4
Demografie

Year 2021
Volume 63

Review for Population Research

Jaroslav Kraus
Statistical Disclosure Control Methods for Harmonised Protection of Census Data: a Grid Case

Hana M. Broulíková – Matěj Kučera – Markéta Arltová
The Timely Diagnosis and Treatment of Alzheimer's Disease: Microsimulating Cost-Effectiveness in the Czech Republic

CZECH
STATISTICAL
OFFICE

ARTICLES

REPORTS

DIGEST

DATA

BIBLIOGRAPHY

The opinions of the authors do not necessarily reflect those of the Editorial Board.

Demografe is a peer-reviewed journal included in the citation database of peer-reviewed literature Scopus, and also in the Emerging Sources Citation Index (part of the Web of Science Core Collection).

STATISTICAL DISCLOSURE CONTROL METHODS FOR HARMONISED PROTECTION OF CENSUS DATA: A GRID CASE

Jaroslav Kraus[1)]

Abstract
The 2011 Population and Housing Census in the Czech Republic was accompanied by a significant change in the technology used to prepare course of the fieldwork, along with changes in how the data are processed and how the outputs are disseminated. Grids are regular polygon networks that divide the territory of country in a grid-like way/pattern into equally large territorial units, to which aggregate statistical data are assigned. The disadvantage of grids is that these are territorially small units that are often minimally populated. This mainly has implications for the protection of individual data, which is associated with statistical disclosure control (SDC).

The research question addressed in this paper is whether data protection (perturbation methods) leads to a change in the characteristics of the file either in terms of statistics of the whole file (i.e. for all grids) or in terms of spatial statistics, which indicate the spatial distribution of the analysed phenomenon. Two possible solutions to the issue of grid data protection are discussed. One comes from the Statistical Office of the European Communities (Eurostat) and the other from Cantabular, which is a product of the Sensible Code Company (SCC) based in Belfast.

According to the Cantabular methodology, one variant was processed, while according to the Eurostat methodology, two variants were calculated, which differ by the parameter settings for maximum noise D and the variance of noise V. The results of the descriptive statistics show a difference in absolute differences when Cantabular and Europstat solutions are compared. In the case of other statistics, the results are fully comparable. This paper is devoted to one specific type of census output. The question is to what extent these results are relevant for other types of census outputs. They differ fundamentally in the number of dimensions (grids have only two dimensions). It would therefore be appropriate to use SDC procedures that allow greater flexibility in defining SDC parameters.

Keywords: population and housing census, statistical disclosure control (SDC), grids

https://doi.org/10.54694/dem.0285 **Demografie,** 2021, **63 (4): 199–215**

INTRODUCTION

Population censuses are a fundamental demographic and statistical task that have long been organised in almost every country in the world. Census programmes are undoubtedly evolving, but the basis remains the same: a census is a survey

1) Czech Statistical Office, contact: jaroslav.kraus@czso.cz.

of the population, houses, flats, and households. The last Population and Housing Census in the Czech Republic took place in 2011 and was conducted in conformity with Regulation No. 763/2008 of the European Parliament and the Council of the European Union. Based on a proposal from the Czech government, the Parliament of the Czech Republic ordered a census by Act No. 296/2009 Coll.

Martin (*Martin*, 2011) evaluated gridded population models using the 2001 Northern Ireland census. He noted that there is a growing interest in the use of gridded population models, which potentially offer the advantages of stability over time and ease of integration with non-population data sources. High-resolution global gridded data for use in population studies were provided in (*Lloyd et al.*, 2017). Recent years have seen substantial growth in openly available satellite and other geospatial data layers, which represents a range of metrics relevant to mapping the global human population at fine spatial scales. Such datasets are vital for measuring the impacts of population growth, monitoring change, and planning policy interventions. (*Lloyd et al.*, 2019) mention the use of global spatio-temporally harmonised datasets to produce high-resolution gridded population distribution datasets. Multi-temporal, globally consistent, high-resolution human population datasets have been used to produce consistent and comparable population distributions to help map sub-national heterogeneities in health, wealth, and resource access, and monitor change in these areas over time. Finally, (*Doxsey-Whitfield et al.*, 2015) took advantage of the improved availability of census data to provide a first picture of the gridded population of the world.

Compared to past censuses, the 2011 Population and Housing Census introduced a relatively significant change in the procedure for preparing the census and in the actual course of the fieldwork, along with changes in how the data were processed and the outputs disseminated. Some methodological approaches have also changed and become more aligned with international recommendations (*CZSO*, 2011; 2013). Although a number of changes have been relatively widely discussed in the literature, one type of output remains somewhat overlooked: census results in a grid network.

In 2012 and 2013, the Czech Republic participated in a project of the European Communities (Eurostat) called Representing Census Data in the European population grid (Geostat). The aim of the project was to create a prototype of the European population grid compiled from national data sets of the results of censuses held around 2010 (in the Czech Republic in 2011) in all participating and cooperating countries and to describe the methodology for generating and displaying these data in the grid.

Three different methods were used to calculate statistical (attribute) data in grids. Because of its high accuracy and the quality of its outputs, the 'aggregation method' is the preferred approach. It is based on the assumption that georeferenced statistical microdata are widely available (provided with X, Y coordinates), with accuracy to the level of buildings and these data are then aggregated within individual grids. In the absence of such spatially localised statistical data, the values for individual squares are derived from the lowest territorial units for which the relevant statistical variables are still available (e.g. municipalities or census tracts); this method is called disaggregation. Finally, if georeferenced microdata are available for only a part of the studied area, then the 'hybrid method' is usually applied, which is based on a combination of the two previously described methods (*Kraus et al.*, 2014).

Grids are regular polygon networks that divide the territory of a country into equally large territorial units, to which aggregate statistical data are assigned (*Klauda*, 2011). In the case of a census, these are squares with an edge of 1 km and aggregations of mean data on the population, although there is nothing to prevent the assignment of data on houses, flats, or households as well.

It is the regular identical shape and thus the identical size of all the cells that is one of the main advantages of grids, which facilitates their mutual comparability in space – for example, across states. Another advantage is long-term stability over time, which contrasts with frequent changes in the definition of administrative units. Networks of squares enable the presentation of statistical data in a very detailed spatial resolution, which brings the advantage of an easier and more accurate analysis of territorial structures (*Kraus et al.*, 2014). However, each method has its advantages and disadvantages. In the case of grids, it is mainly that they do not coincide with territorial administrative boundaries. This is, of course, solvable, but always only to a certain extent. The second disadvantage is that these are territorially small units, which are often minimally populated, and this is primarily an issue for the protection of individual data - which is associated with statistical disclosure control (SDC).

The research question addressed in this paper is whether data protection (perturbation methods) leads to a change in the characteristics of the file:

- either in terms of the statistics of the whole file (i.e. for all grids), or
- in terms of spatial statistics, which indicate the spatial distribution of the analysed phenomenon.

The issue of SDC is relatively extensive and has been addressed by a number of authors. In this paper, the author often refers to the proceedings of (*Domingo-Ferrer et al.*, 2018), which contain documentation on this issue in relation to the census. A large amount of information, including legal aspects, can be found in (*Hunderpool et al.*, 2012), including calculation procedures for frequency tables. An illustrative way of measuring SDC results, including other useful information, is contained in (*Domingo-Ferrer et al.*, 2006). (*Templ*, 2017) has written a work that is devoted to methods and applications in R in the field of SDC. And (*Thijs et al.*, 2021) have written a practical guide that also deals with applications in R. There is, therefore, sufficient information available for anyone to create own approach to the issue.

Nevertheless, in this paper, two possible solutions to the issue of grid data protection will be discussed. One comes from the Statistical Office of the European Communities (Eurostat) and the other from Cantabular, which is a product of the the Sensible Code Company (SCC) based in Belfast. SensibleCode was involved as a partner in the UK's 2021 population census (*Company*, 2021).

METHODS AND METHODOLOGY

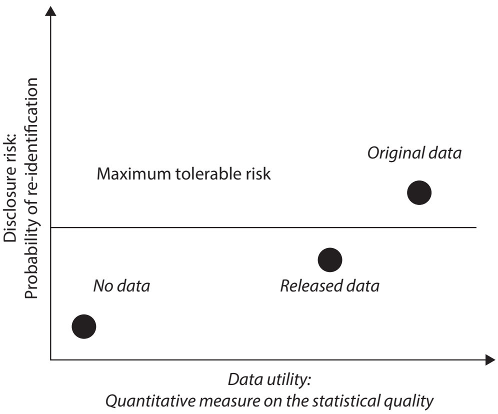
Statistical disclosure control (SDC) is a statistical field that has been developing dynamically in recent years and on which there is already enough good-quality literature. There are many reasons for this development. Disclosure control thinking has to keep up with increases in computing power, developments in matching software, and the proliferation of public and private databases. Statistical offices need to find the right balance between the need to inform society as much as possible, on the one hand, and the need to safeguard the privacy of the respondents on the other (*Hunderpool et al.*, 2012, p. xi). There are several reasons why statistical data protection should be respected. Above all, there are legal regulations that deal directly with the issue of SDC, such as Commission Regulation (EC) No. 831/2002 (*Eurostat*, 2002) of 17 May 2002 implementing Council Regulation (EC) No. 322/97 (*Eurostat*, 1997) on Community Statistics concerning access to confidential data for scientific Purposes, or Commission Regulation (EC) No. 223/2009 of the European Parliament and Council of 11 March 2009 on European statistics. However, there are also a number of other legally relevant documents that focus on this issue (*Domingo-Ferrer et al.*, 2012, pp. 23–35).

There are several ways to address the issue of SDC: traditional methods include tabular data protection or the protection of the output of statistical analyses, and modern methods include microdata protection. This paper is devoted to the latter, and specifically with respect to census output in a grid network. If you work with microdata, i.e. with individual records, then the methods for protecting these data can be divided into several groups. The purpose of all these efforts is to strike a balance between the risk of publishing detailed information and the usefulness of publishing that information.

When assessing SDC methods and their parameters for statistical outputs, an iterative process is carried out. For each method and its parameters, quantitative disclosure risk and information loss measures are calculated. These points can then be plotted on a *Disclosure Risk - Data Utility (R-U) Confidentiality Map*. The optimal SDC method to choose is the one that reduces the disclosure risk to tolerable risk thresholds while ensuring high quality data that are fit for purpose (*Shlomo et al.*, 2006, p. 69).

In the case of microdata, it is possible to define the principles for managing the confidentiality

Figure 1 R-U confidentiality map



Source: *Hunderpool et al.*, 2012, p. 5.

address this issue of confidentiality: Regulation 1588/90 or Regulation 322/97. For statistical disclosure control in the European Union, the following two laws are currently of importance: Commission Regulation (EC) No. 831/2002 and 322/97 on Community Statistics, concerning access to confidential data for scientific purposes.

The purpose of SDC for microdata is to prevent confidential information from being linked to specific respondents when a microdata file is being released. More formally, we can say that, given an original microdata set V, the goal of SDC is to release a protected microdata set **V'** in such a way that:

- the disclosure risk (i.e. the risk that a user or an intruder can use **V'** to determine confidential variables on a specific individual among those in **V**) is low;
- user analyses (regressions, means, etc.) on **V** and **V'** yield the same or at least similar results (*Hunderpool et al.*, 2012, p.23).
- There are two methods to create a protected microdata set **V'**:
- either by masking original data, i.e. generating a modified version **V'** of the original microdata set **V**;
- or by generating synthetic data **V'** that preserve some of the statistical properties of the original data **V**.
- Regarding masking methods, these can in turn be divided into two categories depending on their effect on the original data:
- *Non-perturbative masking:* Non-perturbative methods do not distort data; rather, they produce partial suppressions or reductions of detail in the original data set. Global recording, local suppression and sampling are examples of non-perturbative masking.
- *Perturbative masking:* The microdata set is distorted before publication. In this way, unique combinations of scores in the original data set may disappear and new unique combinations may appear in the perturbed data set; such confusion is beneficial for preserving statistical confidentiality. The perturbation method used should be such that statistics computed on the perturbed data set do not differ significantly from the statistics that would be obtained on the original data set (*Hunderpool et al.*, 2012, p. 33). The whole process of work also depends on whether they are continuous or discontinuous variables.

Random noise is defined by noise probability distributions and by a mechanism to draw from the noise distributions. In its basic form, random noise is generated independently and identically distributed with a mean of zero and a positive variance, which is determined by the statistical agency. A zero mean ensures that no bias is introduced into the original variable. The random noise is then added to the original variable. Adding random noise to a continuous variable will not alter the mean value of the variable for large datasets but will introduce more variance depending on the variance parameter used to generate the noise (*Shlomo*, 2010, p. 3).

Measuring information loss and utility for the SDC decision problem is a more subjective matter. It depends on the users, the purpose of the data, the required statistical analysis, and the type and format of the statistical data. Therefore, it is useful to have a wide range of information loss measures with which to assess the impact of SDC methods on statistical data. These measures include:

- effects on the bias and variance of point estimates and other sufficient statistics,

- distortions to the rankings of variables and univariate and joint distributions between variables,
- changes to model parameters and goodness of fit criteria when carrying out statistical analysis (*Shlomo et al.*, 2006, p. 69).

When assessing SDC methods and their parameters for statistical outputs, an iterative process is carried out. For each method and its parameters, quantitative disclosure risk and information loss measures are calculated. An optimal SDC method is chosen, which reduces the disclosure risk to tolerable risk thresholds, while ensuring high quality data that are fit for purpose (*Shlomo et al.*, 2006, p. 69).

Information loss measures can be classed into two research areas: information loss measures for use by data suppliers so that they can make informed decisions about optimal SDC methods and information loss measures aimed at users so that they can make adjustments to the statistical analysis on modified disclosure controlled statistical data (*Shlomo et al.*, 2006, p. 69).

DATA PROTECTION SOLUTIONS

Eurostat's solution is described in detail in (*Eurostat*, 2017). The relationship to census grid data is also mentioned here. This new geographical variable (e.g. grid id) also needs to be considered from the viewpoint of statistical disclosure control, especially with regard to already existing and used geographical variables. Grid data are particularly useful because they are easy to interpret.

Many grid data will presumably contain zero frequencies. A statistical disclosure control solution cannot alter the spatial distribution of grid data too much. This means that if a few grid cells contain non-zero frequencies in a certain geographical area, they should not be changed very much, and not too many zero grid frequencies should be changed to positive frequencies.

The disclosure risk of statistical data can be quantified using disclosure risk measures. Disclosure risk measures make notions and concepts operational and help to make decisions about the data release. If the disclosure risk is low, a statistical institute might release the data without any change. However, if the disclosure risk is unacceptably high, the statistical institute has to protect the data carefully (*Eurostat*, 2015, chap. 3.1. I, p. 3). The aim is both to protect grids that contain low frequencies of absolute numbers, and to protect low frequencies of attribute values, such as gender, age, marital status, etc. Eurostat's solution is based on the pre-tabular method of targeted record swapping and the post-tabular random noise method. Record swapping is a pre-tabular SDC method, and as such, it is applied to microdata. Some pairs of records are selected in the microdata set. The paired individuals/households are matched on some variables in order to maintain the analytical properties and to minimise the bias of the perturbed microdata set as much as possible. Record swapping exchanges some of the non-equal variable-values between paired individuals/households (*Eurostat*, 2015, chap. 3.1. I, p. 7). The exchanged variables are often geographical variables, and in the case of this paper the grids are used.

Random noise, as a post-tabular method, is defined by noise probability distributions and by a mechanism that draws from the noise distributions. The implementation of random noise as outlined below may involve three 'modules':

- the cell key module,
- the module for determining noise based on cell key and the noise distribution parameter matrix,
- the module to restore additivity (*Eurostat*, 2015, chap. 3.1. I, p. 8).

Cell keys should be drawn from a discrete uniform distribution defined on some integer values (for example, integers between 1 and 100). The process that defines the cell keys has to be consistent, i.e. it must guarantee that the same cell always gets the same key in any hypercube or grid cell or tabulation (*Eurostat*, 2015, chap. 3.1. I, p. 8).

The performance of a random noise method can easily be controlled in a flexible way by means of parameter settings that define the probability distributions. In a typical implementation, the following properties will be required and/or controlled by the parameters:

- noise expectation/unbiasedness property;
- noise variance;
- the property that certain frequencies (e.g. 1s and 2s) should not appear in the perturbed data;

- the property that (structural) zero cells will never be perturbed (*Eurostat*, 2015, chap. 3.1. I, p. 8).

When consistent cell keys are used then the perturbation step leads to consistently perturbed data sets. The ptable files for various settings have been provided by Eurostat for testing. The settings are mainly defined by the maximum perturbation parameter D and by the noise variance parameter V. The ptable provides lists of every combination of cell value and cell key and determines a perturbation value for that cell. The 'p-value' is added to the original cell value, (although most of these changes will be +0) to create the final post perturbation cell value (*Eurostat*, 2015, chap. 3.1. I, pp. 8–9).

Cantabular adds noise to tabular outputs, using the cell-key method, in the same way as the Eurostat methods. Tables are produced dynamically from microdata in real-time in response to a user's query and noise is added deterministically based on a computed cell-key and a perturbation table. Zeros can also be perturbed without affecting any structural zeros found in the data for each query.

The maximum value and variance of perturbation applied are completely configurable via the use of a perturbation table lookup, so different noise distributions can be applied to outputs. In addition to cell-key, Cantabular also includes a disclosure rules language that allows for the real-time checking of table outputs for disclosive cells and the subsequent suppression of outputs per geographic area.

While the Eurostat approach includes a module to restore additivity, Cantabular does not, as this is not possible with a flexible table builder. This loss of additivity can to a small and statistically insignificant degree affect the utility of data for users. This can be avoided by always querying Cantabular for the population counts that are required instead of using Cantabular to create multidimensional hypercubes, which are then themselves queried.

The benefit of taking this approach is that it allows real-time queries for arbitrary cross-tabulations to be made. This is also facilitated by the disclosure rules language, which allows for tables that are still disclosive after the application of cell-key to be automatically suppressed (*Cantabular*, 2021).

INFORMATION LOSS MEASURES

The starting point for measuring the loss of information due to the use of SDC is the evaluation of frequency tables, i.e. the analysis of the differences between the original and the perturbed value. For perturbative methods, we typically measure the maximas, means, medians, and some percentiles of:

- the absolute differences (AD),
- the relative differences (RAD) between original and altered counts in a table, and
- the (squared) differences of the square roots between the original and altered counts.

Counts may be altered because a perturbative protection method has been applied to the data, or because of the effect of cell suppression. The most straightforward way in which to take suppression into account is to impute zeroes for the suppressed count (Eurostat, 2015, chap. 3.1. I, p.10).

According to (*Domingo-Ferrer et al.*, 2006, p. 72), let D^k represent a row (i.e., a distribution) k in a table, and let $D^k(c)$ be the cell frequency c in the row. Let n_r be the number of rows in the comparison. The absolute distance (AD) is then defined as

$$\mathrm{AD(k, c)} := \left|D^k_{pert}(\mathrm{c}) - D^k_{orig}(\mathrm{c})\right|$$

and the summary statistics per aggregate k mean is defined as

$$\overline{AD(k)} := \frac{\sum_{c\in k} AD(k, c)}{n_k}$$

The relative absolute distance (RAD) is defined as

$$\mathrm{RAD(k, c)} := \frac{\left|D^k_{pert}(c) - D^k_{orig}(c)\right|}{D^k_{orig}(c)}$$

and the summary statistics per aggregate k sum is defined as

$$\sum_{RAD}(k) := \sum_{c\in k} RAD(k, c)$$

Finally, the difference of the square roots is defined as

$$\mathrm{D\_R(k, c)} := \left|\sqrt{D^k_{pert}(\mathrm{c})} - \sqrt{D^k_{orig}(\mathrm{c})}\right|$$

and the suggested summary statistics, e.g. Hellinger's distance (HD), is defined as

$$HD(k) := \frac{1}{\sqrt{2}} \|D\_R(k)\|_2 = \sqrt{\frac{1}{2}\sum_{c\in k}(D\_R(k, c))^2}$$

which is used to quantify the similarity between two probability distributions - a namely the original and perturbed datasets. Once these are derived, it is then possible to calculate Hellinger's distance utility (HDU) as

$$1 - \text{HD(orig, perturb)} / \sqrt{(\Sigma\text{orig})}$$

which measures the relative degree of agreement between the original and the perturbed dataset in the interval (0;1).

