STUDY DESIGN

Lib Oncol. 2024;52(2-3):129–141
doi: 10.20471/LO.2024.52.02-03.19

VERIFICATION OF THE FECAL IMMUNOCHEMICAL TEST ON THE OC-SENSOR CERES DEVICE

ANTONIA ŠOIĆ[1], SANJA LANGER[1], ZVJEZDANA ŠPACIR PRSKALO[1], MIHAELA GAĆE[1], MILICA VRBANČIĆ[1], INES SEVER[1] and LJILJANA MAYER[1]

[1]Department of Medical Biochemistry in Oncology, University Hospital for Tumors, Sestre milosrdnice University Hospital Center, Zagreb, Croatia

Summary

Organized screening programs for the detection of colorectal cancer (CRC) in the Republic of Croatia include testing for occult bleeding in the stool. Traditionally, the guaiac fecal occult blood test (gFOBT) has been used. Another well-established method is the fecal immunochemical test (FIT). In the Department of Medical Biochemistry in Oncology laboratory, at the University Hospital for Tumors, Sestre milosrdnice University Hospital Center, FIT verification was performed on the OC-SENSOR Ceres device according to CLSI guidelines. The parameters of precision, trueness, carryover, stability, and a comparison with gFOBT were assessed. Precision was demonstrated with the calculated expanded relative measurement uncertainty, which was 2,54%, 1,24%, and 5,38%. Accuracy was calculated and shown as bias (2,1%, 0,8%, and 1,3%). Hemoglobin stability in the sample was determined for up to 10 days at 2–8°C (4,95%) and even at room temperature (9,57%). The obtained result for analyte carryover was -0,43%. A comparison of gFOBT and FIT found that gFOBT shows a higher number of false-negative results (38%). Considering all the advantages of FIT compared to gFOBT and the successful verification of the method, it is justified to replace gFOBT with FIT.

KEYWORDS: *fecal immunochemical test (FIT); verification; colorectal cancer screening; stability; guaiac fecal occult blood test (gFOBT)*

INTRODUCTION

Colorectal cancer (CRC) is the second most common type of cancer in the Republic of Croatia. One in thirty-five women and one in twenty-two men will develop CRC. The mortality rate from CRC in Croatia is 53 per 100,000 inhabitants, significantly higher than the European Union (EU) average of 32 per 100,000, placing it among the highest rates in the EU. The incidence of colon cancer in Croatia is comparable to that of the Netherlands, which stands at 97 per 100,000 inhabitants; however, their mortality rate is lower at 36 per 100,000. In 2020, life expectancy in Croatia was nearly three years below the European average(1).

Cancer mortality can be reduced through early detection and treatment. Early detection has two key components: screening and early diagnosis. Screening aims to identify individuals whose medical reports suggest a particular cancer or precancer, even before symptoms develop. If abnormalities are found during screening, further testing should be conducted to confirm a diagnosis, followed by a referral for treatment if cancer is present. With early diagnosis, cancer is more likely to respond to treatment, increasing the chances of survival and reducing morbidity. In Croatia, a national program for the early detection of CRC has been implemented every other year for over 15 years, targeting individuals aged 50-74. For early detection through organized screening to be

Corresponding author: *Antonia Šoić, Department of Medical Biochemistry in Oncology, University Hospital for Tumors, Sestre milosrdnice University Hospital Center, Ilica 197, 10000 Zagreb, Croatia. e-mail: antonia.soic@gmail.com*

effective, it is crucial to reach an adequate percentage of the target population. A participation rate of 45% is considered satisfactory, while a rate of 65% is desired(2).

Early detection of colon cancer by testing for occult bleeding in stool (bleeding not visible to the eye) in an asymptomatic, average-risk population has proven to be an effective method for early detection, due to its ease of application and relatively low cost compared to treatment at advanced stages of the disease. Large population-based and controlled studies on fecal occult blood testing for early colon cancer detection have reported a positive fecal occult blood test in 1% to 5% of subjects. A positive result indicates the need for further diagnostic procedures, most often involving colonoscopy, proctosigmoidoscopy, or a double-contrast barium enema examination. The sensitivity of the guaiac fecal occult blood test (gFOBT) for detecting colon cancer is approximately 50%, with a specificity of 98%. The sensitivity to detect advanced adenoma ranges from 6% to 17%(3).

The gFOBT relies on guaiac-impregnated reaction paper and the peroxidase activity of intact hemoglobin. Oxidation of guaiac (*4-[5-(4-hydroxy-3-methoxyphenyl)-3,4-dimethylfuran-2-yl]-2-methoxyphenol*) with the catalyst hydrogen peroxide develops the characteristic blue color *((4Z)-2-methoxy-4-[(5E)-5-(3-methoxy-4-oxocyclohexa-2,5-dien-1-ylidene)-3,4-dimethylfuran-2-ylidene]cyclohexa-2,5-dien-1-one*). The test is specific to free hemoglobin from lysed erythrocytes, as hemolysis in stool is primarily caused by water and salts. However, the guaiac reaction is also sensitive to non-human hemoglobin (from meat and meat products) and vegetable peroxidases. To avoid false positive (FP) results, a restricted diet is required for several days before sample collection. Additionally, iron supplements should be avoided as high doses may cause FP results, as should drugs like aspirin, indomethacin, phenylbutazone, corticosteroids, and reserpine, which can induce gastrointestinal irritation or bleeding. Vitamin C intake over 1 gram per day can lead to false negative (FN) results due to the reducing properties of ascorbic acid. Sample collection typically requires consecutive stool samples over three days(4).

Unlike gFOBT, which is a qualitative test with results displayed as either positive or negative, the fecal immunochemical test (FIT) is a quantitative method that provides numerical analysis values. Additionally, FIT measures hemoglobin concentration in stool without being affected by various interferences, eliminating the need for restrictive diets before sampling.

The FIT is endorsed by leading professional associations, including the European Society for Medical Oncology (ESMO)(5), National Comprehensive Cancer Network (NCCN)(6), American Cancer Society (ACS)(7), American Society of Clinical Oncology (ASCO)(8), World Endoscopy Organization (WEO)(9), United States Preventive Services Task Force (USPSTF)(10), the American College of Physicians (ACP)(11,12), American College of Gastroenterology (ACG)(13). As of December 2022, the European Commission also recommends FIT as a screening method for colorectal cancer(14).

Since 2018, following the initiation of procedures to introduce the diagnostic-therapeutic procedure (DTP) for specialist council healthcare screenings for occult bleeding through fecal immunotesting, FIT has also been available within the public health system of the Republic of Croatia.

