

THE LANCET Oncology

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Vaccarella S, Franceschi S, Zaridze D, et al. Preventable fractions of cervical cancer via effective screening in six Baltic, central, and eastern European countries 2017–40: a population-based study. *Lancet Oncol* 2016; published online Aug 22. [http://dx.doi.org/10.1016/S1470-2045\(16\)30275-3](http://dx.doi.org/10.1016/S1470-2045(16)30275-3).

Appendix

Quantifying potential for cervical cancer prevention in six Baltic and Central/Eastern European countries

Specification of model-based projections

We analysed invasive cervical cancer (CC) incidence data obtained from population-based cancer registries from six countries from the Baltic and Central/Eastern European area: Estonia, Lithuania, Latvia, Belarus, Bulgaria and the Russian Federation. National data were available, by single calendar year of diagnosis and by 5-year age groups (30-34, 35-39, ..., 70-74). The time span of observations varied from 15 years in Bulgaria and the Russian Federation to 40 years in Estonia. The last year of diagnosis available was 2007 in all countries except Bulgaria (2008) and the Russian Federation (2012). Analyses were restricted to ages 30 to 74. Age-period-cohort (APC) models were fitted as follows:

$$\lambda(\text{age, period}) = g^{-1}\{f_A(\text{age}) + f_P(\text{period}) + f_C(\text{cohort})\},$$

where λ represents the incidence rate as a function of age and calendar period, g is the chosen “link” function and f_A , f_P , and f_C are appropriate functions of age, period and cohort, respectively. The canonical link function, i.e., the log function (i.e., $g^{-1}(x) = \exp(x)$) was used in the present analysis.¹⁻⁴

The data were arranged in 5-year age groups and single calendar year of diagnosis. Birth cohorts were obtained by subtracting age (midpoint of 5-year age band) from the calendar period year of diagnosis. Age effects were analyzed as a factor variable. For each 5-year age-class and 1-year period of diagnosis, the number of events and person-years corresponded to 5 x 1-year subsets of a Lexis diagram. An APC model applied on one-year rates would include 9 age groups and a relatively large number of periods and cohorts, varying across countries. Whenever possible, it would be preferable to obtain additional information about one of the three factors that has scientific justification *a priori*, and imposing a constraint reflecting this information that could enable a unique set of parameters to be identified^{1,2,5} thus circumventing the non-identifiability problem. The non-identifiability problem of APC models was circumvented by taking advantage of the consistent relationship between age and CC incidence. It has been observed that in unscreened populations, CC incidence rates increase up to approximately age 45 years and then flatten.⁶ This observation is true also in populations from the Nordic countries in a pre-screening epoch.⁷ We, therefore, constrained incidence rates to be equal at ages 45–49 and 65–69, thus reflecting *a priori* scientific belief and enabling the estimation of a unique set of parameters for the age, period and cohort effects. The same constraint can be used in screened populations, where incidence rates flatten approximately at 35 years of age, that is, when the beneficial effect of screening, that typically targets women aged 25 years or older, starts to manifest.

Forecasts in a scenario without improvements in screening activities

Expected future incidence rates of CC up to 2040 were obtained by projecting period and cohort effects, whereas future age effects were assumed to remain the same. Restricted cubic splines were used to model and project forward period and cohort effects as varying continuously. Cubic splines are flexible and allow for a more realistic scenario as changes over time are expected to vary smoothly rather than with sudden jumps. A key feature of the restricted cubic splines is the estimate of a smoothed curve within the range of the observed data, which is obliged to pass through a pre-defined number of “knots”. There is a tradeoff between the number of knots, which determines flexibility of the spline, and the parsimony of the model. As period and cohort variables were analysed by single year of diagnosis and of birth, knots for the two corresponding functions were placed at equally-spaced 5-year intervals across the observed range. The extrapolation outside the last knot is linear with a larger weight given to the most recent trends. Therefore, in order to avoid too much dependence of the projections on the trend at the end of the observed data, the knots in each country were chosen so that the final knot was placed 3 years within the observed range.