For both AD and RAD simple descriptive statistics like max, mean, and median, the percentiles p60, p70, p80, p90, p95, and p99 would be calculated. In addition, the cumulative distribution function $\mathbf{F}_{\mathbf{RAD}}(\boldsymbol{r})$ proportion of cells with relative absolute difference less than (r) could also be calculated. These measures are based on the idea that if the synthetic and original data are similar, data set membership should be indistinguishable between the two data sets.

Another statistical analysis that is frequently carried out on tabular data are tests for independence between categorical variables that span a table. The test for independence for a two-way table is based on a Pearson Chi-Squared Statistic (*Shlomo*, 2006, p. 214). This statistic defined for i is from 1 to s and the summation for j is from 1 to r, is formulated as

$$Q_P = \sum_{i=1}^{s}\sum_{j=1}^{r}\frac{(n_{ij} - m_{ij})^2}{m_{ij}}$$

where

$$m_{ij} = E\{n_{ij} \mid H_0\}\frac{n_{i+}\,n_{+j}}{n}$$

is the expected value of the frequencies in the *i* th row and *j* th column.

Measures of association when one or both variables are nominally scaled are more difficult to define, since you cannot think of association in these circumstances as negative or positive in any sense. However, indices of association in the nominal case have been constructed and most are based on mimicking R-squared in some fashion. One such measure is the uncertainty coefficient, and another is the lambda coefficient (*Stokes et al.*, 2012, p. 129).

The asymmetric lambda λ (Columns|Rows) is interpreted as the probable improvement in predicting the column variable Y (perturbed data) given knowledge of the row variable X (original data). The range of the asymmetric lambda is $0 \le \lambda(C|R) \le 1$. The asymmetric lambda (C|R) is computed as

$$\lambda(C|R) = \frac{\sum_i r_i - r}{n - r}$$

and its asymptotic variance is

$$\mathrm{Var}(\lambda(C|R)) = \frac{n - \Sigma\, r_i}{(n - r)^3}\left(\sum_i r_i + r - 2\sum_i (r_i \mid l_i = l)\right)$$

The nondirectional lambda (symmetric) is the average of the two asymmetric lambdas, (λ(C | R) and (λ(R | C). Its range is $0 \le \lambda \le 1$. The lambda symmetric is computed as

$$\lambda = \frac{\sum_i r_i + \sum_j c_j - r - c}{2n - r - c} = \frac{w - v}{w}$$

and its asymptotic variance is computed as

$$\mathrm{Var}(\lambda) = \frac{1}{w^4}\left(wvy - 2w^2\left(n - \sum_i\sum_j (n_{ij} \mid j = l_i, i = k_j\right) - 2v^2(n - n_{kl})\right)$$

The uncertainty coefficient U is the symmetric version of the two asymmetric uncertainty coefficients. Its range is $0 \le U \le 1$. The uncertainty coefficient is computed as

$$U = 2(H(X) + H(Y) - H(XY)) / (H(X) + H(Y))$$

and its asymptotic variance is

$$\mathrm{Var}(U) = 4\sum_i\sum_{ij}\left(H(XY)\ln\left(\frac{n_i n_j}{n^2}\right) - H(X) + (H(Y))\ln\left(\frac{n_{ij}}{n}\right)\right)^2$$

where $H(X)$, $H(Y)$, and $H(XY)$ are defined in the previous section. See (*SAS Stat*, 2021) for the completed description.

For each measure, the asymptotic standard error (ASE) has been calculated, which is the square root of the asymptotic variance denoted by the variable. If the sample size is adequate, then the measure of association is approximately normally distributed, and the confidence intervals of interest can be calculated as

$$\text{est} \pm z_{\alpha/2} \cdot \text{ASE}$$

where est is the estimate of the measure, zα/2 is the 100 (1-α/2) percentile of the standard normal distribution, and ASE is the asymptotic standard error of the estimate (SAS Stat). In this case, 95% confidence interval was used.

The Gini index (or Gini ratio) is a measure of statistical dispersion and it is the most commonly used measurement of inequality preferably used in economics. It measures the inequality among values of a frequency distribution. An index of zero expresses perfect equality, where all the values are the same, and an index of 1 (or 100%) expresses maximal inequality among the values. The sample Gini coefficient was calculated using the formula:

$$\mathrm{G} = \frac{1}{2\bar{X}n(n-1)} \sum_{i=1}^{n} (2i - n - 1)X_i$$

where X_i are the sizes sorted from smallest to largest, X1 ≤ X2 ≤ Xn (*Dixon*, 1987).

FROM GRIDS TO SPATIAL STATISTICS

In the case of grid data, it is also necessary to take into account spatial measures, which measure the degree of spatial distribution both original and perturbed data sets:

- the global autocorrelation rate
- the local autocorrelation rates.

Spatial autocorrelation may be a result of unobserved or hard-to-quantify processes, combined in various places, and together the causing spatial structuring of a given phenomenon. If there is a spatial autocorrelation, it is determined by examining whether the variable value for a given (e.g. geolocalised) observation is associated with values of the same variable for neighbouring observations (*INSEE*, 2018, p. 67). Spatial autocorrelation may be positive or negative or there may be no spatial autocorrelation among the given data. Spatial autocorrelation can be measured globally or locally; both ways assess the same thing – i.e. whether there is a spatial correlation of a given phenomenon – but they are not the same.

There are different ways of measuring spatial autocorrelation; Moran's I is often used. The principle of computation is that it takes into account the difference between the value of the variable and the average of values of that variable for a given area (e.g. neighbourhood). Moran's index is the preferred approach (compared to others), because it is more stable against extreme values, and it can be used in two ways (see below). The index can be written in several ways, but it is frequently written as follows:

$$I_W = \frac{n}{\Sigma_i \Sigma_j w_{ij}} \frac{\Sigma_i \Sigma_j w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\Sigma_i (y_i - \bar{y})^2} \quad i \neq j$$

Null hypothesis H0, states that there is no spatial correlation in the given territory. Vice versa, if *Iw* > 0, then there is a positive autocorrelation, which means that high values are neighbouring high ones and low values are neighbouring low ones. In the case of a negative autocorrelation, the reverse would apply. Depending on the distribution of a spatial variable, the calculation of a median value: neighbour high ones and low values neighbour low ones. Depending on the distribution of a spatial variable, the calculation of a median value is

$$E(I_W) = E(c_W) = -\frac{1}{n-1}$$

and the calculation for the testing statistics is

$$\frac{I_W - E(I_W)}{\sqrt{\mathrm{Var}(I_W)}} \sim \frac{c_W - E(c_W)}{\sqrt{\mathrm{Var}(c_W)}} \sim \mathrm{N}(0, 1)$$

A key element for calculating the indices of spatial autocorrelation is to determine the neighbourhood, i.e. to select spatial entities that are neighbours by *definition*. In the case of this study, the neighbourhood was defined by the edges_corners method, i.e. grids that had a common edge or vertex were always taken as neighbours. An explanation of this approach can be found in (*Kraus*, 2019).

The Moran's index is a global statistic, which provides no information about the extent of local variation in spatial variability. For that, there are tools that enable us to assess the local level of spatial autocorrelation (LISA) and to measure the intensity and importance of autocorrelation between the value of the variable in a spatial unit and the value of the same variable in neighbouring spatial units. These indicators examine the following two features:

- for each observation they show the intensity of the clustering of similar/opposite values around that observation;
- the sum of local indices at all observations is proportional to the corresponding global index, e.g. to global Moran's I.

In the case of Moran's I, its local value can be written as follows

$$I_i = (y_i - \bar{y}) \sum_j w_{ij} (y_j - \bar{y})$$

and the value of the global index is as follows

$$I_w = \text{konst} \cdot \sum_i I_i$$

where:

- $I_i > 0$ indicates the clustering of similar values (higher or lower than the average for a given neighbourhood), and
- $I_i < 0$ indicates the clustering of different values.

The spatial clustering of similar or different values is observed as follows: as High-High values (HH), Low-Low values (LL), High-Low values (HL), or Low-High (LH) values. If we mean a high value surrounded by other high values or a low value surrounded by other low values then they are referred to as hot spots or cold spots, respectively. If we mean a high value surrounded by low values or a low value surrounded by high values, then these are spatial outliers (*Anselin*, 1995). The significance of each local indicator is based on a spatial distribution of data and statistics that is asymptotically approaching the normal distribution:

$$z(I_i) = \frac{I - E(I_i)}{\sqrt{\text{Var}(I_i)}} \sim \text{N}(0, 1)$$

Since the global rate of spatial autocorrelation (Moran's I) proved to be distinctively higher in the case where the neighbouring municipalities method is used, local rates of Moran's I were further computed only for this method of neighbourhood determination.

RESULTS

A relatively simple model was chosen for the calculation (it is a test), where the output (i.e. perturbed) variable is the number of people who are usually living according to the grid network. This total is information that can be published without restriction. The constraint occurs when it needs to be published in combination with another variable or variables. For the purpose of this test, two variables that enter perturbation were selected: sex and age. As the combination of age, sex, and individual grid units would create too low a frequency, age was transcoded into ten-year groups in line with Eurostat's recommendation: the output is the number of usually living by sex, age, ten-year age groups, and grids. These combinations were then aggregated again into the number usually living according to the grid network and the result was evaluated.

According to the Eurostat methodology, two variants were calculated, which differ by setting the parameters maximum noise D and variance of noise V. D = 3 and V = 1 are settings recommended on the basis of Eurostat testing. Furthermore, in accordance with the Eurostat methodology, version 4 was calculated with values D = 2 and V = 1, i.e. with a lower level of perturbation parameter D but with the same level of noise variance V. Another option is to keep zero values, i.e. grids with a zero number of habitual residents. They are not subject to perturbations.

Cantabular was configured with a perturbation table designed to replicate Eurostat variant 1, but with a reduced cell-key range, compatible with Cantabular. It had a maximum absolute perturbation parameter D of 3 and a noise variance V of approximately 1 (*Cantabular*, 2021).

Recommended statistics were calculated for the difference in the number of usually living between

Table 1 Simple descriptive statistics for absolute difference (AD) - Eurostat solution

Maximum	Mean	Median	60th pctl	70th pctl	80th pctl	90th pctl	95th pctl	99th pctl	Variance
Variant 1									
3	0.41	0	0	1	1	1	2	2	0.39
Variant 4									
2	0.49	0	1	1	1	1	2	2	0.42

Source: Author's calculation.

Table 2 Simple descriptive statistics for absolute difference (AD) - Cantabular solution

Maximum	Mean	Median	60th pctl	70th pctl	80th pctl	90th pctl	95th pctl	99th pctl	Variance
Variant 1									
49	2.82	0	1	3	5	9	13	22	23.05

Source: Author's calculation.

the original and the perturbed value according to the grid network.

The result of the absolute difference (AD) shows that there is a clear difference between the results according to the Eurostat and Cantabular method in the case of maximum, mean and variance. This difference is not so significant for the median and lower percentiles.

The higher maximum value for absolute difference shown in the table above for Cantabular is caused by a query being run at a high level of detail – age by sex by grid square – before the results are then added up at a lower level of detail – total population by grid square – for a comparison with the original unperturbed data.

This has the effect of compounding perturbation because of the loss of additivity in the marginal totals that is inherent in the cell-key method. If the initial query was done at total population by grid square, the maximum absolute difference would be 3, as set in the perturbation configuration. As discussed above, Cantabular does not attempt to restore additivity in order to provide a larger, more flexible range of outputs (*Cantabular*, 2021).

The cumulative distribution function F_{AD} (proportion of cells with an absolute difference less than d was calculated for = 1 to 15. While in the case of results according to the Eurostat methodology there was a complete enumeration in variant 1 for CDF = 3 and in the case of variant 4 even for CDF = 2, the results according to the Cantabular methodology show a gradual and uniform increase in frequencies up to value 15. This follows from a previous finding of a maximum of ***AD***, which was for Eurostat variants 2 and 3, while for Cantabular was 49.

However, in the case of the relative absolute difference (R_{AD}), the differences between the Eurostat and Cantabular methodologies are blurred. The maximum R_{AD} reaches the value 3 for both variant 1 of the Eurostat methodology

Table 3 Cumulative distribution function (CDF) for absolute difference – Eurostat solution

Variant 1					Variant 4				
CDF	Frequency	Percent	Cumulative frequency	Cumulative percent	CDF	Frequency	Percent	Cumulative frequency	Cumulative percent
0	42,100	65.67	42,100	65.67	0	38,158	59.53	38,158	59.53
1	18,074	28.19	60,174	93.87	1	20,546	32.05	58,704	91.58
2	3,575	5.58	63,749	99.45	2	5,400	8.42	64,104	100.00
3	355	0.55	64,104	100.00					

Source: Author's calculation.

Table 4 Cumulative distribution function (CDF) for absolute difference – Cantabular solution

Variant 1				
CDF	**Frequency**	**Percent**	**Cumulative frequency**	**Cumulative percent**
0	40,427	50.45	40,427	50.45
1	8,345	10.41	48,772	60.57
2	6,028	7.52	54,800	68.39
3	4,504	5.62	59,304	74.01
4	3,501	4.37	62,805	78.38
5	2,786	3.48	65,591	81.86
6	2,249	2.81	67,840	84.66
7	1,856	2.32	69,696	86.98
8	1,621	2.02	71,317	89.00
9	1,369	1.71	72,686	90.71
10	1,120	1.40	73,806	92.11
11	985	1.23	74,791	93.34
12	856	1.07	75,647	94.41
13	692	0.86	76,339	95.27
14	654	0.82	76,993	96.09
15	3,136	3.91	80,129	100.00

Source: Author's calculation.

Table 5 Simple descriptive statistics for relative absolute difference – Eurostat solution

Maximum	Mean	Median	60th pctl	70th pctl	80th pctl	90th pctl	95th pctl	99th pctl	Variance
Variant 1									
3	0.28	0.06	0.14	0.29	0.50	1.00	1.00	2.00	0.39
Variant 4									
2	0.35	0.11	0.22	0.50	1.00	1.00	1.00	2.00	0.42

Source: Author's calculation.

Table 6 Simple descriptive statistics for relative absolute difference – Cantabular solution

Maximum	Mean	Median	60th pctl	70th pctl	80th pctl	90th pctl	95th pctl	99th pctl	Variance
Variant 1									
3	0.19	0.08	0.12	0.18	0.29	0.50	0.86	1.00	23.05

Source: Author's calculation.

and the Cantabular methodology. Similarly, both methodologies yield completely comparable values for both the mean and the percentile values. This is because the denominator of the indicator contains the numbers of original values, so that even in the case of differences between the original and the perturbed value of higher frequencies, the relative differences decrease.

The CDF results for variable R_{AD} show that a higher degree of agreement between the original and the perturbed value exists at lower CDF_RAD levels for the Eurostat method, but with increasing value the situation rotates and for 0.50 the Cantabular method contains 93 percent of all (cumulative values) and while for 0.50 Eurostat methods 1 and 4 contain, respectively, 88 and 85 percent. The results are therefore similar.

The relative Hellinger distance (HDutility) again shows that both methods yield completely comparable

Table 7 Cumulative distribution function (CDF) for relative absolute difference (RAD) – Eurostat solution

Variant 1					Variant 4				
CDF_RAD	Frequency	Percent	Cumulative frequency	Cumulative percent	CDF_RAD	Frequency	Percent	Cumulative frequency	Cumulative percent
0.02	43,422	67.74	43,422	67.74	0.02	41,238	64.33	41,238	64.33
0.05	2,334	3.64	45,756	71.38	0.05	2,304	3.59	43,542	67.92
0.10	3,028	4.72	48,784	76.10	0.10	3,087	4.82	46,629	72.74
0.20	3,391	5.29	52,175	81.39	0.20	3,563	5.56	50,192	78.30
0.30	2,379	3.71	54,554	85.10	0.30	2,428	3.79	52,620	82.09
0.40	1,508	2.35	56,062	87.45	0.40	1,573	2.45	54,193	84.54
0.50	206	0.32	56,268	87.78	0.50	225	0.35	54,418	84.89
0.99	2,547	3.97	58,815	91.75	0.99	2,568	4.01	56,986	88.90
1.00	5,289	8.25	64,104	100.00	1.00	7,118	11.10	64,104	100.00

Source: Author's calculation.

Table 8 Cumulative distribution function (CDF) for relative absolute difference (RAD) – Cantabular solution

Variant 1				
CDF_RAD	Frequency	Percent	Cumulative frequency	Cumulative percent
0.02	47,077	58.75	47,077	58.75
0.05	6,623	8.27	53,700	67.02
0.10	6,753	8.43	60,453	75.44
0.20	7,021	8.76	67,474	84.21
0.30	4,049	5.05	71,523	89.26
0.40	2,178	2.72	73,701	91.98
0.50	1,029	1.28	74,730	93.26
0.99	3,207	4.00	77,937	97.26
1.00	2,192	2.74	80,129	100.00

Source: Author's calculation.

Table 9 Hellinger distance (HD) and related utility measures – Eurostat solution

HD	HDutility	Max difference	Mean Abs Difference	rootMeanSquare
Version 1				
212.13	0.93	3.00	0.34	0.68
Version 4				
294.33	0.91	2.00	0.41	0.74

Source: Author's calculation.

Table 10 Hellinger distance (HD) and related utility measures – Cantabular solution

HD	HDutility	Max difference	Mean Abs Difference	rootMeanSquare
Version 1				
69.53	0.98	49	2.82	5.57

Source: Author's calculation.

results. For the Eurostat method, this agreement is at the level of 0.93, resp. 0.91 and in the case of the Cantabular method even 0.98. This indicates that there are no statistically significant differences between the original and the perturbed values.

All of the measures of ordinal association indicate a positive association. The resulting association rates are comparable for all three methods and the ASE value indicates that they are statistically significant. Slightly higher values obtained by the Cantabular method suggest in favour of this method of data perturbation.

Gini's concentration coefficient is used in geographic surveys because it overcomes the deficiencies of the coefficient of variation depending on the average and is therefore more appropriate for affecting the variability of asymmetric distributions typical of socio-geographical phenomena (*Netrdová et al.*, 2012). An interesting comparison is the one with the result of the GINI index calculation between the original data and the perturbed data. The results show that the value of the Gini index expresses high inequality among values, but at approximately the same level for the original and the perturbed data.

Previous results showed a statistical evaluation of the results, without questioning whether the original and perturbed values are somehow differently distributed in space. The answer to this question is given by the global and local measures of spatial autocorrelation.

Table 11 Measures of association between original and perturbed data

Statistic	Eurostat – version 1				Eurostat – version 4				Cantabular			
	Value	ASE	95% Confidence limits		Value	ASE	95% Confidence limits		Value	ASE	95% Confidence limits	
Pearson correlation (Rank Scores)	0.730	0.001	0.728	0.731	0.730	0.001	0.729	0.731	0.955	0.000	0.954	0.956
Lambda asymmetric C|R	0.329	0.001	0.327	0.330	0.288	0.001	0.286	0.289	0.155	0.002	0.152	0.159
Lambda asymmetric R|C	0.393	0.001	0.392	0.394	0.268	0.001	0.267	0.269	0.176	0.002	0.173	0.180
Lambda symmetric	0.362	0.001	0.361	0.364	0.277	0.001	0.276	0.279	0.166	0.002	0.163	0.169
Uncertainty coefficient C|R	0.701	0.000	0.701	0.702	0.650	0.000	0.649	0.650	0.628	0.001	0.626	0.630

Source: Author's calculation.
C|R – columns|rows, ASE - asymptotic standard errors (*Stokes et al.*, 2012, p. 125).

