Title
Key indicators of organized cancer screening programs: results from a Delphi study

Short title:
Key indicators of organized cancer screening programs

List of authors

Marcell Csanádi
Syreon Research Institute, Budapest, Hungary marcell.csanadi@syreon.eu

Inge de Kok
Erasmus MC, Rotterdam, the Netherlands i.dekok@erasmusmc.nl

Eveline Heijnsdijk
Erasmus MC, Rotterdam, the Netherlands e.heijnsdijk@erasmusmc.nl

Ahti Anttila
Finnish Cancer Registry, Helsinki, Finland ahti.anttila@cancer.fi

Sirpa Heinävaara
Finnish Cancer Registry, Helsinki, Finland sirpa.heinavaara@cancer.fi

János G. Pitter
Syreon Research Institute, Budapest, Hungary janos.pitter@syreon.eu

György Széles
Syreon Research Institute, Budapest, Hungary gyorgy.szeles@syreon.eu

Urška Ivanuš
Institute of Oncology Ljubljana, Ljubljana, Slovenia uivanus@onko-i.si

Jennifer Priaulx
London School of Hygiene & Tropical Medicine, London, United Kingdom jennifer.priaulx@lshtm.ac.uk

Piret Veerus
National Institute for Health Development, Tallinn, Estonia piret.veerus@tai.ee

Carlo Senore
CPO Piemonte and University Hospital "Città della Salute e della Scienza", Turin, Italy carlo.senore@cpo.it

Harry de Koning
Erasmus MC, Rotterdam, the Netherlands h.dekoning@erasmusmc.nl

Zoltán Vokó
Syreon Research Institute, Budapest, Hungary zoltan.voko@syreon.eu

Corresponding author:
Marcell Csanádi
Email: marcell.csanadi@syreon.eu
Postal address: Hungary 1142 Budapest, Mexikoi str. 65/A
Phone: +36 30 6384894

Abstract

To maximize benefits and reduce potential harms of organized cancer screening programs in Europe, monitoring, quality assurance and evaluation of the long-term impact are required. We aimed to identify the most important indicators that are required to collect and report. The study was designed to establish a consensus within a European-level working group and come up with a manageable list of key indicators. We conducted a Delphi study among policymakers, researchers and program coordinators who were experts in breast, cervical or colorectal cancer screening. Study participants evaluated the importance of screening indicators on a 5-point Likert-scale. Interval cancer rate, detection rate, screening attendance, screening coverage, cancer incidence, cause-specific mortality, proportion of persons attending further assessment after a positive screen test result, proportion of persons attending a treatment after diagnosis, invitation coverage, and distribution of cancers by mode of detection were the top 10 indicators by study participants. In general, performance indicators were considered more important than outcome indicators. Sub-group analyses by cancer types showed similar results and only experts of cervical cancer screening had slightly different preferences. Sub-group analyses by experts' roles indicated that policymakers found different indicators important compared to researchers or program coordinators, probably because of their different point of view on screening. The implication of our priority ranking is twofold: it serves as an initial guidance for countries that have not yet established a system to collect data; and as a checklist for those where data collection is already established to assess the comprehensiveness of their system.

Keywords

organized screening program; breast cancer screening; cervical cancer screening; colorectal cancer screening; Delphi study; indicators of screening; European screening experts

Acknowledgements

This study was part of the EU-TOPIA project, funded by the EU-Framework Programme (Horizon 2020) of the European Commission, project reference 634753. The content of this publication reflects only the EU-TOPIA groups' views and the European Commission is not liable for any use that may be made of the information contained herein. We would like to acknowledge all members of the EU-TOPIA consortium for their support in the development of the project and contributions to this publication.

Conflict of interest

The authors of this manuscript have no conflict of interest to be declared.

Introduction

The Council of the European Union (EU) recommended in 2003 that screening for breast, cervical and colorectal cancer should take place only in a form of organized, population-based programs. It also specified the fundamental principles of early detection of cancer and presented a shared commitment by Member States to implement organized screening programs [1]. The 4th Edition of the European Code against Cancer states that organized screening programs are preferred to opportunistic screening as these provide better conditions for quality assurance to achieve the greatest benefit with the least harm [2]. However, there is still a need for better monitoring, quality assurance and evaluation of the long-term impact of organized cancer screening in Europe [3,4,5]. Quality assurance and continuous evaluation are among the key principles of implementing screening programs [6] and the programs' effectiveness highly depends on these procedures [7]. Still, a study conducted among European countries showed that efforts for these activities are carried out to a differing and non-standardized extent, which makes it difficult to compare long-term effects of screening programs [8].

