A Proposal to Standardize Reporting Units for Fecal Immunochemical Tests for Hemoglobin

Callum G. Fraser, James E. Allison, Stephen P. Halloran, Graeme P. Young; on behalf of the Expert Working Group on Fecal Immunochemical Tests for Hemoglobin, Colorectal Cancer Screening Committee, World Endoscopy Organization

Manuscript received November 23, 2011; revised March 2, 2012; accepted March 8, 2012.

Correspondence to: Callum G. Fraser, PhD, Centre for Research into Cancer Prevention and Screening, Ninewells Hospital and Medical School, University of Dundee, Dundee DD1 9SY, Scotland (e-mail: callum.fraser@nhs.net).

Fecal immunochemical tests for hemoglobin are replacing traditional guaiac fecal occult blood tests in population screening programs for many reasons. However, the many available fecal immunochemical test devices use a range of sampling methods, differ with regard to hemoglobin stability, and report hemoglobin concentrations in different ways. The methods for sampling, the mass of feces collected, and the volume and characteristics of the buffer used in the sampling device also vary among fecal immunochemical tests, making comparisons of test performance characteristics difficult. Fecal immunochemical test results may be expressed as the hemoglobin concentration in the sampling device buffer and, sometimes, albeit rarely, as the hemoglobin concentration per mass of feces. The current lack of consistency in units for reporting hemoglobin concentration is particularly problematic because apparently similar hemoglobin concentrations obtained with different devices can lead to very different clinical interpretations. Consistent adoption of an internationally accepted method for reporting results would facilitate comparisons of outcomes from these tests. We propose a simple strategy for reporting fecal hemoglobin concentration that will facilitate the comparison of results between fecal immunochemical test devices and across clinical studies. Such reporting is readily achieved by defining the mass of feces sampled and the volume of sample buffer (with confidence intervals) and expressing results as micrograms of hemoglobin per gram of feces. We propose that manufacturers of fecal immunochemical tests provide this information and that the authors of research articles, guidelines, and policy articles, as well as pathology services and regulatory bodies, adopt this metric when reporting fecal immunochemical test results.


Colorectal (bowel) cancer screening using the traditional guaiac-based fecal occult blood test (FOBT) has been shown to reduce mortality in randomized controlled trials (1,2). As a result of this work, these tests are used in many asymptomatic population screening programs, particularly in Europe. Guaiac-based FOBTs are qualitative tests that simply give negative or positive results. They have advantages and disadvantages (3). A substantial disadvantage is that the cutoff to designate guaiac-based FOBT negative and positive results is set by the chemistry adopted by the manufacturer; thus, the positivity rate and the clinical characteristics cannot be adjusted by the end user unless an algorithm is used based on the number of positive results found on the initial guaiac-based FOBT from the six sample windows (4,5). This process is used in the United Kingdom, but it makes the screening program organization and execution more complex (6).

Fecal immunochemical tests for hemoglobin—sometimes called immunochemical FOBTs—have many advantages over the traditional guaiac-based FOBT and are becoming widely used as their merits have become recognized (7). Although colorectal cancer screening approaches differ around the world (8,9), the use of fecal immunochemical tests is recommended in recent guidelines (4,5,8,9). Fecal immunochemical tests, usually abbreviated as FIT, are available in two rather different formats, namely, qualitative (in which the results are either positive or negative) or quantitative (in which fecal hemoglobin concentration is estimated). Quantitative fecal immunochemical tests allow the end user to choose the cutoff fecal hemoglobin concentration that triggers a follow-up colonoscopy. There are many excellent articles on the use of fecal immunochemical tests in the detection of colorectal cancer and adenoma in screening programs (4,5,10–13).

