The Effect of the Variability in Fecal Immunochemical Test Sample Collection Technique on Clinical Performance

Erin L. Symonds[1,2], Callum G. Fraser[3], Dawn Bastin[2], Grace Berwald[4], and Graeme P. Young[2]

ABSTRACT

Background: Fecal immunochemical test (FIT) performance can be affected by post-collection variables. Collection technique might also affect fecal hemoglobin concentration (f-Hb). Variation in quantity of feces collected in samples returned in a colorectal cancer detection program, and the effects of under-sampling, were assessed.

Methods: Collection devices obtained from patients undergoing FIT were assessed for the color (in five classes) of the feces in buffer, mass, and f-Hb. Associations between these were examined in an *in vitro* study on Hb-spiked feces. Variables possibly associated with under-sampling were investigated using multivariable logistic regression. The effect of low sample mass on clinical performance (false-negative results) was determined.

Results: Of 6,898 samples collected by 3,449 individuals (46.9% male, median age: 65.3 years), the buffer was lightest in color in 362 (5.2%), and darkest in 420 (6.1%). Samples with the lightest color had a significantly lower f-Hb compared with all darker classes ($P < 0.001$). Mass was recorded for 650 devices: The lightest colored samples had significantly lower mass ($P < 0.05$). The correlation between mass and f-Hb was confirmed *in vitro* ($r = 0.897$, $P < 0.001$). Low mass was not associated with age, sex, or technical factors ($P > 0.05$). Under-sampling related to the lightest color was not associated with false-negative results for colorectal cancer and advanced adenoma, but was for all neoplasia and inflammatory bowel disease.

Conclusions: Wide variation existed in the amount of feces collected. Under-sampling results in lower measured f-Hb and may increase false-negative results.

Impact: Color of sample buffer could be used to identify inadequate sampling.

Introduction

Regular screening with fecal occult blood tests has proven effective at reducing the mortality associated with colorectal cancer, as well as reducing the incidence (1–3). Many countries use fecal immunochemical tests (FIT) in screening programs (4). FIT detect human hemoglobin (Hb) in feces, which allows the targeting of those who most merit diagnostic colonoscopy, because they have an elevated risk of advanced neoplasia [colorectal cancer or advanced adenoma (AA)] and potentially other significant conditions such as inflammatory bowel disease (IBD).

Completion of most brands of FIT involves collecting one or two samples of feces, with optimal performance assuming that an appropriate quantity of feces is introduced into the device by the collection probe, that the device meters the amount placed into the buffer in the collection device, and that it is then stored appropriately before laboratory analysis. For quantitative FIT, latex agglutination immunoassay is then performed on a small volume of the feces in buffer solution to give a fecal Hb concentration (f-Hb).

There are multiple steps from collection to analysis that can adversely affect the diagnostic accuracy of f-Hb. Most studies have focused on potential variables after collection of the fecal sample, showing that storage temperature, ambient temperature, and days from the collection until analysis can result in degradation of Hb within the collection device (5–8), resulting in false-negative results, missing the opportunity for disease detection. However, factors related to the collection itself may also influence the diagnostic accuracy and subsequent clinical performance, such as quantity of feces collected in the device. Results from previous studies on the effect of the addition of different quantities of fecal samples are inconsistent, with one study showing that the addition of feces to Hb-spiked buffer changed the measured f-Hb (9), whereas another study showed that varying the quantity of Hb-spiked feces collected in the FIT devices did not significantly change f-Hb (10).

Although these laboratory-based studies showed that pre-analytical conditions are important, the variability in sample collection in everyday clinical practice has not been assessed for FIT. It is important to determine whether sampling variability could adversely affect clinical performance. The aims of this study were to determine the variability in sample collection between individuals returning FIT devices, to assess the patient factors associated with inadequate sample collection, and to determine whether low sample collection adversely affects clinical performance, particularly the proportion of false-negative results for advanced neoplasia and other bowel disease.

