EUnetHTA Assessment OTJA10

STOOL DNA TESTING FOR EARLY DETECTION OF COLORECTAL CANCER

– Scientific Subsumption & Additional Considerations –

1) Introduction

The European network for Health Technology Assessment (EUnetHTA) has evaluated the performance characteristics of 2 novel stool DNA tests – ColoAlert and ColoGuard – for the early detection of colorectal cancer (CRC), comparing them to the currently most used non-invasive method (fecal immunochemical testing; FIT) and the invasive gold standard, the colonoscopy.

The main reason is clear: The screening colonoscopy is – from a medical standpoint – a sensitivity tool without false-positives which is in practice strongly suffering from its acceptance in large parts of the population, rarely exceeding 20 % participation rates. The FIT, being based on a simple stool sample and thereby able to reach 73 %[1] and more in this regard, in contrast oversees up to 37 % of all affected patients when meeting recommended specificity goals.[2] Therefore, more sensitive stool tests are needed, especially improving detection rates in the early stages which show 5-year-survival-rates of 85 % and more.

As a Gastroenterologist with a decades-long experience and a track record in consulting Austrian health politicians, I was happy to be invited to advise to the authors in this special field for decades.

The current version 1.4, although being a comprehensive subsumption to the topic, neglects some factors and novel scientific evidence I want to elaborate on in the following.

With this document, I aim to deliver further important information to health political deciders, health insurances and others to whom it may concern.

2) Quantitative Aspects

a. Specificity of ColoGuard

Problem: Although even Exact Sciences Corp, the manufacturer of ColoGuard communicates a specificity value of 87 % (rounded up from 86.6 %) on its website[3], in EUnetHTA's assessment a specificity value of 89.8 % is used[4].

Effect: This decreases the amount of (useless) colonoscopies by 24 %, vastly undermining the latter part of the Benefit-Harms-Analysis (BHA) for ColoGuard.

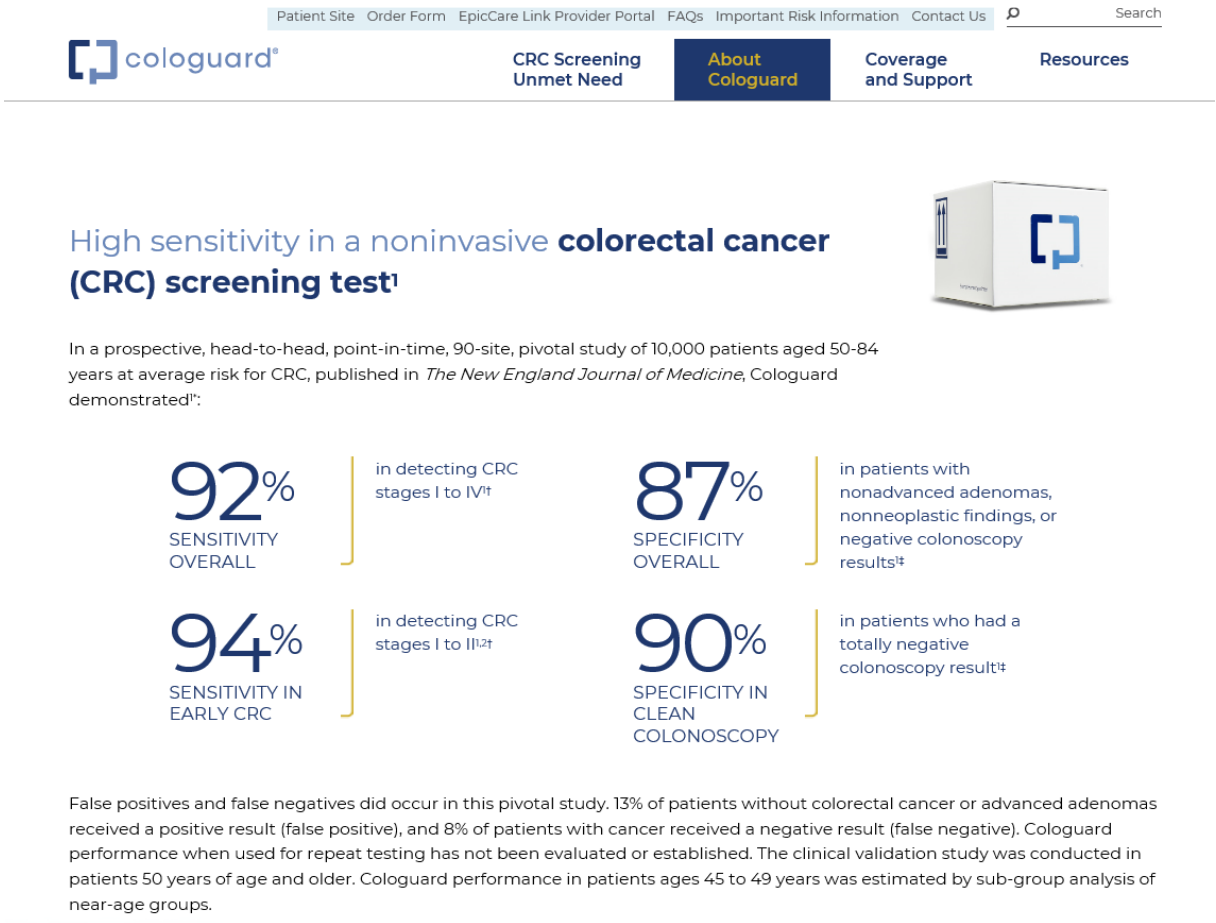
Additional note: In Europe, only screening tests with a specificity of over 90 % can be recommended in general.

