
Upper 95% Mean	4.6360903
Lower 95% Mean	4.6212502
N	389

Quantiles

100.0%	maximum	242.30
99.5%		240.11
97.5%		229.16
90.0%		209.98
75.0%	quartile	193.70
50.0%	median	179.77
25.0%	quartile	170.37
10.0%		163.90
2.5%		154.92
0.5%		145.53
0.0%	minimum	144.54

Moments

Mean	183.30764
Std Dev	18.071078
Std Err Mean	0.91624
Upper 95% Mean	185.10905
Lower 95% Mean	181.50622
N	389

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) (CLSI I/LA28-A2: Quality Assurance for Design Control and Implementation of Immunohistochemistry Assays, 2nd edition, Appendix A1: Statistical Points to Consider When Evaluating the Precision of Immunohistochemistry Assays). The sponsor calculated 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 is shown in Table 36.

Table 36: Inter-Site Agreement

Site Comparison	Number Agreed	Total Compared	Agreement Rate	95% CI Lower Bound***
ANA* – Site 100 and Site 200	768	777	0.988	0.978
APA** – Site 100 and Site 200	1026	1035	0.991	0.983
Site Agree – Site 100 and Site 200	897	906	0.990	0.982
ANA – Site 100 and Site 300	744	746	0.997	0.990
APA – Site 100 and Site 300	1012	1014	0.998	0.993
Site Agree – Site 100 and Site 300	878	880	0.998	0.992
ANA – Site 200 and Site 300	756	764	0.990	0.979
APA – Site 200 and Site 300	1004	1012	0.992	0.984
Site Agree – Site 200 and Site 300	880	888	0.991	0.982

Descriptive statistics were separately calculated for all marker/sample combinations. The percent coefficient of variation (%CV = 100 * standard deviation / mean) was calculated only for samples expected to have a positive result (> cut-off 183 in *Cologuard* score). For completeness, FDA also calculated the %CV for samples expected to have a negative result. Inter-site descriptive statistics are provided in Table 37.

Table 37: Inter-Site Descriptive Statistics for the *Cologuard* Score

Sample	Variable	N	Mean	Lower 95% CL for Mean	Upper 95% CL for Mean	Standard Deviation	Total %CV
Negative Stool Pool	*Cologuard* Score	387	9.98	9.65	10.31	3.31	33.17
High Negative Stool Pool		394	62.92	60.24	65.61	27.14	43.13
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	14.17
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

For samples expected to have positive results, inter-site %CVs were all less than 20%. Percent CVs tend to be larger for the samples expected to have a negative result because the mean *Cologuard* score is relatively small.

The *Cologuard* composite score is a many-to-one function of the methylation, mutation, and hemoglobin assay results. Many combinations of assay results are possible that will produce the same *Cologuard* score. The precision of the assays will depend on the levels being measured. Thus, for any given *Cologuard* score a range of precisions is possible depending on the assay results producing that score. FDA asked the sponsor to investigate the distribution of precision at values of the *Cologuard* score. Using 953 samples from a cut-off study, the sponsor simulated 1000 combinations of assay results, the resulting *Cologuard* scores, and the precision of the scores based on %CVs at the assay values obtained from previously performed precision studies. The simulated distribution of precision at any given *Cologuard* was submitted to FDA and found to be acceptable (not shown).

Lot-to-Lot Reproducibility

Lot-to-Lot reproducibility was evaluated for *Cologuard* based on recommendations 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).*

The objective of this study was to assess the lot-to-lot reproducibility of 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 38). Percent 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 38: Descriptive Statistics for Lot-to-Lot *Cologuard* Score

Sample Name	N	Median	Mean	Lower 95% CL for Mean	Upper 95% CL for Mean	Standard Deviation	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) (Table 39). 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%.

Table 39: Agreement

Lot Comparison	Number Agreed	Total Compared	Agreement Rate	95% CI Lower Bound***
ANA* - Lot1 and Lot2	142	142	1.0000	0.9744
APA** - Lot1 and Lot2	188	188	1.0000	0.9806
Lot Agree - Lot1 and Lot2	165	165	1.0000	0.9779
ANA - Lot1 and Lot3	140	142	0.9859	0.9501
APA - Lot1 and Lot3	180	182	0.9890	0.9609
Lot Agree - Lot1 and Lot3	160	162	0.9877	0.9561
ANA - Lot2 and Lot3	142	144	0.9861	0.9507
APA - Lot2 and Lot3	184	186	0.9893	0.9617
Lot Agree - Lot2 and Lot3	163	165	0.9879	0.9569

NOTE: Proportion values are point estimates used to determine the Clopper-Pearson 2 sided Confidence Interval. Only Clopper-Pearson Lower Limit values are shown in the above table.
*ANA = Average negative agreement
**APA = Average positive agreement
***Clopper-Pearson Confidence Interval

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

Robustness

The objective of this study was to assess *Cologuard* assay performance in response to defined variable conditions at specific steps in the test procedure. Exact Sciences assessed the robustness of *Cologuard* performance 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.

