Acta Oncologica, 2013; 52: 1667–1675

informa
healthcare

ORIGINAL ARTICLE

Comparative evaluation of nine faecal immunochemical tests for the detection of colorectal cancer

SHA TAO[1], CHRISTOPH M. SEILER[2], ULRICH RONELLENFITSCH[3] & HERMANN BRENNER[1]

[1]*Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Germany,* [2]*Department of General, Visceral, and Trauma Surgery, University Clinic Heidelberg, University of Heidelberg, Heidelberg, Germany and* [3]*Department of Surgery, University Hospital Mannheim, University of Heidelberg, Mannheim, Germany*

Abstract

Background. Faecal immunochemical tests (FITs) for haemoglobin are increasingly used for non-invasive screening for colorectal cancer (CRC) but large scale comparative studies of different FITs for detection of CRC, overall and by stage, are sparse. We aimed to determine and compare performance of different FITs for the detection of CRC, and to assess their stage-specific sensitivities. *Material and methods.* We assessed sensitivity, specificity and their corresponding 95% confidence intervals for six qualitative FITs among 74 CRC cases (59% stage I or II cancers) and 1480 controls free of colorectal neoplasm. Overall and stage-specific receiver operating characteristic curves were derived for three quantitative FITs. The areas under the curves (AUCs) were calculated and compared. *Results.* Pairs of overall sensitivity and specificity of the qualitative FITs ranged from 66% and 96% to 92% and 62%, respectively. For the three quantitative tests, AUCs ranged from 0.90 to 0.92, with sensitivities ranging from 80% to 87% at cut-offs yielding 90% specificity. AUCs ranged from 0.85 to 0.92, 0.94 to 0.96, and 0.86 to 0.93 for stage I, stage II and advanced stages (stage III and IV) cancers, respectively. At a specificity of 90%, the tests detected 65%–85% of stage I cancers. *Conclusion.* The diagnostic performance of FITs regarding detection of CRC is promising, even though the pre-defined cut-offs of some of the qualitative FITs need to be adjusted to limit false-positive rates in screening setting. At cut-off levels yielding 90% specificity, the quantitative tests detected the vast majority of CRCs, even at early stages.

Screening for colorectal cancer (CRC) using faecal occult blood tests (FOBTs) has been shown to reduce both mortality and incidence of CRC [1]. However, the guaiac-based FOBT (gFOBT), which has been used for decades, has a number of limitations, in particular low sensitivity and the effect of diet [1]. In recent years, both qualitative (dichotomous) and quantitative faecal immunochemical tests (FITs) for haemoglobin have been developed and propagated for population-based CRC screening [2–8]. They use specific antibodies against human blood components and are not affected by diet. Within the group of FITs, there are differences in the measurement devices, such as the use of different antibodies or different detection limits, which may influence the diagnostic performance of these tests. We have previously compared different qualitative and quantitative FITs regarding their potential to detect colorectal adenomas in the screening setting [7–10]. The small numbers of CRC detected in the screening setting did not allow meaningful analyses of the sensitivity for detecting CRC. Several recent studies found a quantitative FIT to have a higher sensitivity than gFOBTs regarding detection of adenoma and CRC at an acceptable level of specificity in either the screening setting or the hospital-based setting [5,6,11,12]. However, direct comparison of the diagnostic performance of different FITs for detection of CRC, overall and by stage, are sparse.

We meanwhile substantially enlarged our study population of screening participants and furthermore recruited additional cases of CRC from a pre-treatment clinical setting to enable comparative analyses on performance characteristics of different

Correspondence: H. Brenner, German Cancer Research Centre (DKFZ), Division of Clinical Epidemiology and Aging Research (C070), Im Neuenheimer Feld 581, D-69120 Heidelberg, Germany. Tel: +49 6221 421300. Fax: +49 6221 421302. E-mail h.brenner@dkfz.de

(Received 24 January 2013; accepted 19 March 2013)

ISSN 0284-186X print/ISSN 1651-226X online © 2013 Informa Healthcare
DOI: 10.3109/0284186X.2013.789141

qualitative and quantitative FITs with regard to detection of CRC, with a particular focus on detection of early stage CRC.

Material and methods

Study design and population

The study population consisted of two subgroups. Subgroup A (screening setting) consisted of participants undergoing screening colonoscopy, which is offered as a primary screening examination to the whole resident population aged 55 years or older in Germany since 2002. Subgroup B (clinical setting) consisted of patients who were recruited after being diagnosed with CRC, but before hospitalisation for surgery.

Participants of subgroup A were recruited between 2005 and December 2009 in the context of the BliTz study (Begleitende Evaluierung innovativer Testverfahren zur Darmkrebsfrüherkennung), an ongoing prospective study conducted in cooperation with 20 gastroenterology practices in southwestern Germany. Detailed information on the BliTz study has been provided elsewhere [7,8,10,13,14]. Briefly, participants were recruited at a preparatory visit of a screening colonoscopy. Several exclusion criteria were applied to make sure that the study population represented the average risk population of CRC screening (Figure 1a). Furthermore, to minimise potential misclassification, only participants with adequate bowel preparation and complete colonoscopy were included. For this analysis whose focus was on sensitivity for detection of colorectal carcinomas as well as on specificity among subjects free of neoplasms, we further excluded participants with colorectal adenomas. According to recent methodological work, we abstained from matching of CRC patients and neoplasm free subjects by age and sex [15].