The OC-SENSOR Ceres® (Eiken Chemical, Japan) device measures hemoglobin concentration in feces using the latex turbidimetric method. This automated immunochemical approach is based on a specific antigen-antibody reaction, which occurs between an antigenic determinant of hemoglobin A_0 (HbA_0) and the active groups of the anti-human HbA_0 antibodies. The extent of binding depends on the concentrations of both the antigen and antibody. Antibodies are attached to polystyrene latex particles via inactive groups, and during the antigen-antibody reaction, agglutination of latex particles occurs, resulting in a change in the intensity of light transmitted through the reaction mixture. A dose-response curve of the absorbance unit *vs*. concentration is generated using the results obtained from the calibrators. The concentration of hemoglobin in the patient sample is determined from this curve(15).

Verification components in medical biochemistry laboratory

To prove that a method serves its intended purpose, it is necessary to conduct validation or verification. These processes ensure accurate, reproducible, and precise results. Validation is the

process of thoroughly examining the characteristics of a method to provide objective evidence that a particular element is suitable for its intended purpose. It is carried out by the manufacturers of the test and includes a series of measurements of various parameters, which establish acceptance criteria. Verification is a procedure that provides objective evidence that a method meets all the requirements of its intended purpose. It requires financial, time, and human resources, so it must be carefully planned. The fundamental principles of verification have been applied in medical-biochemical laboratories since their inception. Since 1968, the Clinical and Laboratory Standards Institute (CLSI) has been developing and actively promoting processes for global harmonization and creating formal consensus for standardization in laboratory diagnostics. It regularly publishes continuously improved documents, standards, and guidelines, and conducts ongoing certification and educational programs. Through its efforts, CLSI ensures the essential prerequisites for the development of clinical laboratories, improves the comparability of test results, and encourages excellence in laboratory medicine.

Verification of the FIT was performed on the OC-SENSOR Ceres device in the laboratory of the Department of Medical Biochemistry in Oncology at the University Hospital for Tumors, Sestre milosrdnice University Hospital Center, Zagreb, Croatia. The parameters evaluated included precision (which encompasses repeatability, intermediate precision, and reproducibility), trueness, carryover, stability, and comparison with the gFOBT. Linearity does not apply to this method, as FIT is an immunochemical approach that uses six calibrators to establish a calibration curve described by a higher-order (non-linear) mathematical relationship. The limit of detection (LOD), limit of quantification (LOQ), and cut-off value are determined by the manufacturer and are used as such in clinical practice. LOD is 6 ng/mL, LOQ is 20 ng/mL, measure range is 20 – 1000 ng/mL, cut-off is 100 ng/mL.

Precision is a quantitative value that indicates the degree of agreement (concordance) between the results of a series of measurements conducted on the same homogeneous sample under precisely defined conditions. It reflects the ability of an analytical method to yield identical results in repeated measurements of the same sample. The distribution of random errors influences precision and is not a direct measure or indicator of accuracy.

The measure of trueness is represented by the deviation (bias), which is the difference between the mean value of the results obtained using the tested method under specific conditions and the expected value of the reference material or the value measured by another method or analytical system.

Carryover is a procedure that assesses the extent to which a sample with an extremely high result can cause falsely elevated values in subsequent samples with extremely low concentrations. This assessment is typically expressed as a percentage.

Sample stability refers to the ability to maintain consistent quantification values over a defined period and at specific value levels during storage under predetermined conditions. Stability needs to be assessed primarily because screening processes often involve analysis with a time delay.

Additionally, we examined the comparability of the FIT and gFOBT methods. This comparison extends beyond the traditional framework of assessing bias between two methods since they differ in their quantitative and qualitative nature. However, it was important to compare their diagnostic accuracy since these methods are meant to be used primarily as screening tests.

MATERIALS AND METHODS

Verification of the OC-SENSOR Ceres® device (Eiken Chemical, Japan), which employs an immunochemical method for the automated and quantitative measurement of hemoglobin concentration in stool samples was conducted from June to October 2024.

To verify precision and trueness, manufacturer-supplied samples intended for internal quality control were used. These samples have declared specific target values and coefficients of variation. The procedure for determining repeatability, inter-precision, intra-laboratory precision, and trueness was conducted according to the CLSI guidelines EP15-A3 *User Verification of Precision and Estimation of Bias*. Commercial control samples were analyzed at different times of the day and by different laboratory personnel to capture the maximum variability of the analytical system. Three

levels of control samples were run in 5 replicates over five consecutive days (N=25). For each day, the mean value of the measurements was calculated, along with the total mean value of all the obtained means, the standard deviation, and the square of the standard deviation.

Intra-laboratory precision and within-run precision were expressed by standard deviation and percentage of the coefficient of variation.

The Grubbs test was utilized to identify outlier values, where any point falling outside the acceptable range has been considered an extreme value and was excluded from statistical processing. Consequently, in such cases, the measurement was repeated.

Grubbs range = mean ± G x SD, where G (Grubbs factor) for 25 measurements is 3,135.

One-way analysis of variance (ANOVA, Table 1) was used for statistical data processing:

k – number of series(5)
n – number of replicates per series(5)
N – total number of measurement results included in the analysis(25)
– mean value of all measurement results

The variances within the series (VW) and between the series (VB) were derived from the data presented above:

$$VW = MS2$$
$$VB = (MS1 - MS2) / n_0$$

where n_0 (mean value of the number of results per series) depends on the number of measurements (for the basic plan of the verification procedure $n_0 = 5$).

Repeatability (in-run precision, s_R), inter-precision (between-run precision, s_B), and intra-laboratory precision (s_{WL}) were calculated based on one-way analysis of variance (ANOVA) using the following equations:

$$s_R = \sqrt{V_W} \qquad s_B = \sqrt{V_B} \qquad s_{WL} = \sqrt{V_W + V_B}$$

The results of the verification were interpreted by comparing the calculated value with the value declared by the manufacturer. Suppose the manufacturer expressed repeatability or intra-laboratory precision as a coefficient of variation. In that case, it is necessary to calculate the standard deviation of the mean value of all results for the tested concentration levels:

$$\sigma_r = CV\%_r \times \overline{\overline{x}}$$

where CV%r is the declared coefficient of variation for repeatability/in-laboratory precision at the approximate tested concentration level, and $\overline{\overline{x}}$ is the mean value of all measurement results obtained over five days.