The use of the canonical (log) link function for APC models requires an exponential transformation that can lead to overestimation of the future number of cases, where long-term predictions are required and trends in the prediction base are increasing or stable,⁸ as is the case for Baltic and Central/Eastern European countries. We therefore modified the forecasts so that the multiplicative components of the estimated period and cohort effects, i.e., $\exp(f_P)$ and $\exp(f_C)$, were replaced by the corresponding first-order approximation of the Taylor series expansion for the exponential, i.e., $(1 + f_P)$ and $(1 + f_C)$. This modification produces more conservative projections as $(1 + x)$ is always lower than $\exp(x)$. Fitted and projected ASRs were obtained by recombining the estimated functions for age, period and cohort and using the female country- and age-specific population estimates (up to 2040) produced by the UN Population Division.

Forecasts in a scenario where screening programmes will improve starting from 2017

These projections were developed via an equivalent extrapolation of period and cohort effects, but with an additional assumption that CC screening programmes will improve starting from 2017 in each of the six BCEE countries. We hypothesised that the gradual impact on the CC rates would be equivalent to that observed in Denmark upon early-implementation of high-quality screening from the late-1960s. In this country, regional CC screening started in 1962, expanding progressively so that the programmes covered approximately 40% of the population in 1967, and national roll-out was completed in 1996.⁹ Relative to the other Nordic countries, the largest declines in period-specific effects were estimated in Denmark.¹⁰⁻¹² The key assumption is that declines in period-specific effects represent the beneficial effect of screening activities on CC incidence rates whereas the absence of such period effects, as observed in the BCEE countries, represents a lack of efficient screening activities. The period-specific effects estimated for Denmark thus represented the maximum achievable impact that mass screening programmes could have in a population, if implemented. National projections were thus based on a scenario where high-quality screening will be implemented in 2017 on extrapolating the age f_A and cohort effects f_C into the future and while obtaining the period effects in a scenario with screening improvements $f_{P,scr}$ by projecting the equivalent period effects only until 2017 and, thereafter, projecting a declining trend with a slope equal to that estimated for period effects in Denmark after the introduction of screening in the period from the late-1960s until 2010 (Appendix 2a,b). As in the scenario without improvement of screening programmes, the multiplicative components of the estimated cohort effects $\exp(f_C)$ were replaced by the corresponding first-order approximation of the Taylor series expansion for the exponential, i.e., $(1 + f_C)$. We have, however, maintained the standard exponential transformation $\exp(f_{P,scr})$ for the corresponding multiplicative components, as this alternative scenario imposes the period effects to decline, instead of increasing or being constant.

Detailed Results.

Figure 1a. Effects of age, period and cohort as estimated by the age-period-cohort models (APC), Denmark

Age effects are shown on a rate per 100,000 scale; cohort and period effects are on a relative risk scale. We assumed that significant declines in period effects estimated by the APC models in Denmark were owing to the effects of cervical cancer screening. Join-point regression was used to identify time points where the significant decline in the slope of period trends occurred. The significant join-point was identified to correspond to year 1965. The slope of the log rate ratios for period effects subsequent to 1965 has been estimated to be equal to -0.1878 (dot-dashed green line) and used to build a scenario for Baltic and Central/Eastern European countries where screening will be implemented. We hypothesized a starting date in year 2017 and a subsequent gradual impact on the cervical cancer incidence rates similar to that observed in Denmark after the beginning of screening activities in the late 1960s.