There have been several attempts to provide guidance on collecting data for quality assurance and monitoring screening programs. Indicators on quality and performance have already been defined by the European quality assurance guidelines for breast, cervical and colorectal cancer [9,10,11,12,13]. These provide an overview of the fundamental points and principles that supports quality assurance and identifies performance indicators to collect. The second report on Cancer Screening in the EU published in 2017, aims to provide the most up-to-date information on the implementation of Council recommendations [14]. In this report, over 80 experts from EU member states provided data on the implementation status and on the performance indicators for breast, cervical and colorectal cancer screening [14]. Similarly to the first report on Cancer Screening in the EU [15], it also provided justification for initiatives at European level and in the member states to expand and improve implementation of population-based programs. It also revealed that the data collection for more complex performance indicators was limited in several countries due to the lack of continuous and systematic evaluation in place. Several other EU initiatives also generated knowledge and evidence for measuring and evaluating the performance of screening programs such as the EUROCOURSE [16], the EUNICE [17] and the EUROSCREEN [18].

These previous initiatives recommend to collect many indicators including more complex ones for monitoring quality and performance. However, some countries with less of a tradition or experience in screening and with fewer resources or without up-to-date registries are unlikely to collect them all. This makes European level comparison on the effect of different programs more challenging. Therefore, a key list of indicators is needed which is relatively short and developed by a consensus among European experts. This study aimed to identify the most important indicators to collect for measuring the impact of organized screening programs. We also investigated the key indicators across different cancer types and took into account different roles of experts in the screening programs.

This study was conducted within the EU-TOPIA (Towards improved screening for breast, cervical and colorectal cancer in all of Europe) H2020 project that aims to systematically evaluate and quantify the harms and benefits of screening programs for breast, cervical, and colorectal cancer in all European countries[1]. It also intends to build capacity for future self-evaluation of screening programs in Europe, for which providing a list of key indicators is important.

[1] http://eu-topia.org/about-eu-topia/about-eu-topia/

Methodology

We developed a set of key indicators to monitor the effects of screening programs in a Delphi study with the involvement of screening experts from Europe who participated at the EU-TOPIA workshop about screening indicators in September 2017 in Budapest, Hungary[2]. Participants were either researchers, screening program coordinators or policymakers related to breast, cervical or colorectal cancer screening programs. Experts were asked to specify their role(s) in the screening program and the cancer site(s) of their expertise. All participants were allowed to select more than one role and more than one cancer site. The Delphi process was carried out according to the steps described in the literature [19,20] with the adaptations detailed below. The current study consisted of three consecutive rounds.

Round 1 was a preparatory phase, where a literature review was carried out to identify an extensive pool of indicators for evaluating organized cancer screening programs.[3] In Delphi studies it is both an acceptable and a common modification to use a structured questionnaire, which is based on a literature review in Round 1 [20]. Based on the review 27 indicators that can be used to estimate the long-term effectiveness of cancer screening programs were defined for this study (Table 1).

In Round 2, a 27-item online questionnaire was sent to EU-TOPIA workshop participants in advance. Participants had to individually evaluate the importance of each listed indicator on a 5-point Likert-scale. The following values were attributed to the scale: strongly disagree: -2; disagree: -1; neutral: 0; agree: 1; strongly agree: 2. Round 2 was completed in August-September 2017. After the data collection, a summary of Round 2 was sent to participants showing aggregated results for each indicator, which were presented on a bar chart showing the frequency of each option on the scale.

In Round 3, results of Round 2 were presented at the EU-TOPIA workshop and discussed by a roundtable panel, which gave all study participants in the audience an opportunity to refine their opinion before final voting round. A mobile phone application was used for voting at the workshop. Participants were also asked whether they filled in the online survey in Round 2. Therefore, a sub-group of Round 3 answers was possible to be analysed among those who participated in both rounds.