Materials and Methods

Assessment of FIT returned within a clinical program

FIT devices that were returned by post to the laboratory by patients at elevated risk for colorectal cancer (the SCOOP surveillance

[1]Bowel Health Service, Flinders Medical Center, Bedford Park, South Australia, Australia. [2]Cancer Research, Flinders Health and Medical Research Institute, Flinders University, Bedford Park, South Australia, Australia. [3]Center for Research into Cancer Prevention and Screening, University of Dundee, Dundee, Scotland, United Kingdom. [4]Department of Medicine, College of Medicine and Public Health, Flinders University, Bedford Park, South Australia, Australia.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Corresponding Author: Erin L. Symonds, Flinders Center for Innovation in Cancer, Bedford Park, South Australia 5042, Australia. Phone: 61-8-8404-2813; Fax: 61-8-8204-5703; E-mail: erin.symonds@sa.gov.au

Cancer Epidemiol Biomarkers Prev 2021;30:175–81

doi: 10.1158/1055-9965.EPI-20-0984

program; ref. 11) were reviewed for color. A color chart was created with five different colors, graded 1 to 5 (lightest to darkest, Supplementary Fig. S1), and the buffer within the collection device containing the fecal sample was reviewed for the color by two individuals (D. Bastin and E.L. Symonds). For the first four months of the study, all samples were reviewed by both to ensure consistency in grading of color and then, with the experience gained, by one. In addition, a subset of collection devices was weighed before being sent to patients and re-weighed following return to the laboratory after the collection had been completed. All colors were recorded before determining final mass to avoid potential bias. Following recording of the color (and mass, where appropriate), all samples were assayed to determine f-Hb.

The FIT kit that was sent to participants included instructions, two collection devices (OC-Sensor, Eiken Chemical Co., Tokyo, Japan), two sample collection sheets, and a pouch and envelope for return post to the laboratory (Flinders Center for Innovation in Cancer, South Australia). The collection device includes a circumferentially grooved probe for sampling feces with the result that feces sits in the groves and on the probe. The probe is inserted into the device which, using a collar, meters the amount of feces reaching the buffer to approximately 10 mg feces. The device contains 2.0 mL of Hb stabilizing buffer. Patients were asked to collect samples at home from two different bowel motions and to collect their samples from the outer surface of the feces kept clear of toilet bowl water. Neither dietary nor medicine restrictions were required.

For all participants, personal data were collected, including sex, age, residential postcode, and reason for being included in the surveillance program. Socioeconomic status (Index of Relative Socio-economic Disadvantage score) was determined from the participant's residential postcode through linkage with the Australian 2016 census data (12) and expressed in quintiles, with 5 being the areas of most advantage. Where available, technical variables relevant to the FIT examination were also recorded, including the number of days from sample collection to analysis, storage condition before postage to the laboratory (room temperature or refrigerator), season of the year and ambient maximum temperature (recorded in Adelaide, South Australia) that sampling occurred, whether the participant had previously collected a sample for FIT, and FIT collection device manufacturing lot number. The number of days between first and second sample collection and whether the sample had been collected before or after midday were assessed to be an indirect measure of the participant's gastrointestinal transit time and the resulting fecal consistency.

A subgroup of participants completed the collection within eight months before a scheduled colonoscopy. All colonoscopy procedures were performed by, or under the supervision of, experienced and hospital-accredited colorectal surgeons or gastroenterologists. Details that were collected from the colonoscopy records included the date of the investigation, quality of procedure (based on bowel preparation score and intubation distance), and pathology findings.

Collection and examination of the collection devices was approved under the Southern Adelaide Human Research Ethics Committee, which follows the Declaration of Helsinki. Written informed consent was given accompanying return of the completed devices.

In vitro study to assess sample mass, color, and impact on f-Hb

A study involving Hb-spiked feces was performed to confirm or refute the relationship between the added sample mass and color, as well as the impact on f-Hb. Hemolysis of erythrocytes from a blood sample was performed through the addition of distilled water. The lysed solution was mixed into a fecal sample from a healthy individual to give a final f-Hb of approximately 80 μg Hb/g feces. This spiked feces was then sampled using the OC-Sensor collection devices, adding varying amounts of sample to the probe- from very small quantities, clearly not filling the grooves in the probe, to obvious overloading of the probe, to approximate the five different colors applied in the clinical examinations. Five different quantities were collected, with 10 collection devices used in each sampling scenario. The collection devices were weighed before and after collection of the sample. The devices were stored at room temperature for 24 hours to allow dispersion of the sample within the collection device buffer.

Analysis for f-Hb

Collection devices were stored at 4°C until analyzed, then assayed for f-Hb using an automated analyzer (OC-Sensor DIANA, Eiken Chemical Co.), following the manufacturer's instructions. If a device had f-Hb above the upper analytical limit (200 μg Hb/g feces), the sample was diluted with the FIT system buffer and re-analyzed.

Data analysis

Participants returned two completed FIT collection devices, but f-Hb were considered individually, except for comparing sample collection technique between the first and second collection per participant. Median and interquartile range (IQR) were calculated for age, mass, days between collection of samples and analysis, and f-Hb. Median f-Hb and device mass were compared between the different colors with an ANOVA on ranks and Spearman Rank correlation. Positivity across a range of criterion values (thresholds) was calculated for each color and compared using a χ^2 test. Variables associated with a less than optimal quantity of collected feces were assessed by univariate analysis, followed by multivariable analysis, including the variables of clinical interest or those that had a P value of <0.05 on univariate analysis.