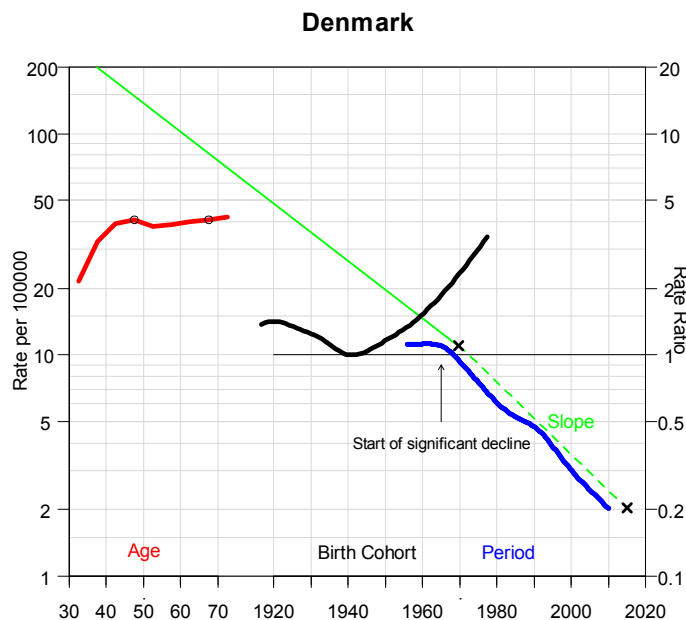


Figure 1b. Effects of age, period and cohort as estimated by the age-period-cohort models

Age effects are shown on a rate per 100,000 scale; cohort and period effects are on a relative risk scale. Under a no screening scenario, restricted cubic splines were used to project period and cohort effects into the future (dotted black and

blue lines). Under a scenario where screening activities will be implemented with a progressive impact similar to that observed in Denmark; starting from 2017, period effects are forced to decline with a slope that is set to be equal to that computed for period effects in Denmark (dot-dashed green lines).

