Head to Head Comparisons of Different FITs

Thomas F. Imperiale, MD
Indiana University Medical Center

WEO CRCSC – DDW San Diego, CA
May 20, 2016

What I'll discuss

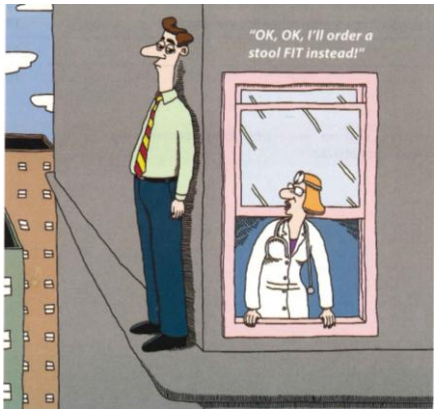
- Technical / clinical features of FIT
- How to compare FIT performance
  - Technical
  - Clinical
    - Individual studies
    - Reviews of multiple studies
- Conclusion

FIT Features

- Immunoassay for human hemoglobin (hapto, albumin)
- More specific (and more sensitive) than gFOBT
- Fewer specimens required – uptake higher than gFOBT
- No dietary restrictions or medication interference (Vit C)
- Can be qualitative or quantitative, manual or automated
- Advantages of quantitative FIT
  - adjust cutoff to "fit" goals and resource capacity
  - Integrate test results with other features (risk stratification)

What makes a FIT good? ? What makes a good FIT?

- Good performance
- Good for a health care system



Technical Metrics / Performance

- Analytical sensitivity – lower limit: mean [Hb] in 20 unused collection devices + 2 s.d.
- Carryover – thoroughness of probe cleansing
- Imprecision – agreement among measurements (s.d./CV)
- Precision profile – serial measurements across a range
  - mean +/- s.d. within the range reported
- Linearity
- Hook/prozone – FN results with very high [analyte]
- Stability -
  - Over a range from lower detection limit to strongly ⊕
  - Temperature, time from collection → testing

Carroll M, et al. Evaluation of quantitative faecal immunochemical tests for haemoglobin. 2013

Figure 3: Linearity: OC-SENSOR DIANA

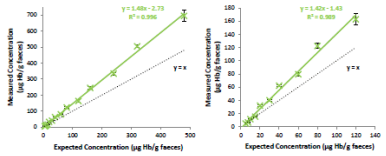
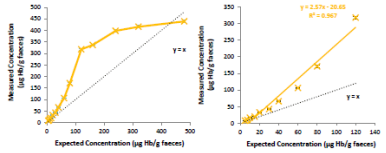


Figure 4: Linearity: FOB Gold/BioMajesty



Technical Evaluations – Guittet 2011

Methods

- Compare analytical performance of 3 FITs
  - Magstream, OC-Sensor, FOB Gold
- Stability, reproducibility at different temps & intervals
  - Stool samples of 10 health volunteers, FIT ⊖, spiked with human blood

Results

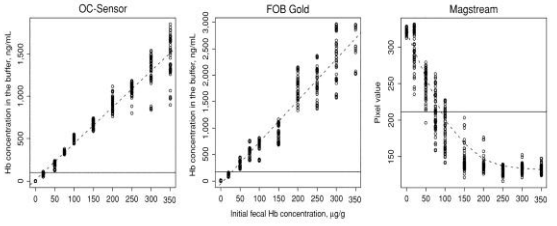
- Reproducibility: OC Sensor > Magstream > FOB Gold
- OC Sensor less affected by higher temperatures

Conclusion

- OC-Sensor > Magstream > FOB-Gold

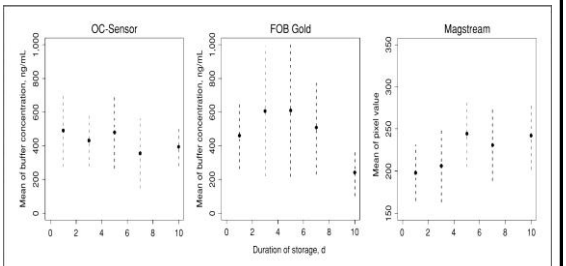
Guittet L et al. Cancer Epid Biomarker Prev 2011

Value returned by the automated analyzers according to test and fecal Hb concentration



Guittet et al. Cancer Epidemiol Biomarkers Prev 2011;20:1492-1501

Comparison of mean measurements and SD according to test: duration of storage.