To determine the effect of sample collection mass on clinical performance, colonoscopy outcomes were assessed for the presence of: (i) advanced neoplasia (colorectal cancer or AA; with AA, including adenomas or sessile-serrated lesions with size ≥10 mm, high-grade dysplasia, villous morphology, or more than three small tubular adenoma), and (ii) IBD or any neoplasia [including colorectal cancer, AA, and non-advanced adenomas (NAA, including small tubular adenomas and low-risk sessile-serrated lesions)]. The f-Hb were assessed for these clinical outcomes, and a result was defined as a "false negative" if any of these clinical findings were present with an f-Hb below the positivity threshold (analysis was done for thresholds of 10, 20, 40, and 80 μg Hb/g feces that covers the main range of positivity thresholds used around the world).

Statistical analysis was done with SigmaPlot (v11, Sysstat Software, Inc.) and Stata (v13, StataCorp). A P value of <0.05 was considered statistically significant.

Results

Assessment of FIT returned within a surveillance program

The characteristics of the 3,449 participants that correctly completed two sample collections from January 2019 to February 2020 are shown in **Table 1**.

Of the collection devices ($n = 6,898$), 362 (5.2%) were judged to be the lightest color (color 1), 1,836 (26.6%) color 2, 2,569 (37.2%) color 3, 1,711 (24.8%) color 4, and 420 (6.1%) were the darkest color (color 5, Supplementary Fig. S2). Of the paired samples ($n = 3,449$ pairs), 858 (24.9%) had a darker color on the second sample collection and 872 (25.3%) were lighter, but the majority ($n = 1,719$, 49.8%) had the same

Table 1. Participant and sample characteristics (n = 3,449).

Characteristic	*n* (%)
Age, years (median, IQR)	65.3 (58.7–70.2)
Male	1,616 (46.9)
Index of relative socio-economic disadvantage	
Quintile 1 (most disadvantaged)	444 (12.9)
Quintile 2	713 (20.7)
Quintile 3	607 (17.6)
Quintile 4	526 (15.3)
Quintile 5 (least disadvantaged)	1,155 (33.5)
Reason for participation in surveillance program[a]	
Family history of colorectal cancer	471 (13.7)
Previous history of adenoma	2,790 (80.9)
Previous history of colorectal cancer	145 (4.2)
Had completed previous FIT	2,838 (82.3)
Days between the collected specimens (median, IQR)	1.0 (1.0–1.0)
Days from collection until analysis (median, IQR)[b]	7.0 (5.0–9.0)
Storage condition[a,b]	
Room temperature	1,864 (27.0)
Refrigerator	4,232 (61.4)
Season of year for specimen collection	
Spring	928 (26.9)
Summer	577 (16.7)
Fall	811 (23.5)
Winter	1,133 (32.9)

Abbreviations: FIT, fecal immunochemical test; IQR, interquartile range.
[a]Not all data available.
[b]Calculated for 6898 FIT devices.

color: 53 (1.5%) participants returned both devices with color 1 and 106 (3.1%) with color 5.

Association of sample mass and feces in buffer color

There were 650 collection devices returned that had mass recorded before and after sample collection. The median mass of feces collected was 33.8 mg (IQR, 17.4–61.4 mg). Color was significantly related to sample mass, with colors 2, 3, 4, and 5 having a significantly greater mass ($P = 0.002$) than that for color 1 (**Fig. 1**).

Association between f-Hb, feces in buffer color, and positivity

Median f-Hb for colors 2, 3, 4, and 5 were all significantly greater than that for color 1 ($P < 0.001$; **Table 2**). Positivity at thresholds of 10 and 20 μg Hb/g feces were significantly different between the colors

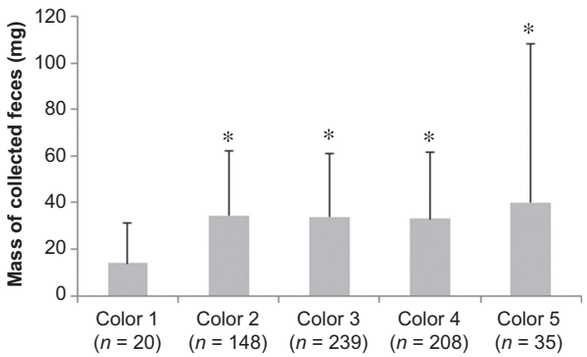

Figure 1.
Median (with 75th percentiles) mass of sample collected for the five colors of feces in buffer. *, $P < 0.05$ compared with color 1.