Table 12 Gini coefficient for original and perturbed data

Gini coefficient	Original data	Eurostat – version 1	Eurostat – version 4	Cantabular perturbed data
		Perturbed data	Perturbed data	
	0.893	0.895	0.895	0.900

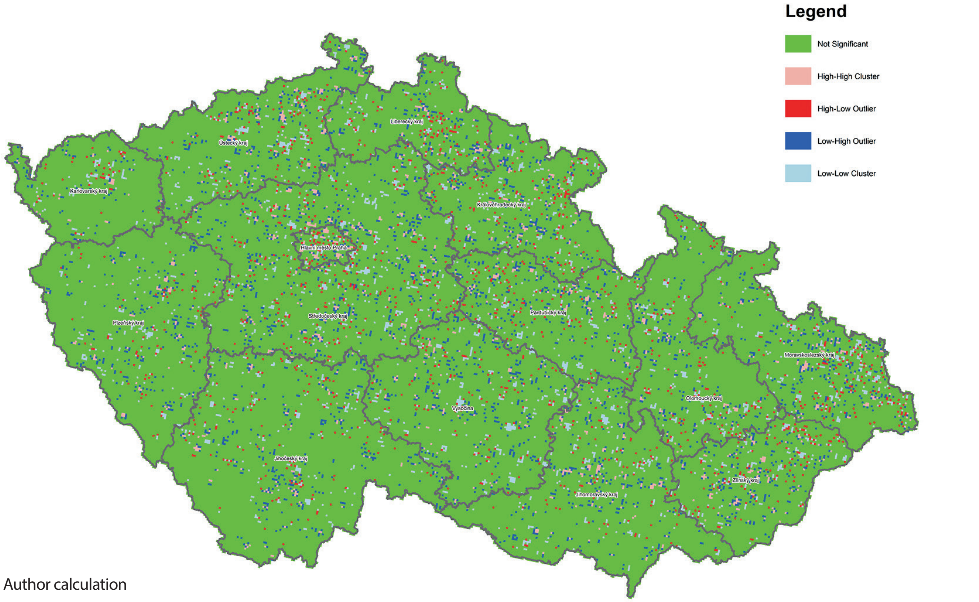
Source: Author's calculation.
C|R – columns|rows, ASE - asymptotic standard errors (*Stokes et al.*, 2012, p. 125).

Table 13 Global Moran's I summary for original and perturbed data

	Original data	Eurostat		Cantabular variant
		Variant 1	Variant 4	
Moran's Index	0.493	0.493	0.492	0.496
Variance	0.000	0.000	0.000	0.000
z-score	277.8	277.8	277.8	280.5
p-value	0.0	0.0	0.0	0.0

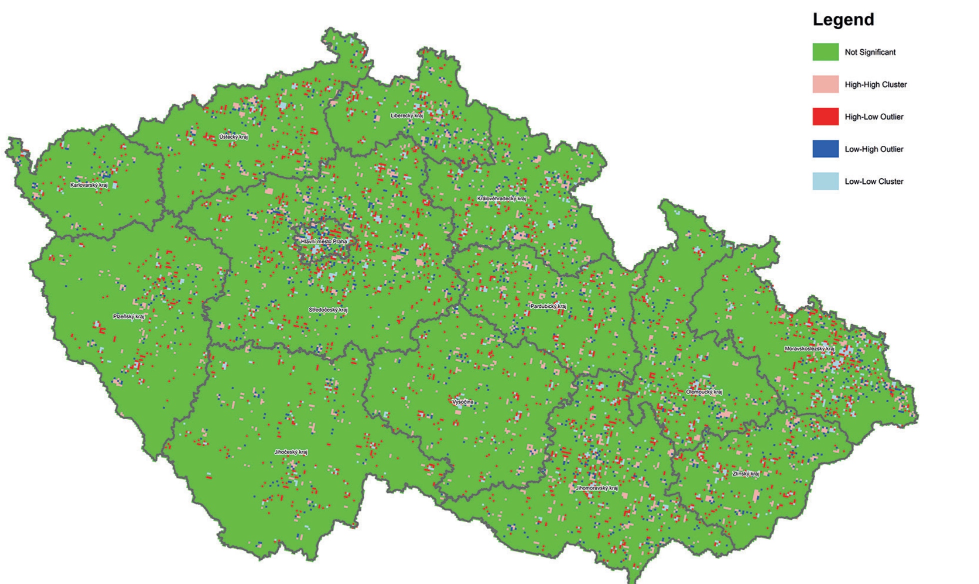
Source: Author's calculation.

Figure 2 Typology of grids according to the difference between the original and the perturbed value calculated according to Eurostat methodology (local Moran's I)



Source: Author calculation

Figure 3 Typology of grids according to the difference between the original and perturbed values calculated according to the Cantabular methodology (local Moran's I)



Source: Author calculation

Note: The above cartograms show that perturbation does not lead to a change in the spatial structure of the observed phenomenon, which in this case is the number of usually living in the individual grids. If there was a change, then the dominant (not significant) value, marked in green, would be replaced in larger areas (i.e. groups of grids) by a different colour than, and thus the structure would be disrupted.
Because this is not the perturbed value of the number of usually living by grids, the derived structures used in this model case (five-year age structures, sex) are guaranteed to yield consistent results (i.e. compared to the original values).

Moran's I calculation was based on the neighbourhood defined by the Edges Corners method, meaning that neighbours are those grids that have either an edge or a corner in common. The results show that the value of the index is practically the same for the original and perturbed data calculated by both the Eurostat and Cantabular methods and differs only to the third decimal place. Given the p-value, the pattern appears to be significantly different from random, and the z-score indicates that all models are very similar.

The local level of spatial autocorrelation (LISA) indicates the local values of Moran's I. This indicator was calculated for the difference between the original and the perturbed value of each grid. From Figures 2 and 3 it is evident that the type Not Significant (the bright green colour) predominates, i.e. perturbation is also a spatially random process that does not change the spatial distribution of usually living.

CONCLUSION

The 2011 Population and Housing Census in the Czech Republic was accompanied by a significant change in the technology used to prepare the census and in the actual course of fieldwork, along with changes in how the data were processed and the outputs were disseminated. Some methodological approaches to processing the data have also changed and are now more aligned with international recommendations. Although a number of changes have been relatively widely discussed in the literature, one type of output remains somewhat overlooked: census results in a grid network.

Working with a network of grids has both advantages and disadvantages, but the main disadvantage is that grids are small territorial units that are often minimally populated. This is mainly a problem in terms of the protection of individual data, which is associated with statistical disclosure control (SDC).

The research question addressed in this paper is whether data protection (perturbation methods) leads to a change in the characteristics of the file either in terms of the statistics of the whole file (i.e. for all grids) or in terms of spatial statistics, which indicate the spatial distribution of the analysed phenomenon.

Two possible solutions to the issue of grid data protection were examined. One comes from the Statistical Office of the European Communities (Eurostat) and the other from Cantabular, which is a product of the Belfast company Sensibile Code Ltd.

In both cases, the data protection solutions are described. One possible solution is to add noise to tabular outputs, using the cell-key method. Tables are produced dynamically from microdata in real-time in response to a user's query and noise is added deterministically based on a computed cell-key and a perturbation table. Zeros can also be perturbed without affecting any structural zeros found in the data for each query.

The starting point for measuring the loss of information due to the use of SDC is the evaluation of frequency tables, i.e. the analysis of the differences between the original and the perturbed value. For perturbative methods, measures of the maximas, means, medians, and some percentiles of absolute differences (AD) and relative differences (RAD) between the original and altered counts in a table and the (squared) differences of the square roots between original and altered counts were calculated.

However, in the case of grids, it was also necessary to focus on spatial measures, which measure the degree of spatial distribution of both the original and the perturbed data sets, e.g. the global autocorrelation rate and the local autocorrelation rates.

The results are based on a relatively simple model for the calculation, where the output (i.e. perturbed) variable is the number of people usually living according to the grid network. The constraint occurs when it should be published in combination with another variable or variables. For the purpose of this test, two variables that enter perturbation were selected: sex and age. As the combination of age, sex, and individual grid units would create too low a frequency, age was transcoded into ten-year groups in line with Eurostat's recommendation: the output is the numbers usually living by sex, age, ten-year age groups, and grids. These combinations were then aggregated again into the number usually living according to the grid network and the result was evaluated.

According to the Eurostat methodology, two variants were calculated, which differ by the parameters set for maximum noise D and the variance of noise V. Another option is to keep zero values, i.e. grids with a zero number of habitual residents. They are not subject to perturbations. Cantabular was configured with a perturbation table designed to replicate Eurostat variant 1, but with a reduced cell-key range compatible with Cantabular.

The results of the descriptive statistics show a difference in the absolute differences compared with the Eurostat methodology, and Cantabular explains the different way of processing microdata. In the case of other statistics, the results are fully comparable.

This paper is devoted to one specific type of census output. The question is to what extent these results are relevant for other types of outputs and in particular for outputs in hypercubes. They differ fundamentally in terms of the number of dimensions (grids have only two dimensions). It would therefore be appropriate to use SDC procedures that allow greater flexibility in defining SDC parameters.

References

- Anselin, L. 1995. Local Indicators of Spatial Association – LISA. *Geographical Analysis*, 27(2), pp. 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x.
- Cantabular. 2021. Internal communication with The Sensible Code Co. (Mike Thompson – Aine McGuire), 5. 5. 2021
- Company The Sensible Code. 2020. Press Release. https://sensiblecode.io/resources/case-study-ons.pdf.
- Český statistický úřad (CZSO), n.d. *Pramenné dílo SLDB 2011.* 2013th edition. Prague: Czech Statistical Office, 461 pp.
- Domingo-Ferrer, J. – Franconi, L. (dir.). 2006, *Privacy in Statistical Databases: CENEX-SDC Project International Conference, PSD 2006, Rome, Italy, December 13-15, 2006, Proceedings*, Berlin Heidelberg, Springer-Verlag, Information Systems and Applications, incl. Internet/Web, and HCI. https://doi.org/10.1007/11930242.
- Domingo-Ferrer, J. – Montes, F. (dir.) 2018. *Privacy in Statistical Databases: UNESCO Chair in Data Privacy, International Conference, PSD 2018, Valencia, Spain, September 26–28, 2018, Proceedings*, Springer International Publishing, Information Systems and Applications, incl. Internet/Web, and HCI. https://doi.org/10.1007/978-3-319-99771-1.
- Dixon, P. M. – Weine, J. – Mitchell-Olds, T. – Woodley, R. 1987. Bootstrapping the Gini Coefficient of Inequality. *Ecology*, 68(5), pp. 1548–1551. https://doi.org/10.2307/1939238.
- Doxsey-Whitfield, E. – MacManus, K. – Adamo, S. B. – Pistolesi, L. – Squires, J. – Borkovska, O. – Baptista, S. R. 2015. Taking Advantage of the Improved Availability of Census Data: A First Look at the Gridded Population of the World, Version 4. *Papers in Applied Geography*, 1(3), pp. 226–234. https://doi.org/10.1080/23754931.2015.1014272.
- Eurostat. 2015. Centre of Excellence on Statistical Disclosure Control, *CROS - European Commission.* https://ec.europa.eu/eurostat/cros/content/centre-excellence-statistical-disclosure-control-0_en.
- Hundepool, A. – Domingo-Ferrer, J. – Franconi, L. – Giessing, S. – Schulte Nordholt, E.– Spicer K. – de Wolf P.-P. 2012. *Statistical Disclosure Control.* John Wiley & Sons, 270 p. https://doi.org/10.1002/9781118348239.
- Klauda, P. 2011. Site-Oriented Statistics and its Geoinformatic Potential. *Statistika - Economy and Statistics Journal*, 2011(2), pp. 107–110.
- Kraus, J. – Moravec, Š. 2014. Prezentace výsledků SLDB 2011 v síti čtverců – projekt GEOSTAT," *Demografie*, 2014(56), pp. 143–146.
- Lloyd, C. T. – Chamberlain, H. – Kerr, D. – Yetman, G. – Pistolesi, L. – Stevens, F. R. – Gaughan, A. E., Nieves, J. J. – Hornby, G. – MacManus, K. – Sinha, P. – Bondarenko, M. – Sorichetta, A. – Tatem, A. J. 2019. Global spatio-temporally harmonised datasets for producing high-resolution gridded population distribution datasets. *Big Earth Data*, 3(2), pp. 108–139. https://doi.org/10.1080/20964471.2019.1625151.
- Lloyd, C. T. – Sorichetta, A. – Tatem, A. J. 2017. High resolution global gridded data for use in population studies. *Scientific Data*, 4(1), p. 170. https://doi.org/10.1038/sdata.2017.1.
- Loonis, V. – de Bellefon, M.-P. 2018. *INSEE-EFGS-Eurostat, Handbook of Spatial Analysis.* 2018th edition. INSEE – Eurostat, INSEE, 394 p.
- Martin, D. – Lloyd, C. – Shuttleworth, I. 2011. Evaluation of Gridded Population Models Using 2001 Northern Ireland Census Data. *Environment and Planning A: Economy and Space*, 43(8), pp. 1965–1980. https://doi.org/10.1068/a43485.

- Netrdová, P. – Blažek, J. Aktuální tendence lokální diferenciace vybraných socioekonomických subjektů v Česku: směřuje vývoj k větší mozaikovitosti prostorového uspořádání? (in Czech) *Geografie*, 2012, 3, pp. 266–288.
- SAS. Stat SAS(R) 9.2 User's Guide [online]. 2nd Ed. [cit. 27.04.2021] https://v8doc.sas.com/sashtml/stat/chap28/sect20.htm.
- Shlomo, N. – Young, C. 2006, Statistical Disclosure Control Methods Through a Risk-Utility Framework. In: Domingo-Ferrer, J. – Franconi, L. (eds). *Privacy in Statistical Databases.* https://doi.org/10.1007/11930242_7.
- Shlomo, N. 2010b. *Measurement Error and Statistical Disclosure Control*, 118 p. https://doi.org/10.1007/978-3-642-15838-4_11.
- Stokes M. E. – Davis Ch. S. – Koch G. G. 2012. *Categorical Data Analysis Using SAS. Third Edition*, 3rd edition. Cary, NC, SAS Institute, 590 p.
- Templ M. 2017. Statistical disclosure control for microdata. *Cham: Springer International Publishing.* https://doi.org/10.1007/978-3-319-50272-4.
- Thijs B. – Welch M. n.d. Statistical Disclosure Control for Microdata: A Practice Guide for sdcMicro — SDC Practice Guide documentation. https://sdcpractice.readthedocs.io/en/latest/.

JAROSLAV KRAUS

graduated from the Prague University of Economics and Business and completed his PhD studies at the Faculty of Science, Charles University. He works at the Czech Statistical Office in the Department of Population Statistics. He specialises in the issue of population censuses and especially in data processing, the statistical protection of census data, and the spatial analysis of demographic phenomena. He has published several papers in this area in recent years.

THE TIMELY DIAGNOSIS AND TREATMENT OF ALZHEIMER'S DISEASE: MICROSIMULATING COST-EFFECTIVENESS IN THE CZECH REPUBLIC

Hana M. Broulíková[1), 2)] – Matěj Kučera[1), 4)] – Markéta Arltová[3)]

Abstract
Using a microsimulation model, this study evaluates the cost-effectiveness of a hypothetical policy that would increase diagnosis and treatment availability for people at an early stage of Alzheimer's disease in the Czech Republic. If widely available, timely diagnosis and treatment would represent a dominant strategy bringing net benefit of EUR 13,751 per lifetime of each person living with this disease.

Keywords: Alzheimer's disease, cost-effectiveness, timely diagnosis, Czech Republic, dementia

https://doi.org/10.54694/dem.0287 **Demografie,** 2021, **63 (4): 216–225**

INTRODUCTION

The Czech population has been significantly ageing. According to the Czech Statistical Office, 26% of the population in the Czech Republic in 2019 was over the age of 60 (*CZSO*, 2020). This number increased by 9% since 1980 and it is expected to grow over time and reaching 37% by 2050 (*UN*, 2018). One of the difficulties closely associated with ageing is the decline in cognitive functions (*Murman*, 2015), which could lead to dementia, a syndrome defined by progressive impairments to memory, thinking, and behaviour that affect people's ability to look after themselves (*WHO*, 2020). Alzheimer's disease (AD) is the most common cause of dementia, accounting for about 60% of all cases (*WHO*, 2020). Owing to the increasing prevalence of dementia and the associated emotional and economic burden it causes, the World Health Organization (WHO) has recognised dementia as a public health priority (*World Health Organization*, 2012).

According to *Mátl et al.* (2016), the number of people living with dementia in the Czech Republic reached 156,000 in 2015. This figure is in line with the upper limit of an estimate published by the Institute

1) National Institute of Mental Health, Klecany, Czech Republic.
2) Faculty of Science, Vrije Universiteit Amsterdam, The Netherlands, contact: h.m.broulikova@vu.nl.
3) Faculty of Finance and Accounting, Prague University of Economics and Business, Czech Republic.
4) Second Faculty of Medicine, Charles University, Prague, Czech Republic.

for Health Metrics and Evaluation (2019), which reported that there were 135,738 cases in 2015 with a corresponding 95% confidence interval (CI), of 118,821 – 154,936, 139,042 cases in 2016 (CI 121,407 – 159,145), and 142,442 cases in 2017 (CI 124,351 – 163,151). The prevalence of dementia has been estimated to reach 250,000 by 2050 (*Holmerová – Hort et al.*, 2017).

The last two decades witnessed a sharp drop in the share of people living with undiagnosed dementia. According to *Waldemar et al.* (2007), the rate of diagnosis in 2004 was only 9%. *Mátl et al.* (2016) report that 6% of people living with dementia received inpatient and 24% outpatient care in 2015. Most recently, the official health statistics derived from the national health-care register revealed that there were 102,000 people living with dementia in 2017, suggesting that up to 72% of people living with dementia receive a diagnosis (*World Health Organization*, 2019, Ministry of Health of the Czech Republic 2021). Although these figures suggest that there has recently been a dramatic decrease in the treatment gap – from 91% in 2004 to 28% in 2017 – part of the drop has likely been brought about by an improvement in the availability of reliable data on the number of treated cases.

Nevertheless, the diagnosis gap has two dimensions: one is *whether* a person has received a diagnosis and the other is the *timing* of the diagnosis. In contrast to the improvements in obtaining a diagnosis, late diagnosis remains a major problem in the country. According to a case study, 56% of people admitted to hospital with dementia had received no diagnosis or treatment prior to being hospitalised for this reason. At the same time, 50% of these hospitalised patients had already progressed to a moderate and 42% to a severe stage of dementia (*Lužný et al.*, 2014). The late diagnosis hypothesis is further documented by the short survival of people with dementia: 44% die within one year, and only 16% live longer than five years from the dementia diagnosis (*Broulíková et al.*, 2020). The unavailability of timely diagnosis and post-diagnostic support is recognised in the recently adopted National Action Plan for Alzheimer's Disease and Related Illnesses (*Ministry of Health of the Czech Republic*, 2021), the first strategic objective of which focuses on remedying this situation.

Dementia is associated with substantial health and social care costs. Care for Alzheimer's and related forms of dementia was estimated to cost USD 818 billion worldwide in 2016 (*Prince et al.*, 2015) and 44.7 billion Czech Koruna (USD 2 billion) in the Czech Republic in 2015 (*Mátl et al.*, 2016). The high costs associated with the disease are mainly driven by informal caregiving and social care (*Winblad et al.* 2016), but people with dementia also face costly adverse health events leading to high hospitalisation rates (*Bernardes et al.*, 2018). Modelling studies suggest that the timely treatment of Alzheimer's disease could increase patients' utility while decreasing their lifetime costs. *Weimer and Sager* (2009) estimated that timely detection and treatment in the United States resulted in net social benefits of USD 94,000 and governmental fiscal savings of USD 15,000 per patient's lifetime (*Weimer – Sager*, 2009). For the United Kingdom, *Getsios et al.* (2012) suggest more modest but still substantial societal savings of GBP 5,700 (USD 8,400) and a decrease in medical costs of GBP 2,100 (USD 3,100). In the Czech Republic, a study focusing on the effect of timely diagnosis on lifetime costs estimated that the savings from timely diagnosis could amount as much as EUR 26,000, depending on the person's age at the disease's onset and the person's cognitive score at the time of diagnosis (*Broulíková et al.*, 2018).