Guittet et al. Cancer Epidemiol Biomarkers Prev 2011;20:1492-1501

Analytical Summary of 11 FOBTs

Device	Sensitivity	Imprecision	Analysis time 1 sample	Analysis time 32 samples	Temperature	Sunlight	High Hb	Loading
Units	mg Hb/g		Minutes		Reliability			
Hema-screen	0.8-1.0	S	1·1	11	U	U	U	A
MonoHaem	0·5	S	7	16-17	A	a	A	A
Hema-Chek	0·85	S	1·3	13	U	A	U	A
Magstream	0·02-0·1	S	14·5	40	U	U	U	U
Health Check N° 2	0·06	S	11·5	21	a	U	U	A
Q-OB	0·02-0·15	S	10·5	30	A	U	A	U
Hem-Check-1	0·01	S	12	50-55	A	U	A	A
BM-Test Colon Albumin	0·5-0·65	P	15·5	26	A	U	U	A
HemaWipe	0·5	S	2·1	15	U	U	U	A
Haemoccult	0·6-0·8	S	0·9	9·5	U	a	U	A
Haemoccult Sensa	0·15-0·6	S	1·3	10	U	a	U	A

S = Satisfactory, P = Poor, A = Affected, a = Slightly affected, U = Unaffected
Pearson, Bennitt, Halloran. Faecal occult blood tests. Evaluation Report MDA/2000/05

FITs & Pathology Proficiency Testing

- 14 FIT brands evaluated by ≥ 1 of 4 proficiency testing programs – 8 FITs had ≥ 25 results
- All but 2 (automated FITs) were CLIA-waived
- All testing involves samples spiked with human blood and control samples – results were pooled
- 5 FITs performed well and similarly
  - Sensitivity: 98.1% - 98.8%
  - Specificity: 98.1% - 99.6%
- Conclusion – many FITs performed acceptably, others "probably should not be used for…screening"

Daly JM et al. J Pri Care Comm Health 2013; 4:245-50

Proficiency Test Program Results

FIT	# Specimens	Sensitivity	Specificity
Hemoccult ICT*	1633	98.2	98.1
Hemosure	1711	94.4	91.2
Hema-Screen Specific*	699	98.1	99.4
iFOB Complete	357	87.2	80.8
Ultra FOB	836	91.6	94.1
OC-Light*	976	98.4	98.3
OC-Auto Micro 80*	586	98.8	99.6
Quick Vue iFOB*	1321	98.3	99.0

Daly JM et al. J Pri Care Comm Health 2013; 4:245-50

Clinical Metrics / Performance

- Uptake / completion
- (Adherence)
- Detection rate for CRC/AA
- **Sensitivity / Specificity**
- Efficiency (NN screen, scope)
- (Efficacy / effectiveness)
- (Cost-effectiveness)

FITs Available in the U.S.

Name	Manufacturer
Hemoccult ICT	Beckman-Coulter
HemoSure	Alere
InSure	Quest Diagnostics
QuickVue iFOB	Quidel
HemaScreen	Immunostics
OC-Light (manual)	Polymedco
FOBT-Chek (automated)	Polymedco

Ways of Comparing FITs

- Large-scale head to head comparison
- Model
- Published literature
  - 2 or more FITs
  - Systematic reviews

Ideal Study

- Study population – 50-70 ± 5 years, average-risk → screening colonoscopy
- Intervention – several FITs completed from same stool specimen
- Target lesions – CRC, AA
- Test metrics
  - Sensitivity, specificity
  - Completion rate
- Required sample size = 18-25K

Lots of Variation to Consider

How FITs vary
- Manual vs automated
- Qualitative vs quantitative
- # specimens
- Cutoff/threshold

How studies vary
- Colonoscopy for all tested
- Colonoscopy for ⊕ FIT / follow-up or sigmoidoscopy for ⊖ FIT
- Case-control

Metrics
- Sensitivity / specificity
- Participation (uptake)
- Positivity rate
- Detection rate
- # needed to screen/scope