($P < 0.01$), but no significant differences were seen when the threshold was set at 40 or 80 μg Hb/g feces (**Table 2**).

Variables that could predict suboptimal sample collection (color 1)

Demographic, physiological, and technical variables were assessed for association with color 1. Significant and clinically relevant variables that were used in the multivariable logistic regression analysis model are shown in **Table 3**. The only patient-related variable that was a predictor for suboptimal sample collection was if the participant was undergoing surveillance for a family history of colorectal cancer (and had not had previous neoplasia). There were no technical variables (such as the number of days between sample collection and analysis, or storage condition) that were associated with color 1 (**Table 3**).

Does less than optimal sampling affect detection of relevant lesions?

Following exclusion of colonoscopies with poor bowel preparation, that did not reach the cecum, and with unclear pathology, there were 934 samples associated with the 467 colonoscopies that occurred soon after sample collection (median time from first sample collection until colonoscopy: 6.0 days, IQR, 4.0–12.0 days. Demographics reported in Supplementary Table S1). Six, 208, 280, and 28 samples were collected before a diagnosis of colorectal cancer, AA, NAA, and IBD, respectively.

Of the 934 samples, the number of samples with f-Hb less than the thresholds of 10, 20, 40, and 80 μg Hb/g feces, was 657 (70.3%), 738 (79.0%), 823 (88.1%), and 877 (93.9%), respectively. The false-negative results (missed cases) for advanced neoplasia and for detection of all neoplasia and IBD according to color are shown in **Table 4**. The percentages of false-negative results for advanced neoplasia were not different between color 1 and the darker colors. However, when considering the detection of all neoplasia and IBD (as when using FIT in symptomatic patient to prioritize for colonoscopy), the percentage of missed pathology was higher in color 1 than the darker colors. This shows that collection of less than the ideal sample (10 mg) does not affect the detection of advanced neoplasia but does lead to false-negative results when considering IBD and all neoplasia.

In vitro study: sample mass and f-Hb variation using Hb-spiked samples

The mass of the collection devices before fecal sampling had low variability. The mean mass (± standard deviation) was 7.67 ± 0.01 g (coefficient of variation, 0.13%). There was a clear relationship between quantity of feces collected into the device and color ($r = 0.964$, $P < 0.001$), as well as f-Hb ($r = 0.897$, $P < 0.001$, **Fig. 2**). Low f-Hb were observed when less than 10 mg of sample was loaded on to the collection device probe, with a plateau in concentration for sample mass greater than 10 mg.

Discussion

FIT are commonly used in colorectal cancer screening programs, in assessment of patients presenting with lower bowel symptoms (13), and in surveillance programs for colorectal neoplasia (11, 14). Post-collection factors, such as time between collection and sample analysis, and storage conditions before analysis, can cause degradation of Hb (7, 15, 16), reducing positivity and clinical performance of FIT. However, no previous study has examined the variability in the quantity of sample collected and the impact that this has on clinical

Table 2. Test positivity at different fecal hemoglobin concentration thresholds for samples with different colors.

	Color					
	1	2	3	4	5	
n	**362**	**1,836**	**2,569**	**1,711**	**420**	***P***
Median Hb (IQR), μg Hb/g feces	0.6 (0–2.0)	1.8 (0.6–3.8)	2.4 (1.2–5.1)	2.8 (1.4–5.8)	3.2 (1.8–6.0)	$P < 0.001$
Positivity at 10 μg Hb/g feces threshold	5.0%	9.1%	12.6%	15.0%	13.8%	$P < 0.001$
Positivity at 20 μg Hb/g feces threshold	2.8%	5.0%	6.7%	7.6%	6.9%	$P = 0.001$
Positivity at 40 μg Hb/g feces threshold	2.2%	2.7%	3.4%	3.9%	3.3%	$P = 0.245$
Positivity at 80 μg Hb/g feces threshold	1.7%	1.0%	2.0%	1.9%	1.7%	$P = 0.152$

Abbreviations: Hb, hemoglobin; IQR, interquartile range.

outcomes within a clinical program. We found that the color of sample buffer in the returned collection device correlated with the mass of feces in the device, and that a low mass affected the measured f-Hb. This did not affect the clinical performance for detection of advanced neoplasia, with all sample colors having a similar false-negative proportion but it did, however, increase the false-negative rate in participants with other significant pathologies due to missing detection of NAA and IBD.