Participants of subgroup B were recruited in a satellite sub-study (DACHS+) to the DACHS study (Darmkrebs: Chancen der Verhütung durch Screening), an ongoing case-control study focusing on the role of colonoscopy in colorectal cancer prevention [16–18]. Colorectal cancer patients referred by general practitioners or gastroenterologists for surgery to one of four participating hospitals were informed about the study. Stool samples were collected after diagnosis and before surgery. Patients who were recruited and provided stool samples between 2006 and December 2009 were included in this analysis. Patients diagnosed with CRC due to visible blood in stool were excluded in order to avoid over-optimistic estimates of sensitivity of tests designed for use in the screening setting.

The studies were approved by the Ethics committee of the University of Heidelberg and of the physicians' chambers of Baden-Württemberg, Rheinland Pfalz and Hessen.

Sample collection

After giving written informed consent, participants were asked to collect a stool sample from one bowel movement with a small container (60 ml) and keep it in a provided plastic bag in the freezer or, if not possible, in the refrigerator at home until colonoscopy (for participants from the screening setting) or hospital admission (for participants from the clinical setting). For the participants from the clinical setting, 75% of the samples were collected more than one week after the last colonoscopy. There was no specific recommendation for diet or medication restrictions. The stool-filled containers were received at the gastroenterology practice on the day of colonoscopy (screening setting) or on the day of hospitalisation for surgery (clinical setting). The samples were stored at −20°C, then shipped on dry ice to a central laboratory an average of seven days after collection and stored at −20°C until analysis. Other procedures were the same for both settings, and identical standard operational procedures were applied throughout.

After colonoscopy, colonoscopy and histology reports were collected from all participants from the screening setting. For participants from the clinical setting, we collected the medical reports from the hospital after surgery. Relevant information was extracted by two independent, trained research assistants who were not aware of the stool test results.

Laboratory analyses

The stool-filled containers were thawed at a median interval of four days on arrival at the central laboratory. Six qualitative FITs, Bionexia FOBplus (DIMA, Göttingen, Germany); PreventID CC (Preventis, Bensheim, Germany); immoCARE-C (CAREdiagnostica, Voerde, Germany); FOB advanced (ulti med, Ahrensburg, Germany); QuickVue iFOB (Quidel, San Diego, CA, USA); and Bionexia Hb/Hp Complex (DIMA, Göttingen, Germany), two ELISA-based quantitative FITs (RIDASCREEN® Haemoglobin and RIDASCREEN® Haemoglobin-Haptoglobin, R-Biopharm AG, Bensheim, Germany) and one agglutination-based quantitative FIT (OC- SENSOR, Eiken chemical Co., Tokyo, Japan) were performed on each stool sample. For each test, a defined amount of stool was collected from each stool sample and dissolved in a defined volume of buffer using the manufactures' sample