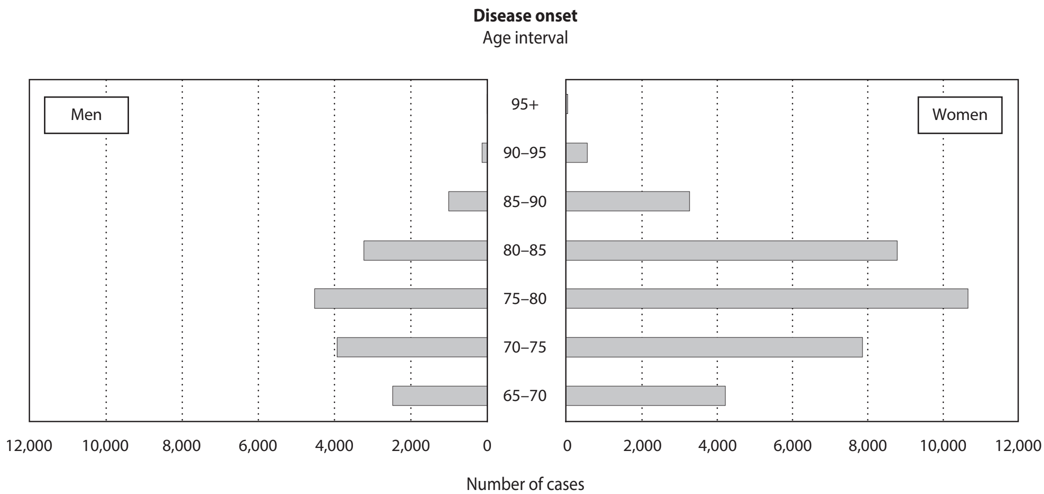
The present microsimulation study builds on the previous model by *Broulíková et al.* (2018) and provides comprehensive insight into the cost-effectiveness of the timely diagnosis and treatment of AD in the Czech Republic. Unlike the previous study, we derive the demographic composition of the Czech population living with AD from the national health-care registers. This step provides valuable information about the demographics of people living with AD in the country, and, importantly, allows us to appropriately address heterogeneity of the population as well as the uncertainty of the results. Moreover, to evaluate cost-effectiveness, we study both the effects of the timely diagnosis on costs and the quality of life of people living with AD.

DATA AND METHODS

People living with AD in the Czech Republic

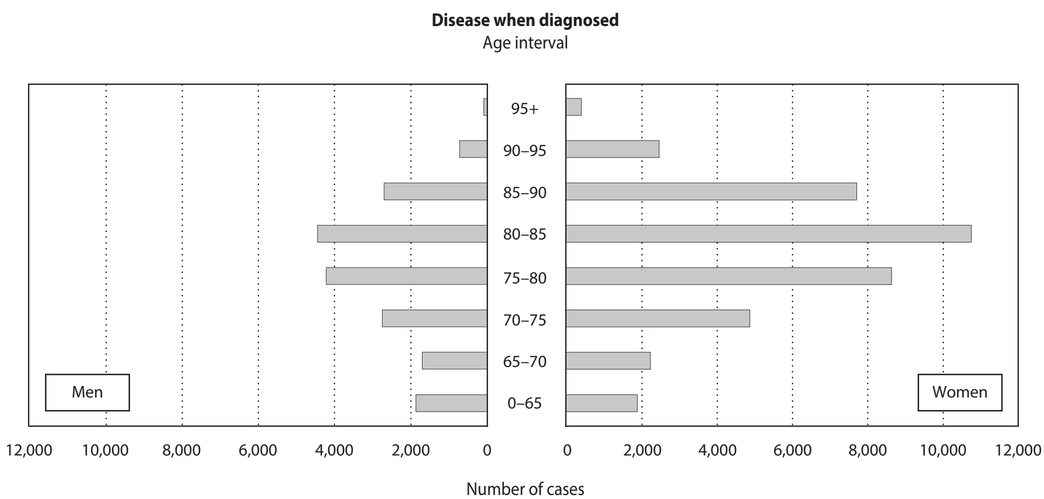
Our study modelled a heterogeneous cohort of 100,000 people with incident AD. Given that no official data stratified by age and gender on the incidence of AD in the Czech Republic have been published, we derived this information by combining data from the national health-care register with estimates of the time between onset of the disease and diagnosis in the country. First, all patients with an AD diagnosis that was made between 1994 and 2014 were filtered from the National Register of Hospitalised Patients maintained by the Institute of Health Information and Statistics. The data source and filtering strategy are described in detail elsewhere (*Broulíková et al.*, 2020). Second, to account for the discrepancy between

Figure 1a Age and gender structure of patients with Alzheimer's disease at the disease onset



Source of data: Nationwide Register of Hospitalized Patients, author's calculation.

Figure 1b Age and gender structure of patients with Alzheimer's disease when diagnosed



Source of data: Nationwide Register of Hospitalized Patients, author's calculation.

the time of the disease's onset and its official diagnosis, we simulated the mean time of untreated patients' cognitive decline from point of the disease's onset to moderate cognitive deficit and shifted the age of the identified these people accordingly, i.e. by 3.8 years. The moment of the disease's onset is conceptualised as a drop of the Mini mental examination score (*Folstein et al.*, 1975) to 28 points (*Getsios et al.*, 2012), and we assume that in the Czech Republic a diagnosis is made when the score drops to 15 points, which is the middle of the moderate disease phase. Third, to retain only the ages typical for AD's manifestation, we excluded 6,715 patients under the age of 65 at the approximate time of the disease's onset.

Figures 1a and 1b present the age and gender structure of patients with Alzheimer's disease at the moment of the disease's onset and at the time of diagnosis (i.e. when first recorded in the register). There were 57,559 (68% women) people living with AD identified in the register. The mean and median age was 78.5 and 80 years, respectively. After the adjustment for the delay in diagnosis and after removing those under the age of 65 years at the disease's onset, the number of people included decreased to 50,844 (70% women). The mean and median age decreased to 76.8 and 77 years, respectively. Out of this population, we sampled 100,000 people to receive a cohort mirroring age and gender profile of the Czech population with an incident AD that entered the model.

The model

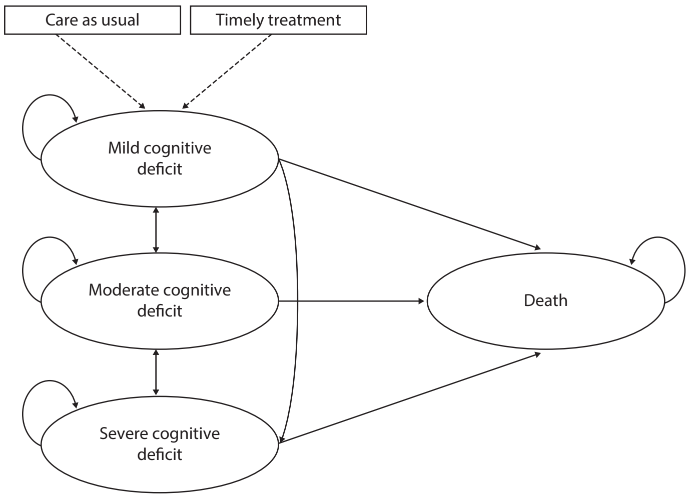
This microsimulation (i.e. patient-level) model has two branches: one branch represents the care usually provided (CAU) in the Czech Republic and the other a hypothetical case with timely treatment (TT). There are two main differences between the two branches of the model. First, as the purpose of the considered intervention is to ensure timely treatment, treatment is initiated in the TT branch when the cognitive deficit is mild. In the CAU branch, diagnosis only occurs when there is already a moderate cognitive deficit, which is a somewhat conservative assumption reflecting the Czech situation. Second, the probability of a patient being diagnosed and treated is higher in the TT branch than in the CAU (1 and 0.5, respectively). After accounting for mortality that occurs before diagnosis, two-thirds of people are diagnosed (timely) in the TT branch, and one-quarter of people are diagnosed (when they already have a moderate cognitive deficit) in the CAU branch of the model. The TT figure represents an ambitious but feasible goal (*Dementia Statistics Hub*, 2019), while the CAU figure is in line with the Czech estimates (*Mátl*, 2016).

Transitions among health states

Regardless of the branch, the patient migrates through four health states in discrete cycles lasting one year. The four states of this model are 'mild cognitive deficit' (MMSE 28 – 21), 'moderate cognitive deficit' (MMSE 20 – 11), 'severe cognitive deficit' (MMSE <11), and 'death' (see Figure 2). The transition between health states depends on the cognitive score of the patient in a given cycle as measured by the MMSE score. The backbone of the model are two established decline schemes representing the progression of the disease in an individual: the Mean decline scheme (*Weimer – Sager*, 2009) and a decline scheme defined by *Lopez et al.* (2005). The Mean decline scheme assumes that the annual decrease of the MMSE score is a random variable with a negative truncated normal distribution. The parameters of the Mean decline scheme differ for treated and untreated people, with the cognition of those who are treated declining slower. The decline scheme by Lopez assumes an annual decline expressed as a random variable with uniform distribution. This time the parameters differ according to a disease progression pace (slow and fast progressors), with the treated people having higher chance of slower progression than those who remain untreated. The parameters of both the Mean decline scheme and Lopez's decline scheme are summarised in Table 1. In the model, each patient has an equal chance (i.e. 50%) of declining according to the Mean scheme and Lopez's scheme.

Everyone enters the model untreated with an MMSE score of 28. The MMSE score in the current cycle is the annual cognitive decline as given by disease progression scheme subtracted from the score in the previous cycle. Cognition is tested every cycle and the new score determines whether the person has made any transition between health states

Figure 2 The structure of the model



Source: Author's illustration.

Table 1 Parameters of the decline schemes

Patient	Distribution	Distribution parameters	Unit
	Mean decline scheme		
Untreated	truncated normal	(0, ∞); μ 3.5; σ 1.5	MMSE per year
Treated	truncated normal	(0, ∞); μ 1.5; σ 1.5	MMSE per year
	Lopez's decline scheme		
Untreated fast	Uniform	(3, 6.8)	MMSE per year
Treated fast	Uniform	(3, 5)	MMSE per year
Untreated/treated slow	Uniform	(−1, 2)	MMSE per year

Source: Weimer and Sager 2009, authors' summary.

or has declined to a score when treatment is supposed to be initiated. The time of diagnosis is a random number from uniform distribution with an interval corresponding to moderate cognitive impairment for the CAU and mild cognitive impairment for the TT branch of the model.

Outcomes

The two main outcomes of this model are costs and quality of life (measured in quality-adjusted life years, QALYs). They are both discounted to the value of the year 2017 by a discount rate of 3% per annum. Four different cost categories are considered: the costs of 'informal care', 'timely diagnosis', costs of 'medication', and 'outpatient care'. While the costs of informal care capture the burden of care usually provided by family caregivers, the other three categories jointly make up the costs of providing medical treatment (health-care costs). The costs of informal care are periodically incurred from the start of the model's run until the patient's death regardless of whether she has actually been diagnosed by a doctor. The costs of medication and outpatient care are periodically incurred by treated patients from the cycle in which treatment is initiated until their death. The costs of timely diagnosis are one-time costs incurred only by treated patients in the TT branch of the model in the cycle in which their treatment is initiated.

With the costs of informal care and medication being conditional on the patient's health state, there are separate figures for mild, moderate, and severe cognitive deficit. These figures are based on Czech published research and a medical cost database (*SUKL*, 2015; *Holmerová et al.*, 2017). Outpatient care consists of an identical series of medical checks regardless of the patient's state of health (*Mohelská et al.*, 2015). The costs of timely diagnosis were estimated by costing a series of diagnostic procedures used in a foreign study (*Boustani et al.*, 2005) on the basis of reimbursements paid by Czech health insurers (*General Health Insurance Fund*, 2017). Specifically, a diagnostic scheme consisting of a visit to a general practitioner, visits to neurologists, a sampling of blood and cerebrospinal fluid, and – for only 5% of the individuals examined – an MRI or CT scan. According to *Boustani et al.* (2005), thirty-one people over the age of 65 need to be screened in order to diagnose one patient with AD. Consequently, the cost of timely diagnosis equals thirty-one times the costs of the described procedures per patient.

The quality of life enjoyed by a person with AD depends in a model cycle in each model cycleon the MMSE score. In particular, the patient's QALY amounts to 0.408 + 0.01MMSE– 0.159institutionalised – 0.004NPI + 0.051partner. These values are derived from a published regression equation (*Jönsson et al.* 2006). As in *Barnett et al.* (2014), the last three parts of the equation, *being institutionalised*, the *neuropsychiatric inventory instrument* score, and whether the patient lives with a *partner-caregiver*, go beyond the level of detail of this model and are omitted here.

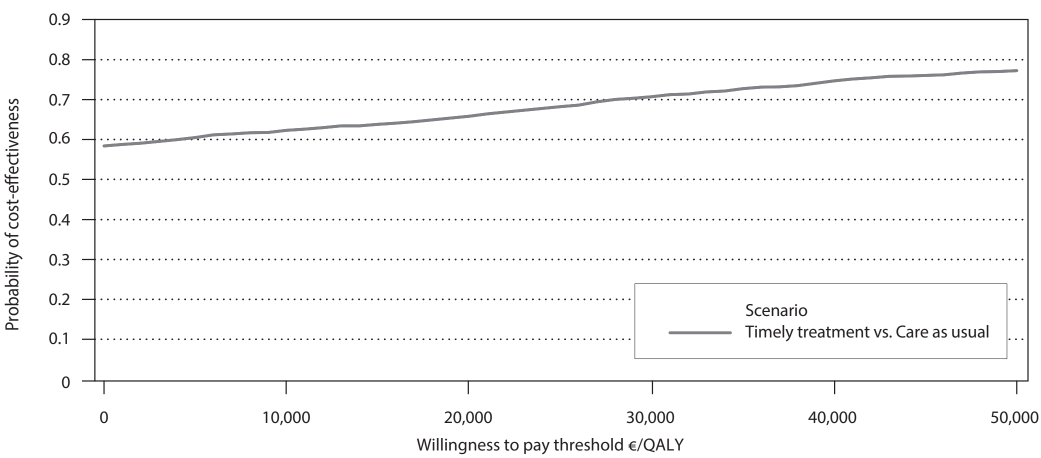
Sensitivity analyses

Parameters that may arguably have the biggest impact on the cost-effectiveness of the TT were chosen for the probabilistic sensitivity analyses (PSA) using the following distributions: the probability of being diagnosed and treated in the CAU branch (uniform (0.25; 1)), the MMSE scores at treatment initiation in the CAU branch (uniform (21; 28)) and the TT branch (uniform (21; 28)), and the costs of informal care (mild: gamma (0.86; 17,797), moderate: gamma (4.89; 5,212), severe: gamma (2.51; 12,092)). The results of the PSA are depicted using the cost-effectiveness acceptability curve (Briggs, 2000; Fenwick et al., 2001).

RESULTS

The results suggest that timely treatment of AD would represent a dominant strategy in the Czech Republic and would yield net benefit of EUR 13,751 per patient. The costs of the care usually administered have now reached EUR 122,430 per lifetime of an average patient, whereas the lifetime costs

Figure 3 Cost-effectiveness acceptability curve



Source: Author's calculation.

of an average patient who receives timely treatment amount to EUR 117,380. In terms of health effects, TT was observed to slightly improved quality of life from 3.67 cumulative lifetime QALYs for an average patient to 3.87 QALYs. Consequently, intervention results in higher quality at a lower cost. The corresponding ICER is –26,121.

Both the decrease in lifetime costs and the increase in lifetime quality is achieved by shifting a part of the time spent in moderate and severe health states to time spent in mild and moderate health states, respectively. In numerical terms, an average patient lives 0.55 year longer with a mild cognitive deficit and 0.54 years longer with a moderate cognitive deficit when TT is implemented than if CAU is provided. The difference in costs comes from the savings on informal care on the one hand and the increase in healthcare costs on the other. The health-care category consists of three sub-categories: diagnosis, medication, and outpatient care. Diagnosis (screening) is the main driver of the increase in health-care costs per patient in the TT branch, accounting for EUR 2,776 of the total increase of EUR 3,616. Nevertheless, these additional health-care costs are outweighed by savings on informal care amounting to EUR 8,666 (121,742 – 113,076) per patient.

The outputs of the PSA are summarised in the cost-effectiveness acceptability curve in Figure 3. In 58% of a thousand repetitions, timely treatment is cost-effective even for the willingness to pay for QALY equal to EUR 0. The probability of an intervention being cost-effective further grows to 76% for the willingness to pay EUR 45,000, which is the standardly used threshold in the Czech Republic (*SUKL*, 2017).

DISCUSSION

Using a microsimulation model, we found that timely treatment would represent a dominant strategy in the Czech Republic. Our results are consistent with the findings of previous studies in the United States (*Weimer – Sager*, 2009) and United Kingdom (*Getsios et al.*, 2012). In the same vein, *Handels et al.* (2017) found that the use of biomarkers in the cerebrospinal fluid improves a patient's prognosis by 11% and results in an average QALY gain of 0.046 and EUR 432 additional costs per patient, with an ICER of EUR 9,400.

The subgroup analysis by *Broulíková et al.* (2018) illustrated the role of gender and age in the cost-effectiveness of timely treatment in the Czech Republic. Generally, women with disease onset at the age of 70 and 80 enjoy higher benefits than men because of their longer life expectancy, with a net benefit of up to EUR 25,969 for the average woman who gets AD at the age of 70. For disease onset at the age of 90, the net benefit is slightly higher for men because, according to the Czech life tables, from this point their life expectancy becomes higher than that of women of the same age. As expected, the net benefit decreases with the degree of cognitive deficit at the time of diagnosis and with the person's age at the time of the disease's onset. The former effect is given by the opportunity for people who receive timely treatment to retain their independence for a longer period of time. The latter effect again follows from a longer lifetime period during which patients can enjoy the effects of treatment; i.e. people who get the disease at a very old age likely die before declining to a severe stage of the disease regardless of treatment. The difference between CAU and TT thus diminishes with age. However, the results showed a positive net benefit for all subgroups except patients who are over the age of 90 and are diagnosed with an MMSE score below 23. Even in this case, the opportunity loss from indicating costly treatment is negligible and amounts to tens of euros per patient lifetime.

The present study has several limitations. First, although we innovatively derive the profile of the incident cohort of people get AD from the Nationwide Register of Hospitalised Patients, this source does omit people who were diagnosed and treated in outpatient care and were never hospitalised (for dementia or other diagnosis). Consequently, the age and gender composition of the Czech population living with AD might be biased. The solution to this problem in future research is to use the newly established National Register of Reimbursed Health Services, which also contains diagnoses made in outpatient care. This source might be used further to track the health-care consumption of people living with dementia and, thus, further improve the unit costs used in the model. Second, reliable data on the time of diagnosis is missing. Our assumption regarding the current diagnostic timing in the phase of moderate cognitive impairment was rather conservative because available studies suggest that dementia is diagnosed late, usually shortly before death. However, more specific information is needed on the share of people who are diagnosed in the mild, moderate, and severe phase of the disease. Finally, dementia progression is better captured by a multidimensional progression scheme, such as the one recently suggested by *Green et al.* (2019).

This article provides an important contribution to the ongoing debate around dementia management in the Czech Republic. Our results generally support the effort to increase access to a timely diagnosis and to post-diagnostic support, which has been declared as a priority in the National Action Plan for Alzheimer's Disease and Related Illness (*Ministry of Health of the Czech Republic*, 2021). Future research should overcome the limitations mentioned above by incorporating more country-specific data from the registers, but also, specifically, by evaluating the effect of the policies introduced by this government document. An example of such a policy is the cognition screenings provided in the office of general practitioners followed by referral to a specialist for those with suspected cognitive impairment.

