SHORT-TERM PREDICTION OF CORONARY HEART DISEASE MORTALITY IN THE CZECH REPUBLIC BASED ON DATA FROM 1968–2014

Jindra Reissigová[1], Miroslav Zvolský[2]

[1]Institute of Computer Science, Czech Academy of Sciences, Prague, Czech Republic
[2]Institute of Health Information and Statistics, Prague, Czech Republic

SUMMARY

Objectives: The aim was to explore the patterns of the coronary heart disease (CHD) mortality rates over the past almost 50 years (1968–2014) in the Czech Republic, and to predict the mortality rates in 2015–2019.

Methods: The number of deaths from CHD and the population size were stratified by sex and age. The mortality rates were age-standardized to European population. Their values in 2015–2019 were estimated using the joinpoint log-linear regression, local log-linear regression and negative binomial log-linear regression, separately for males and females.

Results: A positive change in the trend of the age-standardized mortality rates from CHD was detected after the collapse of communism in 1989. In 1991–2000, the mortality trend was sharply downward, with an annual percent change of −5.8 % for males and −5.2 % for females. In 2000–2014, the decreasing trend was not so sharp (−1.3 % for males and −0.7% for females), yet it should continue in 2015–2019. The crude mortality rates for females are slightly higher than those for males since 2007, however, they are increasing for both sexes. The mortality rates are rising mainly in the age group of 85+ years (in 2014, 25.4% of CHD deaths of males and 54.4% of females occurred at the age of 85+ years).

Conclusions: The age-standardized mortality rates are predicted to decrease in 2015–2019, but the crude mortality rates should increase due to increase in average life expectancy. The burden of deaths is moving to the age group of 85 years and older, mainly in females. A total of 26,039 CHD deaths were registered in the Czech Republic in 2014, and 29,653 are predicted for 2019, if the current trends continue.

INTRODUCTION

National health registries and statistical information systems provide epidemiological surveillance of population health status. Essential management tools lie in monitoring and evaluation of occurrence of diseases, deaths and related events. Data are used not only to calculate basic population indicators of morbidity and mortality, but also to predict their future development, either short-term (usually up to five years) or long-term (usually more than five years). Predictions are useful for planning in healthcare (e. g. preventive programmes, acute care, long-term care, rehabilitation), but also in social security, pension, and the pension fund plan. Predictions may also have scientific importance when exploring the consequence of, e.g., changes of a population risk profile and screening programme.

The national health registries in the Czech Republic, defined in the Act No. 372/2011 Coll., on health services and conditions of their provision, are under the administration of the Institute of Health Information and Statistics of the Czech Republic (IHIS CR) that also uses demographic and mortality data processed by the Czech Statistical Office (CZSO). IHIS CR issues each year a series of publications (recently mainly electronic), in which pro-

vides absolute and relative indicators of morbidity and mortality, classified by age, sex and diagnosis in the Czech Republic and so on, but without predictions of the future development of morbidity and mortality. However, the Czech Society for Oncology (CZO) in cooperation with the Institute of Biostatistics and Analyses (IBA) uses those data for short-term predictions of treatment costs of selected tumors, which among other things include the prediction of mortality, incidence and prevalence of the diseases (1). On the other hand, the future development of certain diseases has also been published by international organizations not only for the Czech Republic but also for other countries. Cancer mortality predictions both up to five years and more than five years are provided by the International Agency for Research on Cancer (2), predictions of mortality by cause for years 2015 and 2030 by the World Health Organization (3), and predicted numbers of CHD deaths up to 2020 by the European Heart Network (4).

This article draws on data from IHIS CR and CZSO and focuses on more detailed analysis of the trend in the mortality rates from coronary heart disease (CHD) in the Czech Republic (1968–2014) and its predictions to 2019 (short-term prediction) for males and females. Due to data collection process the latest mortality data available to the public are two or three years old not only in the

Czech Republic but in many countries. In fact, the short-term predictions usually estimate mortality data in the current year. We present not only the age-standardized and age-specific mortality rates but also the crude mortality rate. The crude mortality rate does not reflect the age structure of the population, and therefore its values have meaning for the management, organization and planning. The age-standardized and age-specific mortality rates suppress the influence of age on mortality, and thus allow comparisons of populations that differ in age structures.