Selected Screening Studies Comparing ≥ 2 FITs

1st Au, Yr	Subject N	FITs Studied	Target lesion	Reference Std	Conclusions
Hundt, 2009	1319	Bionexia FOB-plus immoCARE-C FOB advanced PreventID CC Quick Vue & iFOB Bionexia Hb/Hp Hemoccult	Advanced Adenoma	Colonoscopy for all	• FITs > gFOBT • FIT performance varies widely • ImmoCARE & FOB advanced were "best"
Faivre, 2012	85,149	FOB-Gold Magstream OC-Sensor (Hemoccult II)	CRC AA	Colonoscopy if any FIT ⊕	• FIT > gFOBT • 3 FITs equal
Raginel, 2013	19,797	Magstream OC-Sensor (Hemoccult II)	CRC AA	Colonoscopy if any FIT ⊕	OC-Sensor > Magstream for CRC
Brenner, 2013	2235	OC-Sensor RIDASCREEN-Hb RIDASCREEN-Hb/Hp	CRC AA Any neoplasia	Colonoscopy for all	• FITs > gFOBT • 3 FITs equal

Selected Screening Studies Comparing ≥ 2 FITs

1st Au, Yr	Subject N	FITs Studied	Target lesion	Reference Std	Conclusions
Tao, 2013	74 CRC (10 screen detected) 1480 controls	6 qualitative 3 quantitative	CRC	Colonoscopy	• At 90% specificity, qualitative FITs = • Most CRC detected • Cutoffs for some FITs need adjustment
Zubero, 2014	37,999	FOB Gold (18k) OC-Sensor (19k) Both @ 100 ng/ml	CRC AA NAA	Colonoscopy for FIT positive only	• OC-Sensor superior (error rate 0.2% vs 2.3%) • % stage I-II CRC 80% vs 57% (no difference)
Chiang, 2014	956,005	OC-Sensor (747k) HM-Jack (209k) Both at 20ug/g	CRC AA	⊕ FIT-colonoscopy ⊖ FIT-2 year f/u	OC-Sensor had higher PPV and lower interval cancer rate

gFOBT vs. 3 Different FITs

- 85,149 average-risk adults age 50-74 years
  - 3rd round of screening
- Hemoccult II & 1 of 3 FITs: (two specimens)
  - FOB-Gold (Beckman Coulter – USA)
  - Magstream (Fujirebio – Japan)
  - OC-Sensor (Eiken – Japan)
- Outcomes – detection rates, (ratios for sensitivity and false positivity)
  - FITs vs. gFOBT

Faivre J, et al. Eur J Cancer 2012

gFOBT vs 3 FITs

	FOB-Gold		Magstream		OC-Sensor	
	gFOBT	FIT	gFOBT	FIT	gFOBT	FIT
Positive Test	2.2%	5.2%	2.3%	4.6%	1.7%	3.7%
Colonoscopy (%)	92.4%	91.7%	92%	93%	94%	94%
Detection of CRC	1.5%	2.8%	1.7%	3.3%	1.1%	2.7%
Adv adenoma	**2.7%**	**9.8%**	**3.5%**	**10.9%**	**3.0%**	**12.0%**

Faivre J, et al. Eur J Cancer 2012

Brenner & Tao: 3 FIT Comparison

Test characteristic	Outcome (n)	RIDA-Hb	RIDA-Hb/Hp	OC-Sensor
Sensitivity (%)	CRC (15)	60	53	73
	AA (111)	23	20	26
Specificity (%)	CRC	95	95	96
	AA	97	97	97
Positive PV (%)	CRC	8.1	7.3	10.0
	AA	47	41	52
Positive LR	CRC	13	11.6	16
	AA	8	6.3	9.8
Negative LR	CRC	0.42	0.49	0.28
	AA	0.79	0.82	0.76