Comment GÖG/UMIT 2a: Table 5.4 transparently specifies extracted values for sensitivity and specificity. Of all study participants without colorectal cancer (CRC) and without advanced precancerous lesions (APL) 86.6 % have been detected as negative by

[1] hkk Gesundheitsreport Darmkrebsfrüherkennung (2015)
[2] Gies et al. Gastroenterology 154/2018
[3] Exact Sciences (2019) [https://www.cologuardtest.com/hcp/the-science/pivotal-study; as of 11/06/2019]
[4] Table A21

Cologuard®. Of all participants with negative colonoscopy 89.8 % have been detected as negative by Cologuard®. These two figures are also depicted on the Exact Sciences website cited by Dr. Lexer (see also screenshot below) and refer to the study by Imperiale et al.[5]. In the decision analytic model we applied the specificity of 89.8% as reported by Imperiale et.al. as the "negative results on colonoscopy".

Screenshot (source: https://www.cologuardtest.com/hcp/the-science/pivotal-study, accessed on 09.06.2020)



b. Classifying non-advanced adenomas as Colorectal Cancer

Problem: The assessment classifies non-advanced adenomas (NAA) as an early stage of Colorectal Cancer (CRC) and includes it in the BHA.

This is an assumption that currently certainly can't be considered a consensus in the scientific community and is also agreed on by the main author of the ColoGuard study[6]. The authors of the assessment confirm this on several occasions as well:

LINE 473ff (page 20): …the main target…is to yield…persons with AA or CRC.

LINE 1989ff (page 84): …with regard to disease progression…main target conditions…CRC and AA

---

[5] Imperiale TF, Ransohoff DF, Itzkowitz SH, et al. Multitarget stool DNA testing for colorectal-cancer screening. N Engl J Med. 2014;370(14):1287-1297

[6] Robertson and Imperiale (2015); Gastroenterology 149:1286-1293

This can also be confirmed by looking into the prevalence rates in the population: With ~ 30 % of the screening population falling into this group[7], otherwise we would see way higher prevalence than the current ~ 6 % for CRC in general.[8]

Effect: Including NAA in the data set for the BHA leads to advantages for those tests that excel in this area, which mainly is true for ColoGuard.

Comment GÖG/UMIT 2b: In the BHA, we follow up a target population of individuals at average CRC-risk in Austria. Consequently, individuals can be healthy, adenomas can develop and potentially progress to advanced adenomas and they could further progress to cancer. The model has been calibrated to Austrian epidemiological data. Because the test would be provided on a population level to average risk individuals these health states were agreed on. We did not assume non-advanced adenomas being cancer stages. Cancer stages were modeled separately from non-advanced adenomas.

c. Sensitivity of FIT

Problem: A sensitivity of 73 % for the detection of CRC is used in the BHA. The largest study in this field shows 66 % sensitivity[9]. This is in line with a recent performance comparison for lab-based, quantitative FITs showing additionally the general problem with all FITs being the determination of an optimal cut-off value.[10]

Effect: Using the – in practice unrealistic high – sensitivity of 73 % for FIT for the BHA leads to a roughly 10 % overestimation of the true-positive detection rates of FITs.