Cardiovascular diseases have been the leading cause of death not only in the Czech Republic but in many other countries for many years, and coronary heart disease is responsible for more than half of the deaths (5). There are few studies predicting cardiovascular diseases trends in comparisons with a number of studies dedicated to predictions of cancer trends.

MATERIALS AND METHODS

Study Population

The numbers of deaths from CHD in 1968–2014 in the Czech Republic and corresponding the mid-year (1 July) population estimates for 1968–2014 and the beginning (1 January) population estimates for 2015–2020 were obtained from IHIS CR and from the CZSO publications on population movements* (accessed 28 February 2018). The data in the individual years was stratified by sex and the five-year age groups (0–4, 5–9, ..., 80–84, 85 and over). Coronary heart diseases in the Czech Republic were coded according to the International Statistical Classification of Diseases and Related Health Problems (ICD-revision number) as D410-D414 in 1968–1978 (ICD-8), 410-414 in 1979–1993 (ICD-9) and I20-I25 from 1994 to the present (ICD-10).

Statistical Analysis

The mortality rates were age-standardized by the direct method to the European population (6). The age-standardized mortality rates in 2015–2019 were estimated on the basis of a joinpoint log-linear regression (Appendix 1) and a local log-linear regression (i.e. a state-space model, Appendix 2). The age-specific and crude mortality rates were predicted using a negative binomial log-linear regression (Appendix 3), which was also used to derive the age-standardized mortality rates. Together with the point estimates of the mortality rates, the 95% prediction intervals (PI) are presented for 2015–2019.

The suitability of the individual models, which were derived separately for females and males, was measured on the basis of residual analysis, residual sum of squares or deviance and the Akaike and Bayesian information criterions. The accuracy of the methods was also compared using retrospective predictions. It means that the estimated age-standardized mortality rates in 1978–2014 were compared with those observed and the absolute relative differences (ARD) between the observed and estimated mortality rates (ARD = observed-estimated/observed) were calculated. The Friedman test was used to test for statistical difference in medians among the models. The Bonferroni method was used for multiple comparisons. All the numerical operations were done using R-software (7).

RESULTS

Figure 1 shows the age-standardized mortality rates from CHD in the Czech Republic from 1968–2014 and their predicted values in 2015–2019, separately calculated for males and females. In Figure 1A, the mortality rates were approximated by the joinpoint regression models. Those are piecewise log-linear regressions
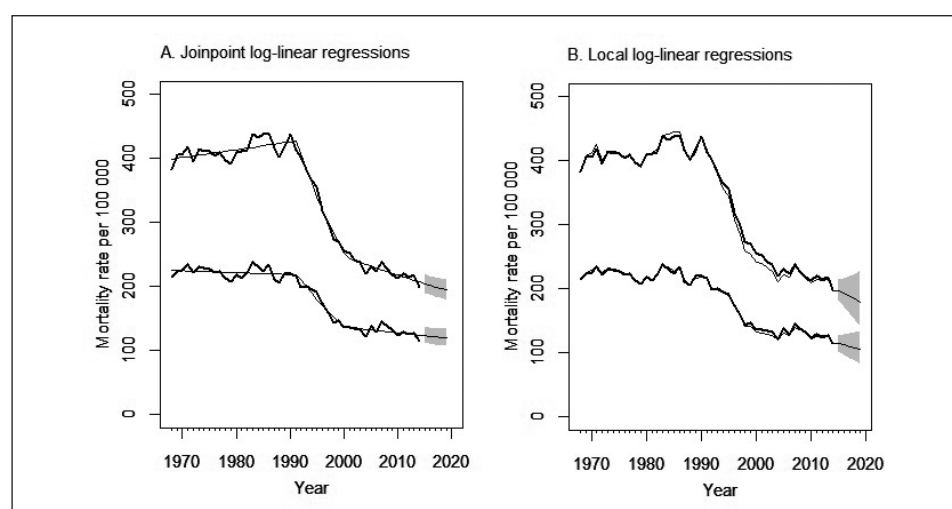