To develop the priority list, indicators were ranked according to the mean value of survey answers. Indicators, which many participants strongly supported had high mean value because of the abovementioned coding of answers. The symmetrical values on the scale ensured that “strongly disagree” has the same negative weight as positive weight “strongly agree” had. This holds for the categories of “disagree” and “agree”, as well. To characterize the level of consensus related to each indicator a previously published consensus measurement was used [21,22]. This measure is typically applied to the Likert-scale to determine degrees of agreement among ordinal-ranked categories. The value of consensus ranges from 0 to 1, the latter meaning perfect consensus. If two indicators had the same mean value, then the one with larger consensus measure score got higher rank on the list. In the current paper the top 10 indicators from the full list are considered as key indicators.

<Table 1 about here>

[2] http://eu-topia.org/workshops/workshop-1-monitoring/

[3] EU-TOPIA Deliverable 2.2 - Definition of indicators (available at: http://eu-topia.org/downloads/)

Results

A total of 143 and 73 respondents filled in the survey in Round 2 and Round 3, respectively. For the further analysis only those who completed all questions were included, resulting in 89 and 61 participants, respectively. Respondents represented 31 European countries. In Round 2 none of the listed indicators had a negative mean value, while in Round 3 one indicator had a mean value below 0, indicating the lowest importance (Table 1).

The list of top 10 indicators for Round 3 was identical to Round 2 (hereafter TOP10), only the order of the indicators changed (Table 2). The TOP10 indicators were *interval cancer rate*, *detection rate*, *screening attendance*, *screening coverage*, *cancer incidence*, *cause-specific mortality*, *proportion of persons attending a further assessment after positive screen test result*, *proportion of persons attending a treatment after diagnosis*, *invitation coverage*, and *distribution of cancers by mode of detection*. Results also show that indicators measuring the performance of the screening programs (e.g. interval cancer rate, detection rate, screening attendance or coverage) were graded higher than indicators measuring the outcomes of the programs (e.g. cause specific mortality).

In Round 2 there was a relatively minor difference between the 10th and the 11th ranked indicators (mean scores 1.247 vs. 1.236, respectively) but a larger difference was observed in Round 3 (mean scores 1.262 vs. 1.066, respectively) (Table 1). The average consensus score considering all indicators also increased in Round 3 (0.711) compared to Round 2 (0.701).

The first sub-group analysis was conducted among Round 3 participants who also filled in the Round 2 online survey (n=49). The list of top 10 indicators in this sub-group was also identical to the TOP10, only the order was different (Table 2). Similarly to the overall results of Round 3, we observed a relatively large difference between the 10th and the 11th ranked indicators (mean scores 1.265 vs. 1.061, respectively). The average consensus score considering all indicators was similar (0.710) to the overall result of Round 3 (0.711).

<Table 2 about here>

The second sub-group analysis was conducted according to the participants' fields of expertise (breast, cervical or colorectal). Of the workshop participants, 42 indicated breast, 29 indicated cervical and 33 colorectal cancer as a field of expertise. The top 10 indicators were identical across breast and colorectal cancer experts, only the order was different (Table 3). Their list was also identical with the TOP10. Cervical cancer experts included two indicators which were not included for breast and colorectal expects (namely, complications related to the screening test; and complications related to referral examinations). The two indicators replaced by these items were invitation coverage, and compliance with treatment (marked in Table 3 with hyphen).

<Table 3 about here>

The third sub-group analysis was conducted according to the participants' roles in cancer screening programs. Of the workshop participants 36 indicated researcher, 29 program coordinator and 10 policymaker role. The top 10 indicators were identical for researchers and program coordinators, only the order was different (Table 4). Their list was also identical with the TOP10. Policymakers included five indicators, which were not included by researchers and program coordinators (proportion of persons falsely referred for further assessment; complications related to the screening test; complications related

to referral examinations; proportion of cancer patients treated with any treatment; probability that a person with a positive screen test result has cancer (Table 4).

<Table 4 about here>

Discussion

Despite the considerable progress on the implementation and organization of cancer screening programs in the EU member states since 2003, much work remains to be done and procedures for continued monitoring, regular feedback and periodic reporting should be ensured [23]. Our study adds to this recommendation by providing a consensus of screening experts on the most important indicators that are required to evaluate organized screening programs. Previous initiatives monitoring and quality assurance have not attempted to prioritize among organized screening program indicators.

