Editorial

Callum G. Fraser, Stephen P. Halloran, James E. Allison and Graeme P. Young

Making colorectal cancer screening FITTER for purpose with quantitative faecal immunochemical tests for haemoglobin (FIT)

Colorectal cancer (CRC) is a major cause of disease worldwide and a leading cause of mortality. Fortunately, CRC is a disease that meets the criteria for successful early detection and treatment through population screening. Moreover, CRC incidence is decreasing in some countries which, at least in part, may be due to the establishment and expansion of screening programmes. Screening tests can be applied either in one-step screening using colonoscopy or flexible sigmoidoscopy, or in two-step screening using simpler non-invasive tests such as faecal tests for occult blood. Tests for the assessment of blood in faeces are widely considered to be the best currently available for two-step screening and are very suited to large scale population-based screening programmes [1]. Traditionally, guaiac-based faecal occult blood tests (gFOBT) have been used and four randomised controlled trials, in the setting of population-based programmes, showed a modest intention-to-treat decrease in mortality [2] that has been demonstrated to occur also in practice [3, 4]. However, gFOBT have many well-known limitations [5] that do not apply to the newer faecal immunochemical tests (FIT) for haemoglobin (Hb) [6]. FIT are now being widely adopted in new screening programmes and existing programmes using gFOBT are moving to FIT.

The merits of FIT over gFOBT are summarised in this issue of the journal by Huang and colleagues [7], who collected faeces in a single specimen collection device from a representative random population in China. The specimens were then analysed using one quantitative and two qualitative FIT. The authors documented the cut-off haemoglobin concentration which they used for referral of the participants for colonoscopy as 20 \( \mu \text{g Hb/g faeces} \) for the quantitative FIT and 20 \( \mu \text{g Hb/g faeces} \) and 40 \( \mu \text{g Hb/g faeces} \) for the two qualitative FIT. The appropriate use by the authors of \( \mu \text{g Hb/g faeces} \) for reporting FIT concentrations is consistent with good practice, and our recent recommendations [8], since these units facilitate more objective comparison of data across different FIT, a process impeded by the current wide use of proprietary ng Hb/mL buffer units.

The study demonstrated different positivity rates for the two qualitative FIT, evidence of the different faecal Hb cut-off concentration needed for a colour change on the immunochromatographic test strips. This is an important finding which confirms previous evidence that available qualitative FIT are not all the same [9]. The cut-off faecal Hb for qualitative FIT are set by the manufacturers and, as a consequence, it is the manufacturers that determine the important clinical outcomes. As the cut-off faecal Hb concentration is decreased, the positivity rate and colonoscopy demand increase, the clinical sensitivity increases, specificity decreases, positive-predictive value decreases and number needed to scope increases. Surely this critical clinical variable, the cut-off faecal Hb concentration, should be set by screening programme organisers so that effective use of the endoscopic resources is enabled and clinical expectations met.

Huang et al. provide a novel solution to the setting of the cut-off faecal Hb concentration for qualitative FIT [7]. The criteria for a positive result for the qualitative FIT were set according to the density of the colour appearing in the test strip. The authors created 10 positive criteria for the qualitative FIT by measuring the density of the colour in the strips of the qualitative FIT, producing a colour ladder from faint pink to dark red with the different grades corresponding to increasing faecal Hb concentrations. Faecal specimens with a test strip colour darker than the designated cut-off colour grade were regarded as positive. Setting the cut-off subjectively at a specific colour rather than the simple presence or absence of colour for one of the qualitative FIT did improve the comparability of results across the quantitative and both quantitative FIT. The authors state that it would be necessary to intensify quality control and refine positive criteria for qualitative FIT to be used for CRC screening and suggest that, otherwise, quantitative FIT might be a better choice [7]. We agree and have recently documented
some of the many reasons why quantitative FIT analyses using automated analytical systems are considered superior to qualitative FIT [10]. Quantitative FIT allow high throughput with high analytical quality, but their most important advantage is that organisers can set one or more cut-off faecal Hb concentrations to fulfil the pre-set objectives of their particular programme.

