

Advancing the evaluation of cervical cancer screening: Development and application of a longitudinal adherence metric

Kine Pedersen¹, Emily A. Burger^{1, 2}, Suzanne Campbell³, Mari Nygård³, Eline Aas¹, Stefan Lönnberg³

¹University of Oslo, Department of Health Management and Health Economics, Oslo, Norway

²Harvard T.H. Chan School of Public Health, Center for Health Decision Science, Boston, MA, USA

³Cancer Registry of Norway, Oslo, Norway

Corresponding author:

Kine Pedersen

University of Oslo

Department of Health Management and Health Economics

PO Box 1089 Blindern

Oslo, 0317, Norway

E-mail: kine.pedersen@medisin.uio.no

Telephone number: +47 22845277

Funding source

K Pedersen and EA Burger was funded by the Research Council of Norway (grant number 238042).

ABSTRACT

Background: Attendance to routine cancer screening at repeated intervals is essential for reducing morbidity and mortality of targeted cancers, yet currently-defined quality-assurance metrics evaluate coverage within a defined period of time (e.g. 3.5 years).

Methods: We developed a longitudinal adherence metric that captures attendance to cancer screening at repeated intervals, and applied the metric to population-based data from the Cancer Registry of Norway that captures two decades of organised screening, including all screening tests and cervical cancer diagnoses for women living in Norway at any time during years 1992-2013 and eligible for at least two screening rounds (1 round=3.5 years, N=1, 391,812). For each woman, we calculated the proportion of eligible screening rounds with at least one registered cytology test, and categorised women into one of five longitudinal adherence categories: never-screeners, severe under-screeners, moderate under-screeners, guidelines-based screeners, and over-screeners. For each category, we evaluated cancer outcomes such as cancer stage at diagnosis.

Results: Only 46% of screen-eligible women were consistently screened at least once every 3.5 years, and the majority of these were over-screened. In contrast, 29% were moderately under-screened, 17% were severely under-screened, and 8% had never attended screening. Screening behaviour was associated with cancer outcomes; for example, the proportion of cancers diagnosed at Stage I increased from 21% among never-screeners to 70% among over-screeners.

Conclusion: The longitudinal adherence metric evaluates screening performance as a succession of screening episodes, reflecting both guidelines and the fundamental

principles of screening, and may be a valuable addition to existing performance indicators.

Key words: Mass screening, cervical cancer, guideline adherence, quality assurance

INTRODUCTION

Routine cytology-based screening has contributed to reducing cervical cancer (CC) incidence and mortality worldwide (1, 2). Adherence to screening guidelines is considered one of the most important factors for screening effectiveness and efficiency (3-6), and the European guidelines for quality assurance of CC screening have recognized screening coverage as a key performance indicator (4). The current adherence metric is calculated as the proportion of women in screening target age who received at least one screening test within a defined period of time. For example, in Norway, 67% of screen-eligible women have attended screening within the last 3.5 years (7). Similar metrics are used to evaluate screening performance in the United Kingdom and other European countries (8). However, due to the low sensitivity of a single cytology-based screen, the effectiveness of a screening programme relies on attendance at repeated intervals (9, 10). Moreover, efforts to improve participation are emerging with limited information about which individuals to target. Although the European guidelines encourage evaluating coverage over longer follow-up periods (4), to our knowledge, there is no methodological framework for screening metrics that can identify different patterns of adherence (e.g. under- and over-screening) and screening behaviour over more than one screening round.

We developed a longitudinal adherence metric that evaluates repeated attendance to cancer screening over recommended intervals, and applied the metric to population-based registry-data that captures two decades of organised CC screening in Norway.

METHODS

The longitudinal adherence metric

We defined a longitudinal adherence metric by identifying and/or evaluating the following components: 1) Observation period (e.g. based on available data), 2) relevant screening interval length to define a single “round” of screening, 3) relevant study population, 4) classification of screening and cancer variables, 5) the observation time and number of eligible screening rounds for each individual, 6) a longitudinal adherence score, calculated as the proportion of attended screening rounds (i.e., a screening round with at least one registered cytology test) out of the total number of eligible screening rounds, and 7) longitudinal adherence score cut-off values to specify the relevant longitudinal adherence categories.

Application to cervical cancer screening in Norway

Study setting

Since 1995, the Norwegian Cervical Cancer Screening Programme reminds women aged 25 to 69 years to attend cytology-based screening every three years. The Cancer Registry of Norway (CRN) is responsible for managing, monitoring and refining the programme (7). The CRN databases include information on the morphology of all cytology tests performed since 1992, with no distinction of indication (i.e. “opportunistic” or “organised”). CC diagnoses (by stage and morphology) are available since 1953. The cancer database is considered nearly 100% complete due to statutory and systematic reporting of results by all pathology laboratories in Norway, and continuous monitoring of received data (11). All

Norwegian citizens have a unique personal identification number, allowing individual data linkage across registries.