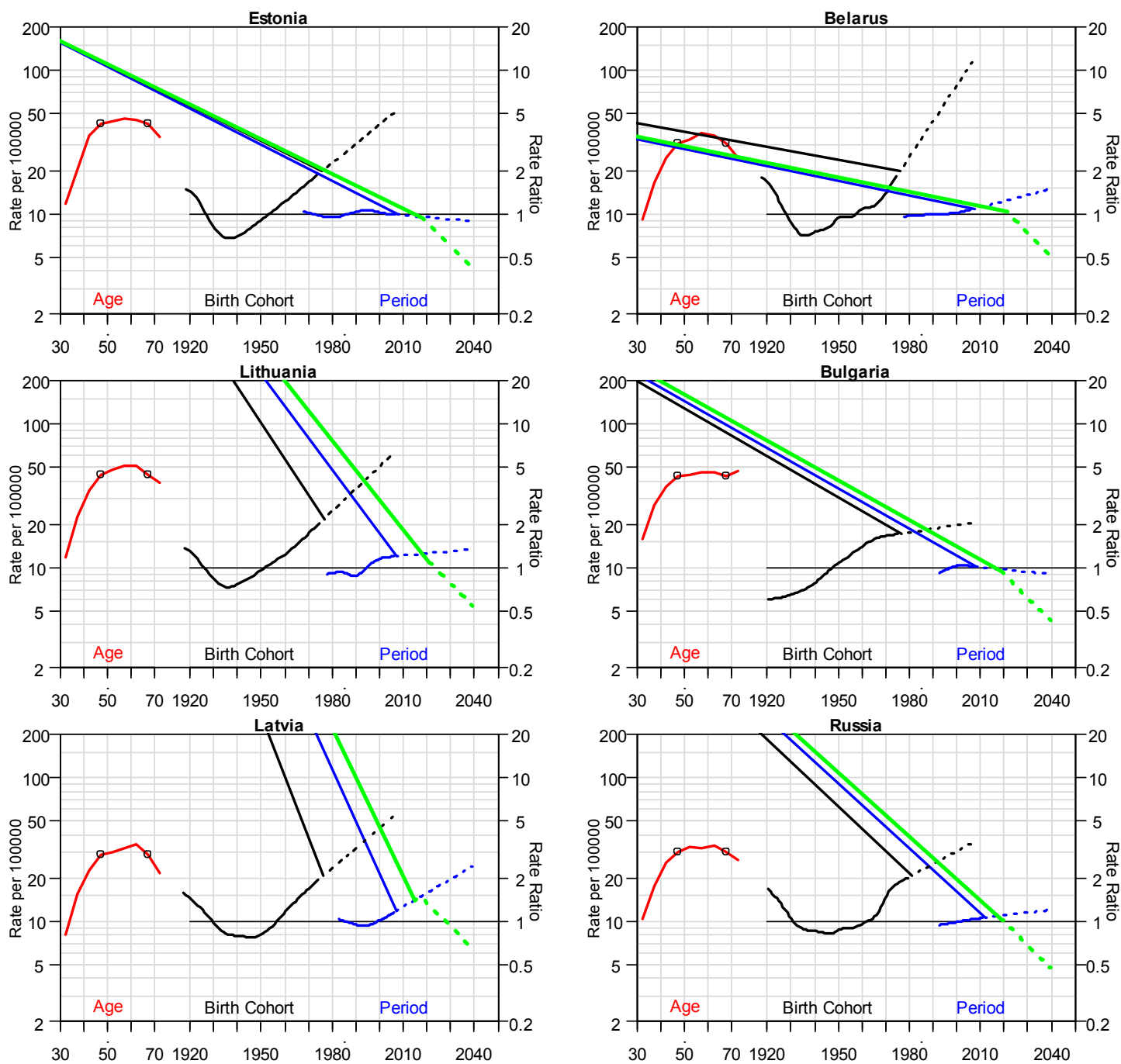