Instructions for sample collection vary, depending on the brand of FIT, and might be adapted by screening program organizers and for other applications. For the OC-Sensor brand, and the instructions provided in this study, participants are advised to fill the grooves in the probe. The integral collar in the device aims to meter the amount of feces transferred into the buffer by scraping off excess feces and delivering approximately 10 mg into the 2.0 mL of Hb stabilizing buffer. The numbers in the five color groups had a normal type of distribution on visual inspection and 5.2% of devices were classified as color 1 (the lightest color of feces in buffer). This corresponded with under-sampling because, of the 650 devices weighed before and after sample collection, color 1 was associated with the lightest mass and in many of these the actual amount of feces collected was <10 mg. Consequently, the optimal 10 mg could not have been delivered into the buffer. Inadequate sample collection could not be explained by participant age or socioeconomic status, nor was it related to lack of experience in previously completing a FIT. It is possible that color could be associated with fecal consistency. A more liquid bowel motion

Table 3. Univariate and multivariable logistic regression analysis documenting association of variables with color 1 of feces in buffer.

	Univariate analysis		Multivariable analysis	
Variable	**Odds ratio (95% CI)**	***P***	**Odds ratio (95% CI)**	***P***
Age (reference: 50–64 years)				
<50 years	0.59 (0.35–1.00)	0.050	0.64 (0.37–1.10)	0.107
65–79 years	0.93 (0.75–1.15)	0.486	1.03 (0.81–1.31)	0.810
Sex (reference: female)				
Male	0.87 (0.70–1.08)	0.209	0.81 (0.64–1.02)	0.075
Index of relative socio-economic disadvantage (reference: Quintile 1, most disadvantaged)				
Quintile 2	1.11 (0.75–1.62)	0.606	—	—
Quintile 3	1.02 (0.68–1.53)	0.916	—	—
Quintile 4	1.29 (0.87–1.92)	0.202	—	—
Quintile 5 (least disadvantaged)	1.05 (0.73–1.50)	0.796	—	—
Reason for surveillance (reference: family history of colorectal cancer)				
Previous history of adenoma	0.67 (0.51–0.89)	0.005	0.67 (0.50–0.90)	0.009
Previous history of colorectal cancer	0.51 (0.26–0.97)	0.041	0.43 (0.20–0.93)	0.031
Physiology				
Sample collected after midday (reference: before midday)	0.93 (0.72–1.19)	0.561	—	—
Number of days between collection of specimen 1 and 2 (reference: 1 day)				
Same day collection	0.90 (0.66–1.21)	0.481	—	—
≥2 days between specimens	1.06 (0.78–1.43)	0.714	—	—
Technical factors				
Previous completion of FIT	1.14 (0.85–1.51)	0.386	1.06 (0.76–1.47)	0.728
Number of days between sample collection and analysis (continuous variable)	1.02 (0.98–1.06)	0.294	1.02 (0.98–1.06)	0.360
Storage of collection device at room temperature (reference: refrigerator)	1.17 (0.92–1.49)	0.197	1.17 (0.92–1.50)	0.201
Ambient maximum temperature (continuous variable)	1.00 (0.99–1.02)	0.709	—	—
Season of year of specimen collection (reference: winter)				
Spring	1.10 (0.84–1.14)	0.500	—	—
Summer	1.04 (0.75–1.44)	0.828	—	—
Fall	0.95 (0.71–1.27)	0.741	—	—
Manufacturing batch of collection devices (reference: Batch A)				
Batch B	0.56 (0.13–2.41)	0.437	—	—
Batch C	0.52 (0.12–2.28)	0.390	—	—

Table 4. Percentage of false-negative results for advanced neoplasia and all neoplasia and inflammatory bowel disease according to color at fecal hemoglobin concentration thresholds of 10, 20, 40, and 80 μg Hb/g feces.