Fig. 1. *Age-standardized mortality rates from CHD (1968–2014) in the Czech Republic and their predictions with 95% PI (2015–2019).*
Males: higher level, females: lower level, observed: bold lines, estimated: thin lines

*https://www.czso.cz/csu/czso/czech-demographic-handbook-2014

that changed the slope in 1991 (p<0.001) and 2000 (p<0.001) for both sexes. While for males, the trend significantly increased in 1968–1991 (p=0.004), for females, the trend non-significantly decreased (p=0.306), with the estimated annual percent changes of 0.3% for males and −0.1% for females. For both sexes, the trends significantly declined in both 1991–2000 and 2000–2014, with the estimated annual percent changes of −5.8% (p<0.001) for males and −5.2% (p<0.001) for females in 1991–2000 and those of –1.3% (p<0.001) for males and −0.7% (p=0.018) for females in 2000–2014.

The development of mortality rates was also predicted by the local log-linear regression (Fig. 1B). The term "local" means that the parameters of the local log-linear regression model varied depending on the time (they were not constant for the entire period).

Both above methods estimated a decline in the age-standardized mortality rates in 2015–2019, however, the higher mortality rates were predicted by the joinpoint regression (on average, by 6.0% for males, by 10.6% for females). The predictions were more stable for the joinpoint regression (the narrower prediction intervals) and the accuracy were rapidly decreasing with time for the local regression (the widths of the intervals increased with year).

The numbers of CHD deaths by age and gender from 2015 to 2019 were estimated using the negative binomial log-linear regression models (Fig. 2). In 2014, 25.4% of 12,603 deaths in males were in the age-group of 85 years and over, and 54.4% of 13,436 in females. The mortality rates are increasing in the age group of 85 years and over from the end of the nineties of the last century, and the increase should continue in 2015–2019 for both sexes. Consequently, the total number of CHD deaths increased by 11.4% from 2000 (23,384) to 2014 (26,039), and a further increase by 13.9% is predicted from 2014 (26,039) to 2019 (29,653), if the current 10-year trends continue and the predicted population structure in 2019 is correct.

The crude mortality rates in females are slightly higher than those of males since 2007, but increasing in both sexes (Fig. 3).

The age-standardized mortality rates in 2015–2019 estimated using negative binomial log-linear regressions were nearly
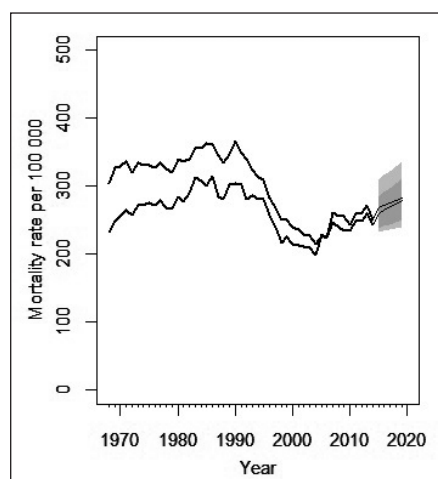
Fig. 3. *Crude mortality rates from CHD (1968–2014) in the Czech Republic and their predictions with 95% PI (2015–2019) calculated using negative binomial log-linear regressions.*
Males: higher level until 2007, females: lower level until 2007, observed: bold lines, estimated: thin lines

identical to those in Figure 1A, and for this reason, they are not presented here.