Depending on the ratio of calculated repeatability (s_R) to declared repeatability (σ_R) and calculated intra-laboratory precision (s_{WL}) to declared intra-laboratory precision (σ_{WL}), we can determine whether the method meets the established criteria. Suppose the verification results do not satisfy the primary criteria. In that case, it is necessary to examine the results about the extended criteria, which are based on a comparison of the data with the upper verification limit (UVL – the upper limit of the 95th percentile of the expected estimated values for imprecision). If the verification results still do not meet the selected criteria, it is essential to contact the manufacturer to identify the exact source of the issue and explore potential solutions (Table 2).

The upper verification limit (UVL) was calculated using the following equation:

$$UVL = F \times \sigma \text{ or } UVL = F \times \%CV$$

where the F-factor depends on the number of tested samples (a minimum of two concentration levels) and the degrees of freedom (DF) related to the total number of measurements.

Intra-laboratory precision represents the standard measurement uncertainty (u). When ex-

Table 1.

Calculation of one-way analysis of variance (ANOVA) parameters

Source of variation	SS	DF	MS
Between series	$SS1 = \sum_{i=1}^{k} n_i(\overline{x}_t - \overline{\overline{x}})^2$	DF1 = k – 1	MS1 = SS1 / DF1
Within the series	$SS2 = SS_{uk.} - SS1$	$DF2 = DF_{uk.} - DF1$	MS2 = SS2 / DF2
In total	$SS_{uk.} = SS1 = \sum_{i=1}^{N} n_i(\overline{x}_t - \overline{\overline{x}})^2$	$DF_{uk.} = N - 1$	

SS – sum of squares; DF – degrees of freedom; MS – the mean value of the sum of squares

Table 2.

Conclusion of the precision verification based on the obtained results

Possible outcomes	Conclusion
$S_R < \sigma_R$ i.e. $S_{WL} < \sigma_{WL}$	The estimated repeatability is lower than the manufacturer's declared repeatability, thus meeting the acceptance criteria.
$S_R > \sigma_R$ i.e. $S_{WL} > \sigma_{WL}$	The estimated repeatability is higher than the manufacturer's declared repeatability. It is necessary to calculate the upper verification limit (UVL) to determine whether this difference is statistically significant.
$S_R < UVL_R$ i.e. $S_{WL} < UVL_{WL}$	The estimated repeatability is lower than the verification value, indicating that it meets the acceptance criteria.
$S_R > UVL_R$ i.e. $S_{WL} > UVL_{WL}$	The estimated repeatability is higher than the verification value, indicating that it does not meet the acceptance criteria.

S_R – calculated repeatability; σ_R – declared repeatability; s_{WL} – calculated intralaboratory precision; σ_{WL} – declared intra-laboratory precision; UVL – upper verification limit

pressed as a coefficient of variation, intra-laboratory precision reflects the relative standard uncertainty of measurement (u_{rel}). To calculate the expanded relative measurement uncertainty at a confidence level of 95%, it is necessary to include a coverage factor of k = 2.

The expanded relative measurement uncertainty (U_{rel}) was calculated based on the formula:

$$U_{rel} = u_{rel} \times 2$$

The calculated expanded relative measurement uncertainty was compared with the selected acceptance criterion, and a conclusion was drawn based on the results obtained where Urel below the acceptance criterion was satisfactory and above was considered unacceptable.

In cases where commercial control materials with a known target value (TV) are used for bias testing, and it is impossible to determine the standard error, it can be considered equal to 0. Consequently, the verification interval in this case becomes narrower, increasing the probability that the mean value of all measurement results ($\overline{x}$) falls outside the verification interval.

The mean value and standard deviation (s_x) of the results for each concentration level were calculated using the following formulas:

$$\overline{x} = \frac{\Sigma_i^n x_i}{n} \qquad s_x = \sqrt{\frac{\Sigma_{i=1}^n (x_i - \overline{x})^2}{n-1}}$$

The standard error of all measurement results was calculated using the following equation:

$$se_{\overline{x}} = \sqrt{\frac{1}{nRun}\left[s_{WL}^2 - \left(\frac{nRep-1}{nRep}\right)s_R^2\right]}$$

sR – precision in series
sWL – intralaboratory precision,
nRun – number of series of measurements
nRep – number of measurement replicates

The verification interval (95% CI) was calculated using the following expression:

$$TV \pm t_{0,975,nRun-1} \times se_{\overline{x}}$$

If the mean value of all measurements falls within the verification interval, the bias is considered statistically insignificant. A determination can be made regarding clinical acceptability by comparing the calculated bias results with the acceptance criteria. The established acceptance criterion was 10% or 12% for low values, as specified by external quality control (Labquality, Finland).

The stability of the analyte in the sample was assessed according to the guidelines outlined in *A protocol for testing the stability of biochemical analytes. Technical document*(16). A refrigerator (Gorenje, Croatia) served to store the samples at a temperature of 2-8°C, while a sterilizer (Instrumentaria ST-01/02, Croatia) was used for samples stored at +35°C. To demonstrate a wide range of hemoglobin concentrations hemolyzed serum was spiked into the original manufactured sample tubes. The initial hemoglobin concentration was determined in 14 samples for each series. Samples with an initial concentration below the limit of quantification (20 ng/mL) were excluded from further processing. One batch was stored at room temperature, the second at 2-8°C, and the third at +35°C. Hemoglobin concentrations were measured again after 24 hours, 72 hours, 7 days, and 10 days. The percentage deviation (PD) at different time intervals from the initial concentration was calculated for each sample according to the following formula:

$$PD = \frac{result\ at\ time\ x - initial\ result}{initial\ result} x100 .$$

Next, the mean value of the absolute deviation of the samples in each series was calculated after each time interval (24 hours, 72 hours, 7 days, and 10 days). The acceptable coefficient of variation, based on the external quality control (Labquality, Finland) criterion was 10% or 12% for low-value samples (up to 100 ng Hb/mL).

The assessment of analyte carryover in the samples was conducted following CLSI guideline H26-A2 *Validation, Verification, and Quality Assurance of Automated Hematology Analyzers; Approved Standard – Second Edition.* Immediately after running high-concentration samples in triplicate (H1, H2, H3), a low-concentration sample was measured in triplicate as well (L1, L2, L3). The percentage of carryover was calculated using the following formula:

$$carry\ over = \frac{L1 - L3}{H3 - L3} x100 .$$

According to the manufacturer's specifications, the acceptance criterion for carryover was set at 1.5%.