References

- Barnett, J. H. – Lewis, L. – Blackwell, A. D. – Taylor, M. 2014. Early intervention in Alzheimer's disease: a health economic study of the effects of diagnostic timing. *BMC Neurology* 14:101. https://doi.org/10.1186/1471-2377-14-101.
- Bernardes, C. – Massano, J. – Freitas, A. 2018. Hospital admissions 2000–2014: A retrospective analysis of 288 096 events in patients with dementia. *Archives of Gerontology and Geriatrics*, 77, pp. 150–157. https://doi.org/10.1016/j.archger.2018.05.006.
- Boustani, M. – Callahan, C. M. – Unverzagt, F. W. – Austrom, M. G. – Perkins, A. J. – Fultz, B. A. – Hui, S. L. – Hendrie, H. C. 2005. Implementing a screening and diagnosis program for dementia in primary care. *Journal of General Internal Medicine*, 20(7), pp. 572–577. https://doi.org/10.1111/j.1525-1497.2005.0126.x.
- Briggs, A. H. 2000. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics*, 17(5), pp. 479–500. https://doi.org/10.2165/00019053-200017050-00006.
- Broulíková, H. M. – Arltová, M. – Kuklová, M. – Formánek, T. – Čermáková, P. 2020. Hospitalizations and mortality of individuals with dementia: evidence from Czech national registers. *Journal of Alzheimer's Disease*, 75(3), pp. 1017–1027. https://doi.org/10.3233/JAD-191117.
- Broulíková, H. M. – Sládek, V. – Arltová, M. – Černý, J. 2018. The Potential Impact of Alzheimer's Disease Early Treatment on Societal Costs of Care in Czechia: A Simulation Approach. *Journal of Mental Health Policy and Economics*, 21(4), pp. 147–161. https://doi.org/10.1016/j.jval.2018.09.1664.
- CZSO. 2020. *Demografická příručka 2019. Obyvatelstvo podle hlavních věkových skupin a pohlaví v letech 1920–2019* [Population by main age group and sex: 1920–2019]. https://www.czso.cz/documents/10180/121739374/130055200109.xlsx/2c3f262b-50e3-45ec-9137-1bcd19ab5bc4?version=1.1.
- Dementia Statistics Hub. 2019. *Alzheimer's Research UK*. https://www.dementiastatistics.org/statistics/diagnoses-in-the-uk/.
- Fenwick, E. – Claxton, K. – Sculpher, M. 2001. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Economics Lette*r, 10(8), pp. 779–787. https://doi.org/10.1002/hec.635.
- Folstein, M. F. – Folstein, S. E. – McHugh, P. R. 1975. Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), pp. 189–198. https://doi.org/10.1016/0022-3956(75)90026-6.
- General Health Insurance Fund. 2017. *Reimbursed Health Services*. Prague, General Health Insurance Fund.
- Getsios, D. – Blume, S. – Ishak, K. J. – Maclaine, G – Hernández, L. 2012. An economic evaluation of early assessment for Alzheimer's disease in the United Kingdom. *Alzheimer's & Dementia*, 8(1), pp. 22–30. https://doi.org/10.1016/j.jalz.2010.07.001.
- Green, C. – Handels, R. – Gustavsson, A. – Wimo, A. – Winblad, B. – Sköldunger, A. – Jönsson, L. 2019. Assessing cost-effectiveness of early intervention in Alzheimer's disease: An open-source modeling framework. *Alzheimer's & Dementia*, 15(10), pp. 1309–1321. https://doi.org/10.1016/j.jalz.2019.05.004.

- Handels, R. L. H. – Wimo, A. – Dodel, R. – Kramberger, M. G. – Visser, P. J. – Molinuevo, J. L. – Verhey, F. R. J. – Winblad, B. 2017. Cost-Utility of Using Alzheimer's Disease Biomarkers in Cerebrospinal Fluid to Predict Progression from Mild Cognitive Impairment to Dementia. *Journal of Alzheimer's Disease*, 60(4), pp. 1477–1487. https://doi.org/10.3233/JAD-170324.
- Holmerová, I. – Hort, J. – Rusina, R. – Wimo, A. – Šteffl, M. 2017. Costs of dementia in the Czech Republic. *The European Journal of Health Economics*, 18(8), pp. 979–986. https://doi.org/10.1007/s10198-016-0842-x.
- Institute for Health Metrics and Evaluation. 2019. *Global Health Data Exchange.*
- Jönsson, L. – Andreasen, N. – Kilander, L. – Soininen, H. – Waldemar, G. – Nygaard, H. – Winblad, B. – Jönhagen, M. E. – Hallikainen, M. – Wimo, A. 2006. Patient- and proxy-reported utility in Alzheimer disease using the EuroQoL. *Alzheimer Disease & Associated Disorders*, 20(1), pp. 49–55. https://doi.org/10.1097/01.wad.0000201851.52707.c9.
- Lopez, O. L. – Becker, J. T. – Saxton, J. – Sweet, R. A. – Klunk, W. – DeKosky, S. T. 2005. Alteration of a Clinically Meaningful Outcome in the Natural History of Alzheimer's Disease by Cholinesterase Inhibition. *Journal of the American Geriatrics Society*, 53(1), pp. 83–87. https://doi.org/10.1111/j.1532-5415.2005.53015.x.
- Lužný, J. – Holmerová, I. – Wija, P. – Ondrejka, I. 2014. Dementia Still Diagnosed Too Late - Data from the Czech Republic. *Iranian Journal of Public Health*, 43(10), pp. 1436–1443.
- Mátl, O. – Mátlová, M. – Holmerová, I. 2016. *Zpráva o stavu demence 2016.* Praha: Česká alzheimerovská společnost.
- Ministry of Health of the Czech Republic. 2021. *National Action Plan for Alzheimer's Disease and Related Illnesses.* Czech Republic.
- Mohelská, H. – Maresova, P. – Valis, M. – Kuca, K. 2015. Alzheimer's disease and its treatment costs: case study in the Czech Republic. *Neuropsychiatric Disease and Treatment*, 11, pp. 2349–2354. https://doi.org/10.2147/NDT.S87503.
- Murman, D. L. 2015. The Impact of Age on Cognition. *Seminars in Hearing*, 36(3), pp. 111–121. https://doi.org/10.1055/s-0035-1555115.
- Mátl, O. – Holmerová, I. 2016. *Zpráva o stavu demence 2016.* Česká alzheimerovská společnost, o.p.s.
- Prince, M. – Wimo, A. – Guerchet, M. – Ali, G. – Wu, Y. – Prina, M. 2015. *The global impact of dementia: an analysis of prevalence, incidence, cost and trends.* World Alzheimer Report.
- SUKL. 2015. http://www.sukl.cz/modules/medication/search.php.
- SUKL. 2017. *Postup pro posuzování nákladové efektivity* [Guidelines for the cost-effectiveness evaluation]. Prague.
- UN. 2018. *World Population Ageing 2017 Highlights UN*, London.
- Waldemar, G. – Phung, K. T. – Burns, A. – Georges, J. – Hansen, F. R. – Iliffe, S. – Marking, C. – Rikkert, M. O. – Selmes, J. – Stoppe, G. 2007. Access to diagnostic evaluation and treatment for dementia in Europe. *International journal of geriatric psychiatry*, 22(1), pp. 47–54. https://doi.org/10.1002/gps.1652.
- Weimer, D. L. – Sager, M. A. 2009. Early identification and treatment of Alzheimer's disease: Social and fiscal outcomes. *Alzheimer's & Dementia*, 5(3), pp. 215–226. https://doi.org/ 10.1016/j.jalz.2009.01.028
- Winblad, B. – Amouyel, P. – Andrieu, S. – Ballard, C. – Brayne, C. – Brodaty, H. – Cedazo-Minguez, A. – Dubois, B. – Edvardsson, D. – Feldman, H. – Fratiglioni, L. – Frisoni, G. B. – Gauthier, S. – Georges, J. – Graff, C. – Iqbal, K. – Jessen, F. – Johansson, G. – Jönsson, L. – Kivipelto, M. – Knapp, M. – Mangialasche, F. – Melis, R. – Nordberg, A. – Rikkert, M. O. – Qiu, C. – Sakmar, T. P. – Scheltens, P. – Schneider, L. S. – Sperling, R. L. – Tjernberg, O. – Waldemar, G. – Wimo, A. – Zetterberg, H. 2016. Defeating Alzheimer's disease and other dementias: a priority for European science and society. *The Lancet Neurology Commission*, 15(5), pp. 455–532. https://doi.org/10.1016/S1474-4422(16)00062-4.
- World Health Organization. 2012. *Dementia: a public health priority*, World Health Organization.
- World Health Organization. 2019. *WHO Global Dementia Observatory: Provisional Country Profile 2017.* Department of Mental Health and Substance Use, WHO.
- World Health Organization. 2020. *Dementia.* from https://www.who.int/news-room/fact-sheets/detail/dementia.

HANA MARIE BROULÍKOVÁ

Hana M. Broulíková is a researcher in the Department of Health Sciences at Vrije Universiteit Amsterdam. In 2019, she finished her PhD on the cost-effectiveness of timely dementia diagnosis at the Faculty of Informatics and Statistics, Prague University of Economics and Business. She worked at the Ministry of Health of the Czech Republic as a coordinator of the National Action Plan for Alzheimer's Disease and Related Illnesses 2020–2030, and she continues to advise the ministry on its implementation. She also collaborates with the National Institute of Mental Health on an economic evaluation of the ongoing reform of mental health care in the Czech Republic.

MATĚJ KUČERA

Matěj Kučera studied General Medicine at the Third Faculty of Medicine at Charles University in Prague in 2021 and is currently PhD student under joint supervision at Charles University Prague and Vrije Universiteit, Amsterdam. The topic of his thesis is "Prevention of dementia and associated costs in the Czech Republic". Since 2018 he has been working as a researcher in the Department of Social Psychiatry at the National Institute of Mental Health and a data specialist in the Destigmatisation Project, which is a part of psychiatric care reform in the Czech Republic. In his research work he focuses on topics such as the prevalence of depression among seniors, the destigmatisation of people with mental illness, and the psychological and emotional well-being of gifted children.

MARKÉTA ARLTOVÁ

Markéta Arltová studied economic statistics at the Prague University of Economics and Business. She currently works as an associate professor in the Department of Public Finance, Faculty of Finance and Accounting at the Prague University of Economics and Business. She specialises in econometric time series analysis and in demographic analyses.

The 13th Conference of Young Demographers Will Take Place in February 2022

The Conference of Young Demographers traditionally serves as an exceptional opportunity to spend four days discussing current demographic issues. It gives students and young researchers a chance to learn and receive opinions and advice in a friendly environment from their peers, who come to the conference from all over the world. It is our pleasure to announce that this year the conference is co-organised in cooperation with the Association for Young Historical Demographers.

The 13th annual Conference of Young Demographers will take place from **1 to 4 February 2022** in Prague at the Faculty of Science, Charles University. Although the conference is mainly aimed at PhD students in the field of (historical) demography, all young researchers – and researchers who are young at heart – from various fields of population studies are welcome to attend. The working language of the conference is English.

Keynote lectures have kindly been promised by Dr Mariona Lozano Riera and Dr José Manuel Aburto. Mairona holds a PhD in sociology and works as a researcher at the Centre for Demographic Studies at the Autonomous University of Barcelona (CED UAB). Her research focuses on the labour market and the position of women in it. José Manuel is currently working at the Leverhulme Centre for Demographic Science (LCDS) at the University of Oxford. His work centres on the study of lifespan and life-expectancy inequalities using methods of mathematical demography.

Following the success of past years, a workshop will be included in the conference programme. The workshop will be organised by Dr Ilya Kashnitsky, DataViz magician and researcher at the Centre for Population Dynamics at the University of Southern Denmark (CPop SDU).

This year's conference will focus more on posters and work-in-progress. We would therefore like to especially encourage submissions to the poster session.

The final programme of the conference will be announced in January 2022.

For more information, including information on being a passive conference participant, please visit our website (youngdemographers.github.io). Registration for passive participation will open in January and will close on 30 January. If you have any questions please feel free to contact us at: yd.demographers@gmail.com. We look forward to meeting you in Prague!

Anna Altová – Klára Hulíková – Barbora Janáková – Kateřina Maláková – Tim Riswick – Jitka Slabá – Martin Vondrášek

Conference RELIK 2021

In 2021 the Reproduction of Human Capital - Mutual Links and Connections conference (RELIK), which seeks to connect demography with other areas, was held at the University of Economics and Business in Prague. This was the 14th year of the RELIK international conference, but this year the conference took place for the first time in a hybrid format on 4 and 5 November 2021.

A large number of papers were submitted to the conference in 2021 and this was reflected in the final programme, which included more than 80 papers. The topics of papers presented in person included, for example, the issue of debt execution in relation to pensions, the sociodemographic structure of the Roma population in the 2011 population census in Slovakia, the impact of education on reproductive behaviour, the financing of the Czech health-care system, beta regression used to analyse the US presidential election, or the experience of employees working from home.

Presentations dedicated to topics such as the payment of the childcare allowance, social quality, post-productive economic activity, labour migration, motivational factors, employee satisfaction and benefits, or tertiary education and digital competencies were presented in the online forum. Many of the papers also focused on the evaluation of the Covid-19 pandemic, such as the change in human capital in informal networks and organisational performance, changes in the number of self-employed people, the trend in unemployment during the pandemic, and the importance of hygiene during the pandemic.

As in previous years, the young scientists' sections, intended for researchers just starting out, were included again this year. Papers presented by young scientists dealt with the effects of technostress on higher education, industry 4.0, the role of social networks and social capital in the labour market, or mortality according to education levels. The theme of Covid-19 resonated here as well in topics such as research on the expansion of remote working in V4 countries, the Covid-19 pandemic media image, and security risks in communication or methods of education during the pandemic.

The conference proceedings from this year and previous years of the RELIK conference are available online at: https://relik.vse.cz/archive/.

Proceedings from 2017 to 2020 were included in the Web of Science database, and the organisational team will work to get the proceedings from 2021 included in the Web of Science database, too.

The conference took place in a friendly atmosphere. Many interesting questions were raised and there were many stimulating discussions.

Jana Vrabcová

LIFE EXPECTANCY IMPROVEMENTS BY AGE, CLASS, AND MORTALITY CHAPTER IN FRANCE, CZECHIA, AND THE UNITED STATES

Onofre Alves Simões[1)] – Andrey Ugarte Montero[2)]

Abstract
Using the decomposition method, this article examines the dynamics of life expectancy. Three developed countries with relevant differences, Czechia, France, and the United States, were chosen for analysis in order to highlight similarities and differences. The analysis covers more than 40 years, 12 age groups, and 20 mortality chapters. The results reveal a pattern: first, mortality at birth improves; then the survival of lives under 65 increases; finally, improvements come from extending the life of seniors.

Keywords: life expectancy, decomposition method, mortality chapters, France, Czechia, United States

https://doi.org/10.54694/dem.0286 **Demografie,** 2021, **63 (4): 228–245**

1 INTRODUCTION

Understanding and explaining the sources of changes in demographic indicators such as life expectancy at different ages has been in the interest of researchers for a long time. However, the topic has gained relevance in recent decades owing to the financial difficulties that have affected, or are expected to affect, pension fund schemes, and social security systems in general. Because of this, attention has increasingly turned to obtaining a better understanding of mortality, the patterns that existed in the past, and how these patterns have been evolving over time, as a way of enhancing the scientific knowledge that will enable the community to better predict the future. Contributions of age and causes of death to life expectancy at birth (LE) can be calculated with decomposition methods (*Andreev – Shkolnikov – Begun*, 2002), (*Arriaga*, 1984; 1989), (*Das Gupta*, 1978). This approach has been widely used for various purposes, in particular, to research the effects on mortality of inequalities in socioeconomic conditions and access to health care, in different countries and regions (*Agyepong et al.*, 2017; *Bergeron-Boucher – Ebeling – Canudas-Romo*, 2015; *Khang et al.*, 2010; *Liu et al.*, 2016; *Martikainen – Valkonen – Martelin*, 2001; *Martikainen – Makela – Peltonen – Myrskyla*, 2014; *Mondal – Shitan*, 2014; *Murwirapachena – Mlambo*, 2015; *Preston – Stokes*, 2012; *Shkolnikov – Andreev – McKee – Leon*, 2013; *Tarkiainen – Martikainen – Laaksonen – Valkonen*, 2012; *Wang et al.*, 2016; *Yang et al.*, 2012). Many other studies focus on the gender gap or other life expectancy issues (*Al-Ramadhan*, 2008; *Auger et al.*, 2014; *Auger et al.*, 2012; *Hosseinpoor et al.*, 2012;

1) Universidade de Lisboa, Lisbon School of Economics & Management and REM, CEMAPRE. Contact: onofre@iseg.ulisboa.pt.
2) Université de Lausanne, Department of Actuarial Sciences. Contact: andrey.ugarte@unil.ch.

Le et al., 2015; *Rosella et al.*, 2016; *Simmons*, 2018; *Trovato – Heyen*, 2006; *Trovato – Lalu*, 1997; *Trovato – Odynak*, 2011; *Vaupel – Romo*, 2002; *Waldon*, 1983; *Waldon – McCloskey – Earle*, 2005; and again *Yang et al.*, 2012).

These works decompose the contributions of age and cause-specific mortality to changes in life expectancy. In our work, we apply the same methodology to obtain deeper knowledge of the dynamics behind life expectancy changes, not only because the analysis covers a very long period of time, but mostly because more than 12 age groups and 20 mortality chapters, as defined by the International Classification of Diseases, are under study. Three developed countries with significant historical differences with respect to mortality (France, Czechia, and the United States) have been chosen here as case studies in order to enable direct comparisons. These countries were selected for three main reasons: they belong to the small group of countries with cause-of-death (CoD) information available in the Human Mortality Database (HMD); they have nearly the same range of the gross national income of this group (available at https://data.worldbank.org/indicator/ny.gnp.pcap.pp.cd); and they have populations with very different sizes and quite different histories. Other choices could have been made, but these specific ones seemed the most interesting to us.

We can find a number of studies on mortality and longevity that have applied the decomposition method and other methods as well that give particular attention to some (or sometimes all) of these countries (for more, see, e.g., *Ho – Hendi*, 2018; *Hulíková Tesárková – Kašpar – Zimmermann*, 2015; *Leon*, 2011; *Meslé*, 2004; *Meslé – Vallin*, 2017; *Meslé – Vallin – Pyrozhkov*, 2012; and *Woolf – Schoomaker*, 2019). This is important, as it sometimes allows for a comparison of results.

Our research focuses on quantifying the changes in LE and on estimating the contributions attributable to each of the different age groups and mortality chapters in an exhaustive and in depth way across a period of over four decades. In some of the cited works, this type of exercise has also been carried out, but our research design contributes to a better understanding of mortality and longevity differences and their origins, the evolution of observed patterns over time, and the historical change in the decomposition of these patterns.

In the following sections, we analyse *in depth and for a 44-year period* the changes in LE. To achieve this, mortality variations are decomposed according to age, cause of death, and the relationship between them.

The information available in the Cause of Death Database (CoDD) covers a period from 1959 to 2013 for the United States, 1958 to 2015 for France, and 1968 to 2015 for Czechia. In order to make the results comparable, a common time interval is necessary. Furthermore, it was important to break the timeline into three shorter periods, to better capture the evolution of the phenomenon studied here. Although other criteria could have been used, a 'neutral' rule was that they should be approximately the same length. As a result, the analysis goes from 1970 to 2013, dividing the timeline from 1970 to 1984, from 1985 to 1998, and from 1999 to 2013.

2 BACKGROUND, DATA, AND METHOD

In general, life expectancy is one of the most popular concepts to use to analyse mortality. It is normally taken as an indicator of human health owing to its capacity to 'summarize mortality in a single measure' (*Augern et. al*, 2014). The picture of the indicator internationally shows a range of shades that have been changing through the years, and they are of value for understanding the situation today. Three aspects of longevity are worth mentioning.

First, global LE for both sexes went from 66.5 years in 2000 to 72 years in 2015 (*World Health Organization*, 2019). According to the United Nations (2019), it increased from 64.2 years in 1990 to 72.6 years in 2019 and it is expected to increase further to 77.1 years in 2050. While considerable progress has been made in closing the longevity differential between countries, large gaps remain. In 2019, LE in the least developed countries was 7.4 years below the global average, largely because of persistently high levels of child and maternal mortality, as well as high levels of violence and conflict and the continuing impact of the HIV epidemic. Second, the number of people in the world aged 60 and older has doubled compared to the number in 1980, and it is expected that by 2050 16% of the world's population will be aged 65 and over – compared to 9% in 2019 (cf. *United Nations*, 2019). Third, there seems to be a growing interest in what are known as the 'Blue Zones' (see *Poulant – Buettner – Pes*, 2019), a selected group of places with an extraordinarily high concentration of people living beyond the age of 100. According to *Buettner and Skemp* (2016), members of this select group include Loma Linda in California (USA), Nicoya in Costa Rica, Sardinia in Italy, Icaria in Greece, and Okinawa in Japan.