Comment GÖG/UMIT 2c: The sensitivity for the detection of CRC was taken from the study by Imperiale et al. 2014[11] (73.8%). The sensitivity of stool tests depends on the tests themselves. The mentioned study (Morikawa T et al. (2005)) applying Megastream has been included later on in a meta-analysis by Launnois 2014 showing the sensitivity of Megastream of 0.668 (95% confidence interval 0.589-0.739) but for OC sensor of 0.97 (95% confidence interval 0.725-0.947). The sensitivity of 73.8% stated in Imperiale et al (applying OC FIT-CHEK, Polymedco) is, therefore, not unrealistic. Looking at the confidence interval (CI) of the sensitivity for both Megastream and OC sensor, 73.8% is at the upper bound of the CI of Megastream and the lower bound of OC sensor.

d. Sensitivity of ColoAlert & ColoGuard

Problem: As KRAS or BRAF mutation indicates that a tumor inhibitor gene was deactivated, patients with one (or both) of those being positive will sooner or later suffer from CRC (or in rare cases: cancer elsewhere in the GI tract) which is confirmed by a follow-up study[12] to the ColoGuard study and a case study[13].

Background: In the precursor stages of CRC, the colonoscopy itself (being the comparator goldstandard in current study designs) fails to confirm CRC in as many as 31 % percent of all

---

[7] Imperiale et al. N Engl J Med 4/2014
[8] Page 19: difference between left & ride side of table 0.2 (37 % vs 11 % CRC + adenoma prevalence)
[9] Morikawa T et al. (2005); Gastroenterology Volume 129, Issue 2
[10] Gies et al. Gastroenterology 154/2018
[11] Imperiale, T. R., DF.; Itzkowitz, SH., Levin, TR; Lavin, P; Lidgard, GP; Ahlquist, DA; Berger, BM (2014). "Multitarget stool DNA testing for colorectal-cancer screening." The New England Journal Of Medicine 370(14): 1287-1297.
[12] Cooper GS et al. Dig Dis Sci 03/18
[13] Ogreid et al. 2007. Eur J Gastroenterol Hepatol 19

cases which is confirmed by the EUnetHTA authors.[14] The additional sensitivity is likely to be gained especially in those early stages that have a major impact on 5-year-survival-rates and therefore the performance of the related screening programs.

In the case of ColoAlert, 12 cases of KRAS & BRAF currently attributed to the control group will sooner or later be in the CRC group. This will lead to a change from a false-positive to a true-positive test result and by this to an increased sensitivity of 87.50 % (56 out of then 64 (partly future) CRC cases). Specificity increases to 96.25 % (231 of then 240 cases).[15]

ColoAlert should profit even more due to it using a second (direct) tumor DNA marker (BRAF) which gives a clear yes/no-answer while the methylation markers used by ColoGuard are indirect.

Effect: As shown, in screening scenarios which run on average for 2 to 3 decades considering them starting at latest 50 and an average life expectancy of 80.9 years[16], tumor DNA stool tests will have significant better performance values than in a common study setting.

Additional note: As of September 20th 2019, the FDA now recommends ColoGuard screening to start from the age of 45 and therefore even earlier than for colonoscopy. Main Reason is the above mentioned "true early detection" that renders such an early beginning effective.[17]

Comment GÖG/UMIT 2d: Cooper et al. 2018 (the study Dr. Lexer cited regarding new data for true positive stool DNA test results) found in 12 selected, reevaluated patients with positive stool DNA test (Cologuard®) and an initial negative colonoscopy that 5/12 had a repeat positive stool DNA test about 1 year later and of these, 3/5 had positive findings on repeat colonoscopic examination. However, as this study has relevant limitations (e.g. small sample size, lack of knowledge as to which component of the stool DNA test panel was positive on initial and repeat evaluation, no definitive exclusion of occult malignancy in the study sample) further and larger studies are needed to validate these findings. Thus the suggested adjustment of sensitivity and specificity data for ColoAlert® based on this study using Cologuard® is not applicable. Nevertheless these findings may indicate that repeat stool DNA tests could be helpful to identify a subset of patients with missed or occult colorectal neoplasia after a negative colonoscopy. As well as these findings emphasize the relevance of high-quality colonoscopies, an aspect that has been addressed in the EUnetHTA report.