Study population

We evaluated longitudinal adherence to the Norwegian Cervical Cancer Screening Program from January 1st 1992 (initiation of the pilot screening programme and nationwide registration of screening tests) until December 31st 2013. In order to capture variation in the timing of cytology tests relative to the three-year interval recommendation for CC screening in Norway (4), we used an interval length of 3.5 years to define one screening round. We included women born between years 1936-1981, and registered as living in Norway at any time during the study period (years 1992-2013), as they were eligible for at least two screening rounds based on age and the year of organised screening implementation (**Supplementary Table 1**). We excluded women who experienced a CC diagnosis, emigration or death before the end of two screening rounds.

Classification of screening and cancer data

We used data from the Norwegian Population Registry and databases at CRN to compile information on screening and cancer data, including population demographics, cytology tests (date of visit and morphology) and CC diagnoses. We differentiated cytology tests performed as part of routine screening (i.e. screening tests) from cytology tests used to follow-up previously abnormal results (i.e. follow-up tests). A cytology test was considered a follow-up test if the woman had any cervical

abnormalities or complementary tests (e.g. cervical biopsy or human papillomavirus (HPV) test) in the two preceding years. We identified all incident cases of invasive CC (International Classification of Diseases and Related Health Problems 10th Revision code C53) from 1953-2013, including date of diagnosis, morphology (i.e. squamous cell carcinomas (SCC), adenocarcinomas and other or undefined CCs, as previously defined (12)), and stage at diagnosis (reported using the International Federation of Gynecology and Obstetrics (FIGO) classification system (stages I-IV) (13)).

Longitudinal adherence categories

We categorised each woman into one of five longitudinal adherence categories based on individual attendance over repeated screening rounds. We defined each woman's observation time in the screening programme as the number of years she lived at screening target age (i.e. ages 25-69 years) during 1992-2013 until a CC event, end of screen-eligibility, emigration, death, or end of follow-up (December 31st 2013), whichever came first. We included an additional +/- 2 years before and after screening target ages for those attending screening outside recommended ages (i.e. ages 23-24 and 70-71 years). For each woman, we identified the number of eligible screening rounds given her observation time in screening, and rounded down to the number of complete 3.5-year screening rounds (with a maximum of six eligible rounds available in our data). We subsequently calculated a longitudinal adherence score, defined as the number of attended screening rounds (i.e., containing at least one registered cytology, either screening or follow-up test) out of the total eligible screening rounds (**Figures 1 and 2**).

We constructed five longitudinal adherence categories using the array of longitudinal adherence scores (**Figure 1**): 1) never-screeners, 2) severe under-screeners (attended half or less of eligible screening rounds), 3) moderate under-screeners (attended more than half of eligible screening rounds, but not all), 4) guidelines-based screeners (attended all eligible screening rounds), 5) over-screeners (attended all eligible screening rounds, and had excessive screening tests within an interval). To distinguish between guidelines-based screeners and over-screeners (who both have a longitudinal adherence score of 100%), we calculated a screening intensity score, defined as the total number of screening tests (excluding follow-up tests to avoid misclassifying women undergoing follow-up testing as over-screeners) divided by the number of eligible screening rounds. Subsequently, we categorized over-screeners as those women who had a screening intensity score >1.33 (i.e., more than one-third excessive screening tests), while the remaining women were categorized as guidelines-based screeners (i.e. a screening intensity score ≤ 1.33).

Analysis

We tabulated the number and proportion of women in each longitudinal adherence category, and the mean year of birth, age of screen-eligibility, observation time within the screening programme, number of eligible screening rounds and the number of screening tests. To evaluate the discriminatory power of the longitudinal adherence metric, we estimated cancer outcomes among women within each longitudinal adherence category, including the number of CCs stratified by morphology and stage. We also calculated truncated age-standardised CC incidence rates (TASRs) for 5-

year age-groups using a world standard population (14). Person-years were accumulated from after the second screening round (i.e. the study inclusion criteria) until the end of follow-up (December 31st 2013, CC event, emigration, or death, whichever came first). We also calculated TAsRs for FIGO stages II+ (i.e. stage II or more severe), which more closely relates to cervical cancer mortality. All statistical analyses were performed using STATA statistical software, version 14.