The retrospective short-term predictions of the age-mortality rates are shown in Figure 4. In males, the median of ARD ranged (from the first to the fifth year) from 3.5–8.9% for the joinpoint regression, 3.9–10.2% for the local regressions, and 4.2–11.2% for the negative binomial model, and, in females, from 5.0–11.3%, 4.6–14.9%, and 4.6–11.2%, respectively. The accuracy of the predictions was dependent on the year of predictions; the most reliable estimates were in the first three years after the base period. Overall, the differences in the median of ARD among the models were small, but statistically significant for males (p=0.043 for males, p=0.056 for females). For males, the least accurate model was the negative binomial model in the fourth and the fifth year, and for females, the local regression in the fifth year.
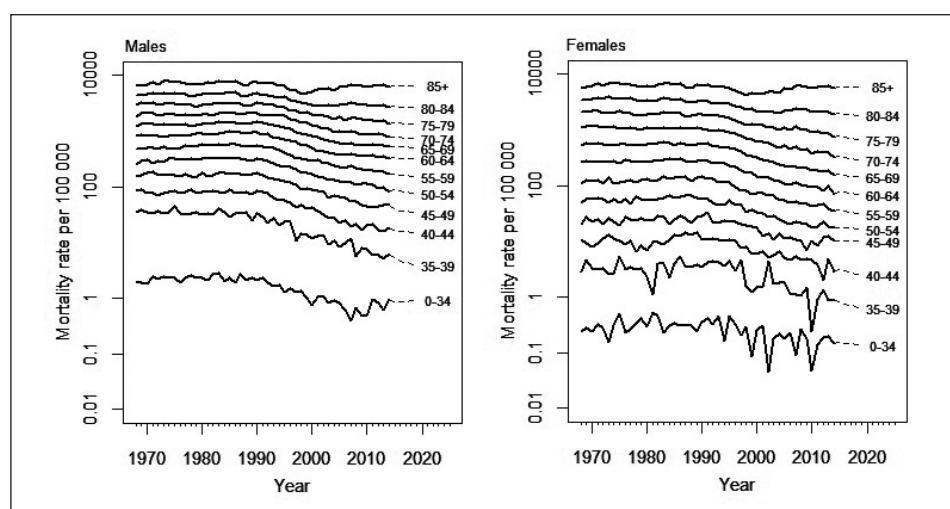
Fig. 2. *Age-specific mortality rates from CHD (1968–2014) in the Czech Republic and their predictions (2015–2019) calculated using negative binomial log-linear regressions.*
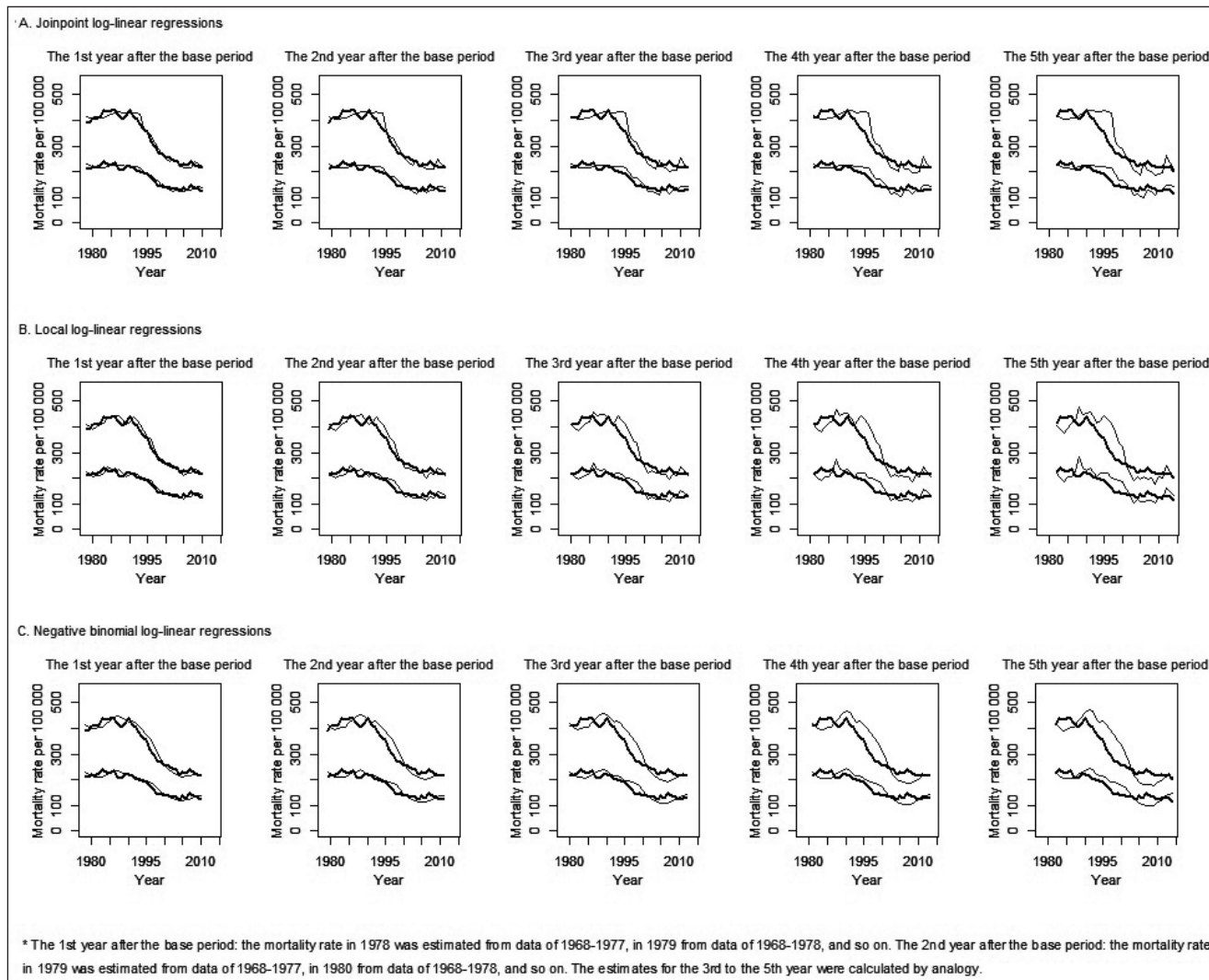95% PI omitted due to lack of space in the picture

