Commentary

Comparison of quantitative faecal immunochemical tests for haemoglobin (FIT) for asymptomatic population screening

Callum G. Fraser

Centre for Research into Cancer Prevention and Screening, University of Dundee, Ninewells Hospital and Medical School, Dundee, Scotland
Correspondence to: Professor Callum G. Fraser. Centre for Research into Cancer Prevention and Screening, University of Dundee, Ninewells Hospital and Medical School, Dundee DD1 9SY, Scotland. Email: callum.fraser@nhs.net.
Comment on: Grobbee EJ, van der Vlugt M, van Vuuren AJ, *et al*. A randomised comparison of two faecal immunochemical tests in population-based colorectal cancer screening. Gut 2016. pii: gutjnl-2016-311819. [Epub ahead of print].

Submitted Sep 02, 2016. Accepted for publication Sep 10, 2016.
doi: 10.21037/tcr.2016.10.22
View this article at: http://dx.doi.org/10.21037/tcr.2016.10.22

Introduction

Although colonoscopy, despite disadvantages, remains the "gold standard" for investigation of suspected colorectal disease, faecal immunochemical tests for haemoglobin (FIT) are now considered the best non-invasive investigation for use in asymptomatic population-based programmatic screening for colorectal neoplasia (1). FIT come in two formats, qualitative tests, based upon immunochromatography, and quantitative tests, based on immunoturbidimetry. Qualitative FIT have some advantages, including potential as point-of-care tests, but have major disadvantages (2), particularly that they have different analytical detection limits and thus give very different clinical outcomes (3). Quantitative FIT are much more suitable for larger-scale screening programmes and have many advantages (4), especially that estimates of faecal haemoglobin concertation (f-Hb) are obtained, allowing considerable flexibility in programme design.

A number of FIT analytical systems are now available and the spectrum continues to increase. In consequence, selection of a system is problematic for those currently planning to introduce FIT-based screening programmes and those evolving programmes from guaiac-based faecal occult blood tests (gFOBT). At present, there is little evidence that any one FIT system has major benefits and, thus, a recently published study on a randomised comparison of two FIT (5) is of timely interest.

Comparison of quantitative FIT analytical systems

Comparisons of FIT analytical systems have been performed for some time and there are a number of possible strategies to determine their characteristics. Some, for example the comparison of four FIT systems commissioned by the NHS Bowel Cancer Screening Programme in England (6), are essentially evaluations of the analytical performance characteristics. This evaluation used both faecal and aqueous samples containing added haemoglobin (Hb). These two matrices were applied to assess analytical sensitivity, carryover, imprecision, precision profiles, linearity, and Hb stability in the sample collection devices. Although much information is generated, such an evaluation gives limited information about performance on the real matrix analysed in screening programmes, that is, native faeces. Realisation of the limitations of this approach led Rubeca *et al*. (7) to propose two protocols to discriminate pre-analytical from analytical variation and investigate overall clinical performance. These protocols were used to compare two FIT systems available in Europe, OC-Sensor Diana and HM-JACKarc, but are again based on artificial samples spiked with Hb. Studies using faecal samples from participants in screening that were frozen and thawed so that a number of FIT could be assessed "simultaneously" have been used to determine data on clinical outcomes (8), but faecal Hb is unstable (9) and

freezing and thawing might lead to erroneous results. Thus, while the matrix may have been somewhat more appropriate than samples spiked with Hb, the strategy seems flawed. Probably the best approach to comparison of FIT systems is exemplified by that performed in Florence (10) in which the HM-JACK was compared with the OC-Hemodia by participants collecting samples for both systems from single passed bowel motions, which were then analysed with both FIT. A similar design was used in studies in France in which gFOBT and a number of FIT were compared in large populations recruited from well-established screening programmes (11,12). However, persuading participants in such programmes.to collect samples from a single bowel motion with more than one FIT collection device—which are similar, but significantly different—must pose significant logistical challenges and some difficulties for participants.

A single sample per participant strategy for comparison of FIT

The approach used in the very recently published study performed in The Netherlands by Grobbee *et al*. (5), using one only sample per participant, has advantages. The study aimed to compare OC-Sensor and FOB-Gold FIT systems with regard to participation rate, usability, positivity rate and diagnostic yield. The study involved comparison of the two FIT in a fourth-round population-based screening cohort. Randomly selected, on a 1:1 basis prior to invitation, individuals received one only specimen collection device as required for either one of two OC-Sensor µ systems or FOB-Gold on a single SENTiFIT system. Interestingly, to avoid confusion, participants in the one household, received the same FIT.