We found a strong consensus among European screening experts on the key indicators of organized screening programs. In the third round of the Delphi panel we observed a larger difference between the 10th and the 11th ranked indicators compared to the previous round and also we observed an increase in the average consensus value. This indicates that workshop participants had a better consensus during the face-to-face meeting towards defining a list of key indicators.

Screening experts had a strong consensus that interval cancer rate, screening attendance, screening coverage are the most important indicators of screening programs. Interestingly, they ranked performance indicators higher than outcome indicators, probably due to the closer temporal and causal association of screening efforts with performance indicators than with outcome indicators. On the other hand, as effective mobilization of the population, good organization of the screening programs and good quality of the screening tests are all prerequisites of health benefits, thus this ranking is not that surprising.

Although indicators on screening benefits were generally graded with higher priority, indicators on potential harms should not be neglected. There is still an ongoing debate over the benefits and harms of screening programs [24]. Therefore, when countries evaluate their programs data collection should include for instance overdiagnosis since it causes overestimations of the sensitivity, specificity, and positive predictive value of screening tests and the incidence of disease [25].

Sub-group analyses by cancer sites and roles in the screening program showed similar results, except for experts in cervical cancer and for policymakers. Experts in cervical cancer ranked indicators on complications related to the screening test and to referral examinations higher compared to the other two cancer sites. This indicates that the quality assurance of these examinations and monitoring of side effects should be prioritized even though complications for the cervical screening itself is low and it can be conducted safely even by trained health visitors [26].

We also found that policymakers in general have different priorities compared to other stakeholders, which may have different reasons. First, our sample size (n=10) was low compared to researchers (n=36) and program coordinators (n=29). Therefore, data on policymakers were less robust to define the top 10 indicators. Second, policymakers tended to put more emphasis on the potential harms of screening activities (i.e. proportion of persons falsely referred for further assessment; complications related to the screening test; complications related to referral examinations) since these are politically more sensitive issues. In general, they seem to prioritize the close monitoring of these undesired events that could place screening into a negative context. However, these data are often not routinely available from the screening databases.

The other two stakeholder groups had the same top 10 indicators, but the individual rankings of the indicators were somewhat different. Most notably cause-specific mortality was ranked as the second most important indicator for researchers and it was only the 7th most important for program coordinators.

Giving higher priority to performance indicators (e.g. detection rate, interval cancer rate, screening coverage or screening attendance) was even more noticeable in case of program coordinators compared to the other two stakeholder groups. We do believe, that this is because these indicators are more closely relate to their responsibilities in the screening programs.

It is important to note that due to the specific approach of early detection with screening, the evaluation of long-term outcomes is not feasible for many years after program implementation. Still, indicators on the impact of screening should be closely and systematically monitored to give an early indication of whether the program could lead to the expected benefits. Based on this approach, health economic modeling could be applied to estimate long-term outcomes and the cost-effectiveness of a program, which has been illustrated in previous examples with even considerable influence on screening policy [27,28,29,30].

This study has some limitations due to its design. First, Delphi studies are usually carried out in a smaller number of experts from the same field. However, representation should be assessed by the qualities of the expert panel rather than its numbers [31]. Since we aimed to achieve a consensus among European screening experts our sample size included experts in three cancer sites from relevant institutes of cancer screening throughout 31 European countries. Second, although we attempted to cover as wide range of indicators as possible we had to limit these to a manageable list in order to perform the Delphi exercise. Therefore, there might be some aspects of organized screening programs that were not captured in the analysis. Nonetheless, our initial list is based on a comprehensive review of the literature which was made publicly available by the EU-TOPIA project before the study.

Our study did not cover practical and feasibility issues of data collection or timely availability of relevant data. These issues however, should be handled by individual countries, according to their institutional settings (e.g. national registries, screening coordination organizations), legal environment on personal data (e.g. data linkage between organizations) and screening related know-how. Another issue might be the definition of more complex indicators such as the interval cancer [32]. Therefore, further efforts for standardizing the calculation of key indicators are required such as the example of NordScreen project [33].