A. Joinpoint log-linear regressions

The 1st year after the base period The 2nd year after the base period The 3rd year after the base period The 4th year after the base period The 5th year after the base period

B. Local log-linear regressions

The 1st year after the base period The 2nd year after the base period The 3rd year after the base period The 4th year after the base period The 5th year after the base period

C. Negative binomial log-linear regressions

The 1st year after the base period The 2nd year after the base period The 3rd year after the base period The 4th year after the base period The 5th year after the base period

* The 1st year after the base period: the mortality rate in 1978 was estimated from data of 1968–1977, in 1979 from data of 1968–1978, and so on. The 2nd year after the base period: the mortality rate in 1979 was estimated from data of 1968–1977, in 1980 from data of 1968–1978, and so on. The estimates for the 3rd to the 5th year were calculated by analogy.

Fig. 4. *Age-standardized mortality rates from CHD in the Czech Republic and their retrospective predictions.*
Males: higher level, females: lower level, observed: bold lines, estimated: thin lines

DISCUSSION

Morbidity and mortality prediction methods have changed over the years as both data quality and statistical models improved. When choosing a prediction method, it must be taken into account what is predicted (rates, numbers of cases), how much historical data is available, whether the numbers of cases are low or high, and how many years are predicted (short-term predictions, long-term predictions). A series of articles comparing different prediction methods have been published (8–10).

To our knowledge, excluding short-term predictions of treatment costs of selected tumors and their corresponding indicators such as mortality and incidence (using Poisson modelling) and prevalence (using multi-component modelling) (1), there are neither short-term nor long-term predictions of morbidity and mortality regularly published in the Czech Republic. In contrast, for example, the American Cancer Society (ACS) has been publishing the estimated number of new cancer cases and deaths in the current year since 1960. The prediction methods of ACS were changed from linear predictions to autoregressive quadratic time trend models, and further to state-space models and to joinpoint

regression models (11, 12). Another example are the Nordic cancer registries (Denmark, Finland, Iceland, Norway, and Sweden) that have been reporting long-term cancer predictions based on age-period-cohort generalized regression models (9).

Our aim was to predict the mortality rates in the Czech Republic for 2015–2019. We focused on the short-term predictions that tend to be more accurate than long-term predictions. The calculations were based on the number of deaths stratified by age and year of death and the corresponding population size, and not taken account of cardiovascular risk factors. Common for here applied statistical methods is the assumption that the current trends in the mortality rate will continue in the future. It is conditional that trends in factors associated with mortality of CHD (e.g., risk and protective cardiovascular factors, treatment methods) will continue by 2019. We age-standardized the mortality rates to European population, which is still being used, by, e.g., IHIS CR and WHO, to provide comparisons among populations with different age structures and over time (the choice of standard population affects the values of standardized mortality rates).