Condition and context for false-negativity	Color		*P*
	1	2, 3, 4, 5	
Number of cases with FIT and colonoscopy	35	899	
Cutoff 10 μg Hb/g feces			
n (%) of false-negative results for advanced neoplasia at <10 μg Hb/g feces	7/30 (23.3)	143/627 (22.8)	0.947
n (%) of false-negative results for neoplasia and IBD at <10 μg Hb/g feces	23/30 (76.7)	347/627 (55.3)	0.021
Cutoff 20 μg Hb/g feces			
n (%) of false-negative results for advanced neoplasia at <20 μg Hb/g feces	7/32 (21.9)	154/706 (21.8)	0.993
n (%) of false-negative results for neoplasia and IBD at <20 μg Hb/g feces	23/32 (71.9)	389/706 (55.1)	0.062
Cutoff 40 μg Hb/g feces			
n (%) of false-negative results for advanced neoplasia at <40 μg Hb/g feces	7/33 (21.2)	178/790 (22.5)	0.859
n (%) of false-negative results for neoplasia and IBD at <40 μg Hb/g feces	24/33 (72.7)	435/790 (55.1)	0.045
Cutoff 80 μg Hb/g feces			
n (%) of false-negative results for advanced neoplasia at <80 μg Hb/g feces	7/34 (20.6)	187/843 (22.2)	0.827
n (%) of false-negative results for neoplasia and IBD at <80 μg Hb/g feces	24/34 (70.6)	463/843 (54.9)	0.072

Abbreviations: FIT, fecal immunochemical test; Hb, hemoglobin.

will have a reduced mass in the collection device and be less likely to cause colonic neoplasia to bleed during passage. Participants were not surveyed for consistency of feces with the Bristol Stool Chart. Thus, as an alternative approach, collected data that could be related to gastrointestinal transit were assessed in this study, namely, the timing of specimen collection. This, however, was also not significantly associated with suboptimal collection.

A number of features of the FIT collection device, that vary between brands, influence the amount of feces collected (17). This includes the design of the sampling probe and the collar that excludes transfer of excess collected feces into the buffer, and the volume and composition of the buffer. We observed that 6.1% of returned FIT collection devices were classified as having the darkest color, even though the collar within the device is designed to meter feces being delivered into the buffer. A plausible explanation is that color is not simply only a reflection of amount of feces but may also relate to other factors such as composition of diet. The amount of feces collected by the probe will also depend on fecal texture.

A previous *in vitro* study suggested that the addition of feces to the buffer causes changes in color, pH and ionic strength, leading to potential changes in the chemical reactions in the analysis (9). Low sample mass was associated with lower f-Hb in the current study. This was also supported by our *in vitro* analysis that sample collections of less than 10 mg caused low f-Hb. This was probably because the grooves in the probe were not completely filled with feces. An *in vitro* study that investigated different sampling techniques found that, if the collection device probe was only inserted into the passed feces once, or if the probe was not visually inspected for adequate sampling, f-Hb was reduced (18). Some brands of FIT promote collection of feces from multiple sites and, although this was not shown to alter diagnostic accuracy compared with sampling from just one site (19), instruction for multi-site sampling might reduce the risk of under-sampling. Our results, and that of a published *in vitro* study that loaded up to 74 mg on to the collection device probe (10), confirm that over-sampling does not adversely affect diagnostic accuracy and clinical outcomes, provided that at least 10 mg is collected and inserted into the device.

Despite a lower f-Hb found with low sample collection, importantly, this did not lead to a significant increase in the false-negative results for advanced neoplasia. This finding is supported by an *in vitro* study that showed that the lower f-Hb on inadequate sampling had only small effects on clinical performance (18). Previous studies have shown that the average (or median) f-Hb exceeds 160 μg Hb/g feces in patients

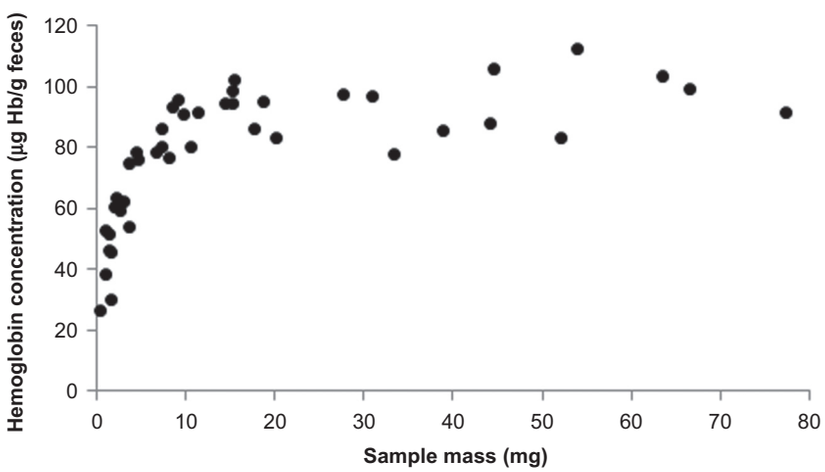

Figure 2.
Association between fecal sample mass and fecal hemoglobin concentration: $n = 50$.

with colorectal cancer, and 110 μg Hb/g feces in those with AA (20, 21). These high concentrations are likely to remain above the positivity threshold, regardless of adequacy of sampling. However, particularly when FIT are used in the assessment of patients presenting in primary care with lower bowel symptoms, where detection of NAA and IBD are considered of importance (22), an increase in the false-negative results for NAA and IBD was observed for color 1. A cogent explanation is that NAA and IBD have lower levels of bleeding.