To evaluate the impact of inclusion/exclusion criteria and definitions of the longitudinal adherence metric on the distribution of longitudinal adherence and ability to differentiate cancer outcomes, we performed four separate sensitivity analyses: 1) Varying the number of eligible screening rounds required for study inclusion (i.e. at least three or four versus two rounds), 2) including only screening tests (not follow-up tests) to calculate the longitudinal adherence score, 3) censoring observation time in the screening programme at the time of the first abnormal cytology test (i.e. atypical squamous cells of undetermined significance or more severe), and 4) varying the intensity score cut-off to define over-screening (i.e. a screening intensity score of 1 or 1.5). Finally, to compare our sample population with the current coverage metric reported by CRN, we estimated a cross-sectional coverage rate as the proportion of women in our sample with an interval cut-off in year 2012 or 2013 and who were registered with at least one cytology test within that screening round.

RESULTS

Study population

We identified 1,493,942 women born between 1936 and 1981 and living in Norway at any time between 1992 and 2013. We excluded a total of 102,130 women with

less than two eligible screening rounds in the period 1992-2013. Among the excluded women, 4,596 had a CC diagnosis, of which 2,405 occurred prior to the study period (<year 1992) or screening initiation age, and the remaining 2,191 occurred among women with less than two complete rounds of screening. The analysed cohort comprised of 1,391,812 women with 2,716 cases of CC, of which 74% and 21% were SCC and adenocarcinoma, respectively (**Table 1**). On average, women in the analysed cohort entered screening at age 35 years, and had an average of 17.5 years and 5 rounds of observation time in screening. The proportions of women in the analysed cohort with two, three, four, five or six eligible screening rounds were 8%, 12%, 17%, 42%, and 21%, respectively. The mean age at CC diagnosis was 47.6 years (range: 30-77 years).

Longitudinal adherence and cancer outcomes

Using the longitudinal adherence metric to characterize patterns in longitudinal screening adherence, we found that 8% of women were never-screeners, 17% were severe under-screeners, 29% were moderate under-screeners, 19% were guidelines-based screeners, and 27% were considered over-screeners (**Table 1**). Among women who consistently screened at least once per eligible screening round (i.e. guidelines-based screeners and over-screeners), 59% were over-screeners.

Cancer outcomes varied considerably across the longitudinal adherence categories. For example, while the TAsR (for all CCs) per 100,000 woman-years was 18.9 overall, it was highest among never-screeners (50.4), almost three times higher than for guidelines-based screeners (16.9) (**Table 1**). Moderate under-screeners had the lowest incidence at 9.4, while the incidence was slightly higher

among over-screeners (18.1) than guidelines-based screeners. When calculating TAsRs for FIGO stages II+ only, the incidence was lowest among over-screeners. The mean age at cancer diagnosis was highest among never-screeners (55.8 years) and lowest among over-screeners (43.9 years). The histological cancer type and stage distribution at cancer diagnosis also varied by category; the proportion of SCC decreased from 83% to 67% with more frequent screening, while the proportion of adenocarcinoma increased from 10% to 27%. The proportion of other types was similar across categories (i.e. 5% to 7%). In addition, the proportion of cancers diagnosed at Stage I increased with more frequent screening (ranging from 21-70%) (**Table 1 and Figure 3**), while never-screeners had the highest proportion of Stage IV cancers (21%).

The distribution of participants between longitudinal adherence categories and patterns in cancer outcomes remained consistent across different inclusion/exclusion criteria and definitions of longitudinal adherence (**Supplementary Table 2**). However, when we assumed an intensity score cut-off of >1 to define over-screening, guidelines-based screeners constituted only 7% of women in the analysed cohort, and had higher cancer incidence than over-screeners. Furthermore, over-screeners had lower cancer incidence than guidelines-based screeners when we required at least three or four eligible screening rounds for study inclusion, or when the intensity score cut-off was increased to >1.

Finally, the cross-sectional coverage rate for years 2012-2013 in our sample was 66%, similar to the 67% coverage rate estimated for years 2012-2014 in Norway.

DISCUSSION

Using population-based registry data that captures two decades of organised screening in Norway, we developed a screening adherence metric that evaluates adherence to screening over multiple rounds and reflects screening behaviour on a more granular level (including over- and under-screening). As population-based cancer screening is a continuing process (15), longitudinal adherence may reflect the fundamental principle of organised screening more accurately than the current coverage metric. In contrast to the 67% coverage rate observed in Norway (7), we found that only 46% of screen-eligible women consistently had at least one screening test every 3.5 years, of which more than half were over-screening. However, a larger proportion of screen-eligible women receive the benefits of routine screening, reflected by the 75% of women who attended screening at least once every six years.