3) Qualitative Aspects

a. Quality of Control Colonoscopies

Problem: The assessment doesn't factor in the quality of the control colonoscopies that follow positive results from any stool test, although both DNA tests have the additional advantage compared to FIT that those colonoscopies are way more precisely executed when tumor DNA was found due to this being a known absolute risk factor by gastroenterologists.[18]

Effect: Both DNA tests will prevent CRC better in practice than the assessment foresees because of the increased quality of potential control colonoscopies. ColoAlert should profit

[14] Table A21
[15] Table 2 in Dollinger et al. Clin Lab 10/2018
[16] https://de.statista.com/statistik/daten/studie/954/umfrage/lebenserwartung-bei-geburt-in-ausgewaehlten-laendern-der-europaeischen-region/
[17] FDA PMA Supplement 2019; www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P130017S029
[18] Koliani-Pace et al. 2017, Gastrointestinal Endoscopy, Vol 85 (3): 666-668

even more due to it using a second (direct) tumor DNA marker (BRAF) which gives a clear yes/no-answer while the methylation markers used by ColoGuard are indirect.

Comment GÖG 3a: This issue is discussed within research question [D0026] "How does the test modify the effectiveness of subsequent interventions?" and subsequently mentioned in the discussion of the assessment. The study cited in [D0026] (Johnson et al. 2017[19]) is the one which is also cited by the Editorial[20] Dr. Lexer refers to (and is the original source of that statement regarding an effect of positive DNA test results on the adenoma detection rate of colonoscopy).

b. Improved Screening Schemes

Problem: The assessment doesn't factor in the explanatory power for the following screening scheme for a patient tested positive by a stool tests. Both DNA tests offer an array of biomarkers of which the direct ones must lead – in case of being positive – to an adapted screening scheme for the patient, at least shortening the interval until the next stool test is performed.[21]

Effect: Both DNA tests will prevent CRC better in practice than the assessment foresees due to the personalized screening scheme arising out of positive test reports.

Comment GÖG/UMIT 3b: In the model it is assumed that a positive test would always lead to a colonoscopy. Based on the results of the colonoscopy, this could lead to shorter screening intervals as stated in the methods part: „Individuals with identified advanced adenomas continue screening after removal by polypectomy and start 3-yearly surveillance with colonoscopy. If an advanced adenoma (again) is found in the surveillance colonoscopy (and removed by polypectomy), patients continue with the 3-yearly surveillance. If only non-advanced or no adenomas are found in the surveillance colonoscopy, individuals enter a 5-yearly surveillance. Five-yearly surveillance is continued as long as no advanced adenomas are detected. A detection of advanced adenomas again leads to the shorter 3-yearly surveillance. Individuals with detected non-advanced adenomas continue with colonoscopy every 10 years after removal of the adenoma by polypectomy. Surveillance examinations are considered in all strategies until the age of 74." Hence, we did account for risk-adapted screening intervals according to practice guidelines (Hassan 2013[22]).

c. Adherence Rates

Problem: The assessment considers the adherence rates to all stool tests the same. This is an unrealistic assumption, as the tests differ especially regarding their screening interval. With FIT normally being tests annually to biannually compared to a 3-yearly interval for the DNA tests, adherence for the former will suffer.

A recent customer survey performed by PharmGenomics GmbH, the manufacturer of ColoAlert, showed a 97 % customer satisfaction[23], being in line with this.

---

[19] Johnson DH, Kisiel JB, Burger KN, Mahoney DW, Devens ME, Ahlquist DA, et al. Multitarget stool DNA test: clinical performance and impact on yield and quality of colonoscopy for colorectal cancer screening. Gastrointestinal Endoscopy. 2017;85(3): 657-65.e1.