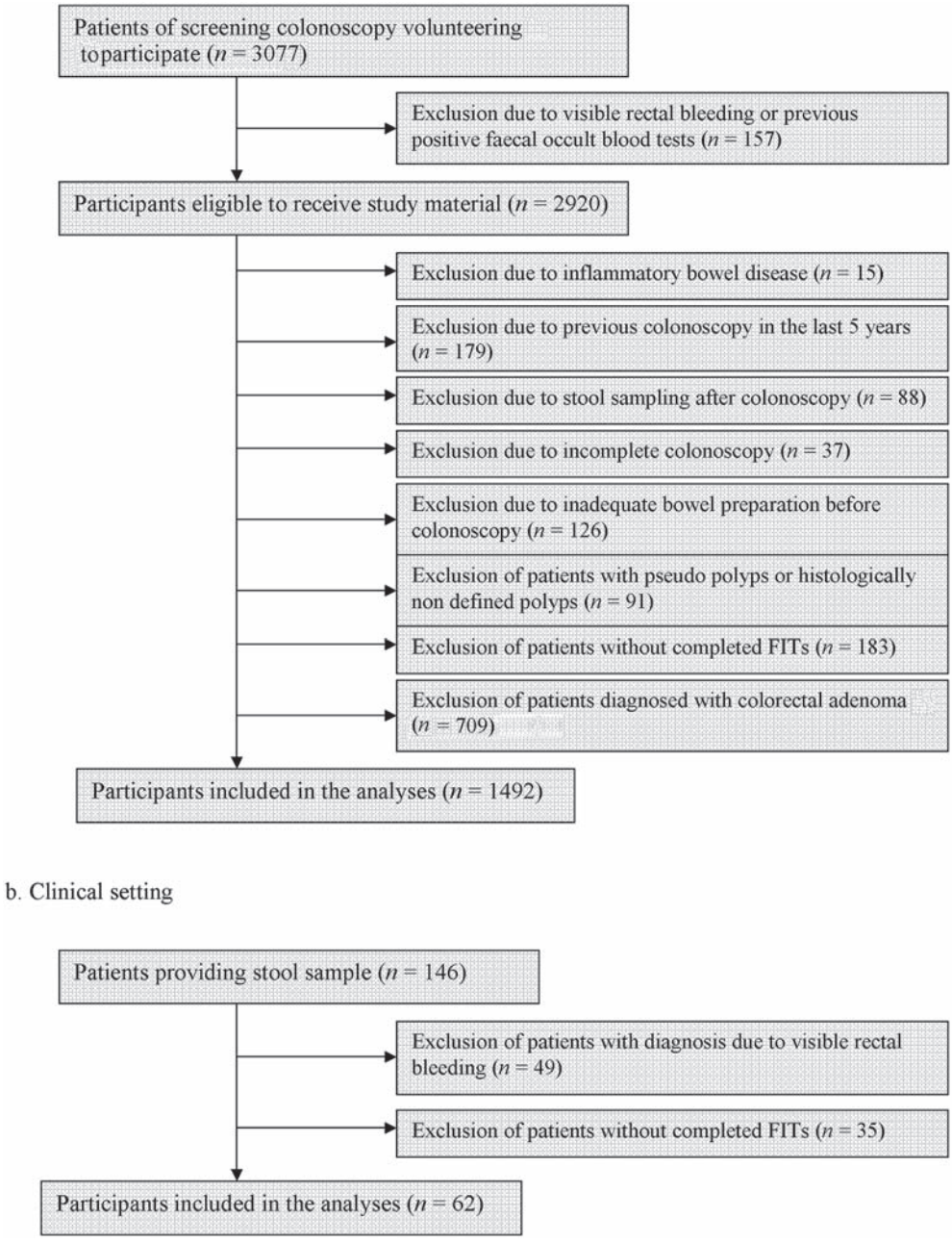

Figure 1. Standards for Reporting of Diagnostic Accuracy (STARD) flow diagram for selection of participants. a. Screening setting. b. Clinical setting.

collection devices. All qualitative tests are based on immunochromatographic technology and test results are rated as positive or negative for each stool sample respectively. All measurements were conducted according to the manufacturers' instructions by trained investigators who were blinded to colonoscopy results. In a reliability study including 200 randomly selected samples, inter-observer reliability of the two independent investigators was very high with kappa coefficients ranging from 0.72 to 0.94 for the six qualitative FITs.

Statistical analyses

Test performance characteristics, including sensitivity and specificity, were calculated for qualitative tests. Note that throughout this manuscript the term sensitivity refers to the proportion of CRC patients identified by FITs (not to the analytical sensitivity of the tests). Specificity was calculated among participants recruited in the screening setting who were found to be free of colorectal neoplasm. For the quantitative tests, we calculated sensitivity at cut-off

points yielding 90% and 95% specificity, which are commonly considered as specificities required for population-based screening. We calculated 95% confidence intervals (CIs) for sensitivity and specificity based on the exact binomial distribution. Furthermore, receiver operating characteristic (ROC) curves analyses were conducted for the three quantitative FITs. The area under the curves (AUCs) and the corresponding 95% CIs were calculated and compared using the method described by DeLong et al. [19]. We performed the above described analyses with and without stratification of CRC cases according to stage at diagnosis. Three categories for stage were assessed: stage I, stage II and advanced stages (stage III and stage IV were combined due to low number of stage IV cases) according to the UICC classification. Furthermore, analyses of sensitivity were stratified according to recruitment of cases in the clinical setting and the screening setting.

To assess the potential impact of neoadjuvant therapy among CRC patients, we further conducted a sensitivity analysis by excluding subjects whose stool samples were collected after neoadjuvant chemotherapy and compared the test characteristics before and after their exclusion. We used MedCalc for Windows, version 9.6.4.0 (MedCal Software, Mariakerke, Belgium) for the ROC analyses and SAS, version 9.2 (SAS Institute, Cary, NC, USA) for the other statistical analyses.

Results

Overall, 1492 and 62 participants were included in the screening and clinical setting, respectively (Figure 1). The distribution of the study participants according to the settings, age and gender and information on cancer stage and location for CRC patients are shown in Table I. In total, the study samples included 74 CRC patients and 1480 controls free of neoplasm.

Table II presents estimates of sensitivity and specificity with their CIs for the different qualitative FITs. Overall sensitivities of the FITs ranged from 66.2% (ImmoCARE-C) to 91.9% (Bionexia Hb/Hp Complex and QuickVue iFOB). The tests tended to show a slightly lower sensitivity for stage I cancers compared with the other stages, while this was not the case for stage II cancers versus advanced stages. For the stage I cancer patients, sensitivity ranged from 55.0% (ImmoCARE-C) to 95.0% (Bionexia Hb/Hp Complex). No consistent differences of sensitivity estimates were seen for CRC patients recruited in the screening setting and the clinical setting, but the wide confidence intervals resulting

Table I. Characteristics of the study population.

Group	Participants of screening colonoscopy (%)		CRC patients recruited in clinical setting (%) n = 62	Total CRC cases (%) n = 74
	No polyps or hyperplasic polyps only n = 1480	CRC n = 12		
Sex				
male	650 (43.9)	8 (66.7)	36 (58.1)	44 (59.5)
female	830 (56.1)	4 (33.3)	26 (41.9)	30 (40.5)
Age				
< 55	82 (5.5)	0	5 (8.1)	5 (6.8)
55–59	532 (36.0)	2 (16.7)	5 (8.1)	7 (9.5)
60–64	337 (22.8)	4 (33.3)	11 (17.7)	15 (20.3)
65–69	326 (22.0)	2 (16.7)	10 (16.1)	12 (16.2)
70–74	153 (10.3)	3 (25.0)	13 (21.0)	16 (21.6)
75+	50 (3.4)	1 (8.3)	18 (29.0)	19 (25.7)
Stage				
I		5 (41.7)	15 (24.2)	20 (27.0)
II		1 (8.3)	23 (37.1)	24 (32.4)
III		5 (41.7)	16 (25.8)	21 (28.4)
IV		0 (0)	7 (11.3)	7 (9.5)
Unknown		1 (8.3)	1 (1.6)	2 (2.7)
Location				
Colon		5 (41.7)	32 (51.6)	37 (50.0)
Rectum		6 (50.0)	29 (46.8)	35 (47.3)
Unknown		1 (8.3)	1 (1.6)	2 (2.7)
Proximal		0 (0)	20 (32.3)	20 (27.0)
Distal		11 (91.7)	41 (66.1)	52 (70.3)
Unknown		1 (8.3)	1 (1.6)	2 (2.7)

CRC, colorectal cancer.