HemoGnost® (BioGnost, Croatia) cards were used for the comparison of FIT with the gFOBT method using 56 adapted samples (homogenized oat porridge spiked with hemolyzed serum). A cut-off of 100 ng/L (as declared by the manufacturer) was used for the qualitative classification of results for FIT and used as a reference measurement in assessing the diagnostic accuracy of gFOBT in comparison. Measures of diagnostic accuracy -sensitivity, specificity, positive (PPV) and negative (NPV) predictive value were calculated as follows:

$$\text{Sensitivity} = \frac{TP}{TP + FN} x100$$

$$PPV = \frac{TP}{TP + FP} x100$$

$$\text{Specificity} = \frac{TN}{FP + TN} x100$$

$$NPV = \frac{TN}{FN + TN} x100$$

RESULTS

Precision and trueness

Data for precision and trueness evaluation for all three control samples are shown in Table 3. The total average of all measurements for samples 1, 2, and 3 were 148,7 ng/mL (SD = 1,86 ng/mL; CV = 1,25%), 449,5 ng/mL (SD = 2,62 ng/mL; CV = 0,58%), 74,8 ng/mL (SD = 1,97 ng/mL; CV = 2,64%), respectively.

The estimated repeatability for samples 1 and 2 (CV = 1,1% and CV = 0,3%) is equal to or lower than the manufacturer's declared repeatability (CV = 1,1% and CV = 0,8 %), thus meeting the acceptance criteria. The estimated repeatability for sample 3 (CV = 2,3%) is higher than the manufacturer's declared repeatability (CV = 1,2%). It is calculated The verification value (V_V = 1,6%) to determine whether this difference is statistically significant. The estimated repeatability is higher than the verification value, indicating that it does not meet the acceptance criteria. The estimated intralaboratory precision for all three samples (CV = 1,3%, CV = 0,6%, CV = 2,7%) is lower than the manufacturer's declared intralaboratory precision (CV = 1,7%, CV = 1,5%, CV = 2,3%), thus meeting the acceptance criteria. The calculated expanded relative measurement uncertainty (Urel) for all three samples (2,54%, 1,24%, 5,38%) was compared with the selected acceptance criterion (10%, 10%, 12%), and a conclusionwas drawn based on the results obtained where Urel below the acceptance criterion is satisfactory.

A determination of trueness is made by comparing the calculated bias results with the acceptance criteria. The established acceptance criterion was 10% or 12% for low values and inaccuracy (%bias) of consecutive measurements of three samples were 2,1%, 0,8%, and 1,3%.

Stability

The stability of samples is shown in Figure 1 for three different storage conditions (at room temperature, at 2-8 °C, and at 35 °C) measured at four different time stops (after 24 h, 72 h, 7 days, and 10 days). After 24 hours, hemoglobin concentration deviated by 5,17% at room temperature, 4,40% at 2-8°C and 6,67% at +35°C. After 72 hours,

Table 3.

Results for precision and trueness estimation of three commercially available control samples

	SAMPLE 1	SAMPLE 2	SAMPLE 3
Pattern Name	OC-FIT Control LV1	OC-FIT Control LV2	OC-FIT Control LV3
N	25	25	25
Target value (ng/mL)	145,7	453	73,9
$\bar{x}$	148,720	449,480	74,840
SD	1,8601	2,6160	1,9723
CV (%)	1,25	0,58	2,64
Declared repeatability (%CV)	1,1	0,8	1,2
Declared intralaboratory precision (%CV)	1,7	1,5	2,3
Acceptance criterion for MN (%)	10	10	12
Estimated repeatability (%CV)	1,1	0,3	2,3
Estimated intralaboratory precision (%CV)	1,3	0,6	2,7
Expanded relative measurement uncertainty (%Urel)	2,54	1,24	5,38
Allowed bias (abs.)	14,57	45,30	8,87
Estimated bias (abs.)	3,02	3,52	0,94
Estimated bias (%)	2,1	0,8	1,3
Acceptable %bias	10	10	12

N – total number of measurements; SD – standard deviation; CV – coefficient of variation; MN – measurement uncertainty

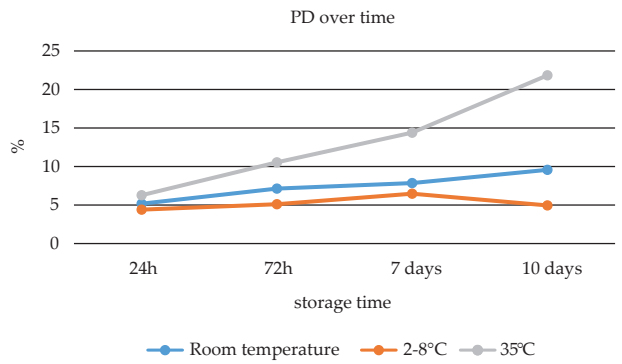

Figure 1. Graphic representation of the mean percentage deviation (PD) of the hemoglobin concentration from the initial value based on the storage time and temperature.

the deviation increased to 7,13% at room temperature, 5,10% at 2-8°C and 10,54% at +35°C. After 7 days, the deviations were recorded as 7.85% at room temperature, 6.47% at 2–8°C and 14.39% at +35°C. After 10 days, the deviations peaked at 9.57% at room temperature, 4.95% at 2–8°C and 21.83% at +35°C.

Carryover

The determined hemoglobin concentrations for replicate high and low samples are as follows: 697 ng/mL, 700 ng/mL, 710 ng/mL, 8 ng/mL, 9 ng/mL and 11 ng/mL.

The obtained result for the analyte carryover was -0.43%.

Comparison of quantitative (FIT) and qualitative (gFOBT) methods of fecal occult blood detection

The FIT method yielded negative results for hemoglobin concentrations below 100 ng/mL in 13 samples, indicating that 23% of the samples analyzed by the gFOBT method were true negative (TN). In 21 samples, the FIT method measured hemoglobin concentrations above 100 ng/mL, classifying them as positive, while the gFOBT method produced negative results. This resulted in 38% of false negative (FN) results with the gFOBT method. Both methods produced positive results in 22 samples, leading to a 39% rate of true positive (TP) results with the gFOBT method (Table 4).

The sensitivity of gFOBT compared to FIT was 51%, with a specificity of 100%. The positive predictive value (PPV) was 100%, while the negative predictive value (NPV) was 38%.

Table 4.