The strength of this study is that it has assessed sample collection *in vivo* with real patients, as well as in an *in vitro* setting. This had the advantages that a subgroup of this cohort underwent colonoscopy (regardless of FIT result), allowing the clinical performance to be assessed with a focus on determining if inadequate sampling led to increased false-negative results. The limitations are that our cohort is not an average-risk population; however, it is expected that similar variability in fecal collection would be observed in other applications of FIT. Another limitation is that assessment has only been made with one brand of FIT, but the strength of this is that this brand is widely used globally (4). In addition, our study did not assess the mass of the sample within the buffer and instead recorded the device mass, which includes feces retained by the collar of the collection device. Although this gives a measure of how much sample was collected, it does not clearly show how much was taken into the buffer. Nonetheless, it clearly showed that sampling less than 10 mg was associated with low f-Hb and higher false-negative rates for detection of neoplasia and IBD.

Sampling of low quantities of feces has an adverse effect on clinical performance and our results on the variability in sample collection show that clear instructions for sampling for FIT are needed. Because under-sampling is the main concern, our results could guide improved strategies, such as through offering replacement FIT to any who returned samples that are analyzed to have f-Hb below the threshold and are deemed to be light in color. Color of the sample buffer could be used to identify potentially inadequate sampling. Ideally, assessment should not be performed manually, but instead would be incorporated as a measurement feature in FIT analyzers.

Authors' Disclosures

E.L. Symonds reports grants and nonfinancial support from Eiken Chemical Company (this company provided a small grant and the fecal immunochemical tests used in the study but had no input into the study design, data analysis, or preparation of this article) during the conduct of the study, as well as grants and nonfinancial support from Clinical Genomics (this company provided small grants and laboratory assays to the institution for separate unrelated studies) outside the submitted work. C.G. Fraser reports personal fees from Hitachi Chemical Diagnostic Systems Co., Ltd. (consultancy) and nonfinancial support from Alpha Labs Ltd. (support to attend relevant meetings) outside the submitted work. D. Bastin reports grants and nonfinancial support from Eiken Chemical Company during the conduct of the study. G.P. Young reports nonfinancial support from Eiken Chemical (this company provided the fecal immunochemical tests used in the study but had no input into the study design, data analysis, or preparation of this article) during the conduct of the study, as well as personal fees from Clinical Genomics (paid consultant for an unrelated ctDNA test for colorectal cancer) outside the submitted work. No disclosures were reported by the other authors.

Authors' Contributions

E.L. Symonds: Conceptualization, data curation, formal analysis, supervision, validation, investigation, methodology, writing–original draft, writing–review and editing. **C.G. Fraser:** Conceptualization, methodology, writing–review and editing. **D. Bastin:** Resources, data curation, investigation, methodology, writing–review and editing. **G. Berwald:** Formal analysis, investigation, writing–review and editing. **G.P. Young:** Conceptualization, resources, supervision, methodology, writing–review and editing.

Acknowledgments

The work received grant funded by the financial support of Cancer Council SA's Beat Cancer Project on behalf of its donors and the State Government of South Australia through the Department of Health together with the support of the Flinders Medical Center Foundation, its donors and partners (to G.P. Young and E.L. Symonds).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received June 30, 2020; revised August 26, 2020; accepted October 6, 2020; published first October 12, 2020.