Table II. Sensitivities and specificities of different qualitative faecal immunochemical tests for haemoglobin.

	Bionexia FOBplus	Bionexia Hb/Hp Complex	PreventID CC	ImmoCARE-C	FOB advanced	QuickVue iFOB
Overall sensitivity						
Patients	62/74	68/74	63/74	49/74	60/74	68/74
% (95% CI)	83.8 (73.8–90.5)	91.9 (83.4–96.2)	85.1 (75.3–91.5)	66.2 (54.9–76.0)	81.1 (70.7–88.4)	91.9 (83.4–96.2)
Sensitivity according to the stage at diagnosis						
Stage I						
Patients	16/20	19/20	17/20	11/20	15/20	18/20
% (95% CI)	80.0 (58.4–91.9)	95.0 (76.4–99.1)	85.0 (64.0–94.8)	55.0 (34.2–74.2)	75.0 (53.1–88.8)	90.0 (69.9–97.2)
Stage II						
Patients	21/24	24/24	22/24	18/24	22/24	23/24
% (95% CI)	87.5 (69.0–95.7)	100 (86.2–100)	91.7 (74.2–97.7)	75.0 (55.1–88.0)	91.7 (74.2–97.7)	95.8 (79.8–99.3)
Advanced stage (III + IV)						
Patients	23/28	23/28	22/28	19/28	22/28	25/28
% (95% CI)	82.1 (64.4–92.1)	82.1 (64.4–92.1)	78.6 (60.5–89.8)	67.9 (49.3–82.1)	78.6 (60.5–89.8)	89.3 (72.8–96.3)
Sensitivity according to the recruited settings						
Screening setting						
Patients	11/12	11/12	11/12	9/12	8/12	11/12
% (95% CI)	91.7 (64.6–98.5)	91.7 (64.6–98.5)	91.7 (64.6–98.5)	75.0 (46.8–91.1)	66.7 (39.1–86.2)	91.7 (64.6–98.5)
Clinical setting						
Patients	51/62	57/62	52/62	40/62	52/62	57/62
% (95% CI)	82.3 (71.0–89.8)	91.9 (82.5–96.5)	83.9 (72.8–91.0)	64.5 (52.1–75.3)	83.9 (72.8–91.0)	91.9 (82.5–96.5)
Sensitivity according to the location of cancers						
Colon						
Patients	31/37	34/37	30/37	28/37	31/37	34/37
% (95% CI)	83.8 (68.9–92.4)	91.9 (78.7–97.2)	81.1 (65.8–90.5)	75.7 (59.9–86.6)	83.8 (68.9–92.4)	91.9 (78.7–97.2)
Rectum						
Patients	29/35	32/35	31/35	19/35	27/35	32/35
% (95% CI)	82.9 (67.3–91.9)	91.4 (77.6–97.0)	88.6 (74.1–95.5)	54.3 (38.2–69.5)	77.1 (61.0–87.9)	91.4 (77.6–97.0)
Specificity among participants free of neoplasms						
Patients	1264/1480	912/1480	1259/1480	1423/1480	1331/1480	1108/1480
% (95% CI)	85.4 (83.5–87.1)	61.6 (59.1–64.1)	85.1 (83.2–86.8)	96.2 (95.0–97.0)	89.9 (88.3–91.4)	74.9 (72.6–77.0)

from the low case number for the screening setting have to be kept in mind. Likewise, no consistent differences of sensitivity estimates were seen between colon and rectum cancer. Furthermore, within the CRC patients recruited in the clinical setting, no consistent differences of sensitivity estimates were seen between CRC patients diagnosed through a screening colonoscopy (n = 17) and others. No relevant change in sensitivities was observed when we excluded CRC cases that provided stool samples after neoadjuvant therapy (n = 14).

Specificities for qualitative FITs ranged from 61.6% (Bionexia Hb/Hp Complex) to 96.2% (ImmoCARE-C). Comparing different tests, higher levels of sensitivities went along with lower levels of specificities, and vice versa.

Figure 2 presents the ROC curves of the three quantitative FITs regarding the detection of CRC. Pairs of sensitivity and specificity of the six qualitative FITs (a-f) are shown in the same figure for comparison. The corresponding AUCs (95% CI) were 0.92 (0.91–0.94), 0.91 (0.89–0.92), and 0.90 (0.88–0.91) for the ELISA-based haemoglobin, the ELISA-based haemoglobin-haptoglobin test, and the agglutination-based FIT, respectively. The small differences between these tests were not statistically significant (p-values > 0.45). Pairs of sensitivity and specificity of the six qualitative FITs were very close to the ROC curves of the quantitative tests, i.e. similar sensitivities of the qualitative FITs were observed compared to quantitative tests at corresponding levels of specificity.

Sensitivities of quantitative FITs at cut-off points yielding 90% and 95% specificity are shown in Table III. At a specificity of 90%, sensitivities (95% CI) were 86.5% (76.9–92.5), 79.7% (69.2–87.3) and 83.8% (73.8–90.5) for the ELISA-based haemoglobin, ELISA-based haemoglobin-haptoglobin and agglutination-based test, respectively. At a specificity of 95%, sensitivities (95% CI) were 82.4% (72.2–89.4), 73.0% (61.9–81.8) and 81.1% (70.7–88.4) for the ELISA-based haemoglobin, the ELISA-based haemoglobin-haptoglobin, and agglutination-based test, respectively.