Many different fecal immunochemical test devices are commercially available, and they vary in a number of fundamental aspects, including fecal collection technique, number of samples collected, hemoglobin stability after collection, device technology, analytical methodology, the technique to determine the analytical result, antibody characteristics, and calibration material and derivation of its assigned hemoglobin concentration. Clinicians, laboratory professionals, and health policy experts, among others, want to know which of the many available fecal immunochemical test devices is best suited for their particular screening program. As we have recently discussed (14), the diversity in fecal immunochemical tests in current use makes both performance data and conclusions from clinical studies that use these tests difficult to compare.

It is reasonable to expect that those involved in colorectal cancer screening understand the analytical performance characteristics of the existing tests and standardize how this information is generated.
and presented to all users. We believe that standardized documentation of performance characteristics will allow those who are responsible for procuring colorectal cancer screening tests or comparing their clinical performance to make more informed decisions on which fecal immunochemical test is best suited for their purposes.

**Qualitative Fecal Immunochemical Tests**

Commercially available qualitative fecal immunochemical tests use collection devices that consist of a card or tube; however, compared with automated analytical approaches, these tests, which mainly use immunochromatographic test cassettes for analyses, are not suited to large organized population-based screening programs. An advantage of qualitative fecal immunochemical tests is that their analysis is generally relatively simple and reliable and not subject to dietary interference, so the dietary restriction usually recommended for the more sensitive guaiac-based FOBT is unnecessary: Usually, qualitative fecal immunochemical tests have integral quality control systems to monitor each test performance (15).

Although few studies have directly compared different qualitative fecal immunochemical tests, some data generated for the Scottish Bowel Screening Programme (16, 17) illustrate the issues that we wish to highlight concerning the comparison of data derived from different tests. During these studies on the potential benefits of a two-tier reflex screening approach using a combination of guaiac-based FOBT and a fecal immunochemical test, the use of two qualitative fecal immunochemical tests was examined, and identical results were obtained with both when 200 fecal samples—112 negative and 88 positive—were analyzed. The manufacturers of both tests claimed an identical hemoglobin concentration cutoff of 50 ng hemoglobin per mL buffer. The fact that identical results were obtained in this case might lead to the belief that all qualitative fecal immunochemical test devices provide the same results; however, this is not the case and qualitative fecal immunochemical tests are not interchangeable.

The fact that qualitative fecal immunochemical tests differ with respect to their analytical and clinical performance has been demonstrated by Brenner et al. (18). The intertest agreement of six qualitative fecal immunochemical tests was studied through simultaneous analyses of the same fecal samples. The positivity rates of the six tests were 6.1%, 11.0%, 22.3%, 24.1%, 35.0%, and 46.8%; earlier, this group documented the respective cutoff hemoglobin concentrations for these six fecal immunochemical tests quoted by the manufacturers as 50, 40, 10, 40, 50, and 25 ng hemoglobin per mL buffer (19). As reported in detail previously (18), clinical sensitivity and specificity for advanced neoplasia also varied widely and, as expected, showed clear positive and negative relationships, respectively, with the positivity rate. Although it was suggested (18) that the very different positivity rates reflected the different test cutoff hemoglobin concentrations, our objective comparison of the relationships between cutoff fecal hemoglobin concentrations and clinical characteristics (14) clearly showed that the positivity rate is not directly related to the stated cutoff hemoglobin concentration, at least when it is expressed as nanograms of hemoglobin per milliliter of buffer. This is potentially confusing because it might be assumed that the fecal immunochemical test with the lowest quoted cutoff hemoglobin concentration would give the highest positivity rate and have the highest clinical sensitivity and the lowest specificity, but this is clearly not so.

There are other disadvantages of qualitative fecal immunochemical tests. In addition to the major disadvantage that the cutoff hemoglobin concentration, and, as a consequence, the positivity rate and clinical characteristics, is set by the manufacturer, qualitative fecal immunochemical tests take longer to analyze compared with guaiac-based FOBTs, require subjective visual reading, and the analyses cannot be automated (15). These disadvantages make qualitative fecal immunochemical tests much less attractive than quantitative automated fecal immunochemical test analytical systems for use in large screening programs.