The longitudinal adherence metric was able to differentiate cancer outcomes according to screening behaviour. As expected, we observed the highest CC incidence among women who never or seldom attended screening (i.e. never- or severe under-screeners), and these women were diagnosed at an older age and at a more advanced stage than women who screened more frequently. We also found that the proportion of cancers diagnosed at Stage I steadily increased with more frequent screening. In contrast, the proportion of SCC decreased with more frequent screening, which is expected as SCC is the morphologic type that is most readily preventable through screening (16).

This is, to our knowledge, the first study to develop and evaluate a longitudinal screening adherence metric. A previous study evaluated adherence to the Swedish screening programme by measuring the proportion of time in screening within a 12-year period, but only reported cancer outcomes for the dichotomized

categories of adherent and non-adherent women (17). While evaluating the proportion of time participating in screening may be helpful to compare mean degree of participation across groups (e.g. immigrant groups), this measure does not stratify different patterns of screening behaviour and does not capture over-screening. Other studies have evaluated longitudinal adherence to other, non-CC screening programmes over time (18-20), but did not consider multiple levels of under-screening, over-screening, nor the ability to differentiate cancer outcomes. We found that longitudinal adherence was lower than cross-sectional adherence within the last screening round; similar results were found for cervical, colorectal, and breast cancer screening elsewhere (17, 21, 22), suggesting that attendance over multiple screening rounds more accurately reflects screening performance.

Limitations

The longitudinal adherence metric developed in this study does not incorporate all dimensions of guidelines-based recommendations, such as compliance to recommended follow-up. We considered attending screening at routine intervals the most important component of longitudinal adherence. Other aspects, such as the maximum length of time between screening tests, may be equally important. For example, screening at the beginning and end of two consecutive intervals could imply nearly 7 years between two screens. However, compared with the current dichotomized coverage metric, the longitudinal adherence metric allows for multiple levels of adherence (i.e. under/over-screening) and is therefore less prone to severe misclassification (e.g. classifying a woman attending screening every four years as a non-attender). Results may vary depending on different conceptualisations of the metric and category thresholds, as well as data availability. For example, we found

that reducing the threshold for over-screening to a screening intensity score of >1 (rather than >1.33) implied a higher cancer incidence among guidelines-based screeners than over-screeners, primarily due to the large proportion of women in our sample with an intensity score of 1.1-1.33 (Supplementary Table 2). However, the distribution of longitudinal adherence proportions was consistent across different definitions of the metric, and we consider the selected conceptualisation transparent and reproducible for other settings. In addition, our sample population showed good validity when compared with both the cross-sectional coverage rate observed in Norway, as well as the TASR reported by NORDCAN (19.3 for ages 30-79 (23) versus 18.9 for ages 30-77 years in our sample). Finally, high-quality registry-data over a period corresponding to at least two screening rounds are required for other screening programmes to utilize the longitudinal adherence metric. Although evaluating longitudinal adherence may not currently be feasible for all cancer screening programmes, application in countries with well-established cancer registries may inform screening policies elsewhere.

Policy implications and future research

The longitudinal adherence metric may inform studies that guide the continued refinement of screening programmes, such as evaluating whether CC specific mortality differ by patterns in screening behaviour. The metric may also help identify groups of women to target for interventions to improve screening adherence.

Improving adherence among women that consistently never- or under-screen is expected to improve the effectiveness and efficiency of organised screening (24, 25).

Home-based testing for HPV has shown promising potential to increase screening participation (26, 27), yet the cost-effectiveness of self-sampling policies relies on

targeting the most under-screened women and those at highest risk of developing CC (28, 29). As such, understanding women's screening history and its association with cancer outcomes is essential.

Within each longitudinal adherence category there is likely heterogeneity in individual characteristics and the underlying risk of developing CC. Self-selection of participants in a screening programme is often referred to as "healthy screenee bias", suggesting that screening participants on average are healthier than non-participants (30). For example, a Danish study found that participants in CC screening had lower all-cause mortality rates than non-participants (31), and a study in Finland showed that women who chose not to attend after receiving an invitation had a higher risk of CC than non-invited women (32). In this descriptive analysis, we found that moderate under-screener had the lowest cancer incidence, while over-screener had higher incidence than guidelines-based screeners. These longitudinal adherence categories may be particularly subject to selection bias if women are aware of individual CC risk based on risk factor exposure. Future studies may use the longitudinal adherence framework to evaluate characteristics of study participants to evaluate potential "healthy screenee bias" and other disparities in participation (e.g. socioeconomic status).

We conclude that the longitudinal adherence metric developed in this study more accurately reflects guidelines-based recommendations by capturing screening behaviour over repeated intervals and differences in cancer outcomes, and may be a valuable addition to the existing performance indicators of organised screening. In turn, the metric may be applied within other prevention programmes that rely on repeated follow-up such as breast and colorectal cancer screening programmes,

where screening coverage is a key performance indicator defined similar to the current adherence metric for CC screening (33).