Although the strategy adopted for calibration and control of the analyses is described (5), disappointingly, no data are given on the actually analytical performance achieved, in terms of analytical CV, for example. Moreover, no data are given on the comparability of the two OC-Sensor systems used or the comparative performance of the two laboratories in which the analyses were performed. It has been recommended by the Expert Working Group on FIT for Screening (EWG), Colorectal Cancer Screening Committee of the World Endoscopy Organization, that all publications involving FIT analyses should comply with the FITTER guidelines (13).

Further, the sample handling prior to analysis was not the same for both analytical systems in that, after arrival at the laboratory, until analysis, the OC-Sensor sample was stored at –20 ℃ and the FOB-Gold FIT sample was stored at 4 ℃. Although it is stated that both FIT samples were stored according to the manufacturer's recommendations, no evidence was provided that freezing and thawing ensures sample stability. Analysis of fresh samples would seem advantageous and is recommended for future studies.

In this well powered study (5), 19,291 eligible invitees were included: the randomisation clearly worked well in that 9669 invitees received OC-Sensor devices and 9622 FOB-Gold devices. Both FIT devices were returned by 63% of invitees. Positivity rate was 7.9% for OC-Sensor, significantly different from the 6.5% for FOB-Gold. However, there was no significant difference in diagnostic yield of advanced neoplasia, defined as cancer plus advanced adenoma (1.4% for OC-Sensor and 1.2% for FOB-Gold, or the positive predictive value (31% and 32% respectively). It was concluded that OC-Sensor and FOB-Gold were equally acceptable to a screening population. Interestingly, very similar conclusions were drawn from a very recently published study done in Latvia comparing these same FIT systems (14), but there were differences, for example, the positivity was 8.3% for OC-Sensor and 12.8% for FOB Gold. Interestingly, in an older Spanish study, FOB-Gold also had a higher positivity rate than OC-Sensor (15), but there are a number of plausible reasons why that particular study differed from those done more recently (5). The very important point is that different quantitative FIT systems report different positivity rates even at seemingly identical cut-off f-Hb.

As recommended by the EWG (13), both studies (5,14) used units of µg Hb/g faeces. These should be used ubiquitously, because such units are required since different specimen collection devices sample different amounts of faeces into different volumes of buffer. Thus, the traditional units used in this field of ng Hb/ml buffer are system-specific and do not facilitate comparison of data across systems, for example, the 10 µg Hb/g faeces used as cut-off f-Hb for referral for colonoscopy with both systems corresponds to 50 ng Hb/mL buffer for OC-Sensor and 59 ng Hb/mL buffer for FOB-Gold. However, as above, there are differences in the positivity rates in spite of the use of these supposedly common units. This is undoubtedly for a number of reasons, as nicely stated by Grobbee *et al*. (5). Firstly, FIT generally sample wet faeces into buffer in the specimen collection devices and, although the assumption is that the volume of faecal material sampled is constant over devices, the amounts sampled can vary substantially in reality. Secondly, different FIT make use of antibodies against different globin epitopes

and this could potentially influence positivity rate. As for many other measurands examined in laboratory medicine, these systems are all said to measure "faecal haemoglobin" but, in reality, they measure faecal haemoglobin plus a range of early degradation products, which probably vary from system to system. Thus, the measurand is not identical across systems. A very important conclusion from the work (5) was that, despite apparent harmonisation using identical f-Hb cut-offs, adjusting for positivity rate instead would result in an equal number of colonoscopies, and a similar diagnostic yield. An earlier study from Taiwan also concluded that even after harmonisation, different FIT perform differently (16). The important recommendation made by Grobbee *et al.* (5) that identical positivity rates rather than the same f-Hb cut-offs should be used in comparison of FIT seems cogent and compelling.

Conclusions and a caveat

Recent studies from The Netherlands (5) and Latvia (14) have both compared two different FIT systems. The results may help those planning to begin FIT-based screening programmes. But, the final germane question is: "how transferable will these data be over time and geography?" As pointed out by Grobbee *et al.* (5), the manufacturers of FIT continually evolve their products and thus, outcomes may not be comparable over time. An example of change over time is that both FIT systems studied (5,14) have had improvements made to their buffers to increase the Hb stability (17,18). Moreover, population characteristics are not the same over geography: for example, there are differences in f-Hb over age and sex amongst countries (19). Thus, an important caveat is clear: before application, users of published data on FIT systems should investigate whether the systems being considered for adoption will actually have the same characteristics as those documented in the literature.

Acknowledgments

Funding: None.