The joinpoint regression proved to be advantageous for short-term predictions since it can account for sharp changes (10).

The local regression is flexible to sudden changes because has time-varying coefficients, allowing the model to adjust to sudden changes (11, 13). Both these methods predicted a decline in the age-standardized mortality rates from CHD in the Czech Republic for both sexes. The local regressions, however, predicted lower values than the joinpoint regressions. This was due to the increased sensitivity of the local regression to recent values: the age-standardized mortality rate declined from 2013 to 2014 by 8.7% in males, and by 10.3% in females. It may (or may not) be random fluctuations in the mortality rates.

A significant positive change in the trend of the age-standardized mortality rates from CHD in the Czech Republic was detected after the collapse of communism in 1989. Since then until the beginning of the 21st century, the mortality trend was sharply downward. The decreasing trend was not so sharp after that, but the downward trend should continue in 2015–2019. The difference between the age-standardized CHD mortality rates of males and females in the Czech Republic has been gradually decreasing over the nearly fifty years.

Unlike the age-standardized mortality rates, the crude mortality rates (i.e. the total numbers of deaths to the total size of the population) should increase in coming years. No essential gender differences in the crude mortality rates were observed in the recent years. The increasing numbers of deaths can be associated with an increase in average life expectancy in the Czech Republic. The average life expectancy was increasing in 1990–2014 from 67.6 to 75.2 years for males, and from 75.4 to 81.1 years for females (14).

Any age-standardization of mortality rates masks potential differences among age-specific mortality rates. Age-period-cohort generalized models have been using the age-specific predictions. Those models have lots of versions, e.g. for increasing trends to avoid unrealistic long-term predictions (8, 9). Age, period and birth cohort serve as surrogate variables for risk and protective factors, which are not directly accounted for. We based the short-term prediction on data from a shorter historical period, in which it is not possible to adequately assess the cohort effect, and an age-period generalized model was applied. The shorter (the last 10 years) historical period was proved to increase prediction accuracy of those models (9). Note that all three applied models (the joinpoint, local and generalized regressions) reflected the trend observed in the recent period, but each method in different way.

All predictions are uncertain, and their results should be always interpreted with caution. When predictions fail, it may be caused by either an inappropriate choice of statistical models or unexpected changes of default (base) conditions, e.g., development of a risk profile, new forms of treatment, preventive programmes, or data quality. The reasons of failure should always be analyzed retrospectively. For example, the median of the absolute relative differences (ARD) between the total observed and predicted number of cancer in the Nordic countries were typically 10–20% depending on the length of the base period, the prediction method and the length of prediction (9). Here presented retrospective five-year predictions showed the biggest inaccuracies in 1990–2000, when there was an unpredictable sharp decline in the mortality rates after the collapse of communism in 1989. Overall, the median of ARD between the observed and those predicted age-specific mortality rates ranged from 3.5–11.3% for the joinpoint regressions, from 3.9–14.9% for the local regressions, and from 4.2–11.2% for the negative binomial regressions. Our results can be also partly af-

fected by quality and methodological changes in coding causes of death, and in particular by shifts from acute to chronic coronary heart diseases. While deaths from heart attack are relatively well recognized, the causes of chronic heart failures are not. A doctor can determine as the cause of death in the range I20-I25 (CHD), or I50 (heart failure), or even any related complications.

There are few studies predicting future cardiovascular mortality or morbidity trends. Mortality rates from stroke and CHD were predicted up to 2003 in Sweden based on data from 1969–1993 (15, 16). The more recent study predicting CHD mortality rates up to 2030 was conducted in Finland (17). The burden of CHD in Finland is moving to older age groups and the total number is increasing similarly as in the Czech Republic. While the overall age-standardized mortality rates in the Nordic countries began to decrease due to cardiovascular prevention programmes, in the Czech Republic as a result of lifestyle and living standard improvements after the fall of communism in 1989, and the gradual integration of the Czech Republic into prevention programmes (e.g. into the WHO Country Integrated Noncommunicable Disease Intervention Program created in 1983).