Presentation of the individual results of hemoglobin detection in samples using FIT and gFOBT methods (N = 56)

FIT (ng/mL)	gFOBT	FIT (ng/mL)	gFOBT	FIT (ng/mL)	gFOBT
0	negative	100	negative	416	positive
4	negative	136	negative	438	positive
8	negative	142	negative	450	positive
18	negative	172	negative	455	positive
20	negative	178	negative	457	positive
26	negative	223	negative	465	positive
30	negative	233	negative	511	positive
37	negative	233	negative	516	positive
38	negative	249	negative	541	positive
40	negative	258	negative	567	positive
47	negative	272	negative	650	positive
66	negative	275	negative	707	positive
73	negative	287	negative	751	positive
		303	negative	779	positive
		310	negative	791	positive
		347	negative	846	positive
		350	negative	862	positive
		369	negative	865	positive
		398	negative	873	positive
		407	negative	874	positive
		409	negative	912	positive
				943	positive
TN	23%	FN	38%	TP	39%

DISCUSSION

The national strategic framework of the Republic of Croatia against cancer until 2030 aims to ensure the effectiveness of the screening program, targeting a response rate of at least 45% of individuals examined for occult blood in the stool, with a desired response rate of 65%. The goal is to reduce mortality by at least 15% five years after the program's initiation and to achieve a coverage rate of 60% through screening. Additionally, the framework aims to detect cancer in its earliest stages, when the disease is more accessible to treat, thereby improving the patient's quality of life and survival rates. It also seeks to provide appropriate diagnosis and treatment for individuals with a positive test for occult bleeding in the stool and to enhance awareness of the early signs and symptoms of the disease(17).

Among European countries with organized CRC screening programs utilizing fecal occult blood tests, only Croatia, Latvia, and Greece continue to employ gFOBT as of 2024. Countries that initially established screening programs adopted gFOBT but have since transitioned to the more advanced FIT. In contrast, countries that began their screening initiatives more recently opted for FIT from the outset, recognizing its superior sensitivity and specificity for detecting occult blood(18).

Eurostat is the statistical office of the European Union that collects, aggregates, and visualizes data related to various topics, including responses to colorectal cancer screening (Figure 2). Recent Eurostat data indicate significant disparities in screening responses among EU countries. The Netherlands and Sweden exceed 65%, Finland approaches 80%, while Croatia's response is only around 25%. One of the repercussions of low response rates and late detection is an adverse financial impact. The disease is often diagnosed at an advanced stage when treatment is more uncertain, prolonged, and more expensive. This may explain why, in recent

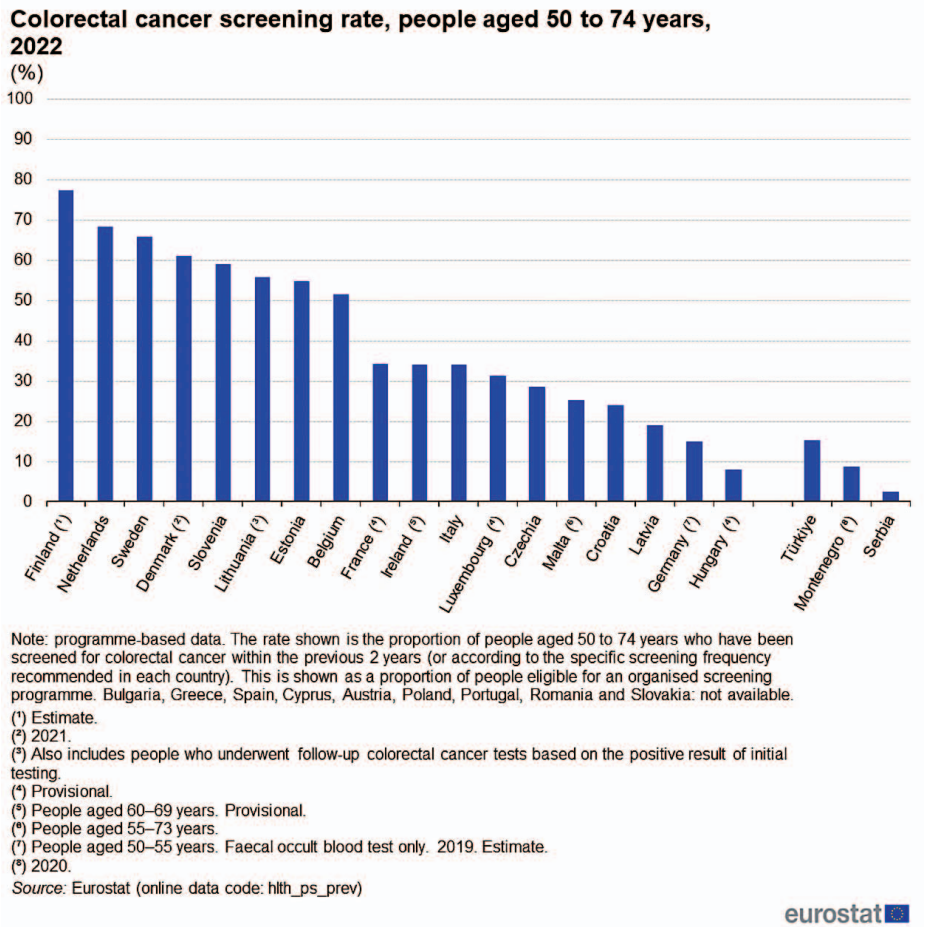

Figure 2. Graphic representation of the turnout for CRC screening by European countries(19).

years, a significant number of research studies have addressed the causes of non-uniformity in cancer screening across EU member states. Identifying the predominant individual reasons for these disparities is considered the only effective way to improve the current situation. Due to Croatia's high mortality rate from late-stage CRC, any intervention leading to an increased response to screening could also enhance chances for an earlier diagnosis and better patient outcomes.

Significant differences in screening turnout were observed based on the degree of urbanization. In urban areas of the Republic of Croatia, approximately 55% of the population aged 50 to 74 reported that they had never been screened for colorectal cancer, whereas in rural areas, this percentage increases to over 70%. Furthermore, a disparity in turnout was noted depending on the educational status of the population. Nearly 60% of individuals aged 50 to 74 with a lower level of education have never participated in screening. In comparison, the percentage of non-attendance drops below 50% for those with a higher level of education. Additionally, the male mortality rate from colorectal cancer is more than twice that of females(20).

Published studies suggest that the implementation of FIT could lead to higher turnout, increased detection of patients in the early stages (thereby improving survival rates), and financial savings. In 2008, Dutch researchers published the first randomized comparison between gFOBT and FIT in an average-risk population. The turnout rate for gFOBT was 47%, while for FIT it was 60%. The detection of advanced adenomas and carcinomas was significantly higher in the group tested with FIT, with 2.5 times more advanced adenomas and carcinomas and 2.2 times more carcinomas detected

compared to gFOBT(21). In a five-year pilot study conducted from 2009 to 2014, they confirmed a 20% higher turnout with FIT compared to gFOBT. Additionally, after one round of FIT screening, they found that a hypothetical individual would gain an average of 0.003 life years and save the health system €27 compared to gFOBT, and 0.003 life years and €72 compared to no screening(22). Since the start of screening in 2014 in the Netherlands, there has been an increase in the detection of stage I colorectal cancer, rising from 18% to over 30%(23). A 2015 Italian study reported that FIT-based screening programs were associated with a significant reduction in colorectal cancer mortality (by 22%) and that this effect occurred much earlier than with gFOBT-based programs(24).