CONFLICTS OF INTEREST

None declared.

ACKNOWLEDGEMENTS

We are grateful for the contributions of Ivar S. Kristiansen, Søren T. Klitkou, Gry B. Skare and the Cancer Registry of Norway. This study was funded by the Norwegian Research Council (238042). Preliminary results of this study were presented at the International Papillomavirus Conference, Lisbon, Portugal, 21st September 2015 and the Society for Medical Decision Making Conference, Vancouver, Canada, 24th October 2016.

ETHICAL APPROVAL

This study was approved by the Data Protection Official at Oslo University Hospital (approval number 2015/17224).

KEY POINTS

- This is the first study to develop a longitudinal screening adherence metric and apply the metric to population-based data that captures two decades of organised cervical cancer screening in Norway.
- The longitudinal adherence metric developed in this study captures different patterns of screening behaviour, including under-screening and over-

screening, and shows an association between screening behaviour and cancer outcomes.

- In Norway, 46% of women were consistently screened at least once every 3.5 years, and the majority of these were over-screened.
- The longitudinal adherence metric may be a valuable addition to existing screening performance indicators, and may inform other studies that guide the continued refinement of screening programmes.

REFERENCES

1. Peirson L, Fitzpatrick-Lewis D, Ciliska D, Warren R. Screening for cervical cancer: a systematic review and meta-analysis. *Systematic reviews*. 2013;2:35.
2. Bray F, Loos AH, McCarron P, Weiderpass E, Arbyn M, Moller H, et al. Trends in cervical squamous cell carcinoma incidence in 13 European countries: changing risk and the effects of screening. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2005;14(3):677-86.
3. Bos AB, Rebolj M, Habbema JD, van Ballegooijen M. Nonattendance is still the main limitation for the effectiveness of screening for cervical cancer in the Netherlands. *International journal of cancer Journal international du cancer*. 2006;119(10):2372-5.
4. Arbyn M, Anttila A, Jordan J. J., Ronco G, Schenck U, Segnan N, Wiener HG, Herbert A, et al. European guidelines for quality assurance in cervical cancer screening. 2nd edition. International Agency for Research on Cancer: Lyon, France. . 2008.
5. Chubak J, Hubbard R. Defining and measuring adherence to cancer screening. *Journal of medical screening*. 2016.
6. Arbyn M, Anttila A, Jordan J, Ronco G, Schenck U, Segnan N, et al. European Guidelines for Quality Assurance in Cervical Cancer Screening. Second edition--summary document. *Ann Oncol*. 2010;21(3):448-58.
7. Cancer Registry of Norway. Annual Report of the Norwegian Cervical Cancer Screening Program, 2013-14 Accessed at https://www.kreftregisteret.no/globalassets/publikasjoner-og-rapporter/livmorhalskreft/arsrapport/livmorhals_2015pdf on October 16, 2015 Oslo: Cancer Registry of Norway, 2015.
8. Elfstrom KM, Arnheim-Dahlstrom L, von Karsa L, Dillner J. Cervical cancer screening in Europe: Quality assurance and organisation of programmes. *European journal of cancer (Oxford, England : 1990)*. 2015;51(8):950-68.
9. Schiffman M, Wentzensen N, Wacholder S, Kinney W, Gage JC, Castle PE. Human papillomavirus testing in the prevention of cervical cancer. *Journal of the National Cancer Institute*. 2011;103(5):368-83.
10. IARC. Screening for squamous cervical cancer: duration of low risk after negative results of cervical cytology and its implication for screening policies. IARC Working Group on evaluation of cervical cancer screening programmes. *British medical journal (Clinical research ed)*. 1986;293(6548):659-64.
11. Bilet EF, Langseth H, Thoresen SO, Bray F. Completeness of invasive cervical cancer at the Cancer Registry of Norway. *Acta oncologica (Stockholm, Sweden)*. 2009;48(7):1070-3.
12. Lonnberg S, Hansen BT, Haldorsen T, Campbell S, Schee K, Nygard M. Cervical cancer prevented by screening: Long-term incidence trends by morphology in Norway. *International journal of cancer Journal international du cancer*. 2015;137(7):1758-64.
13. Pecorelli S. Revised FIGO staging for carcinoma of the vulva, cervix, and endometrium. *International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics*. 2009;105(2):103-4.
14. Doll R, Cook P. Summarizing indices for comparison of cancer incidence data. *International journal of cancer Journal international du cancer*. 1967;2(3):269-79.
15. Wilson JMG, Jungner G. Principles and practice of screening for disease. 1968(Geneva: WHO; 1968. Available from: <http://www.who.int/bulletin/volumes/86/4/07-050112BP.pdf>).
16. Schiffman M, Wentzensen N. Human papillomavirus infection and the multistage carcinogenesis of cervical cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2013;22(4):553-60.
17. Azerkan F, Sparen P, Sandin S, Tillgren P, Faxelid E, Zendejdel K. Cervical screening participation and risk among Swedish-born and immigrant women in Sweden. *International journal of cancer Journal international du cancer*. 2012;130(4):937-47.
18. Jensen CD, Corley DA, Quinn VP, Doubeni CA, Zauber AG, Lee JK, et al. Fecal Immunochemical Test Program Performance Over 4 Rounds of Annual Screening: A Retrospective Cohort Study. *Annals of internal medicine*. 2016;164(7):456-63.
19. Worthington C, McLeish K, Fuller-Thomson E. Adherence over time to cervical cancer screening guidelines: insights from the Canadian National Population Health Survey. *Journal of women's health (2002)*. 2012;21(2):199-208.