Stage-specific ROC curves of the three quantitative tests and pairs of sensitivity and specificity of the six qualitative FITs are shown in Figure 3. AUCs for the ELISA-based haemoglobin, haemoglobin-haptoglobin, and the agglutination-based test ranged from 0.85 to 0.92, 0.94 to 0.96, and 0.86 to 0.93 for

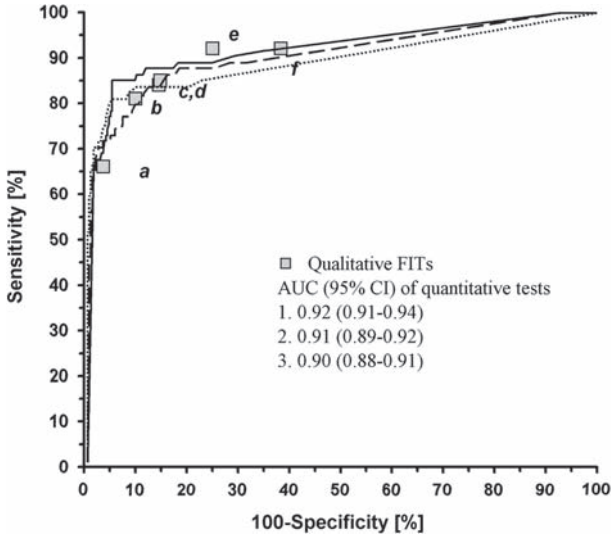

Figure 2. Direct comparison of different immunochemical faecal occult blood tests with regard to sensitivities for detection of colorectal cancer and percentages of false-positive test results (100 %-specificity). Quantitative tests: ——— 1. ELISA-based measurement of faecal haemoglobin; - - - - - 2. ELISA-based measurement of faecal haemoglobin-haptoglobin; ····· 3. Agglutination-based faecal occult blood test. Qualitative tests: (a) ImmoCARE-C (b) FOB advanced (c) Bionexia FOBplus (d) PreventID CC (e) QuickVue iFOB (f) Bionexia Hb/Hp Complex. AUC, area under the curve; CI, confidence interval.

stage I, stage II and advanced stages (stage III or IV) cancers, respectively. As shown in Table III, at cut-off points yielding 90% specificity, sensitivities were 70.0%, 65.0% and 85.0% for stage I, 95.8%, 87.5% and 91.7% for stage II, and 89.3%, 82.1% and 78.6% for advanced stage cancers for the ELISA-based haemoglobin, haemoglobin-haptoglobin, and the agglutination-based test, respectively. At cut-off points yielding 95% specificity, corresponding sensitivities were 65.0%, 60.0% and 80.0% for stage I, 87.5%, 83.3% and 87.5% for stage II, and 89.3%, 75.0% and 78.6% for advanced stage cancers. At the same levels of specificity, qualitative FITs showed similar levels of sensitivity.

Discussion

In this article, we investigated and compared test performance of six qualitative and three quantitative FITs with respect to their ability to detect CRC cases. To our knowledge, this is the first study to directly compare different FITs for the detection of CRC and especially according to the stages at diagnosis. With AUCs ranging from 0.90 to 0.92 for all cancers and from 0.85 to 0.92 for stage I cancers, quantitative tests showed good discrimination of CRC cases and subjects free of colorectal neoplasm. The six qualitative FITs showed levels of sensitivity which were comparable to the quantitative tests at corresponding levels of specificity. However, some of the qualitative tests had low levels of specificity that would limit their use in the screening setting.

The strong variation in sensitivity and specificity between qualitative tests parallels previous findings for colorectal adenomas obtained in a smaller sample of participants recruited in the screening setting up to the end of 2007 [7]. For colorectal adenomas, tests with higher sensitivity likewise had lower specificity and vice versa, and pairs of sensitivity and specificity were essentially located on the same ROC curves obtained from quantitative tests [8]. These patterns, along with results of direct cross-validation of qualitative and quantitative tests [9], strongly suggest that differences in performance between the qualitative tests essentially reflect different thresholds of positivity. Compared to the previously reported sensitivities for detecting advanced and non-advanced adenomas [7,8], the sensitivities for CRC assessed in this study are much higher.

Our findings of overall test performance of FITs for detecting CRC are in line with previous studies [20–22]. The study conducted by Morikawa et al. [21], which included more than 20 000 average risk subjects reported a sensitivity of 65.8% and a specificity of 94.6% for CRC detection (79 cases) for an FIT based on the magnetic agglutination technique (Magstream 1000/Hem SP, Fujirebio Diagnostics, Tokyo, Japan). Park et al. [12] assessed a quantitative agglutination-based FIT, the same test that we investigated in our study, in an average risk population. Although only 13 cases of CRC were included, estimates of sensitivity at comparable levels of specificity were similar to our study. The same test has also been investigated by Levi et al. [22] based on 1000 consecutive ambulatory patients including 17 CRC patients. With 91.5% specificity, 82.4% sensitivity was reported for detection of CRC.