[20] Koliani-Pace et al. 2017, Gastrointestinal Endoscopy, Vol 85 (3): 666-668

[21] ibidem

[22] Hassan C, Quintero E, Dumonceau JM, Regula J, Brandao C, Chaussade S, et al. Post-polypectomy colonoscopy surveillance: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. Endoscopy. 2013;45(10): 842-51. Epub 2013/09/14.

[23] PharmGenomics GmbH, Customer Survey 2019; www.coloalert.de

Effect: Both DNA tests will prevent CRC better in practice than the assessment foresees due to their higher adherence rates.

Comment GÖG/UMIT 3c: A customer survey alone does not provide comprehensive evidence on the comparative adherence between different tests within a screening program. In general, adherence strongly depends on the implementation of the screening program. Apart from a study from 2011[24], including a precursor test of Cologuard®, we are not aware of any comparative prospective survey on adherence to FIT or gFOBT screening strategies versus DNA test screening strategies (taking into account not only the tests itself, but also screening intervals).

4) Summary of Effects

As already explained in the introduction, the current version of the EUnetHTA report evaluating 2 DNA tests for CRC screening (Version 1.4) offers a first comprehensive overview of this novel technology. Due to this topic being comparably new, its scientific background is still surrounded by uncertainties. With this document, I further elaborated the topic and aim to lay a basis for an upcoming version, factoring in newer developments and even more – qualitative and quantitative – aspects of the different screening strategies evaluated.

To summarize, the different factors explained above have the following (qualitative) effects on the performance characteristics of the stool tests evaluated, compared to the results described in the assessment:

Chapter	Perf. Characteristic	Scope	FIT	ColoAlert	ColoGuard
2a	Specificity	Test			--
2b	Sensitivity	Test		-	--
2c	Sensitivity	Test	--		
2d	Sensitivity	Test		++	++
3a	Sensitivity	Program		+	+
3b	Sensitivity	Program		+	+
3c	Sensitivity	Program		++	++
Sum	**Sens. & Spec.**	**Program**	- 0	+ ++	0 +

Table 1: Summary of additional effects

Based on this, FIT-based screening programs should deliver in practice a slightly worse performance than calculated with the EUnetHTA BHA, mainly due to its overestimated sensitivity.

On the other side, both DNA tests – respectively screening programs based on them – should perform better than summarized in the assessment. Reasons are the positive effects on the following processes in the screening scheme and their underrated sensitivity

[24] Calderwood AH, Wasan SK, Heeren TC, Schroy PC, 3rd. Patient and Provider Preferences for Colorectal Cancer Screening: How Does CT Colonography Compare to Other Modalities? International Journal Of Cancer Prevention. 2011;4(4): 307-38

especially in early stages where the benchmark colonoscopies can't confirm positive results. These strong effects are not outweighed even when NAAs – scientifically reasonable – are excluded from the BHA.

While for ColoGuard a questionable specificity is used in the assessment, decreasing the total value of a screening program making use of it, ColoAlert's benefits are overall vastly underrated.

5) Conclusion

All in all, stool DNA testing can be considered the future of CRC screening, offering a superior combination of sensitivity and specificity (Youden Index) than the FIT – while even enhancing patient compliance.

This is already clear when looking at chapter 2d: Factoring in only one of the advantages of this scientific approach already makes ColoAlert reach the specificity of the FIT while vastly outperforming its sensitivity and decreasing the share of overseen patients by over 70 %.

Regarding the choice between ColoGuard and ColoAlert only, the decision is more difficult. Screening programs aiming for the highest sensitivity to detect as many affected patients as early as possible and outweighing this to the increased harm of significantly more useless colonoscopies, should favor ColoGuard – as soon as it is available in Europe. It is then up to health political deciders to evaluate its comparably high costs.

In the meanwhile – or in general when factoring in specificity and/or price – ColoAlert is a promising screening tool for European health systems, delivering very good performance characteristics and a leading patient compliance at a price point where health systems should be able to easily amortize the initial investment.

In general, I agree with my fellow oncology advisory board member Professor Richard Greil (University of Salzburg): The overall evidence for stool DNA testing is sufficient to start public screening, at least in a pilot project. A larger evaluation of its practical benefits and the health-economic impact seems commanded.

Dr. med. Gerfried Lexer

Gastroentrologist & Advisor to the Assessment

Comment GÖG/UMIT Summary 4 and Conclusion 5: see comments above.