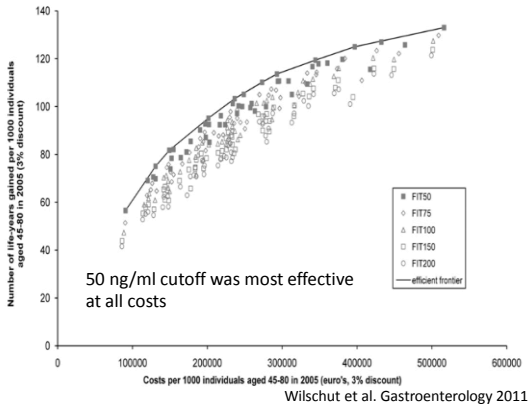
Brenner, Tao. Eur J Cancer 2013

N=2235

Modeling study - Optimizing screening with a quantitative FIT

- MIcrosimulation SCreening ANalysis (MISCAN) model
- Costs and effects estimated for
  - FIT cutoffs of 50, 75, 100, 150, 200 ng/mL
  - Intervals of 1, 1.5, 2, and 3 years
  - Ages 45 to 80
- Outcome: life-years gained

Wilschut et al. Gastroenterology 2011



Wilschut et al. Gastroenterology 2011

Reviews of Studies Comparing ≥ 2 FITs

1st Au, Yr	Study N	FITs Compared	Lesion	Reference Standard	Conclusions
Guittet, 2011	6	OC-Sensor (4) Magstream (2) Hemoccult II (3)	CRC AA	Colonoscopy for FIT +	- For similar PR, PPV with OC-S >Mag - Indirect comparisons lack reliability
Lee, 2013	19	OC-Hemodia OC Micro OC Light Monohaem Heme Select FlexSure OBT Magstream Ridascreen	CRC	Colonoscopy for all (12) Colonoscopy for FIT+, 2-yr f/up (7)	• High overall accuracy for CRC • Performance depends on cutoff
Launois 2014	21	Magstream OC-Sensor (Hemoccult)	CRC AA	Colonoscopy (8) Colonoscopy + f/up or sig (13)	OC-Sensor for CRC detection

Indirect comparison of two quantitative immunochemical faecal occult blood tests in a population with average colorectal cancer risk

L Guittet, L Bailly, V Bouvier and G Launoy

J Med Screen 2011;**18**:76–81
DOI: 10.1258/jms.2011.011012

See end of article for authors' affiliations

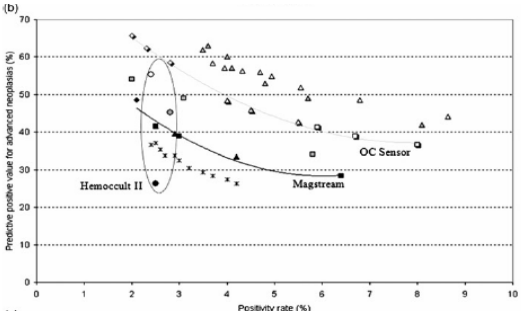
Correspondence to: Lydia Guittet, INSERM ER3 'Cancers & Populations', Faculty of Medicine, Avenue de la Cote de Nacre, F-14032 Caen Cedex, France; guittet@chu-caen.fr

Accepted for publication 23 May 2011

Magstream and OC Sensor quantitative immunochemical faecal occult blood tests (IFOBT) have shown better performances than guaiac (G) tests in colorectal cancer screening, however Magstream and OC Sensor have never been compared. We hypothesized that similar performances could be observed with Magstream and OC Sensors, provided a similar cutoff (expressed in concentration of haemoglobin in the stools) is used. We performed a literature-based indirect comparison between these tests, taking into account the cut-off, the number of samples, and the way they were combined (I_{2+}: at least one positive sample of 2; I_{2++}: both positive samples; I_1: only one sample). Six studies conducted in general average-risk populations were included in this review. For each [test]*[cut-off], positivity rate (PR) decreased and predictive positive value (PPV) increased from I_{2+} to I_1 and I_{2++}. For similar PR, PPV with OC Sensor was greater than with Magstream. This could be due to factors other than the test, because PPVs associated with GFOBT in studies evaluating OC Sensor were greater than PPVs associated with GFOBT in the study evaluating Magstream. Direct comparison between Magstream and OC Sensor is needed to confirm the suspected superiority of OC Sensor.