A tool called the Mortality Analysis Calculator (MAC) was developed to analyse the mortality information from the HMD and the CoDD. The MAC was created by Ugarte for the Department of Research of the Society of Actuaries – actuaries have become players in the search for solutions and contingency plans for countries facing issues associated with population ageing. *The Actuarial Association of Europe* (2019, p. 7) found that because of population ageing 'costs are projected to rise in every country on health and long-term care spending. These projections depend not only on the population projections but also on how life expectancy increases translate into healthy life expectancy and how the demand for health and long-term care services evolve'. One of the main features of the Mortality Analysis Calculator is that it decomposes the changes in a mortality indicator by the contribution to this change that can be attributed to each age group and mortality chapter, when applicable. In order to decompose the variations, the tool uses algorithms that already exist in the literature, relying mostly on the seminal contributions from *Andreev, Shkolnikov and Begun* (2002) and *Arriaga* (1984; 1989).

2.1 Decomposition of the changes in a mortality indicator by age group

Let's consider the changes between years t_1 and t_2 ($t_2 > t_1$) in a mortality indicator estimated for a specific age *a*. To compute the contribution to changes attributable to different ages/age groups, we use the algorithm this is presented in *Andreev, Shkolnikov and Begun* (2002). Their paper presents a formula for decomposing changes between two periods of time, but in general it compares two different 'experiences' of an indicator – for example, whether the change comes from time, gender, or an ethnic group. Although the underlying principles remain the same, it was convenient for the sake of clarity to adjust slightly the mathematical notation of the paper by Andreev, Shkolnikov, and Begun. For this reason, we will briefly describe the method in the paragraphs below.

If we denote a mortality indicator – life expectancy in this case – for age *a* as Ind_a, then the attribution of change in the mortality indicator between years t_1 and t_2 at age *a* attributable to age *x* is

$$ {}^{a}_{x}\delta^{2-1} = \frac{l^2_x}{l^2_a}\underbrace{(Ind^2_x - Ind^1_x)}_{Variation\ age\ x} - \frac{l^2_{x+1}}{l^2_a}\underbrace{(Ind^2_{x+1} - Ind^1_{x+1})}_{Variation\ age\ x+1}, \quad (1)$$

where l^j_x represents the number of survivors aged *x* for year t_j, $j = 1, 2$, and Ind^j_x represents the level of the indicator for age *x* in year t_j.

As has been established by *Andreev – Shkolnikov – Begun* (1982) and *Pressat* (1985), the result of computing ${}^{a}_{x}\delta^{2-1}$ for the decomposition by age does not necessarily have to be the same as that of $-{}^{a}_{x}\delta^{2-1}$, so they suggested replacing ${}^{a}_{x}\delta^{2-1}$ with

$$ {}^{a}_{x}\delta = 0.5({}^{a}_{x}\delta^{2-1} - {}^{a}_{x}\delta^{1-2}) \quad (2)$$

to obtain the attributable contribution coming from age *x* to the change in the indicator for age *a*, so that

$$ (Ind^2_a - Ind^1_a) = \sum_{x=a}^{\omega} {}^{a}_{x}\delta . \quad (3)$$

In the calculations below, (2) and (3) will be used.

2.2 The decomposition of changes attributable to different causes of death

In this paper we also compute the contribution of the evolution of causes of death for mortality indicators. As in the previous section, consider ${}^{a}_{x}\delta$ as the contribution to the change that is attributable to age *x* for the mortality indicator at age *a* between years t_1 and t_2. Assume that the environment is affected by *n* diseases, so that we denote the change in the indicator at age *a* between years t_1 and t_2 that is due to disease *i* ($i = 1, 2, 3, \dots, n$) as ${}^{i}_{a}\alpha^{2-1}$. Following the reasoning in Arriaga's method, the change associated with disease *i* is

$$ {}_{a}^{i}\alpha^{2-1} = \sum_{x=a}^{\omega} {}_{x}^{a}\delta \, {}^{i}\Lambda_{x}^{2-1}, \quad {}^{i}\Lambda_{x}^{2-1} = \frac{{}^{i}q_{x}^{2} - {}^{i}q_{x}^{1}}{q_{x}^{2} - q_{x}^{1}}, \tag{4} $$

where ${}^{i}q_{x}^{k}$ denotes the mortality rate associated with disease i for age x during year t_k. Similarly, q_x^k represents the total mortality rate for age x (i.e. including all n diseases) during year t_k.

Analysing the formula, it becomes evident that it distributes the changes in the indicator that are attributable to different ages using the changes registered in the mortality rates per cause. In this sense, ${}^{i}\Lambda_{x}^{2-1}$ is just the proportion of the overall change in mortality for age x that was registered between times t_1 and t_2 that is attributable to cause i.

The underlying assumption is that the contributions to the changes that can be attributed to a cause are directly proportional to the variations registered in the respective mortality rates. Then, clearly,

$$ \alpha_{a}^{2-1} = \sum_{i=1}^{n} {}_{a}^{i}\alpha^{2-1} = \sum_{i=1}^{n} \sum_{j=a}^{\omega} {}_{j}^{a}\delta \, {}^{i}\Lambda_{j}^{2-1} \tag{5} $$

is the overall change in the mortality indicator for age a between years t_1 and t_2.

3 LIFE EXPECTANCY AT BIRTH FROM 1970 TO 2013

LE shows a remarkable evolution in France, Czechia, and the United States. In all three countries the values of the indicator have increased as mortality dynamics have evolved at different stages of human life and as the effect of diseases have varied over time. We will present next the decomposition of the changes in LE per age group and the contributions to this that are attributable to the different mortality chapters, which we calculate using equations (2)–(5).

3.1 Life expectancy at birth in France

France experienced a sustained increase in LE during the period we are interested in. LE rose from 68.4 years for males and 75.8 years for females in 1970 to 78.8 years and 85.1 years (respectively) in 2013. This represents a change of 10.4 years for males and 9.3 years for females over the entire period. These changes, however, did not occur in a 'uniform' manner and can be explained in different ways across this time interval.

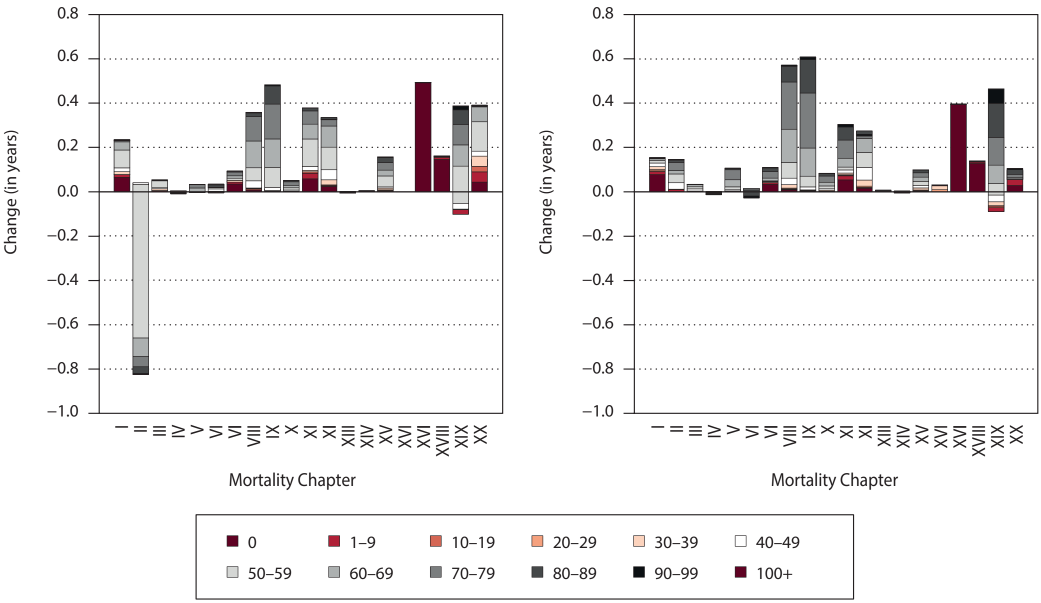
3.1.1 1970–1984

During this period, females experienced a bigger increase in their LE and registered a total change of 3.54 extra years compared to 2.77 extra years for males. By age group, the changes in LE during this period were affected by a very important increase that was caused by mortality changes at very young ages. For example, the contribution to the change that can be attributed to the first year of life alone accounts for around 29% of the total change for males and 19% for females. In general, when we consider the changes that are attributable to ages under 60, the mortality changes in these age groups contributed to almost 54% of the total increase (1.5 years) in the case of males, whereas for females they contributed to 44% (or 1.56 years). In the case of males, estimates show that about 1.3 years (out of the 2.77 years) are attributable to changes in mortality at ages 60 and older, while this increases to 1.98 years when we consider the total change of 3.54 years for females. From this information, it seems that women at older ages saw a much more significant improvement in their mortality prospects than men, who experienced a greater mortality reduction at young ages.

When analysing the results and the estimated changes that can be attributed to the different mortality chapters, the estimates show that some of the most relevant increments in life expectancy, for both genders, occurred as a result of improvements in Cerebrovascular Diseases (Mortality Chapter IX). The changes in mortality between 1970 and 1984 that were due to this group of diseases represented an estimated improvement in LE of about half a year for males and 0.61 years for females. Changes in mortality due to Heart Diseases (Mortality Chapter VIII) also played a central role in improving the indicator for both genders, but the effect is much higher in the case of females since it is estimated that women gained over half a year in LE (in contrast to 0.36 years for males). Similar results are obtained for Mortality Chapter XIX, which is Ill-Defined or Unknown Causes, registering an improvement of 0.46 years for females and 0.38 years for males.

Mortality Chapter XX – External Causes – including death due to accidents, homicides, poisoning, and the like, contributes considerably

Figure 1 Changes in LE by CoD and age, 1970–1984: France, males (left) vs females (right)



Source: HMD, CoDD and MAC.

to the improvement in LE for men (0.39 years), but not for women (0.11 years). Finally, it is worth noting that the results show an important decrease in LE for males due to Chapter II, Malignant Neoplasm, which caused an estimated decrease in life expectancy of 0.82 years for men. Appendix 1 shows all the details of the composition of changes in life expectancy by cause of death.

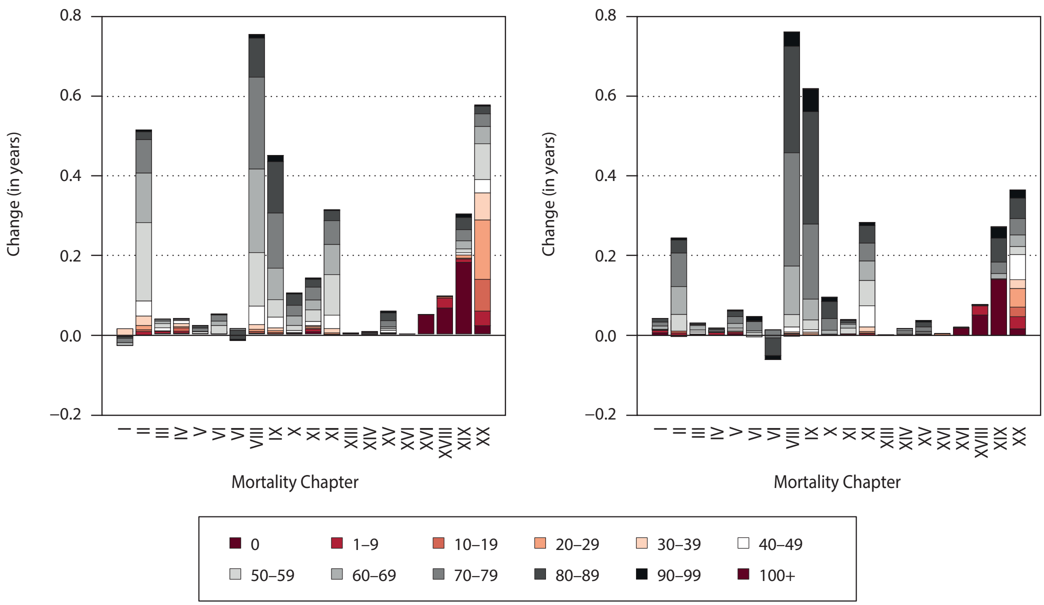
Figure 1 shows the estimated changes in LE by mortality chapter. The overall size of the bars represents the total variations estimated (axis *y*) for the mortality chapters (axis *x*). Every bar is divided into smaller segments with different colours that represent where the overall change comes from in terms of contributing age groups. When parts of a particular bar are in both the positive and negative quadrants, this means that some ages contributed to a decrease in life expectancy for that mortality chapter, whereas other age groups contributed to a gain. An example of this is Mortality Chapter XIX for females (with an overall gain of 0.525 years).

3.1.2 1985–1998

Unlike what happened in the previous period, in the period between 1985 and 1998 LE increased more for males than for females in France (with increments of 3.48 years for men and 2.96 years for women). The variation in LE attributable to each age group underwent several changes in structure: the contribution that is attributable to the first year of life decreased greatly in both absolute and relative numbers and represented about 9% and 8% of the overall change for men and women, respectively. Moreover, the total change in LE for men at ages younger than 60 accounts for 49% (1.72 years) of the total variation, whereas 51% comes from ages older than 60. For women, the difference is more evident, since 29% of the changes (0.87 years) come from age groups younger than 60 and 71% come from seniors.

As regards gains and losses in LE, we can see that France experienced very successful improvements in the treatment of diseases of the Circulatory System, as a result of which Mortality Chapter VIII improved the indicator by 0.76 years for both men and women. This shows that the improvement in this cause of death was much more significant than the one observed in the previous 14-year period. The estimated increases in LE attributable to Cerebrovascular Diseases, which had played a very prominent role in LE increase

Figure 2 Changes in LE by CoD and age, 1985–1998: France, males (left) vs. females (right)



Source: HMD, CoDD and MAC.

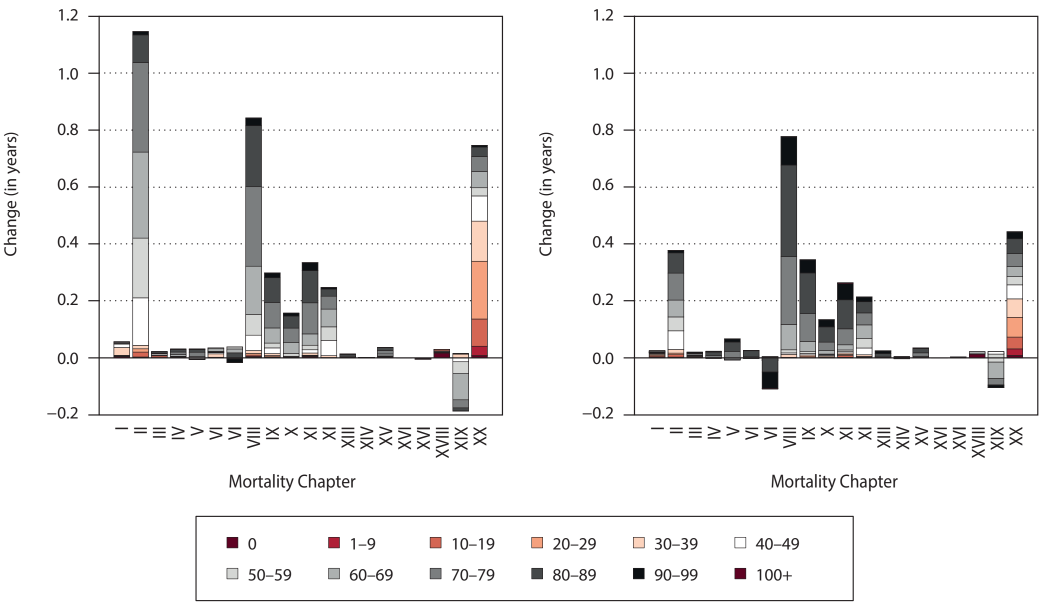
in the previous term for both genders, continue to be relevant, especially for women. In this period, this mortality chapter placed itself as the second and fourth main cause contributing to the enhanced indicator, for women and men. The estimates of the gains that are due to this chapter yield an increment of 0.45 years in LE for males and 0.62 years for females. The mortality experience observed during this period also suggests that there were important improvements in deaths due to external causes, which is the chapter responsible for the second- and third-biggest increments for males and females respectively. In this case, the gain in years of life expectancy is more significant for men than for women. The top three improvements in causes of death are responsible for about 53% and 59% of the overall change in life expectancy during this period. Figure 2 shows this and other results.

3.1.3 1999–2013

From 1999 to 2013, the French population once again saw an important improvement in LE. By the end of this period, like in the previous period, French men experienced a more significant increase than French women – men's LE improved by 3.83 years while women's LE increased by 2.6 years. Not only did men experience a more pronounced improvement in the LE indicator for the second consecutive period, but the difference, when compared with the gains for women, is much more noticeable: 1.23 years greater than that of French women (as opposed to half a year in the previous period). This certainly contributed to a decrease in the gender gap.

The changes in LE during this period confirm a tendency detected earlier: the improvements that are attributable to young ages start to lose their relative importance as mortality improvements start to come from the longevity of the elderly. During this 14-year period, only about 2% of the variation is attributable to changes in the mortality of newborns in both genders, whereas the ages under 60 contributed to 40% of the overall change in the case of males and 25% in the case of females. Moreover, the variations attributable to senior ages (over 60) amount to 2.3 years for males and 1.95 years for females. In this sense, the change in the structure becomes more significant in the case of French men as they seem to keep up better with female mortality at older ages.

Figure 3 Changes in LE by CoD and age, 1999–2013: France, males (left) vs females (right)



Source: HMD, CoDD and MAC.

In this third period (see Figure 3), life expectancy primarily improved as a result of the influence of Mortality Chapter II, Malignant Neoplasms, which generated an improvement of 1.16 years in LE for men. The evolution of mortality caused by Heart Diseases continued to affect positively life expectancy for both genders (with an estimated attribution of an increase of 0.85 years in this indicator for men and 0.78 years for women). Deaths due to external causes (Mortality Chapter XX) also helped to increase LE, with a particularly strong effect in the case of men (0.76 years). Improvements related to Respiratory Diseases started to be more prominent during this term, whereas causes such as Nervous System Disorders (Mortality Chapter VII) caused subtle decreases in life expectancy.

3.2 Life expectancy at birth in Czechia

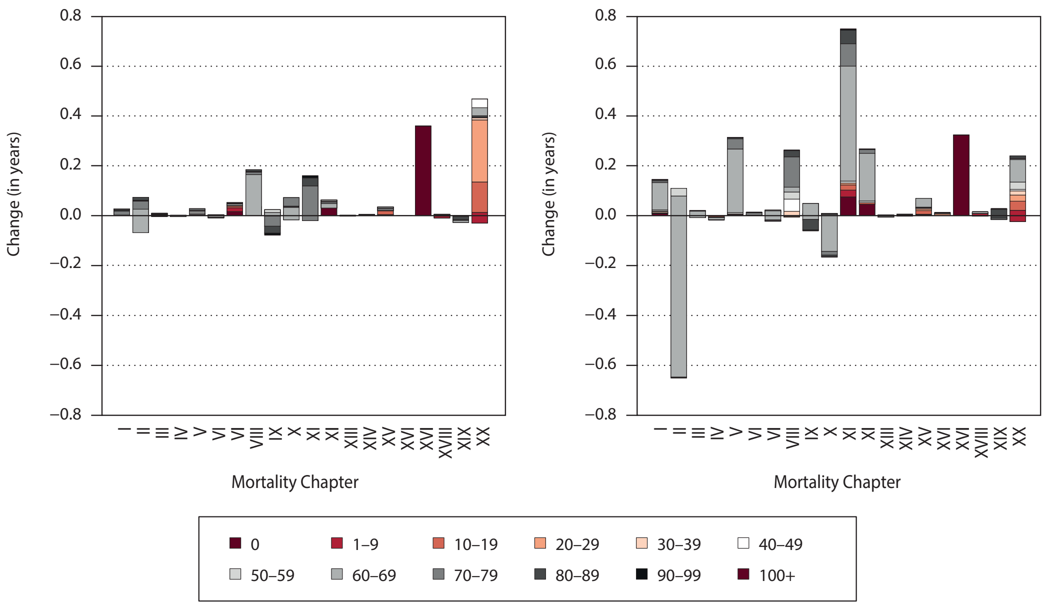
In the case of Czechia, LE increased by 9.11 years for males and 8.16 for females between 1970 and 2013, going from 66.04 years and 72.99 years for males and females, respectively, to 75.15 and 81.15 years. The causes of these variations, as in the case of France, seem to vary in time as mortality evolves by age and cause of death.