20. Cyhaniuk A, Coombes ME. Longitudinal adherence to colorectal cancer screening guidelines. *The American journal of managed care*. 2016;22(2):105-11.
21. Gellad ZF, Stechuchak KM, Fisher DA, Olsen MK, McDuffie JR, Ostbye T, et al. Longitudinal adherence to fecal occult blood testing impacts colorectal cancer screening quality. *The American journal of gastroenterology*. 2011;106(6):1125-34.
22. Jacobsen KK, von Euler-Chelpin M. Performance indicators for participation in organized mammography screening. *Journal of public health (Oxford, England)*. 2012;34(2):272-8.
23. Engholm G, Ferlay J, Christensen N, Kejs AMT, Hertzum-Larsen R, Johannesen TBK, S., et al. NORDCAN: Cancer Incidence, Mortality, Prevalence and Survival in the Nordic Countries, Version 7.3 (08.07.2016). Association of the Nordic Cancer Registries. Danish Cancer Society. Available from <http://www.anccr.nu>, accessed on 7/January/2017. 2016.
24. Kim JJ, Campos NG, Sy S, Burger EA, Cuzick J, Castle PE, et al. Inefficiencies and High-Value Improvements in U.S. Cervical Cancer Screening Practice: A Cost-Effectiveness Analysis. *Annals of internal medicine*. 2015.
25. Burger EA, Kim JJ. The value of improving failures within a cervical cancer screening program: an example from Norway. *International journal of cancer Journal international du cancer*. 2014;135(8):1931-9.
26. Giorgi Rossi P, Fortunato C, Barbarino P, Boveri S, Caroli S, Del Mistro A, et al. Self-sampling to increase participation in cervical cancer screening: an RCT comparing home mailing, distribution in pharmacies, and recall letter. *British journal of cancer*. 2015;112(4):667-75.
27. Enerly E, Bonde J, Schee K, Pedersen H, Lonngberg S, Nygard M. Self-Sampling for Human Papillomavirus Testing among Non-Attendees Increases Attendance to the Norwegian Cervical Cancer Screening Programme. *PloS one*. 2016;11(4):e0151978.
28. Arbyn M, Castle PE. Offering Self-Sampling Kits for HPV Testing to Reach Women Who Do Not Attend in the Regular Cervical Cancer Screening Program. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2015;24(5):769-72.
29. Burger EA, Sy S, Nygard M, Kim JJ. The cost-effectiveness of cervical self-sampling to improve routine cervical cancer screening: The importance of respondent screening history and compliance. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2016.
30. Crowell JM, Ransohoff DF, Kramer BS. Principles of cancer screening: lessons from history and study design issues. *Seminars in oncology*. 2010;37(3):202-15.
31. Kranse R, van Leeuwen PJ, Hakulinen T, Hugosson J, Tammela TL, Ciatto S, et al. Excess all-cause mortality in the evaluation of a screening trial to account for selective participation. *Journal of medical screening*. 2013;20(1):39-45.
32. Lonngberg S, Anttila A, Luostarinen T, Nieminen P. Age-specific effectiveness of the Finnish cervical cancer screening programme. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2012;21(8):1354-61.
33. Anttila A, Lonngberg S, Ponti A, Suonio E, Villain P, Coebergh JW, et al. Towards better implementation of cancer screening in Europe through improved monitoring and evaluation and greater engagement of cancer registries. *European journal of cancer (Oxford, England : 1990)*. 2015;51(2):241-51.

Table 1. Characteristics of study participants and number of cervical cancers by longitudinal adherence category.