Footnote

Provenance and Peer Review: This article was commissioned and reviewed by the Section Editor Xiaoying Zhou (Institute of Gastroenterology, the First Affiliated Hospital of Nanjing Medical University, Nanjing, China).

Conflicts of Interest: CGF undertakes paid consultancy with Immunostics Inc., Ocean, NJ, USA, and Kyowa-Medex Co., Ltd, Tokyo, Japan, and has received support for travel and attendance at conferences from Alpha Labs Ltd, Eastleigh, Hants, UK.

Ethical Statement: The author is accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Young GP, Symonds EL, Allison JE, et al. Advances in Fecal Occult Blood Tests: the FIT revolution. Dig Dis Sci 2015;60:609-22.
2. Allison JE, Fraser CG, Halloran SP, et al. Population screening for colorectal cancer means getting FIT: the past, present, and future of colorectal cancer screening using the fecal immunochemical test for hemoglobin (FIT). Gut Liver 2014;8:117-30.
3. Tao S, Seiler CM, Ronellenfitsch U, et al. Comparative evaluation of nine faecal immunochemical tests for the detection of colorectal cancer. Acta Oncol 2013;52:1667-75.
4. Fraser CG, Allison JE, Young GP, et al. Quantitation of hemoglobin improves fecal immunochemical tests for noninvasive screening. Clin Gastroenterol Hepatol 2013;11:839-40.
5. Grobbee EJ, van der Vlugt M, van Vuuren AJ, et al. A randomised comparison of two faecal immunochemical tests in population-based colorectal cancer screening. Gut 2016. pii: gutjnl-2016-311819. [Epub ahead of print].
6. Carroll MR, Seaman HE, Halloran SP. Tests and investigations for colorectal cancer screening. Clin Biochem 2014;47:921-39.
7. 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.
8. Brenner H, Haug U, Hundt S. Inter-test agreement and quantitative cross-validation of immunochromatographical fecal occult blood tests. Int J Cancer 2010;127:1643-9.
9. Brown LF, Fraser CG. Effect of delay in sampling on haemoglobin determined by faecal immunochemical tests. Ann Clin Biochem 2008;45:604-5.
10. Rubeca T, Rapi S, Confortini M, et al. Evaluation of diagnostic accuracy of screening by fecal occult blood testing (FOBT). Comparison of FOB Gold and OC Sensor assays in a consecutive prospective screening series. Int J Biol Markers 2006;21:157-61.
11. Guittet L, Guillaume E, Levillain R, et al. Analytical comparison of three quantitative immunochemical fecal occult blood tests for colorectal cancer screening. Cancer Epidemiol Biomarkers Prev 2011;20:1492-501.
12. Faivre J, Dancourt V, Denis B, et al. Comparison between a guaiac and three immunochemical faecal occult blood tests in screening for colorectal cancer. Eur J Cancer 2012;48:2969-76.
13. Fraser CG, Allison JE, Young GP, et al. Improving the reporting of evaluations of faecal immunochemical tests for haemoglobin: the FITTER standard and checklist. Eur J Cancer Prev 2015;24:24-6.
14. Santare D, Kojalo I, Liepniece-Karele I, et al. Comparison of the yield from two faecal immunochemical tests at identical cutoff concentrations - a randomized trial in Latvia. Eur J Gastroenterol Hepatol 2016;28:904-10.
15. Zubero MB, Arana-Arri E, Pijoan JI, et al. Population-based colorectal cancer screening: comparison of two fecal occult blood test. Front Pharmacol 2014;4:175.
16. Chiang TH, Chuang SL, Chen SL, et al. Difference in performance of fecal immunochemical tests with the same hemoglobin cutoff concentration in a nationwide colorectal cancer screening program. Gastroenterology 2014;147:1317-26.
17. Dancourt V, Hamza S, Manfredi S, et al. Influence of sample return time and ambient temperature on the performance of an immunochemical faecal occult blood test with a new buffer for colorectal cancer screening. Eur J Cancer Prev 2016;25:109-14.
18. Gnatta E, Zaninotto M, Epifani MG, et al. A new sampling device for faecal immunochemical testing: haemoglobin stability is still an open issue. Clin Chem Lab Med 2014;52:1203-9.
19. Fraser CG, Rubeca T, Rapi S, et al. Faecal haemoglobin concentrations vary with sex and age, but data are not transferable across geography for colorectal cancer screening. Clin Chem Lab Med 2014;52:1211-6.

Cite this article as: Fraser CG. Comparison of quantitative faecal immunochemical tests for haemoglobin (FIT) for asymptomatic population screening. Transl Cancer Res 2016;5(Suppl 4):S916-S919. doi: 10.21037/tcr.2016.10.22