3.2.1 1970–1984

The change in LE in Czechia between 1970 and 1984 (see Figure 4) was very similar for both genders and reached 1.3 years in the case of males and 1.51 for females. Around a year of the change is estimated to have been generated by mortality changes in age groups younger than 60, which means that age groups over 60 contribute in a much less significant manner during these years.

The main cause of death contributing to the indicator's improvement is shared by both genders: mortality changes originating in Mortality Chapter XVII, Conditions of the Perinatal Period, improvements in which led to an increase in life expectancy of 0.36 and 0.33 years for males and females. All this variation is attributable to age 0. Improvements related to deaths caused by accidents, homicides, suicides, poisonings, and the like (as defined in Mortality Chapter XX) also contributed significantly. This chapter is estimated to have generated 0.4 extra years of LE in the case of males (as the main cause explaining the increase in this case) and 0.23 extra years for females. In addition, Mortality Chapter XI, Respiratory Diseases, is estimated

Figure 4 Changes in LE by CoD and age, 1970–1984: Czechia, males (left) vs females (right)



Source: HMD, CoDD and MAC.

to be responsible for an increase of 0.76 years in this indicator in the case of females. Some chapters, however, are estimated to have caused decreases in LE during this period. These include Malignant Neoplasms, which are estimated to have contributed to a decrease in LE of 0.66 years for females.

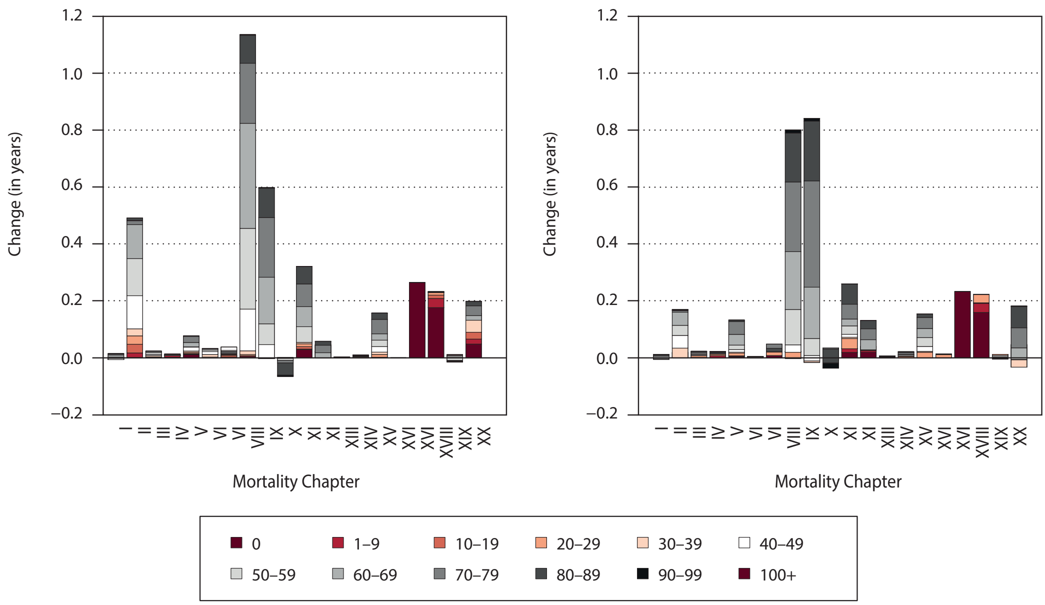
3.2.2 1985–1998

Mortality improvements were much more significant during this 14-year period and led to an increase in life expectancy of 3.56 years for males and 3.23 years for females. Again, mortality improvements among newborns are estimated to be the main driver of these changes, generating an estimated increase in LE of 0.57 years for males and 0.50 years for females. Despite this, age groups under 60 became, in relative terms, less 'important' (for LE changes?) when compared to the previous period. Nevertheless, they continued to be the main driver of change and their absolute contributions increased, generating 1.83 and 1.17 extra years in LE for men and women, respectively. This means that the contributions to changes made by ages 60 and older represented 49% of the variation for men but 64% for women.

When analysing the changes by mortality chapter, the estimated main contributors seem to differ from the ones identified in the previous period. In the case of males, the mortality improvements for causes of death included in Mortality Chapter VIII, Heart Diseases, are the most relevant: they are estimated to have contributed 1.14 years of extra LE. Such improvements are most observed at ages from 40 to 80 years old. The effect of these groups is estimated to explain a year of the overall gains in Mortality Chapter VIII, showing that death rates attributed to these causes mostly improved greatly for adults. Improvements in Cerebrovascular Diseases also seem to be of key importance for the increase in LE. Among men their effect amounts to an increase in LE of 0.60 years. Malignant Neoplasms are also a major contributor in this case, generating 0.48 extra years of LE for males.

In the case of women, mortality improvements in Cerebrovascular Diseases are found to be the main driver of the improvement with an estimated effect of 0.84 extra years in LE. It is worth noting that 0.77 years of the total improvement are due to mortality changes registered at ages over 60. Heart

Figure 5 Changes in LE by CoD and age, 1985–1998: Czechia, males (left) vs females (right)



Source: HMD, CoDD and MAC.

Diseases comes second and generated an estimated 0.80 extra years. Respiratory Diseases also played a central role and caused an estimated increase in LE of over a quarter of a year. Figure 5 presents a detailed picture.

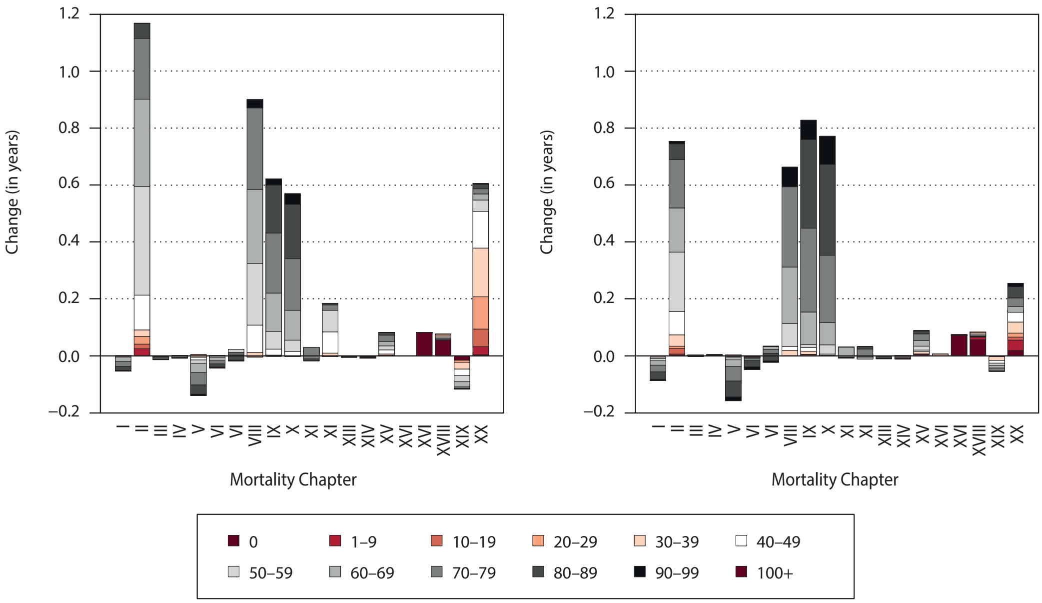
3.2.3 1999–2013

From 1999 to 2013 LE in Czechia took another big leap (in fact for males it was the biggest increase throughout the entire period). The increase in LE during this time was 3.82 years for males and 3.07 years for females. Despite a smaller improvement in mortality at birth during this period, the contributions of age groups under 60 continued to be relevant in the case of males, representing a total gain of 1.85 years or around 48% out of the total change (see Figure 6). In the case of females, the changes attributable to these age groups became less important and amounted to 28% (0.88 years) of the total variation. This shows that women were already experiencing major mortality improvements in senior ages while men were still experiencing very significant changes in younger age groups. In general terms, Czech men seem to be slowly shifting to a pattern that should be more aligned with that of women in the years to come, but this shift seems to be happening at a much slower pace than in the case of French men.

Effects by cause of death show subtle differences based on gender. The increase in men's LE was mostly due to improvements in mortality related to Malignant Neoplasms (generating 1.17 extra years, 50% of which was caused by age groups under 60). Heart Diseases also contributed greatly (0.87 extra years), whereas Cerebrovascular Diseases came third in importance (0.62 extra years).

The chapter contributing most to changes in the LE of women was that of Cerebrovascular Diseases (0.83 additional years), followed closely by Mortality Chapter X, Other Circulatory Diseases (0.77 years). Malignant Neoplasms completes the group of three major contributors, generating an estimated increase in LE of 0.75 years. Together these three Mortality Chapters explain 77% of the overall improvement in the indicator for males and 82% for females. Appendix 2 presents the results in detail.

Figure 6 Changes in LE by CoD and age, 1999–2013: Czechia, males (left) vs females (right)



Source: HMD, CoDD and MAC.

3.3 Life expectancy at birth in the United States

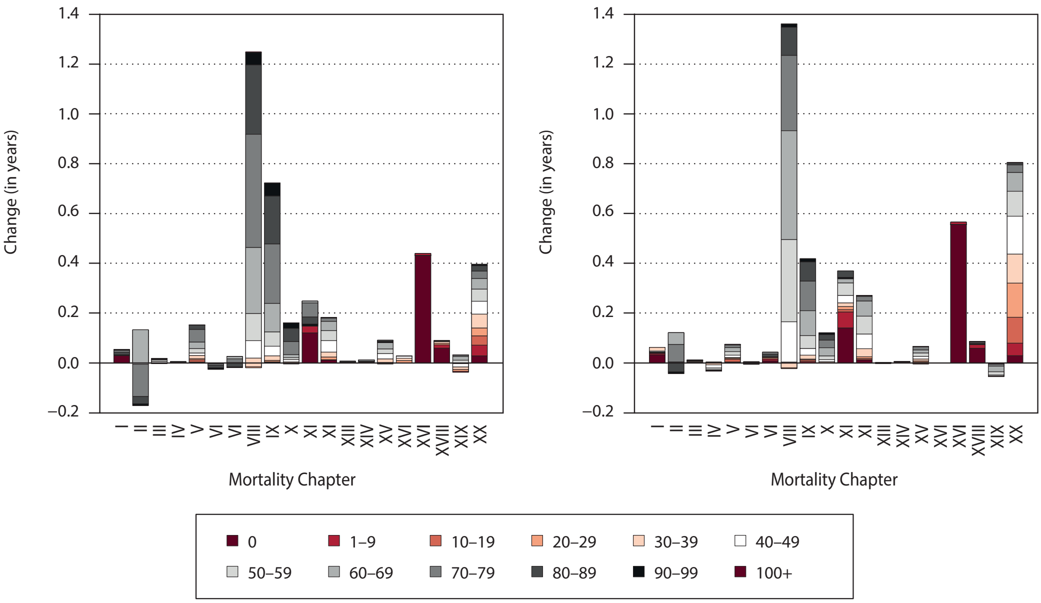
The United States saw LE grow from 67.02 years for males and 74.65 for females in 1970 to 76.6 years and 81.3 years in 2013. One remarkable aspect of these changes is that, out of the three countries being studied, the United States was the country in which the gender gap decreased the most during this period. For now, as in the cases of France and Czechia, the focus will be placed on the sources of changes in LE.

3.3.1 1970–1984

LE improved greatly in the United States between 1970 and 1984. The indicator increased 4.1 years for males and 3.52 years for females. The most relevant change in mortality attributable to a single age group was the change observed at age 0, which was responsible for an estimated increase in LE of 0.8 and 0.66 years for males and females, respectively. Mortality changes at age groups under 60 are responsible for most of the improvement in the indicator for this period in both genders, representing a total increase of 2.74 and 1.95 extra years of LE for men and women. In the case of senior ages, the development was less pronounced during this period and amounted to an improvement in LE of 1.36 and 1.57 years, respectively.

A big part of the positive developments in LE in the United States during this period is estimated to have come from the country's success in fighting Heart Diseases, generating a reduction in deaths high enough to increase LE by 1.36 and 1.25 years for males and females. In the specific case of US women, the mortality reductions in Cerebrovascular Diseases and in Conditions of the Perinatal Period contributed to an increase in LE of 0.72 and 0.43 years, respectively. For males, a very relevant improvement comes from a reduction of deaths resulting from accidents, homicides, suicides, poisonings, and the like, which generated an additional 0.80 years of LE. Changes in mortality from conditions associated with the perinatal period also became a major contributor in the case of men (0.56 years). The estimated decomposition of the changes in LE by Mortality Chapter, for all chapters, is shown in Figure 7.

Figure 7 Changes in LE by CoD and age, 1970–1984: United States, males (right) vs females (left)



Source: HMD, CoDD and MAC.

3.3.2 1985–1998

This period was characterised by more modest increases in LE (an increase of 2.73 years for males and 1.25 extra years for females). According to Figure 8, out of the 2.73 additional years estimated for males, 1.35 years are attributable to mortality improvements at ages below 60. For females, these age groups accounted for 0.66 additional years of LE, out of the total 1.25 years. Because of this, an important part of the improvements in LE during this time period are estimated to be generated by these younger age groups, accounting for 49% and 52% of the overall change among men and women, respectively.

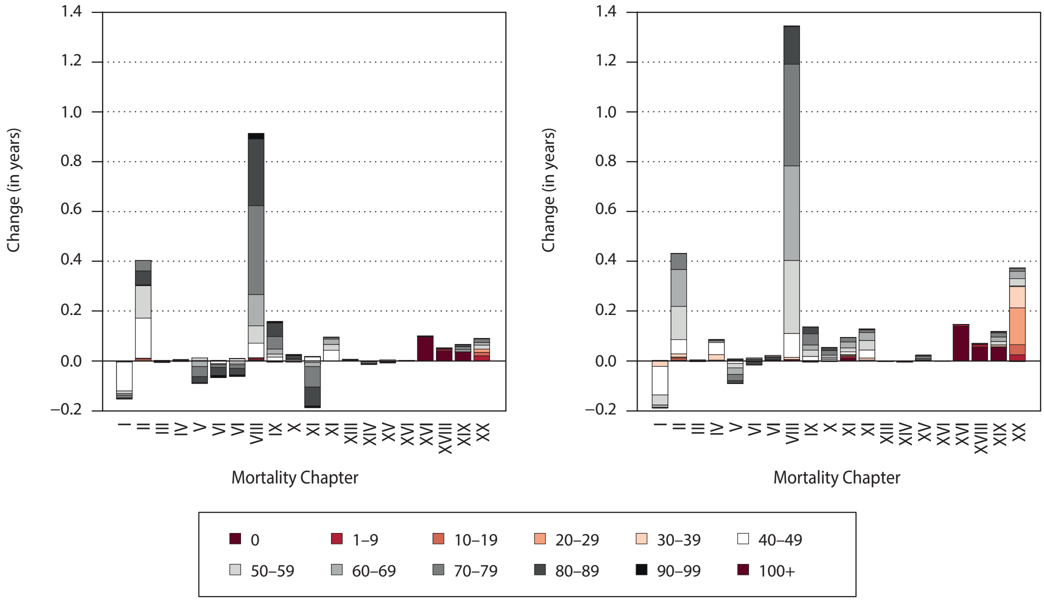
In general terms, the biggest improvement in LE related to a mortality chapter was observed among men, and this was due to the reduction of deaths from Heart Diseases (generating an increase of around 1.35 years). For females, the attributable effect of Mortality Chapter VIII was the most relevant (0.91 years for the period). Another significant increase in LE is estimated to have happened in the case of men, due to a reduction in the death rates associated with Mortality Chapter II.

It is worth noting that Mortality Chapter XI, Respiratory Diseases, was responsible for a decrease in LE for females in the United States, generating a decrease in LE of around 0.19 years. This phenomenon explains a part of the advantage that was observed for males during this period, which, together with the stronger improvement in mortality related to heart disease, explains about 0.70 years of the additional LE gained by US men during this period when compared to women.

3.3.3 1999–2013

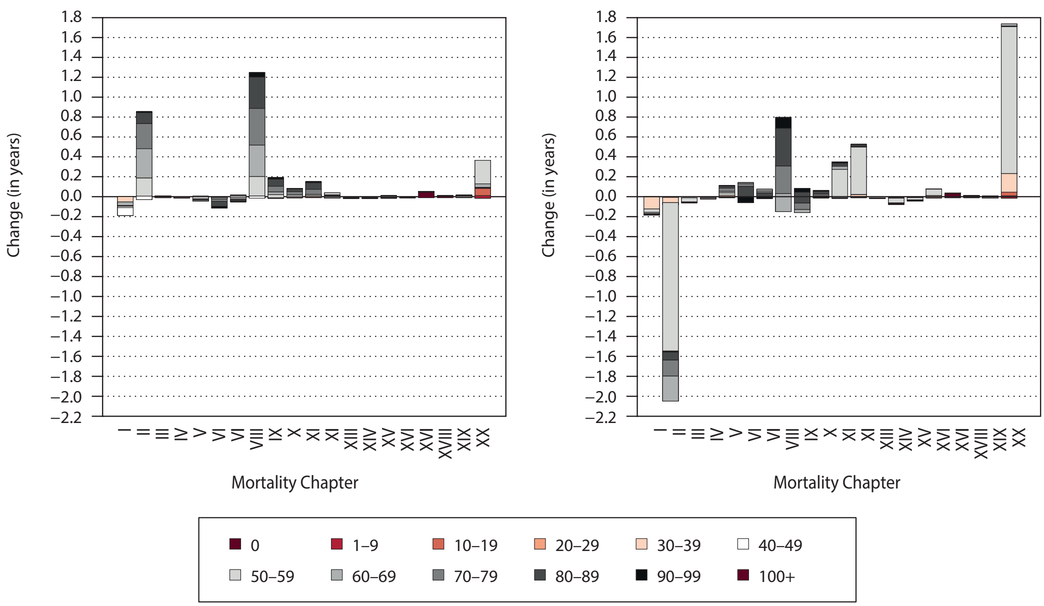
In 1999–2013, increases in LE reached 2.63 and 1.99 years for males and females. These changes are much closer to the ones observed in 1970–1984. However, during these years, the United States seemed to enter a new stage in LE dynamics, one in which LE was driven by mortality improvements in senior ages (cf. Figure 9), as treatments and prevention began to focus more on retirees and older individuals. Mortality improvements in senior age groups (60+) are estimated to have contributed 1.96 years of additional LE for males and 1.62 years for females. This represents 75% and 82% of the overall change and shows a total shift in the pattern previously observed. The fight against Heart Diseases and Malignant Neoplasms

Figure 8 Changes in LE by CoD and age, 1985–1998: USA, males (left) vs females (right)



Source: HMD, CoDD and MAC.

Figure 9 Changes in LE by CoD and age, 1999–2013: United States, males (left) vs females (right)



Source: HMD, CoDD and MAC.

seems to have made these two CoDs the main drivers of the increase in LE for males during this new phase. For females, Heart Diseases and External Causes resulted in the biggest increases. In the case of Heart Diseases, they are estimated to have increased LE by 1.27 for US males and 0.82 years for US females. In the case of Mortality Chapter II, Malignant Neoplasms, the effect is estimated to have caused a gain of 0.88 years for males and a decrease in LE of 1.56 years for females. As noted, the effect of Mortality Chapter XX for females is very relevant in this period, registering a gain of 1.75 years.