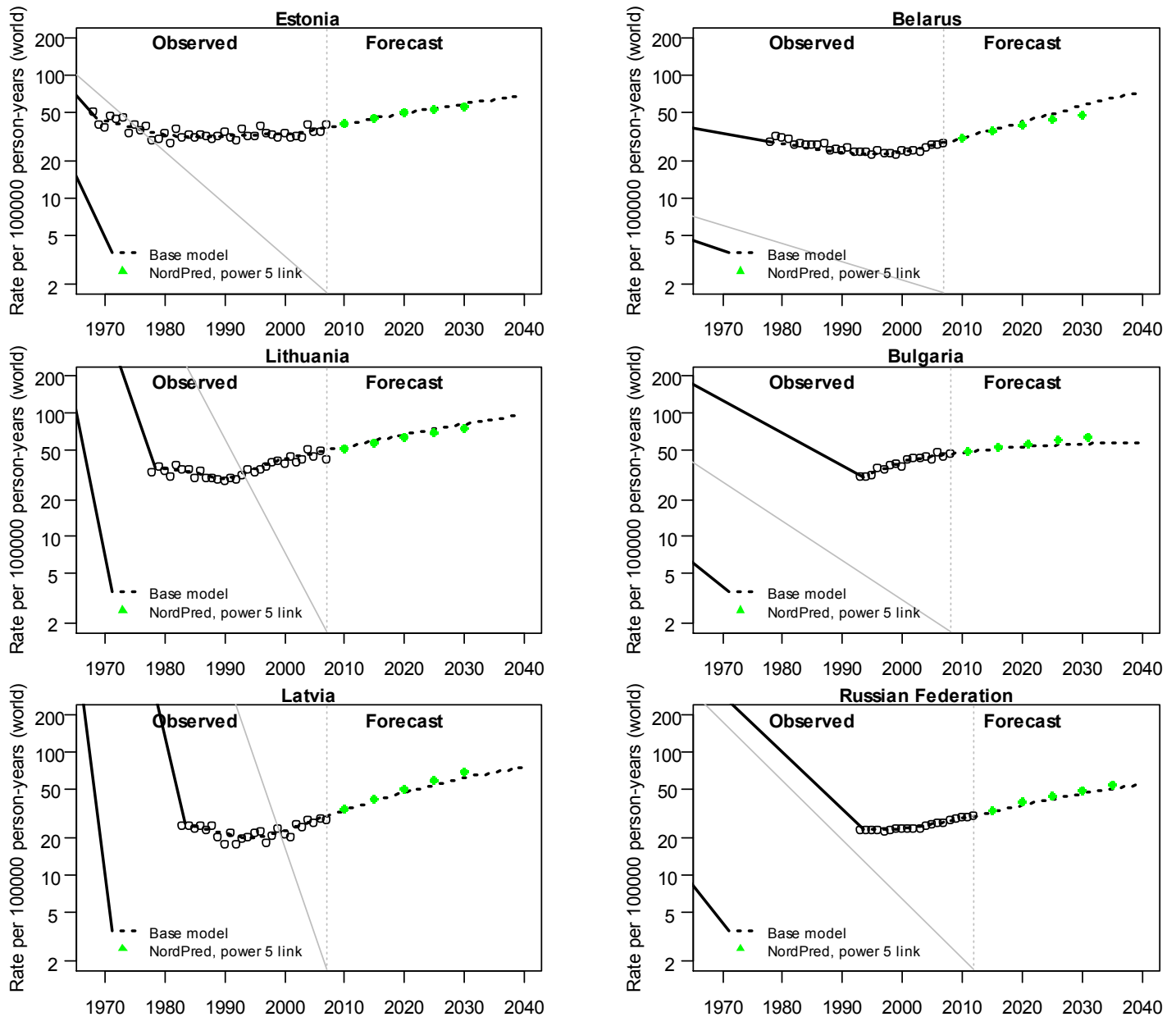