**Quantitative Fecal Immunochemical Tests**

Each quantitative fecal immunochemical test uses a different fecal sample collection device. Many use a stick or probe that picks up a quantity of feces that varies between individual devices. The collected fecal material is then placed in a tube that contains a buffer, which varies among devices from different manufacturers with respect to volume, buffer composition, and preservative properties. It is surprising that, unlike other newer fecal markers such as calprotectin (20), fecal hemoglobin concentrations are reported in a variety of units. These differences make simple comparisons of data between publications that report results of different quantitative fecal immunochemical tests impossible.

The cutoff hemoglobin concentration for fecal immunochemical tests is often given in nanograms of hemoglobin per milliliter of buffer, mainly because that is how a number of available analytical systems report the results. Many studies on the use of quantitative fecal immunochemical tests have used the OC-SENSOR series of analyzers (Eiken Chemical Co, Tokyo, Japan). Although a number of investigations using this particular analytical system have provided valuable insight into the effect using different cutoff hemoglobin concentrations and numbers of test samples on test positivity, clinical sensitivity, and specificity (3, 4), many investigators have simply adopted the manufacturer’s suggested cutoff concentration of 100 ng hemoglobin per mL buffer and have tested a single sample. Results are expressed as either positive or negative, without reference to the hemoglobin concentrations found in the subjects studied; that is, this quantitative test is most often used as a qualitative test despite the fact that there are potential advantages in using the fecal hemoglobin concentrations found in individuals (21–24). Because the OC-SENSOR is currently the most widely adopted quantitative fecal immunochemical test, the commonly used cutoff of 100 ng hemoglobin per mL buffer, appropriate perhaps for this particular analytical system, mistakenly seems to have become viewed as being an appropriate cutoff for other devices and, by default, the standard cutoff. This cutoff simply reflects the concentration of hemoglobin in the sample buffer and not the feces, and thus cannot be translated to other tests. Moreover, recent publications sometimes do not even give the units that are used for test results and express the hemoglobin cutoff concentrations using cryptic terms [eg, “a FIT 50 strategy” (25)]. Such shorthand expressions have no clear meaning and
simply add to the confusion. Results of quantitative measurements made in medicine should always include the matrix, the analyte, the numerical result, and the units (eg, in the fecal immunochemical test setting, fecal hemoglobin = 50 ng hemoglobin per mL buffer).

Given that many countries have a public health need to control the number of colonoscopies because of limited local capacity and to choose the preferred cutoff hemoglobin concentration, rather than adopt that suggested by a manufacturer (26), many (3,4) consider a quantitative fecal immunochemical test that enables such flexibility and facilitates transferability of fecal hemoglobin concentration data over time and geography to be the ideal for screening programs. However, when fecal immunochemical test results are expressed in nanograms of hemoglobin per milliliter of buffer, the cutoff used cannot be applied to other devices or systems. For example, the manufacturer of one quantitative fecal immunochemical test—FOB Gold (Sentinel Diagnostics SpA, Milan, Italy)—suggests that the user set the cutoff hemoglobin concentration, which may lead to different clinical interpretations, depending on the cutoff concentration selected (5). Moreover, the cutoff hemoglobin concentrations recommended by the manufacturers of two other quantitative fecal immunochemical tests (ie, RIDASCREEN Hemoglobin [R-Biopharm AG, Bensheim, Germany] and HM-Jack [Kyowa Medex Co Ltd, Tokyo, Japan]) are expressed in different units (ie, 2 and 33 μg hemoglobin per g feces, respectively) (18,27). Thus, a variety of units are used to report fecal hemoglobin concentrations when measured using different fecal immunochemical test devices and systems. Those devices that express the test results in nanograms of hemoglobin per milliliter of buffer have different cutoff hemoglobin concentrations because they use sample collection devices that pick up different masses of feces and deliver these into differing volumes of buffer. As a consequence, such cutoff hemoglobin concentrations are therefore unique to the device or system and cannot be compared with other devices, regardless of how similar they appear.