Conclusion

This study was the first that attempted to provide a small set of key indicators in over three different cancer sites by relying on the opinion experts. We found a strong consensus among European screening experts on the key list of screening indicators. Sub-group analyses revealed some differences, more notably for policymakers who had slightly different perspective. The implication of our priority ranking depends on the countries' current practice of systematic data collection and regular monitoring. We recommend that countries without regular and systematic approach of monitoring should primarily design their system to collect at least the defined key indicators. Countries where a systematic approach for data collection is already in place, our priority ranking should be considered as a checklist by which monitoring procedures can be verified or if necessary further updated.

References

1. European Council Recommendation of 2 December 2003 on Cancer Screening (2003/878/EC). OJ L 327: 34-38 Available from: http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32003H0878 Accessed: 2018 July.
2. Armaroli P, Villain P, Suonio E, Almonte M, Anttila A, Atkin WS, Dean PB, de Koning HJ, Dillner L, Herrero R, Kuipers EJ, Lansdorp-Vogelaar I, Minozzi S, Paci E, Regula J, Törnberg S, Segnan N. European Code against Cancer, 4th Edition: Cancer screening. Cancer Epidemiol. 2015. 39 Suppl 1:S139-52.
3. Anttila A, von Karsa L, Aasmaa A, Fender M, Patnick J, Rebolj M, Nicula F, Vass L, Valerianova Z, Voti L, Sauvaget C, Ronco G. Cervical cancer screening policies and coverage in Europe. Eur J Cancer. 2009. 45(15):2649-58.
4. Giordano L, von Karsa L, Tomatis M. et al., Mammographic screening programmes in Europe: organization, coverage and participation. J Med Screen. 2012. 19 Suppl 1:72-82.
5. Anttila A, Lönnberg S, Ponti A, Suonio E, Villain P, Coebergh JW, von Karsa L. Towards better implementation of cancer screening in Europe through improved monitoring and evaluation and greater engagement of cancer registries. Eur J Cancer. 2015a. 51(2):241-51.
6. Lynge E, Törnberg S, von Karsa L, Segnan N, van Delden JJ. Determinants of successful implementation of population-based cancer screening programmes. Eur J Cancer. 2012. 48(5):743-8.
7. Rabeneck L, Lansdorp-Vogelaar I. Assessment of a cancer screening program. Best Practice & Research Clinical Gastroenterology. 2015. 29: 979-985
8. Elfström KM, Arnheim-Dahlström L, von Karsa L, Dillner J. Cervical cancer screening in Europe: Quality assurance and organisation of programmes. Eur J Cancer. 2015. 51(8):950-68.
9. Anttila, A., Arbyn, M., De Vuyst, H., et al. European guidelines for quality assurance in cervical cancer screening. Second edition, Supplements. Office for Official Publications of the European Union, Luxembourg. 2015b. 69-108.
10. Arbyn M, Anttila A, Jordan J, et al., editors. European Commission. In. European Guidelines for Quality Assurance in Cervical Cancer Screening. 2nd edition. Luxembourg: Office for Official Publications of the European Communities. 2008. pp. 1–291
11. Perry N, Broeders M, de Wolf C, Törnberg S, Holland R, von Karsa L. European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition--summary document. Ann Oncol. 2008. 19(4):614-22
12. Perry, N., Broeders, M., de Wolf, C., Törnberg, S., Holland, R., & von Karsa, L. *European guidelines for quality assurance in breast cancer screening and diagnosis. Fourth edition, supplements.* Luxembourg: European Commission, Publications Office of the European Union. 2013
13. Segnan N, Patnick J, von Karsa L. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First edition. Luxembourg: Publications Office of the European Union, 2010. Available from: http://www.kolorektum.cz/res/file/guidelines/CRC-screening-guidelines-EC-2011-02-03.pdf Accessed: 2018 July.
14. Ponti A, Anttila A, Ronco G et al., Cancer Screening in the European Union. Report on the implementation of the Council Recommendation on cancer screening (2017). Prepared by International Agency for Research on Cancer. Lyon, France
Available from:
https://ec.europa.eu/health/sites/health/files/major_chronic_diseases/docs/2017_cancerscreening_2ndreportimplementation_en.pdf Accessed: 2018 July.
15. Von Karsa L, Anttila A, Ronco G et al., Cancer Screening in the European Union Report on the implementation of the Council Recommendation on cancer screening. First Report. European Communities, Brussel, 2008