Variable	Never- screeners	Severe under- screeners	Moderate under- screeners	Guidelines- based screeners	Over- screeners	Total
Number of women (%)	117 887 (8)	231 957 (17)	396 889 (29)	265 738 (19)	379 341 (27)	1 391 812 (100)
Mean year of birth [SD]	1957 [14]	1959 [13]	1960 [11]	1961 [13]	1961 [12]	1960 [13]
Mean age of screen-eligibility [SD]	40.8 [12.2]	36.6 [10.8]	34.7 [9.8]	34.9 [10.7]	33.9 [10.0]	35.4 [10.6]
Mean years of observation-time in screening [SD]	14.4 [3.7]	16.6 [4.2]	18.7 [2.9]	16.9 [4.3]	18.0 [4.1]	17.5 [4.0]
Mean number of eligible screening rounds [SD]	3.7 [1.1]	4.3 [1.3]	4.9 [0.9]	4.4 [1.3]	4.7 [1.2]	4.6 [1.2]
Mean number of screening tests ^a [SD]	0.0 [0.0]	2.1 [1.3]	5.2 [2.0]	5.1 [1.7]	8.5 [3.1]	6.0 [3.4]
Number of cervical cancer cases (%)	481 (18)	578 (21)	444 (16)	449 (17)	764 (28)	2 716 (100)
Truncated ASR (all stages) ^b	50.4	26.5	9.4	16.9	18.1	18.9
Truncated ASR (FIGO II+) ^b	30.4	10.3	3.1	3.8	2.8	6.3
Mean age at diagnosis [SD]	55.8 [10.3]	48.1 [10.5]	47.6 [9.5]	44.6 [9.9]	43.9 [10.2]	47.6 [10.9]
Histological type (% within adherence category)						
Squamous cell carcinoma	401 (83)	463 (80)	310 (70)	314 (70)	513 (67)	2 001 (74)
Adenocarcinoma	47 (10)	79 (14)	112 (25)	113 (25)	209 (27)	560 (21)
Other	33 (7)	36 (6)	22 (5)	22 (5)	42 (5)	155 (6)

^aReported as the average total screening tests for each woman. Some women may have multiple tests within a single screening round. ^bAge-standardised using a world standard population and truncated on ages 30-77 years (i.e. the ages under observation in the data). Abbreviations: ASR; Age-standardised cervical incidence rate, FIGO II+; FIGO stage II, III or IV, SD; standard deviation.

Figure legends

Figure 1. Overview of the longitudinal adherence scores and categories.

Overview of the longitudinal adherence scores and categories. The heat map illustrates all possible combinations of eligible and attended screening rounds for women in the analysed cohort, and the associated longitudinal adherence score, defined as the number of attended screening rounds divided by the number of eligible screening rounds. See figure 2 for details on eligible and attended screening rounds. Based on the array of consistency scores, we constructed the following five adherence categories: never-screeners, severe under-screeners, moderate under-screeners, guidelines-based screeners and over-screeners. The table summarizes the longitudinal adherence score and the implied average screening interval (i.e. longitudinal adherence score divided by 3.5 years) for each longitudinal adherence category. Over-screeners had at least one cytology test in all expected screening rounds, and more than 33% excessive screening tests in total (see Methods for details).

Figure 2. Hypothetical screening histories and associated longitudinal adherence category.

Abbreviations: Jul, july; Dec, December; yo, years old.

Hypothetical screening histories and associated longitudinal adherence category.

Abbreviations: Jul, july; Dec, December; yo, years old. Using an interval length of 3.5 years, each woman (A–E) would have a different number of eligible screening rounds (2–6 rounds) from the time she became eligible for screening (triangle) until the end of observation time in screening due to cervical cancer screening stop age

(ages 69–71 years), emigration, death or end of study period in December 2013 (circle with crossed lines), whichever came first. The number of attended rounds indicates the number of screening rounds a woman had at least one cytology test (solid circle). *For screening history E, the woman has an adherence score=1 and a screening intensity score=1.5 (i.e. 6 screening tests divided by 4 eligible rounds), and is therefore classified as an over-screener. See figure 1 and Methods for details on longitudinal adherence categories.

Figure 3. Stage distribution of cervical cancers for each longitudinal adherence category.

Abbreviations: N, number of cervical cancer cases; TASR, truncated age-standardised incidence rate per 100,000 woman-years.

Stages (I, II, III, IV) follow the International Federation of Gynecology and Obstetrics (FIGO) classification system. Study population indicate all women that met the study inclusion criteria.

Figure 1.