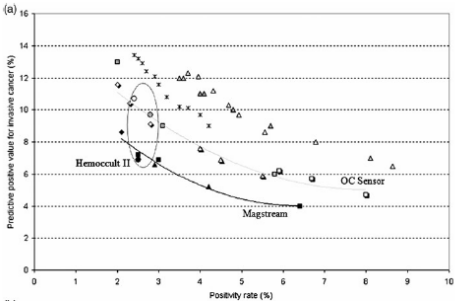
Guittet, et al. 2011 J Med Screen

Positivity rate vs. PPV for Advanced Neoplasia



Guittet, et al. J Med Screen 2001; 18:76-81

Positivity rate vs. PPV for Colorectal Cancer



Guittet, et al. 2011 J Med Screen

Annals of Internal Medicine | Review

Accuracy of Fecal Immunochemical Tests for Colorectal Cancer

Systematic Review and Meta-analysis

Jeffrey K. Lee, MD, MAS; Elizabeth G. Liles, MD, MCR; Stephen Bent, MD; Theodore R. Levin, MD; and Douglas A. Corley, MD, PhD

Background: Performance characteristics of fecal immunochemical tests (FITs) to screen for colorectal cancer (CRC) have been inconsistent.

Purpose: To synthesize data about the diagnostic accuracy of FITs for CRC and identify factors affecting its performance characteristics.

Data Sources: Online databases, including MEDLINE and EMBASE, and bibliographies of included studies from 1996 to 2013.

Study Selection: All studies evaluating the diagnostic accuracy of FITs for CRC in asymptomatic, average-risk adults.

Data Extraction: Two reviewers independently extracted data and critiqued study quality.

Data Synthesis: Nineteen eligible studies were included and meta-analyzed. The pooled sensitivity, specificity, positive likelihood ratio, and negative likelihood ratio of FITs for CRC were 0.79 (95% CI, 0.69 to 0.86), 0.94 (CI, 0.92 to 0.95), 13.10 (CI, 10.49 to 16.35), 0.23 (CI, 0.15 to 0.33), respectively, with an overall diagnostic accuracy of 96% (CI, 93% to 97%). There was substantial heterogeneity between studies in both the pooled sensitivity and specificity estimates. Stratifying by cutoff value for a positive test result or removal of discontinued FIT brands resulted in homogeneous sensitivity estimates. Sensitivity for CRC improved with lower assay cutoff values for a positive test result (for example, 0.89 [CI, 0.80 to 0.95] at a cutoff value less than 20 $\mu g/g$ vs. 0.70 [CI, 0.55 to 0.81] at cutoff values of 20 to 50 $\mu g/g$) but with a corresponding decrease in specificity. A single-sample FIT had similar sensitivity and specificity as several samples, independent of FIT brand.

Limitations: Only English-language articles were included. Lack of data prevented complete subgroup analyses by FIT brand.

Conclusion: Fecal immunochemical tests are moderately sensitive, are highly specific, and have high overall diagnostic accuracy for detecting CRC. Diagnostic performance of FITs depends on the cutoff value for a positive test result.

Primary Funding Source: National Institute of Diabetes and Digestive and Kidney Diseases and National Cancer Institute.

Ann Intern Med. 2014;160:171-181. www.annals.org
For author affiliations, see end of text.

Systematic review of FIT

Trial subgroup	Trial N	CRC Sensitivity (95% CI)	Specificity (95% CI)
Colonoscopy as ref std	12	0.71 (0.58-0.92)	0.94 (0.91-0.96)
< 100 ng/ml	11	0.86 (0.75-0.92)	0.91 (0.69-0.96)
100-250 ng/ml	6	0.63 (0.43-0.79)	0.96 (0.94-0.97)
> 250 ng/ml	4	0.67 (0.59-0.74)	0.96 (0.94-0.98)
OC-Light	4	0.93 (0.83-0.97)	0.91 (0.88-0.92)
OC-Micro / sensor	5	0.86 (0.68-0.95)	0.91 (0.87-0.94)

Lee JK, et al. Ann Intern Med 2014; 160:171-81

Systematic review and bivariate/HSROC random-effect meta-analysis of immunochemical and guaiac-based fecal occult blood tests for colorectal cancer screening

Robert Launois[a], Jean-Gabriel Le Moine[a], Bernard Uzzan[b], Lucia I. Fiestas Navarrete[a] and Robert Benamouzig[b]

Background Current literature evidences higher accuracy of immunological (iFOBT) vis-à-vis guaiac-based (gFOBT) fecal occult blood tests for colorectal cancer (CRC) screening. Few well-designed head-to-head comparisons exist.