Available from: http://www.iccp-portal.org/sites/default/files/resources/cancer_screening.pdf Accessed: 2018 July.
16. Coebergh JW, van den Hurk C, Rosso S et al. EUROCOURSE lessons learned from and for population-based cancer registries in Europe and their programme owners: Improving performance by research programming for public health and clinical evaluation. Eur J Cancer. 2015. 51(9):997-1017.
17. Anttila A and Ronco G. Description of the national situation of cervical cancer screening in the member states of the European Union. Eur J Cancer. 2009. 45(15):2685-708.
18. Paci E, EUROSCREEN Working Group. Summary of the evidence of breast cancer service screening outcomes in Europe and first estimate of the benefit and harm balance sheet. J Med Screen. 2012;19 Suppl 1:5-13.
19. Jones J, Hunter D. Consensus methods for medical and health services research. BMJ. 1995. 5;311(7001):376-80
20. Hsu CC. The Delphi Technique: Making Sense Of Consensus. Practical Assessment, Research & Evaluation. 2007. 12 (10): 1-8.
21. Tastle WJ, Wierman MJ, Dumdum UR. RANKING ORDINAL SCALES USING THE CONSENSUS MEASURE. Issues in Information Systems. 2005. 6(2): 96-102.
22. Tastle WJ. Wierman MJ. An information theoretic measure for the evaluation of ordinal scale data. Behavior Research Methods 2006. 38 (3), 487-494
23. Basu P, Ponti A, Anttila A. Status of implementation and organization of cancer screening in The European Union Member States-Summary results from the second European screening report. Int J Cancer. 2018. 142(1):44-56.
24. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M, and The Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. Br J Cancer. 2013 Jun 11; 108(11): 2205–2240.
25. Black WC. Overdiagnosis: An underrecognized cause of confusion and harm in cancer screening. J Natl Cancer Inst. 2000. 92(16):1280-2.
26. Vajda R, Árváné Egri C, Kovács A, Budai A, Döbrõssy L, Koiss R, Kívés Z, Boncz I. [Quality and performance indicators of the pilot program for cervical cancer screening by health visitors]. Magy Onkol. 2017. 61(4):361-367.
27. Lansdorp-Vogelaar I, Knudsen AB, Brenner H. Cost-effectiveness of colorectal cancer screening. Epidemiol Rev 2011. 33:88-100.
28. de Kok IM, van Rosmalen J, Dillner J, Arbyn M, Sasieni P, Iftner T, van Ballegooijen M. Primary screening for human papillomavirus compared with cytology screening for cervical cancer in European settings: cost effectiveness analysis based on a Dutch microsimulation model. BMJ. 2012. 344:e670.
29. Vokó Z, Nagyjánosi L, Margitai B, Kövi R, Tóth Z, László D, Kaló Z. Modeling cost-effectiveness of cervical cancer screening in Hungary. Value in Health. 2012. 15:39-45.
30. van Ravesteyn NT, Stout NK, Schechter CB, Heijnsdijk EA, Alagoz O, Trentham-Dietz A, Mandelblatt JS, de Koning HJ. Benefits and harms of mammography screening after age 74 years: model estimates of overdiagnosis. J Natl Cancer Inst. 2015. 107(7).
31. Thangaratinam S, Redman CWE. The Delphi technique. The Obstetrician & Gynaecologist. 2005. 7:120–125
32. Sanduleanu S1, le Clercq CM1, Dekker E et al. Definition and taxonomy of interval colorectal cancers: a proposal for standardising nomenclature. Gut. 2015. 64(8):1257-67.
33. The Nordscreen group. Glossary of terms. Available from: http://nordscreen.org/glossary-of-terms/ Accessed: 2018 July.