A potential impact of future risk factors scenarios relating to smoking, physical activity, salt consumption, and saturated fat intakes on future CHD mortality in some European countries is presented in the European Heart Network report (4). Following this report, the number of CHD deaths in people aged 25–74 in 2020 will be 10,598 in the Czech Republic under assumption that trends remain constant since 2007. If mortality will change according to the forecast on current trends there will be 4,890 CHD deaths in that age group. To all these we may add, that a total of 7,354 deaths were observed in that age group in 2014.

Over the past 30 years, age-standardized mortality rates from CHD have been consistently falling not only in most Northern but also Western European countries. From around 2000 to 2005, the mortality rates are also decreasing in the majority of Central and Eastern European countries (18).

CONCLUSION

A decline in the (European age-standardized) mortality rates from CHD should continue in the Czech Republic in 2015–2019 for both sexes, but the number of deaths should be increasing. A total of 26,039 CHD deaths were registered in 2014 and 29,653 predicted for 2019, if the current trends continue.

Conflict of Interests
None declared

REFERENCES

1. Vyzula R, Adámková Krákorová D, Arenberger P, Babjuk M, Belada D, Dušek L, et al. The Blue book of the Czech Society for Oncology. 22nd ed. [Internet]. Brno: Masaryk Memorial Cancer Institute; 2016 [cited 2016 Sep 23]. Available from: https://www.linkos.cz/files/modra-kniha/15.pdf. (In Czech.)

2. International Agency for Research on Cancer. Online analysis, predictions [Internet]. Lyon: IARC [cited 2016 Sep 23]. Available from: http://www-dep.iarc.fr/WHOdb/predictions_sel.htm.
3. World Health Organization. Projections of mortality and causes of death, 2015 and 2030 [Internet]. Geneva: WHO [cited 2016 Sep 23]. Available from: http://www.who.int/healthinfo/global_burden_disease/projections.
4. Andersen K, Aspelund T, Bandosz P, Basak P, Bennett K, Bjorck L, et al. CHD mortality projections to 2020, comparing different policy scenarios: EuroHeart II work package 6 [Internet]. 2014 [cited 2016 Sep 23]. Available from: https://www.escardio.org/static_file/Escardio/EU-Affairs/chd-mortality.pdf.
5. Nichols M, Townsend N, Scarborough P, Rayner M. Cardiovascular disease in Europe 2014: epidemiological update. Eur Heart J. 2014;35(42):2950-9.
6. Doll R, Cook P. Summarizing indices for comparison of cancer incidence data. Int J Cancer. 1967 May 15;2(3):269-79.
7. R Development Core Team. R: a language and environment for statistical computing [Internet]. Vienna: R Foundation for Statistical Computing [cited 2016 Sep 23]. Available from: https://www.R-project.org/.
8. Dyba T, Hakulinen T. Comparison of different approaches to incidence prediction based on simple interpolation techniques. Stat Med. 2000;19(13):1741-52.
9. Møller B, Fekjaer H, Hakulinen T, Sigvaldason H, Storm HH, Talbäck M, et al. Prediction of cancer incidence in the Nordic countries: empirical comparison of different approaches. Stat Med. 2003;22(17):2751-66.
10. Lee TC, Dean CB, Semenciw R. Short-term cancer mortality projections: a comparative study of prediction methods. Stat Med. 2011;30(29):3387-402.
11. Tiwari RC, Ghosh K, Jemal A, Hachey M, Ward E, Thun MJ, et al. A new method of predicting US and state-level cancer mortality counts for the current calendar year. CA Cancer J Clin. 2004;54(1):30-40.
12. Pickle LW, Hao Y, Jemal A, Zou Z, Tiwari RC, Ward E, et al. A new method of estimating United States and state-level cancer incidence counts for the current calendar year. CA Cancer J Clin. 2007;57(1):30-42.
13. Hayat MJ, Tiwari RC, Ghosh K, Hachey M, Hankey B, Feuer R. Age-adjusted US cancer death rate predictions. J Data Sci. 2010;8(2):339-48.
14. Deaths 2013. Health statistics. Prague: Institute of Health Information and Statistics of the Czech Republic; 2015. (In Czech.)
15. Peltonen M, Asplund K. Age-period-cohort effects on stroke mortality in Sweden 1969-1993 and forecasts up to the year 2003. Stroke. 1996;27(11):1981-5.
16. Peltonen M, Asplund K. Age-period-cohort effects on ischaemic heart disease mortality in Sweden from 1969 to 1993, and forecasts up to 2003. Eur Heart J. 1997;18(8):1307-12.
17. Huovinen E, Härkänen T, Martelin T, Koskinen S, Aromaa A. Predicting coronary heart disease mortality-assessing uncertainties in population forecasts and death probabilities by using Bayesian inference. Int J Epidemiol. 2006;35(5):1246-52.
18. Nichols M, Townsend N, Scarborough P, Rayner M, Leal J, Luengo-Fernandez R, et al. European cardiovascular disease statistics. 2012 ed. Brussels: European Heart Network; 2012.
19. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. Stat Med. 2000;19(3):335-51.
20. Commandeur JJF, Koopman SJ. An introduction to state space time series analysis. Oxford: Oxford University Press; 2007.