For the molecular assay component of *Cologuard*, 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; 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.

For the hemoglobin assay component of *Cologuard*, 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.

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. Analysis of this testing demonstrated that the prepared *QuARTS* plate should be processed within 30 minutes and captured DNA could be stored for up to four days without significant degradation in performance.

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

Interference

Cologuard Molecular Assay Interference Testing:

The objective of this study was to evaluate the performance of *Cologuard* molecular assay for potential interference from substances common in stool. Testing was performed using 16 replicates of positive and negative stool homogenate samples, with and without 55 potential 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).

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:

The objective of this study was to evaluate the performance of *Cologuard* hemoglobin assay for potential interference from substances common in stool. Interference 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; and
- Alcohol.

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.

Carry-Over and Cross-Contamination *Cologuard* Testing

The objectives of these studies were to confirm that there is no carry over contamination or cross over contamination of samples in the *Cologuard* assay process, which consists of a molecular assay portion (methylation and mutation) as well as a hemoglobin assay portion.

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.

Analysis of results from the molecular assay and hemoglobin assay carry-over analyses demonstrated that the *Cologuard* assay components and the instruments required for running the assay performed satisfied the acceptance criteria for the study.

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. As in the carry-over study, the high positive samples for this study were also 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.

Analysis of 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 met the study acceptance criteria. In the study, there was one instance of 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 patients (Levi et. al, 2007). The high hemoglobin concentrations tested in this study are much higher than would be expected in use of *Cologuard*. First run results showed a signal in 4 out of 43 negative samples with an average detectable hemoglobin level of 11 ng/mL (0.011%). 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. This indicates that there is no cross-contamination from the automated equipment, but rather operator-induced cross-contamination can occur if procedures are not carefully followed. Data from the combined run passed the pre-specified acceptance criteria described in the protocol.

Stability

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 recommendations 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 above 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:
Exact Sciences conducted a study 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:
Exact Sciences is conducting an on-going study for real-time stability of *Cologuard*, evaluating the functional performance of three reagent lots over a period of 41 weeks. Each lot is comprised of unique batches of reagents, which will be tested at various time points over 41 weeks.

Samples that will be used to evaluate hemoglobin assay reagent stability consist 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 consist of negative stool matrix spiked with oligonucleotides that contain the marker sequences. Oligonucleotides for *NDRG4, BMP3, BTACT, KRAS1, KRAS2,* and *ACT* will be 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 will be tested.

<u>Software Documentation</u>

Complete software documentation, including test results from complete software verification and validation testing were provided and reviewed.

11 REFERENCES

Centers for Disease Control and Prevention (CDC). (2013). Vital signs: colorectal cancer screening test use--United States, 2012. *MMWR Morb Mortal Wkly Rep*, 881-8.

Imperiale TF, R. D. (2010). Understanding differences in the guidelines for colorectal cancer screening. *Gastroenterology*, 1642-47.

Lee JK, L. E. (2014). Accuracy of Fecal Immunochemical Tests for Colorectal Cancer: Systematic Review and Meta-analysis. *Ann Intern Med*, 171-81.

Levin B, L. D.-S. (2008). Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *Gastroenterology*, 1570-95.

Levin TR, C. D. (2013). Colorectal-Cancer Screening — Coming of Age. *N Engl J Med*, 1164-66.

Markowtiz SD, B. M. (2009). Molecular Basis of Colorectal Cancer. *N Engl J Med*, 2449-60.

Nadel MR, B. Z. (2010). Fecal Occult Blood Testing Beliefs and Practices of U.S. Primary Care Physicians: Serious Deviations from Evidence-Based Recommendations. *J Gen Intern Med*, 833-9.

Ransohoff DF, P. M. (2013). Colon cancer screening models: lessons and challenges. *Clin Gastroenterol Hepatol*, 1167-68.

Rex DK, J. D. (2009). American College of Gastroenterology guidelines for colorectal cancer screening 2009. *Am J Gastroenterol*, 739-50.

Shaukat A, M. S. (2013). Long-term mortality after screening for colorectal cancer. *N Engl J Med*, 1106-14.

U.S. Preventive Services Task Force. (2008, October). Screening for Colorectal Cancer: Clinical Summary of U.S. Preventive Services Task Force Recommendation. *AHRQ Publication No. 08-05124-EF-4*. Retrieved from AHRQ Publication No. 08-05124-EF-4: http://www.uspreventiveservicestaskforce.org/uspstf08/colocancer/colosum.htm

Whitlock EP, L. J. (2008). *Screening for Colorectal Cancer: An Updated Systematic Review [Internet]*. Rockville (MD): Agency for Healthcare Research and Quality .

www.asge.org. (n.d.). Retrieved from http://www.asge.org/press/press.aspx?id=5354

www.cancer.org. (n.d.). Retrieved from http://www.cancer.org/cancer/colonandrectumcancer/detailedguide/colorectal-cancer-key-statistics.