Screening using FIT requires stool sampling for a maximum of two consecutive days. A 2013 Dutch study demonstrated that screening intensification with a single-sample FIT yielded equal or higher survival rates at lower costs compared to the two-sample FIT. Consequently, the researchers suggested conducting more rounds of screening with one FIT sample instead of screening every two years with two FIT samples(25). In 2020, Brazilian scientists examined the diagnostic accuracy of tests using one or two samples collected on consecutive days. They concluded that collecting two samples improves the detection rate of advanced adenomas, which are preneoplastic conditions that can lead to CRC(26). It is logical to adhere to the recommendations of manufacturers who suggest using multiple testing methods (2-3 days), primarily for statistical reasons. Since bleeding is often intermittent, analyzing more than one sample reduces the likelihood of failing to detect patients with a positive finding of occult blood in their stool. Additionally, due to the significant variation in fecal hemoglobin concentration observed between samples, many advanced adenomas were missed when only one stool sample was tested, mainly when a higher cutoff for test positivity was applied. Findings from an Australian study indicated that a one-sample test was not necessarily associated with a lower burden of a colonoscopy if a two-sample test had an independently set cutoff that provided comparable sensitivity. If the detection of advanced adenomas is a critical goal of the screening program, and if a colonoscopy burden of over 5% is feasible, the two-specimen test requires fewer colonoscopies for equivalent sensitivity(27).

Every year a significant number of colorectal cancer cases (one-third of the total cases in Croatia) are diagnosed and treated at the University Hospital for Tumors, Sestre milosrdnice University Hospital Center (Zagreb, Croatia) which underscores the justified need for a screening method for occult bleeding. The optimal method for this purpose is FIT, as it allows automated and quantitative measurement of analytes, resulting in fast and reliable laboratory test results. The purpose of thorough verification is to ensure that test results reliably reflect the patient's condition. It confirms that any changes in parameters across a series of measurements accurately represent the patient's condition, rather than being due to instrument bias or imprecision. Verification also ensures that patients do not move between diagnostic groups due to bias and that changes observed in a series of measurements can be clearly distinguished from natural biological variability.

By verifying FIT, we found that the method is precise and has an acceptable bias. However, even though the repeatability at level 3 of the control sample (target value 73.9 ng/mL) was not analytically satisfied, it is not clinically significant considering that these values (ranging from 72 to 78 ng/mL) are below the cut-off value of 100 ng/mL. Young et al. investigated the effect of varying the number of samples (one sample (1-FIT) vs. two samples (2-FIT)) and cut-off values on test sensitivity and colonoscopy workload. For cut-offs ranging from 50 to 400 ng/mL, the sensitivity of 2-FIT for colorectal cancer decreased from 81% to 62%, with an even steeper decline for advanced adenomas, which dropped from 59% to 16%. The rates for 1-FIT declined in parallel. The rate of increase in sensitivity slowed for both 1-FIT and 2-FIT with cut-offs of 100 ng/mL and below. In contrast, at cut-offs of 100 ng/mL and lower, the colonoscopy workload continued to increase even more rapidly, indicating a loss of efficiency in detecting lesions(28).

When a sample with an extremely high concentration precedes an extremely low concentration of the same analyte, it is reasonable to doubt whether the concentration in the low-analyte sample will be falsely increased due to carryover from the high-analyte sample. By calculating the carryover, we demonstrated that the transfer of analytes from the previous sample is negligible.

The results of the Canadian study show that FIT samples exhibit a reduced hemoglobin concentration compared to the baseline value when stored at room temperature. In contrast, storage at refrigerator temperature improved the stability of the FIT samples, although it did not completely prevent the decrease in hemoglobin concentration(29). Our findings demonstrate that hemoglobin degradation was most significant at elevated temperatures (+35°C), with the highest deviation observed after 10 days. Samples stored at 2–8°C exhibited the least variation, suggesting this as the optimal storage condition for preserving hemoglobin stability over time. Taking the acceptance criterion of 10% into account, the samples analyzed in our laboratory demonstrated stability for up to 10 days when stored at room temperature and at a temperature of 2-8°C. However, samples stored at +35°C exceeded the 10% deviation from the initial value after just 72 hours. Marginal hemoglobin concentrations around the cut-off value (122 ng/mL in the room temperature storage group; 109 ng/mL at 2-8°C; and 111 ng/mL at +35°C) exhibited 100% positivity even after 10 days of storage.

In comparing FIT with gFOBT, meta-analyses have shown that FIT demonstrates improved sensitivity in detecting colorectal cancer and advanced neoplasia while maintaining the same specificity(30). A higher sensitivity for colorectal cancer can be achieved with FIT than with gFOBT. A single test with a positivity threshold of 100 ng/mL has been reported to have a sensitivity for colorectal cancer of 87.1%–92.3%, compared to 30.8%–74.2% for traditional gFOBT(31). Our comparison of FIT and gFOBT established a significantly higher sensitivity for FIT; in 38% of the examined samples, gFOBT yielded a negative result while the hemoglobin concentration measured by the FIT method was above 100 ng/mL. Additionally, scientists from the United Kingdom found the sensitivity of FIT to be 92% in a five-year review involving 450 patients diagnosed with colorectal cancer, with 36 results being negative (using a 50 ng/mL cut-off value)(32). Thailand meta-analysis showed sensitivity of OC SENSOR 72,54% (95%CI: 65,82 – 79,25) and specificity of OC SENSOR 89,59% (95%CI: 87,23 – 91,95). The results demonstrated a significant log-linear relationship between FIT concentration and the positive predictive value for predicting colorectal tumors ($R^2 > 0.95$, $p < 0.001$). These findings suggest that higher FIT concentrations are associated with more advanced histological grades. Risk prediction for colorectal neoplasia based on individual FIT concentrations is noteworthy and could enhance the effectiveness of screening programs(33).

The limitations of this study are that stool samples could not be used in certain components of verification. Instead of stool samples with blood, oatmeal was homogenized with hemolyzed serum to obtain a wide range of hemoglobin concentrations in the samples. In the comparison study between FIT and gFOBT, it would have been useful to include colonoscopy findings. Nevertheless, these limitations provide an idea for the next research on the diagnostic sensitivity and specificity of FIT with the patient data.