Figure 1c. Observed and projected incidence rates of cervical cancer, six Baltic and Central/Eastern European countries according to the base model and to Nordpred

Sensitivity analyses have been performed to compare our method, referred to here as the “base model” with the Nordpred model. In this method, commonly used and implemented in the R-package, the default link function is the power 0.2 function (i.e., $g^{-1}(x) = x^5$), instead of the standard log function. The purpose of the power link function is to level off the exponential growth of the projected rate. With this approach, however, as with other approaches using link functions other than the canonical log link, the estimated quantities $\exp(f_P)$ and $\exp(f_C)$ cannot be interpreted as relative risks. A prediction base of the most recent 25-year observed period (five x 5-year periods; in Bulgaria and Russia, however, were only available 15 and 20 years, respectively) was extrapolated to project cervical cancer incidence for 25 years into the future.



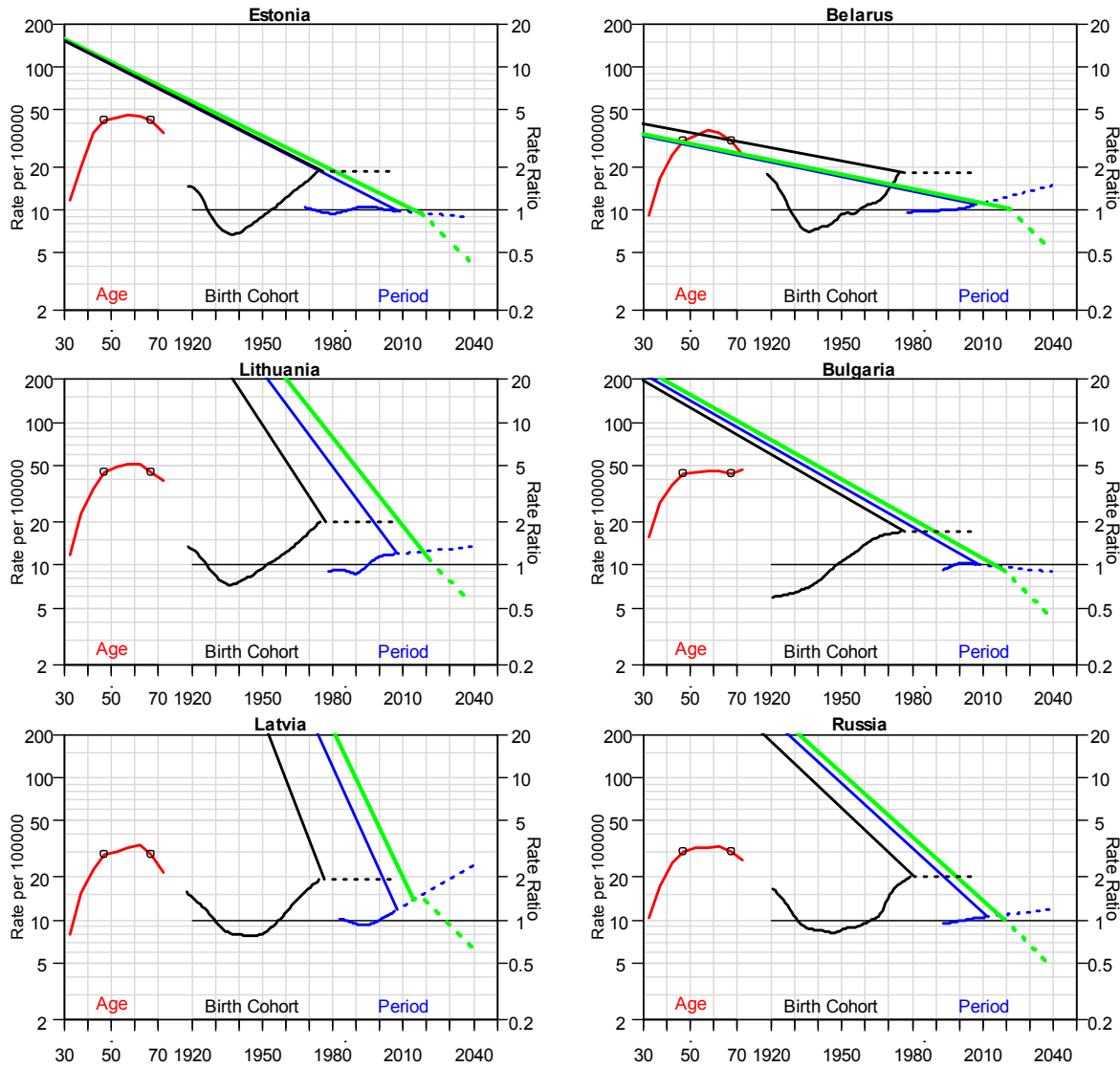
Forecasts of cervical cancer incidence trends using NORPRED were similar to those of obtained using our approach, with percentage differences between the two methods remaining below 10% for most of the predictions at future time points. Possible differences were noted in the long-term (around 2030), where our method produced higher predictions in Belarus and lower projections in Bulgaria, Latvia, and the Russian Federation, although the relative difference between the two methods never exceeded 20%.

Sensitivity analyses

Constant future cohort-effects

A sensitivity analysis, assuming that the cohort risk will remain as estimated for the most recent cohort observed (i.e., midpoints of 1975 for Estonia, Lithuania, Latvia and Belarus; 1976 for Bulgaria; 1979 for the Russian Federation) – as opposed to a continuing increase based on an extrapolation from the most recent generational trend – was undertaken. The age effects were assumed fixed as were period effects in both scenarios (with and without screening). Figure 2a below displays the assumed age, period and cohort effects:

Figure 2a. Age, period and cohort model parameters from an APC model based on the assumptions above



With this alternative assumption, generational risks are constant in the future and, consequently, the rates and the number of cervical cancer cases would be lower in both screening and no-screening scenarios. As expected, the discrepancy is more pronounced in countries where there are strongly positive trends in risk in recent cohorts (e.g., Lithuania, Belarus), as opposed to countries with modest increases (e.g., Bulgaria). The absolute number (and the proportion) of potentially

preventable cases of cervical cancer remains very high, although of a lower magnitude to that assuming continuously increasing cohort effects, as seen in Table 1a below.

Table 1a. Populations included, time span, mean annual number of cases and person-years, observed and projected age-standardised rates under different scenarios, ages 30-74

Country	Projected					
	ASR 2036-2040		Cumulative number of incident cases 2017-2040			
	No screening	Screening from 2017	No screening	Screening from 2017	Potentially preventable by screening	
			N	N	N	%
Estonia	54.1	27.0	4,285	3,189	1,096	25.6
Lithuania	71.2	36.0	14,451	10,246	4,205	29.1
Latvia	57.3	25.6	7,092	4,618	2,474	34.9
Belarus	48.2	22.3	29,067	19,205	9,862	33.9
Bulgaria	53.1	27.0	29,434	21,252	8,182	27.8
Russian Fed.	45.5	21.4	428,117	288,409	139,708	32.6
Total					165,420	

¹Average annual figure. ² Person-years expressed in millions.