A Plan for Standardization

A number of manufacturers of quantitative fecal immunochemical tests have determined the mass of feces that is picked up in their particular collection device (5). Knowledge of the mass of feces collected in any fecal immunochemical test device is a necessary prerequisite for adequate documentation of performance characteristics, and for the generation and use of more appropriate units for reporting results. We recommend that product literature provide details of the mean fecal mass sampled, with confidence intervals. Ideally, the manufacturer should also document the method that was used to derive such data and give additional information, such as the number of samples that was analyzed in each manufacturing lot to derive the mean value and confidence intervals. Moreover, the type(s) of fecal sample examined (eg, loose, soft, firm, hard) should be documented, perhaps, using the well-known Bristol Stool Scale (28). In addition, if tubes are used as collection or sample preparation devices, manufacturers should state the mean volume of buffer, again with confidence intervals derived over a number of manufacturing lots and with details of the method of their derivation.

Improving the Transferability of Data

The way to improve the transferability of data between devices and studies is straightforward. If one knows (or can estimate) the mass, in milligrams, of the sample taken into or onto the collection device and the volume of buffer, in milliliters, into which this mass of feces is delivered, the mass of hemoglobin per mass of feces can be estimated. Results for all fecal immunochemical test devices can then be expressed as the ratio of two masses, that is, micrograms of hemoglobin per gram of feces, as has been done for other fecal markers (20). The conversion of data in previous publications that are expressed nanograms of hemoglobin per milliliter of buffer to the recommended units can be achieved simply according to the formula: μg hemoglobin per g feces = (ng hemoglobin per mL × mL buffer)/(mg feces collected).

For example, according to the manufacturer, the OC-SENSOR delivers 10 mg of feces into in 2.0 mL of buffer; thus, a test result of 100 ng hemoglobin per mL buffer equals 20 μg hemoglobin per g feces. According to the manufacturer, HM-JACK delivers 0.5 mg of feces into 1.25 mL of buffer; thus, a test result of 100 ng hemoglobin per mL buffer with this system equals 250 μg hemoglobin per g feces. According to the manufacturer, SentiFOB (Sentinel Diagnostics SpA, Milan, Italy) delivers 10 mg feces into 1.7 mL buffer; thus, a test result of 100 ng hemoglobin per mL buffer equals 17 μg hemoglobin per g feces. Thus, with consistent of reporting in units of micrograms of hemoglobin per gram of feces, different analytical systems can now be compared objectively.

An analytical system-specific multiplier can be derived to express nanograms of hemoglobin per milliliter of buffer as micrograms of hemoglobin per gram of feces and then be applied to all studies using the particular analytical system. For example, data expressed as nanograms of hemoglobin per milliliter of buffer would be multiplied by 0.2 for the OC-SENSOR and by 2.5 for HM-JACK to convert to micrograms of hemoglobin per gram of feces. Manufacturers could incorporate this multiplication factor into the software of their analytical systems so that users could select their chosen readout (either the preferred μg hemoglobin per g feces or ng hemoglobin per mL buffer). Accurate comparison of different qualitative and quantitative fecal immunochemical tests then becomes more possible when the data are expressed as micrograms of hemoglobin per gram feces.

In addition, having confidence interval data for the estimates of mass and volume would allow objective estimates of the uncertainty of a measurement to be calculated; however, these calculations are somewhat complex, as described by White (29). Although there is a substantial literature on deciding how good laboratory tests must be to facilitate clinical decision making (30), to date, there have been no publications, or even suggestions, as to the best method for setting such analytical quality specifications for the uncertainty of measurement of fecal hemoglobin: This is a matter for future development. Defining how good the tests should be and the actual quality achieved in practice are necessary prerequisites for accreditation of medical laboratories in many countries. Ideally, all laboratories involved in analyses of feces for hemoglobin should be accredited under the international standards based on ISO 15189 (31) and should apply the many relevant quality assurance techniques documented in detail in recent European Guidelines.
for Quality Assurance in Colorectal Cancer Diagnosis and Screening so as to minimize the overall uncertainty of the measurement (5).