Huang et al. [7] performed a comparison of three FIT using specimens of faeces from participants invited for screening. The pre-analytical aspects and analytical performance characteristics are described in some detail. In contrast, the evaluation of Tao et al. [9] compared test performance of six qualitative and three quantitative FIT with respect to their ability to detect CRC. The sensitivities and specificities were reported in considerable detail and it was stated that the cut-off concentration of several qualitative FIT needed to be adjusted to limit false-positive rates in the screening setting. However, nowhere in this study were analytical performance characteristics and, more importantly, the cut-off concentrations of the FIT reported. Similarly, the important study of Raging et al. [11] gives very comprehensive informative data on clinical outcomes but no data on important pre-analytical aspects such as time and storage of collection devices from sample collection to analysis including time and temperature. We recommend a full description of the way specimens are treated from collection through to analysis as we do for the analytical performance characteristics and quality management strategies, both of which also lacking in this report. In a comparison of qualitative and quantitative FIT, Park et al. [12] gave no information on important pre-analytical aspects such as time and storage of collection devices, no information on the analytical methods used, not even the names of the FIT, and no information on analytical performance, or cut-off faecal Hb concentrations. Unfortunately, many other publications on FIT include abundant clinical details but provide scant descriptions of pre-analytical and analytical aspects. For published data to be effectively evaluated, compared and/or reproduced, authors need to record and publish this information.

A decade ago, the Standards for Reporting of Diagnostic Accuracy (STARD) were developed to improve the reporting of studies on diagnostic accuracy [13]. These standards require a full description of the technical specifications of material and methods used, including how and when measurements were taken, a definition of and rationale for the units, cut-offs and/or categories of the results of the index test and the reference standard, the number, training and expertise of the persons executing and reading the index tests and the reference standard and methods for calculating test reproducibility. Recently, the Consortium of Laboratory Medicine Journal Editors lauded the impact of the STARD guidelines but, in view of the lack of compliance, particularly for newer biomarkers, emphasised that authors must report the following: for commercial diagnostic tests: authors must include the actual name and generation of assay, the manufacturer and the instrument used for analyses, performance characteristics, such as the imprecision of the assay in the investigators’ laboratories, the assay’s reportable range, and any reference (normal) range used in the study and must clearly indicate the types of specimens analysed and the storage conditions for these specimens [14]. These items seem insufficient to us for FIT, since certain data required for complete comprehension are not detailed. We believe that the following are desirable in all publications on FIT, although we recognise that journal space is often limited and such information might be made available in other ways, such as in supplemental data files:

- description of specimen collection device;
- details of faecal collection method, number of faecal samples;
- handling and storage of collection devices from sample collection to analysis, including time and temperature;
- analyser make and model and number of systems used;
- number of times each sample was analysed;
- analytical range and whether samples outside this range were diluted and re-assayed;
- analytical precision;
- mode of collection of data;
- units used, with conversion to μg Hb/g faeces if ng Hb/mL used; and
- cut-off concentration(s) with explanation of how assigned.

Assessment of what we consider is required will be elaborated in the near future by dissemination of the Faecal Immunochemical Tests for Haemoglobin Evaluation Reporting (FITTER) standards and check-list currently in the final stages of development by the Expert Working Group on FIT, Colorectal Cancer Screening Committee, World Endoscopy Organization. We trust that wide adoption of FITTER by those preparing, reviewing and editing manuscripts on studies using FIT will enhance understanding of the strengths and limitations of FIT and enable improved clinical outcomes.
Conflict of interest statement

**Author’s conflict of interest disclosure:** CGF consults with Immunostics Inc and Mode Diagnostics and has received travel funding from Alpha Labs Ltd: other authors stated that they have no conflicts of interest regarding the publication of this article. Research funding played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

**Research funding:** Loan of instrumentation and supply of reagents to CGF from Eiken Chemical Co and Mast Diagnostics.

**Employment of leadership:** None declared.

**Honorarium:** None declared.

References


*Corresponding author: Professor Callum G. Fraser, PhD, Centre for Research into Cancer Prevention and Screening, University of Dundee, Dundee DD1 9SY, Scotland, UK, Phone: +44 1382553799, Fax: +44 1382425679, E-mail: callum.fraser@nhs.net*  
Stephen P. Halloran: NHS Bowel Cancer Screening Hub – South of England, Royal Surrey County Hospital NHS Trust and University of Surrey, Guildford, Surrey, UK  
James E. Allison: Division of Gastroenterology, University of California San Francisco, San Francisco, CA, USA  
Graeme P. Young: Flinders Centre for Innovation in Cancer, Flinders University, Bedford Park (Adelaide), South Australia, Australia