ASR = age-standardised (world standard population) incidence rates

Weaker impact of future screening-related period-effects (50% of that assumed in the main analyses)

We have also performed an analysis where the impact of screening-related period effects would expected to be half of that observed in Denmark. Figure 2b shows the assumed age, period and cohort effects. The impact of age and cohort effects would be the same as that used in the main analyses, whereas the slope of the period-effect would be half of that estimated for Denmark (dark green dot line, instead of green dot line).

With this assumption, the number of cases potentially preventable by screening would be lower, overall approximately half of what estimated under the assumption that the impact will be similar to that observed in Denmark, Table 1b.

Figure 2b. Age, period and cohort model parameters from an APC model based on the assumptions above

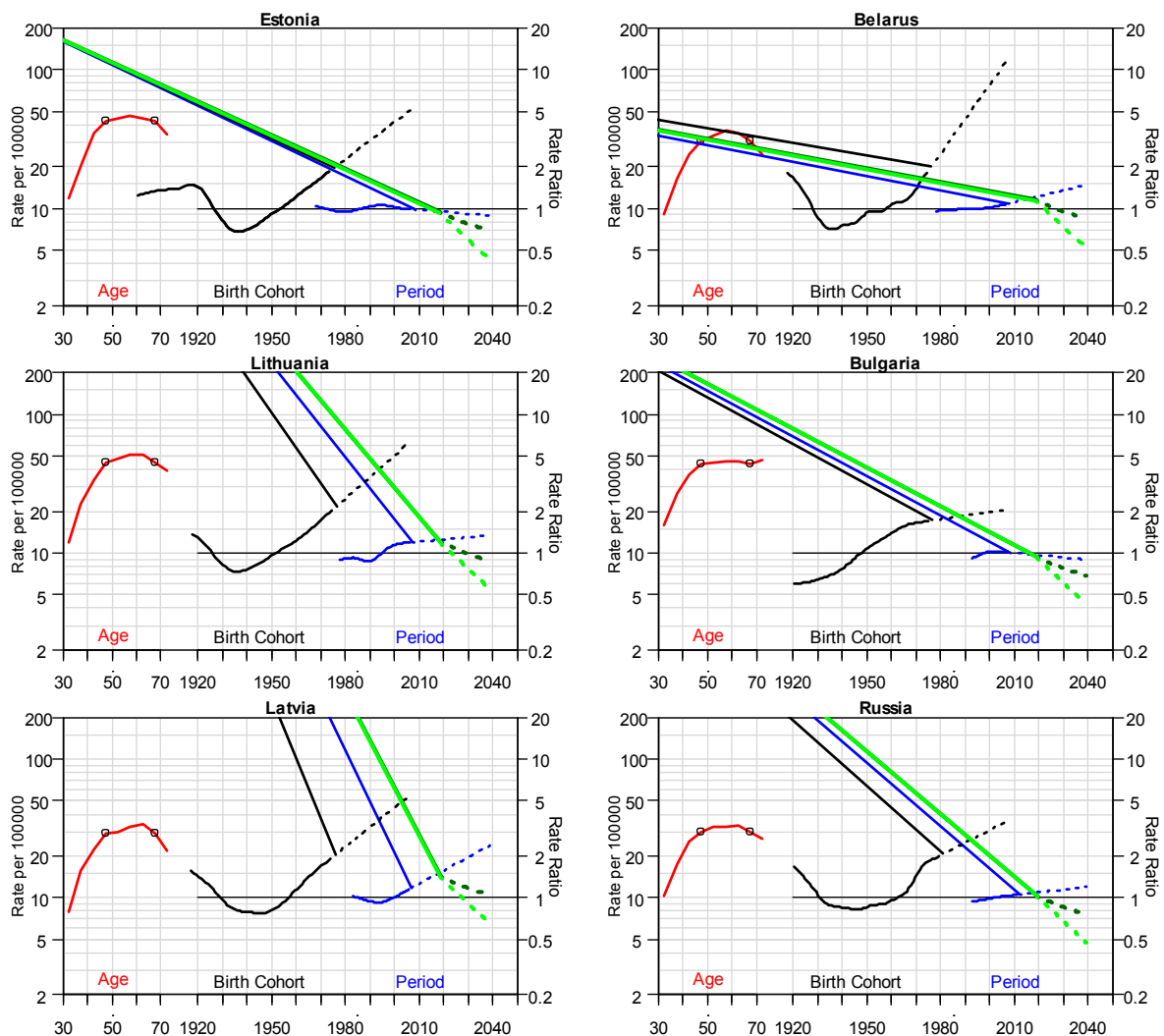


Table 1b. Populations included, time span, mean annual number of cases and person-years, observed and projected age-standardised rates under different scenarios, ages 30-74

Country	Projected					
	ASR 2036-2040		Cumulative number of incident cases 2017-2040			
	No screening	Screening from 2017	No screening	Screening from 2017	Potentially preventable by screening	%
Estonia	64.4	43.2	N	N	N	15.1
Lithuania	87.5	64.4	4,853	4,118	735	13.6
Latvia	68.4	40.3	16,105	13,913	2,192	20.0
Belarus	67.2	46.2	7,773	6,215	1,558	20.4
Bulgaria	55.1	39.0	34,911	27,800	7,111	12.1
Russian Fed.	50.2	33.4	29,967	26,337	3,630	17.6
Total			452,173	372,577	79,596	17.6

¹Average annual figure. ² Person-years expressed in millions.

ASR = age-standardised (world standard population) incidence rates

References

- 1 Clayton D, Schiffers E. Models for temporal variation in cancer rates. I: Age-period and age-cohort models. *Stat Med* 1987; **6**: 449-67.