Attended screening rounds	Eligible screening rounds					Longitudinal adherence category	Longitudinal adherence score	Average screening interval (years)	Description
	2	3	4	5	6				
0	0.00	0.00	0.00	0.00	0.00	Never-screener	0.00	∞	Never attended screening
1	0.50	0.33	0.25	0.20	0.17	Severe under-screener	0.17 - 0.50	7 - 21	Attended half or less of eligible screening rounds
2	1.00	0.67	0.50	0.40	0.33				
3		1.00	0.75	0.60	0.50				
4			1.00	0.80	0.67	Moderate under-screener	0.60 - 0.83	4 - 6	Attended more than half of eligible screening rounds, but not all
5				1.00	0.83	Guidelines-based screener	1.00	3.5	Attended all eligible screening rounds
6					1.00				
Longitudinal adherence score $= \frac{\text{Number of attended screening rounds}}{\text{Number of eligible screening rounds}}$						Over-screener	1.00	< 3.5	Attended all eligible screening rounds, with excessive screening tests

Figure 2.

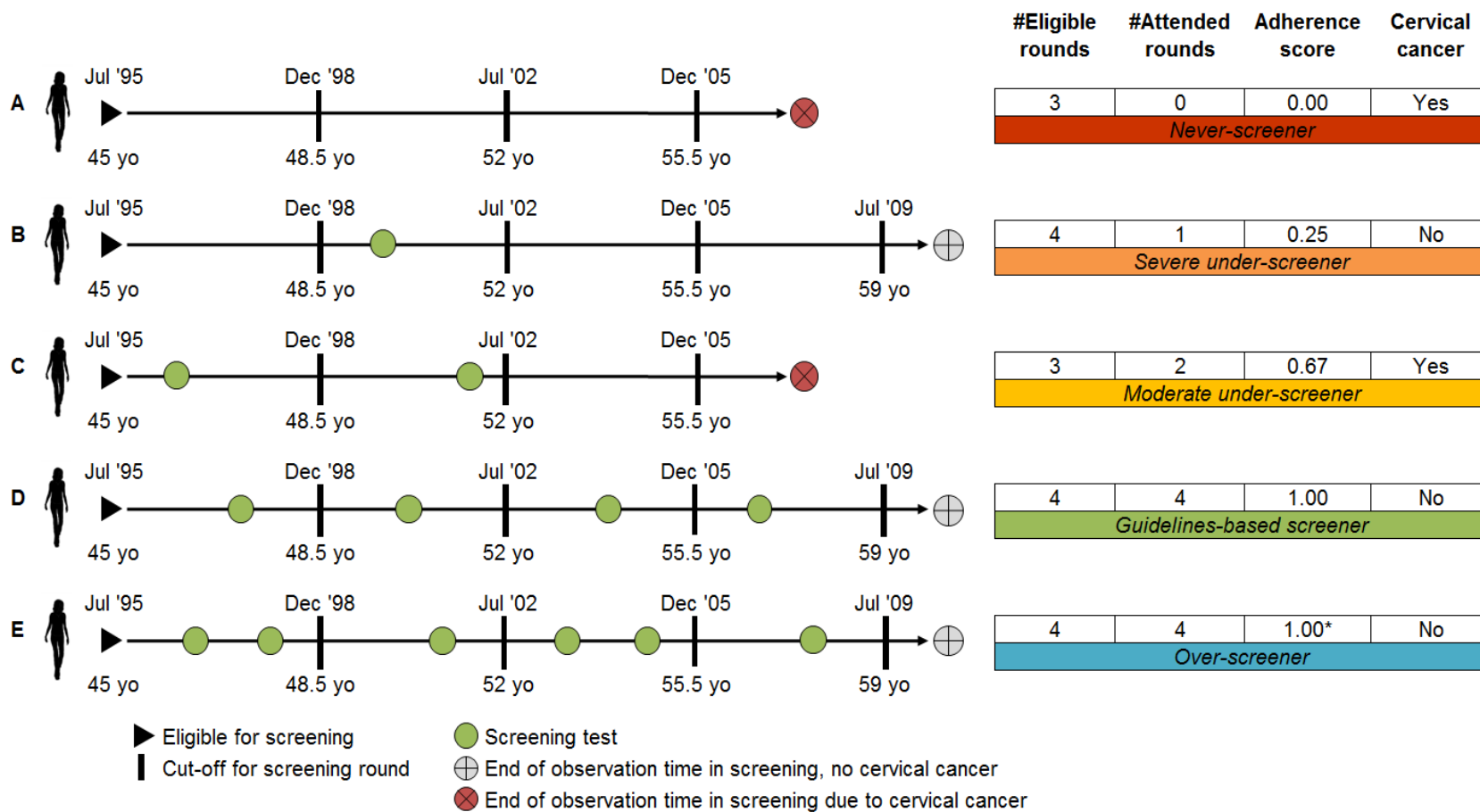


Figure 3.