Appendix 1: Joinpoint log-linear regression

The age-standardized mortality rates (m_t, $t=1968,…,2014$) were modelled by a log-linear continuous function that changes its slope in the so-called change points, i.e.

$$\log(m_t)=\mu+\beta t+\gamma_1(t-\tau_1)^+ + ... + \gamma_n(t-\tau_n)^+ + \varepsilon_t,$$

where μ is an intercept, $\beta, \gamma_1, …, \gamma_n$ are regression coefficients, $\tau_k (k=1, …, K)$ are change points for that pay $(t-\tau_k)^+=(t-\tau_k)$ for $(t-\tau_k)>0$, otherwise $(t-\tau_k)^+=0$, and ε_t is a random error term. We identified two change points ($\tau_1=1991, \tau_2=2000$) using a Monte Carlo permutation method (19). The future values of the mortality rates ($m_t, t=2015, …, 2019$) were obtained by the extrapolation of the trend in the last time segment (2000–2014).

Appendix 2: Local log-linear regression

The age-standardized mortality rates ($m_t, t=1968, …, 2014$) were described by a state-space regression model, specifically by a local log-linear regression model. This model assumes that locally (i.e. at each year), mortality has a log-linear trend, and that its coefficients are time varying. It means that

$$\log(m_t)=\mu+\varepsilon_t,$$
$$\mu_t=\mu_{t-1}+\beta_{t-1}+\eta_t,$$
$$\beta_t=\beta_{t-1}+\xi_t,$$

where μ_t is a local intercept, β_t is a local log-linear slope (drift), ε_t, η_i and ξ_t are independent random error terms (20). The predicted mortality rates ($m_t, t=2015, …, 2019$) were calculated by an extrapolation of trend under assumption that $\beta_t=\beta_{2014}$, $t=2015, …, 2019$.

Appendix 3: Negative binomial log-linear regression

The age-specific mortality rates $m_{at}=\frac{d_{at}}{n_{at}}$ were modelled as

$$\log(d_{at})=\mu+\alpha_a+t\beta_a+\log(n_{at})+\varepsilon_{at},$$

where d_{at} is the number of deaths in the age-group a and in the period t ($a=0–34, 35–39, …, 80–84, 85+, t=2005, …, 2014$), n_{at} is the corresponding mid-year population size, μ is an intercept level, α_a is an age parameter, β_a is a log-linear slope (drift) corresponding to age group a and ε_{at} is a random error term. The only term $t\beta_a$ was extrapolated to estimate the future age-specific mortality rates ($m_{at}, t=2015, …, 2019$) (9). The mid-year (1 July) population sizes n_{at}, $t=2015, …, 2019$ were estimated as the average of the corresponding beginning (1 January) population sizes obtained from CZSO.