CONCLUSION

FIT is a non-invasive test that helps detect serious bowel diseases in asymptomatic patients. Due to limited resources in endoscopy and access to healthcare facilities, FIT is a valuable marker for informed decision-making in clinical management. It enables the rapid identification of patients at the highest risk, thereby reducing the number of unnecessary colonoscopies. FIT is a well-established method with fully defined prerequisites for pre-analytical, analytical, and post-analytical accreditation. A consensus has been reached regarding the standardization of the technique, including the units for expressing results, reporting models, hemoglobin stability, and internal and external analytical quality controls.

Additionally, reputable external quality control providers provide a system of independent external quality assessment. Considering the recommendations of the EU Council that mandate the use of FIT, along with the interruption of gFOBT supplies in the market, there is no choice but to implement FIT as a screening method for colorectal cancer. We plan to assess the impact of FIT implementation on CRC incidence in future studies. Based on the experiences of other countries, we can anticipate better turnout, which is crucial in our country to reduce the mortality rate. It is reasonable to assume that this, combined with time, will favor outcomes in a desirable direction for the Republic of Croatia.

REFERENCE

1. ECIS – European Cancer Information System [Internet] Estimates of cancer incidence and mortality in 2022, for all countries [cited 1.10.2024]. Available from: https://ecis.jrc.ec.europa.eu/explorer.php?$0-0$1-All$4-1,2$3-16$6-0,85$5-2022,2022$7-7,8$2-All$CEstByCountry$X0_8-3$X0_19-AE27$X0_20-No$CEstBySexByCountry$X1_8-3$X1_19-AE27$X1_-1-1$CEstByIndiByCountry$X2_8-3$X2_19-AE27$X2_20-No$CEstRelative$X3_8-3$X3_9-AE27$X3_19-AE27$CEstByCountryTable$X4_19-AE27.
2. Nacionalni program ranog otkrivanja raka debelog crijeva, Republika Hrvatska, Ministarstvo zdrastva i socijalne skrbi, Zagreb 2007.
3. Lin JS, Perdue LA, Henrikson NB, Bean SI, Blasi PR. Screening for Colorectal Cancer: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA. 2021 May 18;325(19):1978-1998. doi: 10.1001/jama.2021.4417.
4. 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 Apr 7;100(7):1103-10. doi: 10.1038/sj.bjc.6604961.
5. Duffy MJ, van Rossum LG, van Turenhout ST, Malminiemi O, Sturgeon C, Lamerz R et al. Use of faecal markers in screening for colorectal neoplasia: a European group on tumor markers position paper. Int J Cancer. 2011 Jan 1;128(1):3-11. doi: 10.1002/ijc.25654.
6. Provenzale D, Gupta S, Ahnen DJ, Markowitz AJ, Chung DC, Mayer RJ et al. NCCN Guidelines insights: colorectal cancer screening, version 1.2018. J Natl Compr Canc Netw. 2018;16(8):939-949. doi: 10.6004/jnccn.2018.0067.
7. Wolf AMD, Fontham ETH, Church TR, Flowers CR, Guerra CE, LaMonte SJ, et al. Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society. CA Cancer J Clin. 2018 Jul;68(4):250-281. doi: 10.3322/caac.21457.
8. Lopes G, Stern MC, Temin S, Sharara AI, Cervantes A, Costas-Chavarri A et al. Early detection for colorectal cancer: ASCO resource-stratified guideline. J Glob Oncol. 2019 Feb;5:1-22. doi: 10.1200/JGO.18.00213.
9. Bresalier RS, Senore C, Young GP, Allison J, Benamouzig R, Benton S et al. Members of the World Endoscopy Colorectal Cancer Screening New Test Evaluation Expert Working Group. An efficient strategy for evaluating new non-invasive screening tests for colorectal cancer: the guiding principles. Gut. 2023 Oct;72(10):1904-1918. doi: 10.1136/gutjnl-2023-329701.
10. Lin JS, Perdue LA, Henrikson NB, Bean SI, Blasi PR. Screening for colorectal cancer: updated evidence report and systematic review for the US preventive services task force. JAMA. 2021 May 18;325(19):1978-1998. doi: 10.1001/jama.2021.4417.
11. Wilt TJ, Harris RP, Qaseem A. High value care task force of the American College of Physicians. Screening for cancer: advice for high-value care from the American College of Physicians. Ann Intern Med. 2015 May 19;162(10):718-25. doi: 10.7326/M14-2326.
12. Qaseem A, Harrod CS, Crandall CJ, Wilt TJ; Clinical Guidelines Committee of the American College of Physicians. Screening for Colorectal Cancer in Asymptomatic Average-Risk Adults: A guidance statement from the American College of Physicians (Version 2). Ann Intern Med. 2023 Aug;176(8):1092-1100. doi: 10.7326/M23-0779.
13. Shaukat A, Kahi CJ, Burke CA, Rabeneck L, Sauer BG, Rex DK. ACG Clinical guidelines: colorectal cancer screening 2021. Am J Gastroenterol. 2021 Mar 1;116(3):458-479. doi: 10.14309/ajg.0000000000001122.
14. European Commission [Internet]. European Health Union: Commission welcomes adoption of new EU cancer screening recommendations [cited 2.10.2024]. Available from: https://ec.europa.eu/commission/presscorner/detail/en/ip_22_7548.
15. Eiken Chemical Co., LTD. [Internet]. OC-SENSOR Ceres [cited 2.10.2024]. Available from: https://www.eiken.co.jp/en/products/fit/ceres/.
16. Gómez-Rioja R, Segovia Amaro M, Diaz-Garzón J, Bauçà JM, Martínez Espartosa D, Fernández-Calle P; Extra-Analytical Quality Commission of the Spanish Society of Laboratory Medicine (SEQCML); Analytical Quality Commission of the Spanish Society of Laboratory Medicine (SEQCML). A protocol for testing the stability of biochemical analytes. Technical document. Clin Chem Lab Med. 2019 Nov 26;57(12):1829-1836. doi: 10.1515/cclm-2019-0586.
17. Zavod za javno zdrastvo Dubrovačko-neretvanske županije [Internet]. Nacionalni program ranog otkrivanja raka debelog crijeva [cited 7.10.2024]. Available from: https://www.zzjzdnz.hr/projekti/nacionalni-program-ranog-otkrivanja-raka-debelog-crijeva.
18. Ola I, Cardoso R, Hoffmeister M, Brenner H. Utilization of colorectal cancer screening tests across European countries: a cross-sectional analysis of the European health interview survey 2018-2020. Lancet Reg Health Eur. 2024 Apr 29;41:100920. doi: 10.1016/j.lanepe.2024.100920.
19. Eurostat Statistics Explained [Internet]. Cancer screening statistics. Colorectal cancer screening rate, people aged 50 to 74 years, 2022 [figure downloaded 7.10.2024]. Available from: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Cancer_screening_statistics#Colorectal_cancer_screening.
20. ECIR – European Cancer Inequalities Registry [Internet]. Uncovering Inequalities Colorectal Cancer Screening in Europe [cited 4.10.2024]. Available from: https://cancer-inequalities.jrc.ec.europa.eu/inequalities-factsheets.
21. Allison JE, Fraser CG, Halloran SP, Young GP. 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 Mar;8(2):117-30. doi: 10.5009/gnl.2014.8.2.117.
22. Van Rossum LG, van Rijn AF, Verbeek AL, van Oijen MG, Laheij RJ, Fockens P et al. Colorectal cancer screening comparing no screening, immunochemical and guaiac fecal occult blood tests: a cost-effectiveness analysis. Int J Cancer. 2011 Apr 15;128(8):1908-17. doi: 10.1002/ijc.25530.
23. Integraal Kankercentrum Nederland (IKNL) [Internet]. Trends in stadiumverdeling dikkedarm– en endeldarmkanker leeftijd 55-75 jaar [cited 17.10.2024]. Available from: https://iknl.nl/kankersoorten/darmkanker/registratie.
24. Zorzi M, Fedeli U, Schievano E, Bovo E, Guzzinati S, Baracco S et al. Impact on colorectal cancer mortality of screening programmes based on the faecal immunochemical test. Gut. 2015 May;64(5):784-90. doi: 10.1136/gutjnl-2014-307508.
25. Goede SL, van Roon AH, Reijerink JC, van Vuuren AJ, Lansdorp-Vogelaar I, Habbema JD et al. Cost-effectiveness of one versus two sample faecal immunochemical testing for colorectal cancer screening. Gut. 2013 May;62(5):727-34. doi: 10.1136/gutjnl-2011-301917.
26. Mattar R, Marques SB, Minata MK, Silva-Etto JMKD, Sakai P, DE Moura EGH. Diagnostic accuracy of one sample or two samples quantitative fecal immunochemical tests for intestinal neoplasia detection. Arq Gastroenterol. 2020 Jul-Sep;57(3):316-322. doi: 10.1590/S0004-2803.202000000-58.
27. Young GP, Woodman RJ, Symonds E. Detection of advanced colorectal neoplasia and relative colonoscopy workloads using quantitative faecal immunochemical tests: an observational study exploring the effects of simultaneous adjustment of both sample number and test positivity threshold. BMJ Open Gastroenterology 2020;7:e000517. doi: 10.1136/bmjgast-2020-000517.
28. Young GP, Woodman RJ, Ang FLI, Symonds EL. Both sample number and test positivity threshold determine colonoscopy efficiency in detection of colorectal cancer with quantitative fecal immunochemical tests. Gastroenterology. 2020 Oct;159(4):1561-1563.e3. doi: 10.1053/j.gastro.2020.05.008.
29. Reid MS, Paul HA, Mostoufi A, Robinson JL, Sadrzadeh SMH. Evaluation of the stability of fecal immunochemical test specimens. Clin Biochem. 2023 May;115:92-96. doi: 10.1016/j.clinbiochem.2022.11.015.
30. Robertson DJ, Lee JK, Boland CR, Dominitz JA, Giardiello FM, Johnson DA et al. Recommendations on fecal immunochemical testing to screen for colorectal neoplasia: a consensus statement by the US multi-society task force on colorectal cancer. Gastroenterology. 2017 Apr;152(5):1217-1237.e3. doi: 10.1053/j.gastro.2016.08.053.
31. Young GP, Symonds EL, Allison JE, Cole SR, Fraser CG, Halloran SP et al. Advances in fecal occult blood tests: the FIT revolution. Dig Dis Sci. 2015 Mar;60(3):609-22. doi: 10.1007/s10620-014-3445-3.
32. Cole E, Narayanan D, Tiam RN, Shepherd J, Hajjawi MOR. Faecal immunochemical test (FIT) sensitivity : a five year audit. Br J Biomed Sci. 2024 May 29;81:12862. doi: 10.3389/bjbs.2024.12862.
33. Phuangrach N, Sarakarn P. the accuracy of fecal immunochemical test in colorectal cancer screening: a meta-analysis. Asian Pac J Cancer Prev. 2022 Mar 1;23(3):759-766. doi: 10.31557/APJCP.2022.23.3.759.