There is only limited evidence regarding the sensitivity of FITs stratified by tumour stage, despite the fact that this is a crucial aspect for a screening test that should detect the disease at an early stage. Oort et al. [11] found a sensitivity of 75% for early stages (I and II, n = 21) and 96.8% for advanced stages (n = 30) cancer cases with a specificity of 91.0% among 1821 hospital-based ambulatory patients for a quantitative FIT. Park et al. [12] also stratified the results by early versus advanced stages and observed a sensitivity of 80% for early stages and 100% for advanced stages at a specificity of 90.1%. However, these estimates were based on rather small case numbers (nine and three cases in the early and advanced stage cancers, respectively). A recent study from the Netherlands [23] tested a quantitative FIT based on 79 CRC patients detected among both screening and

Table III. Sensitivity of quantitative faecal immunological tests at cut-offs yielding 90% and 95% specificity.

	ELISA-based haemoglobin	ELISA-based haemoglobin-haptoglobin	Agglutination-based test
Overall sensitivity (%)			
at 90% specificity (95% CI)	86.5 (76.9–92.5)	79.7 (69.2–87.3)	83.8 (73.8–90.5)
at 95% specificity (95% CI)	82.4 (72.2–89.4)	73.0 (61.9–81.8)	81.1 (70.7–88.4)
Sensitivity according to the stage at diagnosis (%)			
Stage I			
at 90% specificity (95% CI)	70.0 (48.1–85.5)	65.0 (43.3–81.9)	85.0 (64.0–94.8)
at 95% specificity (95% CI)	65.0 (43.3–81.9)	60.0 (38.7–78.1)	80.0 (58.4–91.9)
Stage II			
at 90% specificity (95% CI)	95.8 (79.8– 99.3)	87.5 (56.6–87.3)	91.7 (74.2–97.7)
at 95% specificity (95% CI)	87.5 (69.0–95.7)	83.3 (64.2–93.3)	87.5 (69.0–95.7)
Advanced stage (III + IV)			
at 90% specificity (95% CI)	89.3 (72.8–96.3)	82.1 (64.4–92.1)	78.6 (60.5–89.8)
at 95% specificity (95% CI)	89.3 (72.8–96.3)	75.0 (56.6–87.3)	78.6 (60.5–89.8)
Sensitivity according to the recruited settings (%)			
Screening setting			
at 90% specificity (95% CI)	75.0 (46.8–91.1)	75.0 (46.8–91.1)	83.3 (55.2–95.3)
at 95% specificity (95% CI)	75.0 (46.8–91.1)	66.7 (39.1–86.2)	83.3 (55.2–95.3)
Clinical setting			
at 90% specificity (95% CI)	88.7 (78.5–94.4)	80.7 (69.2–88.6)	83.9 (72.8–91.0)
at 95% specificity (95% CI)	83.9 (72.8–91.0)	74.2 (62.1–83.5)	80.7 (69.2–88.6)
Sensitivity according to the location of cancers (%)			
Colon			
at 90% specificity (95% CI)	81.1 (65.8–90.5)	75.7 (59.9–86.6)	89.2 (75.3–95.7)
at 95% specificity (95% CI)	78.4 (62.8–88.6)	67.6 (51.5–80.4)	89.2 (75.3–95.7)
Rectum			
at 90% specificity (95% CI)	91.4 (77.6–97.0)	82.9 (67.3–91.9)	77.1 (61.0–87.9)
at 95% specificity (95% CI)	85.7 (70.6–93.7)	77.1 (61.0–87.9)	71.4 (54.9–83.7)

symptomatic participants. The authors found 81.6% sensitivity for early stage (I and II, n = 38) cancers at a cut-off point yielding 90% specificity, with an AUC of 0.89. In our study, we observed sensitivities of 65–80% at a specificity of about 90% for stage I cancers.

In addition, we assessed and compared the sensitivities of FITs in different recruitment settings. The main difference between these two settings is the time point of stool sampling before or after diagnosis. In the BliTz study, we recruited participants and collected stool samples prior to diagnosis in a true screening setting. Participants from the clinical setting (DACHS+ study) increased the number of CRC cases in the study, which allowed us to explore stage-specific sensitivities. To avoid over-optimistic estimates of sensitivities of the FITs in the clinical setting, we excluded participants who were diagnosed due to visible rectum bleeding. The similarity of sensitivities observed in the clinical setting and the screening setting suggests that overestimation in the clinical setting was not a relevant issue in our study (sensitivities were even tentatively higher in the screening setting for most tests). We therefore combined participants from both settings when conducting ROC analyses.

In the interpretation of our results, several limitations require consideration. First, we used colonoscopy as a gold standard in the screening setting, although it may not be perfect [24]. However, miss rates of colorectal cancer are mostly very small [25]. Furthermore, only trained and experienced gastroenterologists with high levels of qualification are certified to carry out screening colonoscopies in the German screening colonoscopy program. Second, stool sampling in our study somewhat differed from real-life conditions because the stool was not directly dissolved in a buffer-filled vial but was instead collected in a small container and frozen before testing. This procedure was chosen for practical reasons to enable simultaneous application of multiple tests. The instruction manuals typically suggest directly using fresh stool samples. Potentially lower stabilities of haemoglobin in frozen stool samples compared to buffered stool samples might have led to lower sensitivities and higher specificities at given cut-off points. However, since all stool samples were treated the same way, ROC curves should have remained unaffected. Furthermore, the similarities of our findings with those of studies in which stool samples were collected according to the manufacturers' instruction suggest that sample handling was not a major issue. Third, we only collected stool samples from one bowel movement and could thus not investigate to what extent test performance characteristics would