List of tables

Table 1: Mean scores and consensus values for Round 2 and Round 3

List of indicators	Mean score in Round 2	Mean score in Round 3	Consensus in Round 2	Consensus in Round 3
1) The age distribution of the population.	**0.865**	**0.590**	**0.641**	**0.647**
2) The life-expectancy of population.	**0.775**	**0.311**	**0.641**	**0.624**
3) The number of newly diagnosed cancer cases (i.e. incidence) in a population.	**1.303**	**1.508**	**0.730**	**0.776**
4) The proportion of cancer patients treated with any treatment.	**0.573**	**1.033**	**0.624**	**0.761**
5) The (cumulative) relative survival of cancer patients in a given time frame.	**0.910**	**0.426**	**0.632**	**0.555**
6) The number of deaths from the given cancer (i.e. cause-specific mortality) in a population.	**1.393**	**1.574**	**0.699**	**0.790**
7) The proportion of screening tests outside organized screening (i.e. opportunistic screening).	**0.820**	**0.672**	**0.706**	**0.686**
8) The proportion of complications related to the screening test.	**0.978**	**0.951**	**0.654**	**0.628**
9) The proportion of complications related to the referral examinations.	**0.933**	**0.885**	**0.662**	**0.655**
10) The proportion of complications related to primary treatment for a screen detected disease.	**0.551**	**0.574**	**0.637**	**0.627**
11) The age range of persons targeted to screening.	**1.101**	**0.885**	**0.691**	**0.668**
12) The time between regular invitations (i.e. screening interval).	**0.933**	**0.738**	**0.697**	**0.664**
13) The proportion of persons invited at least once in a given time frame (i.e. invitation coverage).	**1.247**	**1.328**	**0.732**	**0.693**
14) The proportion of persons examined at least once in a given time frame (i.e. screening coverage).	**1.348**	**1.574**	**0.691**	**0.725**
15) The proportion of persons invited or examined at least once in a given time frame (i.e. service provider coverage).	**0.876**	**0.787**	**0.702**	**0.702**
16) The proportion of persons examined in a given time frame (i.e. screening attendance).	**1.438**	**1.525**	**0.754**	**0.750**
17) The proportion of different test results (for example 'positive', 'negative' or 'inadequate').	**1.157**	**0.902**	**0.754**	**0.675**
18) The proportion of persons attending a further assessment after a positive screen test result.	**1.326**	**1.443**	**0.733**	**0.725**
19) The proportion of persons attending a treatment after a diagnosis.	**1.303**	**1.344**	**0.765**	**0.708**
20) The distribution of histological diagnosis (pre-cancers and cancers) (i.e. detection rate).	**1.404**	**1.639**	**0.727**	**0.809**
21) The distribution of cancers by the mode of detection (for example, 'pre-screening', 'screen-detected', 'interval	**1.348**	**1.262**	**0.732**	**0.701**

cancer in participant', 'cancer in non-participant', 'post-screening').				
22) The number of newly diagnosed cancer cases after a negative screening episode (i.e. a negative primary test, or negative follow-up after a positive primary test) in a population (i.e. interval cancer rate).	**1.371**	**1.639**	**0.738**	**0.785**
23) The proportion of screen attenders with at least one biopsy.	**0.326**	**-0.180**	**0.670**	**0.752**
24) Episode sensitivity.	**1.236**	**1.000**	**0.769**	**0.848**
25) The proportion of true negative test results in people without cancer (i.e. test specificity).	**1.213**	**1.049**	**0.731**	**0.739**
26) The probability that a person with a positive screen test result has cancer (or a pre-invasive stage) (i.e. positive predicted value).	**1.180**	**1.066**	**0.735**	**0.767**
27) The proportion of persons falsely referred for further assessment (i.e. false positive rate).	**0.955**	**0.951**	**0.678**	**0.729**

Table 2: Top 10 indicators in Round 2, overall Round 3 and Round 3 of those who filled both rounds (TOP10)

TOP10 Indicators	Rank in Round 2	Rank in Round 3 (overall)	Rank in Round 3 (participants in both rounds)
The number of newly diagnosed cancer cases after a negative screening episode (i.e. a negative primary test, or negative follow-up after a positive primary test) in a population (**i.e. interval cancer rate**).	**4**	**2**	**1**
The distribution of histological diagnosis (pre-cancers and cancers) (**i.e. detection rate**).	**2**	**1**	**2**
The proportion of persons examined in a given time frame (**i.e. screening attendance**).	**1**	**5**	**3**
The proportion of persons examined at least once in a given time frame (**i.e. screening coverage**).	**6**	**4**	**4**
The number of newly diagnosed cancer cases (**i.e. incidence**) in a population.	**9**	**6**	**5**
The number of deaths from the given cancer (**i.e. cause-specific mortality**) in a population.	**3**	**3**	**6**
The proportion of persons attending a **further assessment after a positive screen test result**.	**7**	**7**	**7**
The proportion of persons attending a **treatment after a diagnosis**.	**8**	**8**	**8**
The proportion of persons invited at least once in a given time frame (**i.e. invitation coverage**).	**10**	**9**	**9**
The **distribution of cancers by the mode of detection** (for example, 'pre-screening', 'screen-detected', 'interval cancer in participant', 'cancer in non-participant', 'post-screening'.	**5**	**10**	**10**