**Standardizing Other Aspects of Fecal Immunochemical Tests**

Although the use of appropriate units and knowledge of the mass and volume of the sample would allow a much better comparison of results generated using these devices, it is important to recognize that other analytical differences hinder the transferability of numerical data and cutoff hemoglobin concentrations between different devices. For example, in the context of a screening program, even the same device may respond differently to the variety of geographical, climatic, and social conditions, particularly those associated with sample handling techniques, storage arrangements, and transport protocols (32).

Results obtained with fecal immunochemical tests will also be influenced by the characteristics of the antibodies used to detect different parts of the hemoglobin molecule and the degraded protein. In addition, further work is needed to standardize the documentation on the practicability and reliability performance characteristics that is required for all users to use the chosen test. The information required has been described in large part by the US Food and Drug Administration (FDA) (33). We support the recommendation of the FDA (33) that manufacturers of qualitative fecal immunochemical tests make use of the Clinical and Laboratory Standards Institute document User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline—Second Edition (EP12-A2) (34) when determining the cutoff hemoglobin concentration of their devices and use the terminology therein. There is also a need for documentation and standardization of the hemoglobin material used in determining cutoff hemoglobin concentrations and in calibrating automated systems. It would be of considerable advantage if the hemoglobin material used had its characteristics determined using reference methods and materials (29). The Expert Working Group on Fecal Immunochemical Tests for Hemoglobin, Colorectal Cancer Screening Subcommittee, World Endoscopy Organization, will address these issues in more detail in future planned recommendations and guidelines on fecal immunochemical testing. Other expert working groups will be established in due course to address standardization of many other aspects of screening for colorectal neoplasia, including strategies for follow-up of those who are fecal immunochemical test positive, given that overall test performance involves not only fecal immunochemical testing but also diagnostic follow-up.

**Concluding Remarks**

Health policy experts, guideline-producing bodies, regulatory agencies, and population screening efforts require knowledge that will enable selection of the fecal immunochemical test system that is best suited to meet a priori specified clinical and logistical requirements. Given the complexities inherent in the use of different analytical systems in different conditions, we strongly advocate the adoption of micrograms of hemoglobin per gram of feces as the metric for reporting results of fecal immunochemical tests.

To facilitate this approach, manufacturers and suppliers of fecal immunochemical tests that report results as nanograms of hemoglobin per milliliter of buffer will need to supply detailed validated information on the mass of feces collected and the volume used in the fecal immunochemical test specimen collection device.

**References**


jnci.oxfordjournals.org
Commentary

The authors received no external funding for this study.

Funding

Notes

Dr C G Fraser is a paid consultant to Immunostics, Inc, and has received travel, accommodations, and/or meeting expenses from Alpha Laboratories Ltd, both of which are suppliers of fecal tests. The authors are solely responsible for the writing of this Commentary and the decision to submit the Commentary for publication. The concepts were debated at a meeting of the Colorectal Cancer Screening Committee, World Endoscopy Organization, held in Chicago, IL, on May 6, 2011, and the content was further discussed at a further meeting of the Colorectal Cancer Screening Committee of the World Endoscopy Organization held in Stockholm, Sweden, on October 21, 2011. The concepts described in this Commentary received the imprimatur of those present, including staff of the Quality Assurance Group of the International Agency for Research on Cancer at the latter meeting.

Affiliations of Authors: Centre for Research into Cancer Prevention and Control, Flinders University, Bedford Park (GPY), South Australia, Australia; Kaiser Permanente Northern California Division of Research, Oakland, CA (JEIA); Bowel Cancer Screening Southern Programme Hub, Royal Surrey County Hospital NHS Foundation Trust, Guildford, Surrey, UK (SPH); Centre for Cancer Prevention and Control, Flinders University, Bedford Park (Adelaide), South Australia, Australia (GPY).