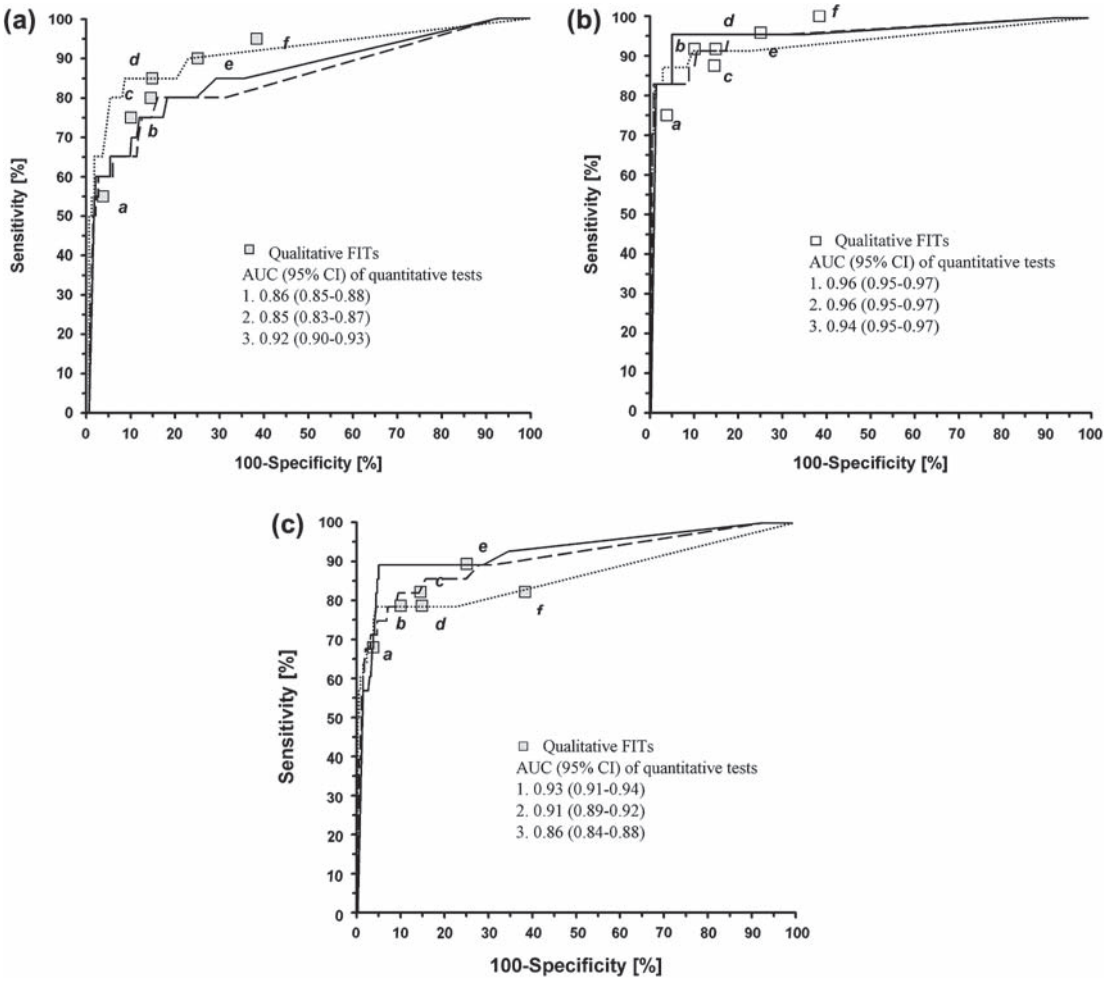

Figure 3. Receiver operating characteristics curves for detecting colorectal cancer stratified by stages at diagnosis for immunochemical faecal occult blood tests. a. Stage I CRC patients. b. Stage II CRC patients. c. Advanced (III and IV) stage CRC patients. Quantitative tests: ——— 1.ELISA-based measurement of faecal haemoglobin; - - - - - 2. ELISA-based measurement of faecal haemoglobin-haptoglobin; ····· 3. Agglutination-based faecal occult blood test. Qualitative tests: (a) ImmoCARE-C (b) FOB advanced (c) Bionexia FOBplus (d) PreventID CC (e) QuickVue iFOB (f) Bionexia Hb/Hp Complex; AUC, area under the curve; CI, confidence interval.

change if stool samples from two or three consecutive bowel moments were considered.

In conclusion, we found the diagnostic performance of different qualitative and quantitative FITs regarding the detection of CRC to be promising and similar, even though the pre-defined cut-off levels of some of the qualitative FITs need to be adjusted to limit false-positive rates in the screening setting. At levels of specificity that are typically required in the screening setting, the quantitative tests detected the vast majority of colorectal cancers, even at early stages.

Acknowledgements

The authors gratefully appreciate the excellent cooperation of physicians conducting screening colonoscopies and clinicians in the study hospitals in patient recruitment, excellent contributions of Isabel Lerch, Dr. Ulrike Haug and Dr. Sabrina Hundt (Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Heidelberg) in data collection, monitoring and documentation and excellent cooperation with Labor Limbach (Heidelberg) in test execution.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

This study was partly funded by a grant from the German Research Foundation, Graduiertenkolleg 793: Epidemiology of communicable and chronic noncommunicable diseases and their interrelationships. The recruitment of CRC cases was in part supported by a starting grant within the intramural funding program at the German Cancer Research Center. All test kits were provided free of charge by

the manufacturers, and additional support was obtained by Eiken Chemical Co. Ltd, Tokyo, Japan, for evaluation of the OC-SENSOR test. The sponsors had no role in the study design, in the collection, analysis, and interpretation of data and preparation, review, or approval of the manuscript.