Table 3: Top 10 indicators by field of expertise

Indicators	Rank in Round 3 (overall)	Breast rank	Cervix rank	Colorectal rank
The number of newly diagnosed cancer cases after a negative screening episode (i.e. a negative primary test, or negative follow-up after a positive primary test) in a population (**i.e. interval cancer rate**).	**2**	**1**	**1**	**1-2***
The distribution of histological diagnosis (pre-cancers and cancers) (**i.e. detection rate**).	**1**	**2**	**2**	**1-2***
The number of deaths from the given cancer (**i.e. cause-specific mortality**) in a population.	**3**	**3**	**4**	**6**
The proportion of persons examined in a given time frame (**i.e. screening attendance**).	**5**	**4**	**7**	**5**
The proportion of persons examined at least once in a given time frame (**i.e. screening coverage**).	**4**	**5**	**5**	**3**
The proportion of persons attending a **further assessment after a positive screen test** result.	**7**	**6**	**6**	**7**
The number of newly diagnosed cancer cases (**i.e. incidence**) in a population.	**6**	**7**	**3**	**4**
The proportion of persons invited at least once in a given time frame (**i.e. invitation coverage**).	**9**	**8**	**-**	**8**
The proportion of persons attending a **treatment after a diagnosis**.	**8**	**9**	**-**	**9**
The **distribution of cancers by the mode of detection** (for example, 'pre-screening', 'screen-detected', 'interval cancer in participant', 'cancer in non-participant', 'post-screening').	**10**	**10**	**8**	**10**
The proportion of **complications related to the referral examinations**.	**-**	**-**	**9**	**-**
The proportion of **complications related to the screening test**.	**-**	**-**	**10**	**-**

*The mean value and the consensus value was the same for these indicators

Table 4: Top 10 indicators by the role of participants in screening programs

Indicators	Rank in Round 3 (overall)	Researcher rank	Program coordinator rank	Policy-maker rank
The number of newly diagnosed cancer cases after a negative screening episode (i.e. a negative primary test, or negative follow-up after a positive primary test) in a population (**i.e. interval cancer rate)**.	**2**	**1**	**2**	**4**
The number of deaths from the given cancer (**i.e. cause-specific mortality**) in a population.	**3**	**2**	**7**	**1**
The distribution of histological diagnosis (pre-cancers and cancers) (**i.e. detection rate**).	**1**	**3**	**1**	**3**
The number of newly diagnosed cancer cases **(i.e. incidence**) in a population.	**6**	**4**	**5**	**2**
The proportion of persons examined at least once in a given time frame (**i.e. screening coverage**).	**4**	**5**	**3**	**-**
The proportion of persons attending a **further assessment after a positive screen test** result.	**7**	**6**	**6**	**9**
The proportion of persons examined in a given time frame (**i.e. screening attendance**).	**5**	**7**	**4**	**-**
The **distribution of cancers by the mode of detection** (for example, 'pre-screening', 'screen-detected', 'interval cancer in participant', 'cancer in non-participant', 'post-screening').	**10**	**8**	**9**	**-**
The proportion of persons invited at least once in a given time frame (**i.e. invitation coverage**).	**9**	**9**	**10**	**8**
The proportion of persons attending a **treatment after a diagnosis**.	**8**	**10**	**8**	**-**
The proportion of persons falsely referred for further assessment (**i.e. false positive rate**).	**-**	**-**	**-**	**5**
The proportion of **complications related to the screening test**.	**-**	**-**	**-**	**6**
The proportion of **complications related to the referral examinations**.	**-**	**-**	**-**	**7**
The proportion of **cancer patients treated with any treatment**.	**-**	**-**	**-**	**10-11***
The probability that a person with a positive screen test result has cancer (or a pre-invasive stage) (**i.e. positive predicted value**).	**-**	**-**	**-**	**10-11***

*The mean value and the consensus value was the same for these indicators