Aim This meta-analysis assesses the performances of two iFOBTs compared with an established gFOBT using colonoscopy as the gold standard.

Methods We mobilized a bivariate and a hierarchical summary receiver operating characteristic (HSROC) model. Positive likelihood ratio (LR^+) and negative likelihood ratio (LR^-) and diagnostic odds ratios were back-calculated. We constructed bivariate credibility ellipses in the HSROC space and calculated areas under the curve to obtain a global measure of test performance. Estimates are presented at 95% credibility levels.

Results We included and analyzed 21 studies. OC-Sensor was the best test for CRC screening, with high sensitivity (0.87; 95% credibility interval: 0.73–0.95) and specificity (0.93; 95% credibility interval: 0.84–0.96), optimal LR^+ (12.01) and LR^- (0.14), and a high diagnostic odds ratio (88.05). Bivariate credibility ellipses showed OC-Sensor's dominance over Hemoccult (sensitivity: 0.47; 95% credibility interval: 0.37–0.58; specificity: 0.93; 95% credibility interval: 0.91–0.95).

Conclusion Our findings support the use of OC-Sensor for CRC detection. The diagnostic estimates obtained may be extended to derive model parameters for economic decision models and to offer insight for future clinical and public health decision making. Our findings could influence the future of FOBTs within the CRC screening arsenal. *Eur J Gastroenterol Hepatol* 26:978–989 © 2014 Wolters Kluwer Health | Lippincott Williams & Wilkins.

European Journal of Gastroenterology & Hepatology 2014, 26:978–989

Keywords: accuracy, advanced adenomas, bivariate, colorectal cancer, diagnostic test, HSROC, literature review, meta-analysis, QUADAS

[a]French Network for Evaluation in Health Economics, REES-France, Paris and [b]Service d'Hépato-Gastro-entérologie, Hôpital Ambroise APHP, Bobigny, France

Correspondence to Robert Launois, PhD, French Network for Evaluation in Health Economics, REES-France, 28 Rue d'Assas, 75006 Paris, France
Tel: +33 1 44 39 16 90; e-mail: launois.rees@wanadoo.fr

Received 16 April 2014 Accepted 16 June 2014

Launois et al 2014 Euro J Gastroenterol Hep

SR & Bivariate / HSROC Meta-Analysis

- Meta-analysis of Magstream, OC-Sensor, HO "using colonoscopy as the gold standard"
- 21 studies included
  - Average-risk population – mean age ≥ 40 years
  - Reference test
    - Colonoscopy for all
    - Colonoscopy for test ⊕, follow-up for test ⊖
    - Colonoscopy for test ⊕, sigmoidoscopy for test ⊖
  - Target lesions of CRC, AA
  - Diagnostic or longitudinal cohort, case-control

Launois R, et al. Eur J Gastro Hep 2014

Bivariate Summary Estimates

	Sensitivity (95% CI)	Specificity (95% CI)	LR +	LR ⊖
Adv Adenoma				
Hemoccult	0.14 (0.09-0.21)	0.95 (0.90-0.97)	2.6	0.91
Magstream	0.48 (0.31-0.66)	0.95 (0.93-0.96)	8.7	0.55
OC-Sensor	0.37 (0.27-0.48)	0.93 (0.90-0.96)	5.6	0.68
Colorectal Cancer				
Hemoccult	0.47 (0.37-0.58)	0.92 (0.84-0.96)	5.9	0.57
Magstream	0.67 (0.59-0.74)	0.93 (0.92-0.95)	9.9	0.36
OC-Sensor	0.87 (0.73-0.95)	0.93 (0.91-0.95)	12.1	0.14