References

[1] Hewitson P, Glasziou P, Watson E, Towler B, Irwig L. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): An update. Am J Gastroenterol 2008;103:1541–9.

[2] Allison JE, Sakoda LC, Levin TR, Tucker JP, Tekawa IS, Cuff T, et al. Screening for colorectal neoplasms with new fecal occult blood tests: Update on performance characteristics. J Natl Cancer Inst 2007;99:1462–70.

[3] Guittet L, Bouvier V, Mariotte N, Vallee JP, Arsene D, Boutreux S, et al. Comparison of a guaiac based and an immunochemical faecal occult blood test in screening for colorectal cancer in a general average risk population. Gut 2007;56:210–4.

[4] Dancourt V, Lejeune C, Lepage C, Gailliard MC, Meny B, Faivre J. Immunochemical faecal occult blood tests are superior to guaiac-based tests for the detection of colorectal neoplasms. Eur J Cancer 2008;44:2254–8.

[5] van Rossum LG, van Rijn AF, Laheij RJ, van Oijen MG, Fockens P, van Krieken HH, et al. Random comparison of guaiac and immunochemical fecal occult blood tests for colorectal cancer in a screening population. Gastroenterology 2008;135:82–90.

[6] Hol L, Wilschut JA, van Ballegooijen M, van Vuuren AJ, van der Valk H, Reijerink JC, et al. Screening for colorectal cancer: Random comparison of guaiac and immunochemical faecal occult blood testing at different cut-off levels. Br J Cancer 2009;100:1103–10.

[7] Hundt S, Haug U, Brenner H. Comparative evaluation of immunochemical fecal occult blood tests for colorectal adenoma detection. Ann Intern Med 2009;150:162–9.

[8] Haug U, Hundt S, Brenner H. Quantitative immunochemical fecal occult blood testing for colorectal adenoma detection: Evaluation in the target population of screening and comparison with qualitative tests. Am J Gastroenterol 2010; 105:682–90.

[9] 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.

[10] Brenner H, Haug U, Hundt S. Sex differences in performance of fecal occult blood testing. Am J Gastroenterol 2010;105:2457–64.

[11] Oort FA, Droste JS, Van Der Hulst RW, Van Heukelem HA, Loffeld RJ, Wesdorp IC, et al. Colonoscopy-controlled intra-individual comparisons to screen relevant neoplasia: Faecal immunochemical test vs. guaiac-based faecal occult blood test. Aliment Pharmacol Ther 2010;31:432–9.

[12] Park DI, Ryu S, Kim YH, Lee SH, Lee CK, Eun CS, et al. Comparison of guaiac-based and quantitative immunochemical fecal occult blood testing in a population at average risk undergoing colorectal cancer screening. Am J Gastroenterol 2010;105:2017–25.

[13] Brenner H, Tao S, Haug U. Low-dose aspirin use and performance of immunochemical fecal occult blood tests. JAMA 2010;304:2513–20.

[14] Tao S, Haug U, Kuhn K, Brenner H. Comparison and combination of blood-based inflammatory markers with faecal occult blood tests for non-invasive colorectal cancer screening. Br J Cancer 2012;106:1424–30.

[15] Brenner H, Altenhofen L, Tao S. Matching of controls may lead to biased estimates of specificity in the evaluation of cancer screening tests. J Clin Epidemiol 2013;66: 202–8.

[16] Brenner H, Chang-Claude J, Seiler CM, Rickert A, Hoffmeister M. Protection from colorectal cancer after colonoscopy: A population-based, case-control study. Ann Intern Med 2011;154:22–30.

[17] Brenner H, Chang-Claude J, Seiler CM, Sturmer T, Hoffmeister M. Case-control study supports extension of surveillance interval after colonoscopic polypectomy to at least 5 yr. Am J Gastroenterol 2007;102:1739–44.

[18] Brenner H, Chang-Claude J, Seiler CM, Sturmer T, Hoffmeister M. Does a negative screening colonoscopy ever need to be repeated? Gut 2006;55:1145–50.

[19] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. Biometrics 1988;44:837–45.

[20] Cheng TI, Wong JM, Hong CF, Cheng SH, Cheng TJ, Shieh MJ, et al. Colorectal cancer screening in asymptomatic adults: Comparison of colonoscopy, sigmoidoscopy and fecal occult blood tests. J Formos Med Assoc 2002; 101:685–90.

[21] Morikawa T, Kato J, Yamaji Y, Wada R, Mitsushima T, Shiratori Y. A comparison of the immunochemical fecal occult blood test and total colonoscopy in the asymptomatic population. Gastroenterology 2005;129:422–8.

[22] Levi Z, Rozen P, Hazazi R, Vilkin A, Waked A, Maoz E, et al. A quantitative immunochemical fecal occult blood test for colorectal neoplasia. Ann Intern Med 2007;146: 244–55.

[23] Terhaar sive Droste JS, Oort FA, van der Hulst RW, van Heukelem HA, Loffeld RJ, van Turenhout ST, et al. Higher fecal immunochemical test cutoff levels: Lower positivity rates but still acceptable detection rates for early-stage colorectal cancers. Cancer Epidemiol Biomarkers Prev 2011; 20:272–80.

[24] van Rijn JC, Reitsma JB, Stoker J, Bossuyt PM, van Deventer SJ, Dekker E. Polyp miss rate determined by tandem colonoscopy: A systematic review. Am J Gastroenterol 2006;101:343–50.

[25] Pohl H, Robertson DJ. Colorectal cancers detected after colonoscopy frequently result from missed lesions. Clin Gastroenterol Hepatol 2010;8:858–64.