- 2 Clayton D, Schiffers E. Models for temporal variation in cancer rates. II: Age-period-cohort models. *Stat Med* 1987; **6**: 469-81.
- 3 Holford TR. Age-period-cohort analysis. In: ARMITAGE P, Colton T. *Encyclopedia of Biostatistics*. 1st ed ed. Chichester, UK: John Wiley & Sons, 1998: 82-99.
- 4 Carstensen B. Age-period-cohort models for the Lexis diagram. *Stat Med* 2007; **26**: 3018-45.
- 5 Holford TR. The estimation of age, period and cohort effects for vital rates. *Biometrics* 1983; **39**: 311-24.
- 6 Plummer M, Peto J, Franceschi S. Time since first sexual intercourse and the risk of cervical cancer. *Int J Cancer* 2012; **130**: 2638-44.
- 7 IARC. CI5 - Cancer Incidence in Five Continents. <http://ci5.iarc.fr/> 2012: Available at <http://ci5.iarc.fr/> Accessed January 2012. Available from: URL: <http://ci5.iarc.fr/>
- 8 Moller B, Fekjaer H, Hakulinen T, et al. Prediction of cancer incidence in the Nordic countries: empirical comparison of different approaches. *Stat Med* 2003; **22**: 2751-66.
- 9 Anttila A, Ronco G, Clifford G, et al. Cervical cancer screening programmes and policies in 18 European countries. *Br J Cancer* 2004; **91**: 935-41.
- 10 Vaccarella S, Lortet-Tieulent J, Plummer M, Franceschi S, Bray F. Worldwide trends in cervical cancer incidence: Impact of screening against changes in disease risk factors. *Eur J Cancer* 2013; **49**: 3262-73.
- 11 Vaccarella S, Franceschi S, Engholm G, Lonnberg S, Khan S, Bray F. 50 years of screening in the Nordic countries: quantifying the effects on cervical cancer incidence. *Br J Cancer* 2014; **111**: 965-9.
- 12 Bray F, Loos AH, McCarron P, et al. Trends in cervical squamous cell carcinoma incidence in 13 European countries: changing risk and the effects of screening. *Cancer Epidemiol Biomarkers Prev* 2005; **14**: 677-86.