Launois R et al. Eur J Gastro Hep 2014

FIT Studies Included in IU Meta-Analysis

Author	N*	Brand of FIT(s)	Test Threshold (μg/g)
Chiu, HM 2016	3889	OC-Sensor	20
Aniwan, S 2015	948	SD Bioline FOB	10
Chen, YY 2014	6096	OC Light	10
Imperiale, 2014	9989	OC-FIT-CHEK	20
Hernandez, V 2014	779	OC-Sensor	20
Stegeman, I 2014	1112	OC-Sensor	10
Ng, SC 2013	4539	FIT Hemosure	50
Chiu, HM 2013	18297	OC Light	10
Brenner, H 2013	2235	OC Sensor	20
		RIDASCREEN Hemo	2
		RIDASCREEN Haemo/Haptoglobin	2
de Wijkerslooth, TR 2012	1256	OC-Sensor	20
Omata, F 2011	1085	OC-Sensor	20
Haug, U 2011	2325	RIDASCREEN Hemo	20
Brenner, H 2010	1330	immoCARE-C	10
		FOB advanced	8
		PreventID	2
		Bionexia	8
		QuickVue iFOB	10
		Bionexia Hb/Hp Complex	5
Park, DI 2010	770	OC-Sensor	20
Parra-Blanco, A 2010	402	OC-Light	10
Nakazato, M 2006	3090	OC Hemodia	16
Morikawa, T 2005	21805	Magstream 1000/Hem SP	67
Sohn, DK 2004	3794	OC Hemodia	20
Cheng, TI 2002	7411	OC Light	10
Nakama, H 2001	4260	OC Hemodia	20

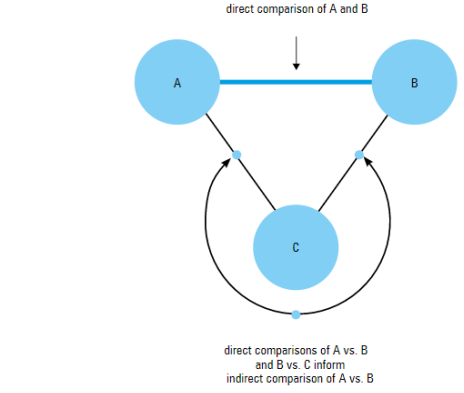
Bivariate Test Characteristics

FIT	Study N	Sensitivity (95% CI) CRC	AA	Specificity (CI)
OC-Sensor* 10-20ug/g	8	0.76 (0.67-0.83)	0.27 (0.24-0.31)	0.94 (0.92-0.96)
OC-Light 10ug/g	4	0.84 (0.63-0.94)	0.42 (0.16-0.73)	0.92 (0.81-0.97)
Germany-based 2-20ug	6	0.76 (0.61-0.86)	0.39 (0.24-0.58)	0.88 (0.74-0.95)
Others: 10-60ug/g	8	0.62 (0.50-0.73)	0.27 (0.17-0.40)	0.95 (0.90-0.98)

*less clinical and statistical heterogeneity for AA, none for CRC

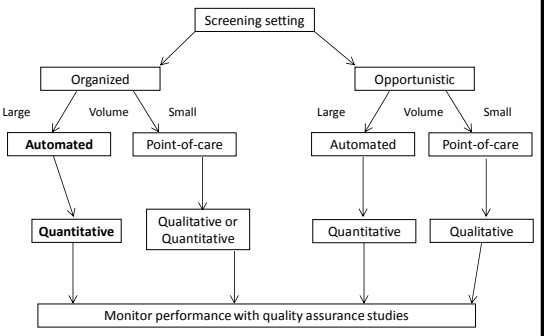
Would network meta-analysis be useful?

- Multiple treatment MA, mixed-treatment comparison
- Combines direct and indirect evidence from all (RCTs) of interventions
- Strengths
  - Compare interventions when few/no head-to-head comparisons
  - Potential for increased certainty of the evidence
- Liabilities
  - Are study characteristics of direct comparisons used to calculate indirect estimates the same/similar among studies?
- Are the methods of NMA adaptable to FIT?



Brignardello-Petersen R, et al. Polskie Archiwum Medycyny Wewnetrznej 2014; 124 (12)

Suggested Approach to Choosing a FIT



Conclusions

- The published literature is replete with head-to-head comparisons of FITs
- Due to variation in study design, test threshold and other features, a comprehensive comparison of FITs remains challenging.
- Several FITs show very good-to-excellent test characteristics
  - OC-Sensor, OC-Light
- Choice of FIT requires consideration of
  --screening setting, volume, and available resources
  --close monitoring to ensure continued performance.

Acknowledgement

- Tim Stump, MA – IU Biostatistics
- James Allison, MD – U.S. FIT guru