Sažetak

VERIFIKACIJA FEKALNOG IMUNOKEMIJSKOG TESTA NA OC-SENSOR CERES UREĐAJU

A. Šoić, S. Langer, Z. Špacir Prskalo, M.Gaće, M.Vrbančić, I.Sever, Lj.Mayer

Organizirani programi probira za otkrivanje kolorektalnog karcinoma u Republici Hrvatskoj uključuju testiranje na okultno krvarenje u stolici. Tradicionalno se koristi gvajak metoda (gFOBT, od engl. *guaiac fecal occult blood test*). Još jedna dobro uspostavljena metoda je fekalni imunokemijski test (FIT, od engl. *fecal immunochemical test*). U laboratoriju Odjela za medicinsku biokemiju u onkologiji Klinike za tumore KBC Sestre Milosrdnice provedena je verifikacija FIT-a na OC-SENSOR Ceres analizatoru prema CLSI smjernicama. Provjereni su parametri preciznosti, istinitosti, prijenosa analita, stabilnosti analita u uzorku i usporedba s gFOBT-om. Preciznost je prikazana izračunatom proširenom relativnom mjerom nesigurnosti i iznosila je 2,54%, 1,24%, 5,38%. Točnost je izračunata i prikazana kao odstupanje (2,1%, 0,8% i 1,3%). Utvrđena je stabilnost hemoglobina u uzorku do 10 dana na 2 – 8°C (4,95%) te čak i na sobnoj temperaturi (9,57%). Dobiveni rezultat za prijenos analita iznosio je -0,43%. Usporedba gFOBT-a i FIT-a pokazala je da gFOBT ima veći broj lažno negativnih rezultata (38%). S obzirom na sve prednosti FIT-a u odnosu na gFOBT i uspješnu verifikaciju metode, opravdano je zamijeniti gFOBT FIT-om.

KLJUČNE RIJEČI: *fekalni imunokemijski test (FIT); verifikacija; probir na kolorektalni karcinom; stabilnost; gvajak test na okultno krvarenje u stolici (gFOBT)*