References

1. Chiu HM, Chen SL, Yen AM, Chiu SY, Fann JC, Lee YC, et al. Effectiveness of fecal immunochemical testing in reducing colorectal cancer mortality from the One Million Taiwanese Screening Program. Cancer 2015;121:3221–9.
2. Cole SR, Tucker GR, Osborne JM, Byrne SE, Bampton PA, Fraser RJ, et al. Shift to earlier stage at diagnosis as a consequence of the national bowel cancer screening program. Med J Aust 2013;198:327–30.
3. Zorzi M, Fedeli U, Schievano E, Bovo E, Guzzinati S, Baracco S, et al. Impact on colorectal cancer mortality of screening programmes based on the faecal immunochemical test. Gut 2015;64:784–90.
4. Schreuders EH, Ruco A, Rabeneck L, Schoen RE, Sung JJ, Young GP, et al. Colorectal cancer screening: a global overview of existing programmes. Gut 2015;64:1637–49.
5. Brown LF, Fraser CG. Effect of delay in sampling on haemoglobin determined by faecal immunochemical tests. Ann Clin Biochem 2008;45:604–5.
6. Grazzini G, Ventura L, Zappa M, Ciatto S, Confortini M, Rapi S, et al. Influence of seasonal variations in ambient temperatures on performance of immunochemical faecal occult blood test for colorectal cancer screening: observational study from the Florence district. Gut 2010;59:1511–5.
7. Symonds EL, Osborne JM, Cole SR, Bampton PA, Fraser RJ, Young GP. Factors affecting faecal immunochemical test positive rates: demographic, pathological, behavioural and environmental variables. J Med Screen 2015;22:187–93.
8. van Rossum LG, van Rijn AF, van Oijen MG, Fockens P, Laheij RJ, Verbeek AL, et al. False-negative fecal occult blood tests due to delayed sample return in colorectal cancer screening. Int J Cancer 2009;125:746–50.
9. Rapi S, Berardi M, Cellai F, Ciattini S, Chelazzi L, Ognibene A, et al. Effects of fecal sampling on preanalytical and analytical phases in quantitative fecal immunochemical tests for hemoglobin. Int J Biol Markers 2017;32:e261–6.
10. Piggott C, John C, Bruce H, Benton SC. Does the mass of sample loaded affect faecal haemoglobin concentration using the faecal immunochemical test? Ann Clin Biochem 2018;55:702–5.
11. Lane JM, Chow E, Young GP, Good N, Smith A, Bull J, et al. Interval fecal immunochemical testing in a colonoscopic surveillance program speeds detection of colorectal neoplasia. Gastroenterology 2010;139:1918–26.
12. Australian Bureau of Statistics (ABS). Census of Population and Housing: Socio-economic Indexes for Areas (SEIFA), Australia 2016. Table 1. State suburb (SSC) index of relative socio-economic disadvantage, distribution of statistical area level 1 (SA1) deciles, 2016 [cited 2019 Jul 20]. Available from: https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/2033.0.55.0012016?OpenDocument.
13. Westwood M, Lang S, Armstrong N, van Turenhout S, Cubiella J, Stirk L, et al. Faecal immunochemical tests (FIT) can help to rule out colorectal cancer in patients presenting in primary care with lower abdominal symptoms: a systematic review conducted to inform new NICE DG30 diagnostic guidance. BMC Med 2017;15:189.
14. Digby J, Cleary S, Gray L, Datt P, Goudie DR, Steele RJC, et al. Faecal haemoglobin can define risk of colorectal neoplasia at surveillance colonoscopy in patients at increased risk of colorectal cancer. United European Gastroenterol J 2020;8:559–66.

15. Symonds EL, Cole SR, Bastin D, Fraser RJ, Young GP. Effect of sample storage temperature and buffer formulation on faecal immunochemical test haemoglobin measurements. J Med Screen 2017;24:176–81.
16. van Rossum LG, van Oijen MG. Different seasons with decreased performance of immunochemical faecal occult blood tests in colorectal cancer screening. Gut 2011;60:1303–4.
17. Rapi S, Rubeca T, Fraser CG. How to improve the performances of fecal immunological tests (FIT): need for standardization of the sampling and pre-analytical phases and revision of the procedures for comparison of methods. Int J Biol Markers 2015;30:e127–31.
18. Gies A, Gruner LF, Schrotz-King P, Brenner H. Effect of imperfect compliance with instructions for fecal sample collection on diagnostic performance of 9 fecal immunochemical tests. Clin Gastroenterol Hepatol 2019;17:1829–39.
19. Amitay EL, Gies A, Weigl K, Brenner H. Fecal immunochemical tests for colorectal cancer screening: is fecal sampling from multiple sites necessary? Cancers 2019;11:400.
20. Ciatto S, Martinelli F, Castiglione G, Mantellini P, Rubeca T, Grazzini G, et al. Association of FOBT-assessed faecal Hb content with colonic lesions detected in the Florence screening programme. Br J Cancer 2007;96:218–21.
21. Digby J, Fraser CG, Carey FA, McDonald PJ, Strachan JA, Diament RH, et al. Faecal haemoglobin concentration is related to severity of colorectal neoplasia. J Clin Pathol 2013;66:415–9.
22. Mowat C, Digby J, Strachan JA, McCann R, Hall C, Heather D, et al. Impact of introducing a faecal immunochemical test (FIT) for haemoglobin into primary care on the outcome of patients with new bowel symptoms: a prospective cohort study. BMJ Open Gastroenterol 2019;6:e000293.