The detailed figures for the decomposition of changes in LE in the United States are presented in the Appendix 3.

4 CONCLUSION

We started by pointing out that, when a comparison is possible, our general results are aligned with the results obtained in the existing literature, as expected. Nevertheless, since we cover more than 40 years, 12 age groups, and 20 mortality chapters, this very complete grid of combinations allows us to shed added light on current knowledge, based on the mortality experiences of the three selected countries.

By using the findings in France, Czechia, and the United States, we showed that the development and increases in LE follow a certain tendency: improving mortality at birth is clearly an essential first step towards increasing LE in any country. Once this is achieved, countries tend to see an improvement in survival at younger ages (e.g. ages under 60) so that mortality is reduced for these age groups. Finally, they have just one way to continue to a more 'advanced stage': once they reach a 'high enough' LE, improvements start to derive from extending the life of seniors and reducing the effects of the diseases that affect them the most. It seems that women are the first segment of the population to reach this final stage in a country, but males are keeping up, and are seeing in general terms bigger improvements in LE. France and the United States already appeared to be at this stage for both genders in 2013. In the case of Czechia, women have reached this level, but men have been moving a little more slowly towards increases in LE due to mortality changes at ages older than 60.

At least 50% of the variations in LE in the three countries could be easily explained by focusing just on four mortality chapters. In fact, in the United States these groups of diseases would explain over 70% whereas in France they account for around 60% of LE improvements. Irrespective of geography or gender, the increasing effectiveness at reducing mortality related to Heart Diseases, Malignant Neoplasms, Cerebrovascular Diseases and External Causes seems to have become key to maintaining increasing levels of LE from 1970 to 2013. In the case of French males and US females, the decrease in mortality rates due to Diseases of the Digestive System has also played a major role in this time interval, adding an estimated 0.89 and 0.82 years during the period, respectively.

It is important to note that the United States was the country in which the gender gap in LE decreased the most, as the difference in male and female LE narrowed by 2.84 years between 1970 and 2013. By contrast, the smallest decrease in the gender gap in LE out of all three countries was observed in Czechia, where there was a change of less than a year in the same period. This phenomenon will be analysed in greater depth in future research.

Acknowledgements

This work was supported by FCT (Fundação para a Ciência e a Tecnologia) [grant number UIDB/05069/2020].

References

- Actuarial Association of Europe, Social Security Sub-Committee of the Pensions Committee 2019. Meeting the Challenge of Ageing in the EU. *AAE Discussion Paper.* https://actuary.eu/wp-content/uploads/2019/03/Meeting-the-challenge-of-ageing-FINAL.pdf.
- Agyepong, I. A – Sewankambo, N. – Binagwaho, A. – Coll-Seck, A.M. – Corrah, T. – Ezeh, A., … and Piot, P. 2017. The path to longer and healthier lives for all Africans by 2030: the Lancet Commission on the future of health in sub-Saharan Africa. *The Lancet*, 390(10114), pp. 2803–2859. https://doi.org/10.1016/S0140-6736(17)31509-X.

- Al-Ramadhan, M. A. 2008. Contributions of age and cause-specific mortality to gains in life expectancies among the Kuwait population. *Genus*, 64, pp. 155–171.
- Andreev, E. – Shkolnikov, V. – Begun, A. 2002. Algorithm for decomposition of differences between aggregate demographic measures and its application to life expectancies, healthy life expectancies, parity progression ratios and total fertility rates. *Demographic Research*, 7(14), pp. 499–522. https://doi.org/10.4054/DemRes.2002.7.14.
- Arriaga, E. 1984. Measuring and explaining the change in life expectancies. *Demography*, 21(1), pp. 83–96. https://doi.org/10.2307/2061029.
- Arriaga, E. 1989. Changing trends in mortality decline during the last decades. In *Differential mortality: Methodological issues and biosocial factors*, pp. 105–129, Oxford: Clarendon Press.
- Auger, N. – Feuillet, P. – Martel, S. – Lo, E. – Barry, A. D. – Harper, S. 2014. Mortality inequality in populations with equal LE: Arriaga's decomposition method in SAS, Stata, and Excel. *Annals of Epidemiology*, 24, pp. 575–580. https://doi.org/10.1016/j.annepidem.2014.05.006.
- Auger, N. – Harper, S. – Barry, A. D. – Trempe, N. – Daniel, M. 2012. Life expectancy gap between the Francophone majority and Anglophone minority of a Canadian population. *European Journal of Epidemiology*, 27, pp. 27–38. https://doi.org/10.1007/s10654-011-9644-8.
- Bergeron-Boucher, M-P. – Ebeling, M. – Canudas-Romo, V. 2015. Decomposing changes in life expectancy: Compression versus shifting mortality. *Demographic Research*, 33, pp. 391–424. https://doi.org/10.4054/DemRes.2015.33.14.
- Buettner, D. – Skemp, S. 2016. Blue Zones. *American Journal of Lifestyle Medicine,* 10(5), pp. 318–321. https://doi.org/10.1177/1559827616637066.
- Cause of Death Database. https://cod.mortality.org/.
- Das Gupta, P. 1978. A general method of decomposing a difference between two rates into several components. *Demography*, 15, pp. 99–112. https://doi.org/10.2307/2060493.
- Ho, J. Y. – Hendi, A. S. 2018. Recent trends in life expectancy across high income countries: retrospective observational study. *BMJ*, 2018, 362, k2562. https://doi.org/10.1136/bmj.k2562.
- Hosseinpoor, A. R. – Lee, J. H. – Lynch, J. – Mathers, C. – Abou-Zahr, C. 2012. International shortfall inequality in life expectancy in women and men, 1950–2010. *Bulletin of World Health Organisatio*n, 90. https://doi.org/10.2471/BLT.11.097378.
- Hulíková Tesárková, K. – Kašpar, D.– Zimmermann, P. 2015. Convergent and Divergent Trends in the European Mortality: What Is the Position of Czechia? *Geografie*, 120(1), pp. 26–49. https://doi.org/10.37040/geografie2015120010026.
- Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de. Last accessed 12.10.2021.
- Khang, Y.-H. – Yang, S. – Cho, H.-J. – Choi-Jung, K. – Yun, S.-C. 2010. Decomposition of socioeconomic differences in life expectancy at birth by age and cause of death among 4 million South Korean public servants and their dependents. *International Journal of Epidemiology*, 39, pp. 1656–1666. https://doi.org/10.1093/ije/dyq117.
- Le, Y. – Ren, J. – Shen, J. – Li, T. – Zhang, C.-F. 2015. The changing gender differences in life expectancy in Chinese cities 2005–2010. *PLoS ONE*, 10, e0123320. https://doi.org/10.1371/journal.pone.0123320.
- Leon, D. A. 2011. Trends in European life expectancy: a salutary view. *International Journal of Epidemiology*, 40, pp. 271–277. https://doi.org/10.1093/ije/dyr061.
- Liu, L. – Oza, S.– Hogan, D. – Chu Y. – Perin, J. – Zhu, J. – … – Black, R. E. 2016. Global, regional, and national causes of under-5 mortality in 2000–15: an updated systematic analysis with implications for the Sustainable Development Goals. *The Lancet*, 388(10063) pp. 3027–3035. https://doi.org/10.1016/S0140-6736(16)31593-8.
- Martikainen, P. – Makela, P. – Peltonen, R. – Myrskyla, M. 2014. Income differences in life expectancy: The changing contribution of harmful consumption of alcohol and smoking. *Epidemiology*, 25, pp. 182–190. https://doi.org/10.1097/EDE.0000000000000064.
- Martikainen, P. – Valkonen, T. – Martelin, T. 2001. Change in male and female life expectancy by social class: Decomposition by age and cause of death in Finland 1971–95. *Journal of Epidemiology Community Health*, 55, pp. 494–499. https://doi.org/10.1136/jech.55.7.494.
- Meslé, F. Mortality in Central and Eastern Europe: long-term trends and recent upturns. *Demographic Research* – Special Collection 2, pp. 45–70. https://doi.org/10.4054/DemRes.2004.S2.3.

- Meslé, F. – Vallin, J. 2017. The End of East-West Divergence in European Life Expectancies? An Introduction to the Special Issue. *European Journal of Population*, 33, pp. 615–627.
- Meslé, F. – Vallin, J. – Pyrozhkov, S. 2012. Impact of Major Groups of Causes on Life Expectancy Trends. *Mortality and Causes of Death in 20th-Century Ukraine*, pp. 173–195. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-2433-4_11.
- Mondal, N. I. – Shitan, M. 2014. Relative importance of demographic, socioeconomic and health factors on life expectancy in low- and lower-middle-income countries. *Journal of Epidemiology*, 24, pp. 117–124. https://doi.org/10.2188/jea.JE20130059.
- Murwirapachena, G. – Mlambo, C. 2015. Life Expectancy in Zimbabwe: An Analysis of Five Decades. *International Business and Economics Research Journal*, 14(3), pp. 417–430. https://doi.org/10.19030/iber.v14i3.9207.
- Poulant, M.– Buettner, D. – Pes, G. 2019. Blue Zones. Reference Module in *Biomedical Sciences*. Elsevier. https://doi.org/10.1016/B978-0-12-801238-3.11437-0.
- Pressat, R. 1985. Contribution des écarts de mortalité par âge à la différance des vies moyennes. *Population*, 40(4–5), pp. 766–770. https://doi.org/10.2307/1532986.
- Preston, S. H. – Stokes, A. 2012. Sources of Population Aging in More and Less Developed Countries. *Population and Development Review*, 38(2), pp. 221–236. https://doi.org/10.1111/j.1728-4457.2012.00490.x.
- Rosella, L. C. – Calzavara, A. – Frank, J. W. – Fitzpatrick, T. – Donnelly, P. D. – Henry, D. 2016. Narrowing mortality gap between men and women over two decades: a registry-based study in Ontario, Canada. *BMJ Open*, 6(11). [e012564]. https://doi.org/10.1136/bmjopen-2016-012564.
- Simmons, S. 2018. The Gender Gap in Life Expectancy in South Africa from 2000 to 2015: The Role of Age-and Cause-Specific Mortality. *International Journal in Sciences: Basic and Applied Research*, 42(5), pp. 24–36.
- Shkolnikov, V. M. – Andreev, E. M. – McKee, M. – Leon, D. A. 2013. Components and possible determinants of decrease in Russian mortality in 2004–2010. *Demographic Research*, 28, pp. 917–950. https://doi.org/10.4054/DemRes.2013.28.32.
- Tarkiainen, L. – Martikainen, P. – Laaksonen, M. – Valkonen, T. 2012. Trends in life expectancy by income from 1988 to 2007: Decomposition by age and cause of death. *Journal of Epidemiology Community Health*, 66, pp. 573–578. https://doi.org/10.1136/jech.2010.123182.
- Trovato, F. – Heyen, N. B. 2006. A varied pattern of change of sex differential in the G7 countries. *Journal of Biosocial Science*, 38, pp. 391–401. https://doi.org/10.1017/S0021932005007212.
- Trovato, F. – Lalu, N. M. 1997. Changing sex differences in life expectancy in Australia between 1970 and 1990. *Journal of the Australian Population Association*, 14, pp. 187–200. https://doi.org/10.1007/BF03029339.
- Trovato, F. – Odynak, D. 2011. Sex differences in life expectancy in Canada: Immigrant and native-born populations. *Journal of Biosocial Science*, 43, pp. 353–368. https://doi.org/10.1017/S0021932011000010.
- United Nations, Department of Economic and Social Affairs, Population Division 2019. World Population Prospects 2019: Press Release. Available at https://population.un.org/wpp/Publications/Files/ WPP2019_PressRelease_EN.pdf.
- Vaupel, J. W. – Romo, V. C. 2002. Decomposing demographic change into direct vs. compositional components. *Demographic Research*, 7, pp. 1–14. https://doi.org/10.4054/DemRes.2002.7.1.
- Wang, H. – Naghavi, M. – Allen, C. – Barber, R. M. – Bhutta, Z. A – Carter, A. – … – Murray, C. J. L. 2016. Global, regional, and national life expectancy, all-cause mortality, and cause-mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, 388(10053), pp. 1459–1544. https://doi.org/10.1016/S0140-6736(16)31012-1.
- Waldon, I. 1983. Sex differences in human mortality. *Social Science Medicine*, 17, pp. 321–333. https://doi.org/10.1016/0277-9536(83)90234-4.
- Waldon, I. – McCloskey, C. – Earle, I. 2005. Trends in gender differences in accidents mortality: Relationships to changing gender roles and other societal trends. *Demographic Research*, 13, pp. 415–454. https://doi.org/10.4054/DemRes.2005.13.17.
- Woolf, S. H. – Schoomaker, H. 2019. Life Expectancy and Mortality Rates in the United States, 1959–2017. *JAMA*, 322(20), pp.1996–2016. https://doi.org/10.1001/jama.2019.16932.
- World Health Organization 2019. *World health statistics overview 2019: monitoring health for the SDGs, sustainable development goals.* World Health Organization.
  Available at https://apps.who.int/iris/bitstream/handle/10665/311696/WHO-DAD-2019.1-eng.pdf?sequence=1 andisAllowed=y .
- Yang, S. – Khang, Y.-H. – Chun, H. – Harper, S. – Lynch, J. 2012. The changing gender differences in life expectancy in Korea 1970–2005. *Social Science Medicine*, 75, pp. 1280–1287. https://doi.org/10.1016/j.socscimed.2012.04.026.

ONOFRE ALVES SIMÕES

is a lecturer at ISEG, Instituto Superior de Economia e Gestão da Universidade de Lisboa (Lisbon School of Economics and Management of Lisbon University), where in the masters programme in actuarial science he teaches survival models and life contingencies and financial mathematics. His fields of interest include life contingencies, pensions and social security and financial mathematics. A few years ago he joined the Mortality Working Group (now Mortality Virtual Forum) of the International Actuarial Association, after developing a stronger interest in the topics of longevity and mortality.

ANDREY UGARTE MONTERO

is a graduate assistant and PhD candidate in actuarial science at HEC Lausanne, Switzerland. He worked as a consultant in actuarial services for Ernst & Young, and has a MSc. degree in actuarial science from Lisbon University. He is also pursuing actuarial certification with the Institute and Faculty of Actuaries of the United Kingdom (IFoA). His fields of interest include longevity risks, health and long-term care insurance, social security systems, and data science.

Appendix

Appendix 1 Decomposition of the changes in LE, in years, by mortality chapter – France

	1970–1984		1985–1998		1999–2013		1970–2013	
Mortality Chapter	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
I – Infectious Diseases	0.2310	0.1491	-0.0059	0.0426	0.0518	0.0161	0.2769	0.2078
II – Malignant Neoplasms	–0.8237	0.1438	0.5098	0.2439	1.1629	0.3777	0.8489	0.7654
III – Other Neoplasms	0.0525	0.0323	0.0368	0.0273	0.0139	0.0156	0.1031	0.0752
IV – Diseases of the Blood	–0.0088	–0.0130	0.0379	0.0179	0.0310	0.0238	0.0601	0.0287
V – Endocrine/Nutritional	0.0324	0.1056	0.0206	0.0612	0.0315	0.0677	0.0844	0.2345
VI – Mental Disorders	0.0189	–0.0268	0.0481	0.0357	0.0183	0.0254	0.0853	0.0343
VII – Nervous System	0.0935	0.1097	–0.0165	–0.0608	–0.0171	–0.1089	0.0599	–0.0600
VIII – Heart Diseases	0.3559	0.5660	0.7551	0.7605	0.8549	0.7783	1.9660	2.1048
IX – Cerebrovascular Diseases	0.4827	0.6078	0.4503	0.6188	0.3030	0.3457	1.2360	1.5723
X – Other Circulatory Diseases	0.0520	0.0829	0.1034	0.0960	0.1591	0.1350	0.3145	0.3139
XI – Respiratory Diseases	0.3782	0.3034	0.1405	0.0395	0.3393	0.2647	0.8580	0.6076
XII – Diseases of the Digestive System	0.3250	0.2521	0.3130	0.2834	0.2504	0.2150	0.8884	0.7505
XIII – Diseases of the Skin	–0.0058	–0.0008	0.0017	0.0017	0.0138	0.0258	0.0098	0.0267
XIV – Diseases of the Musculoskeletal System	0.0016	–0.0049	0.0047	0.0171	–0.0013	0.0050	0.0050	0.0172
XV – Diseases of the Genitourinary System	0.1577	0.0994	0.0571	0.0379	0.0374	0.0355	0.2522	0.1728
XVI – Complications of Pregnancy/Childbirth	0.0000	0.0308	0.0000	0.0045	0.0000	0.0006	0.0000	0.0359
XVII – Conditions of the Perinatal Period	0.4929	0.3954	0.0489	0.0205	–0.0044	0.0028	0.5374	0.4187
XVIII – Congenital Malformations	0.1611	0.1382	0.0950	0.0748	0.0173	0.0142	0.2734	0.2272
XIX – Ill–Defined or Unknown	0.3871	0.4645	0.3023	0.2727	–0.1895	–0.0943	0.4999	0.6429
XX – External Causes	0.3859	0.1049	0.5774	0.3648	0.7577	0.4441	1.7209	0.9138
Total increase	**2.7700**	**3.5404**	**3.4800**	**2.9600**	**3.8300**	**2.5898**	**10.0800**	**9.0902**

Source: HMD, CoDD and MAC.

Appendix 2 Decomposition of the changes in LE, in years, by mortality chapter – Czechia

	1970–1984		1985–1998		1999–2013		1970–2013	
Mortality Chapter	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
I – Infectious Diseases	0.0238	0.1469	0.0146	0.0120	–0.0532	–0.0869	–0.0149	0.0720
II – Malignant Neoplasms	0.0588	–0.6593	0.4826	0.1678	1.1679	0.7533	1.7093	0.2618
III – Other Neoplasms	0.0072	0.0181	0.0234	0.0229	–0.0135	–0.0022	0.0171	0.0388
IV – Diseases of the Blood	0.0007	–0.0172	0.0133	0.0224	–0.0082	0.0005	0.0058	0.0057
V – Endocrine/Nutritional	0.0202	0.3180	0.0774	0.1325	–0.1396	–0.1575	–0.0420	0.2930
VI – Mental Disorders	–0.0067	0.0136	0.0331	0.0047	–0.0424	–0.0483	–0.0160	–0.0300
VII – Nervous System	0.0487	–0.0180	0.0117	0.0212	–0.0176	–0.0226	0.0428	–0.0194
VIII – Heart Diseases	0.1767	0.2662	1.1366	0.8011	0.8707	0.5936	2.1840	1.6609
IX – Cerebrovascular Diseases	–0.0791	–0.0618	0.5983	0.8420	0.6217	0.8275	1.1409	1.6077
X – Other Circulatory Diseases	0.0334	–0.1683	–0.0612	–0.0186	0.5695	0.7713	0.5417	0.5844
XI – Respiratory Diseases	0.1619	0.7587	0.3211	0.2602	–0.0173	–0.0048	0.4657	1.0141
XII – Diseases of the Digestive System	0.0570	0.2711	0.0580	0.1313	0.1846	0.0324	0.2996	0.4348
XIII – Diseases of the Skin	–0.0014	–0.0050	0.0031	0.0068	–0.0049	–0.0101	–0.0032	–0.0083
XIV – Diseases of the Musculoskeletal System	0.0034	0.0002	0.0098	0.0212	–0.0082	–0.0109	0.0050	0.0105
XV – Diseases of the Genitourinary System	0.0279	0.0306	0.1569	0.1544	0.0815	0.0876	0.2663	0.2726
XVI – Complications of Pregnancy/Childbirth	0.0000	0.0122	0.0000	0.0137	0.0000	0.0071	0.0000	0.0330
XVII – Conditions of the Perinatal Period	0.3644	0.3278	0.2651	0.2331	0.0825	0.0715	0.7120	0.6324
XVIII – Congenital Malformations	0.0048	0.0155	0.2307	0.2220	0.0551	0.0664	0.2906	0.3039
XIX – Ill–Defined or Unknown	0.0007	0.0282	–0.0133	–0.0026	–0.1138	–0.0528	–0.1264	–0.0272
XX – External Causes	0.3975	0.2324	0.1988	0.1819	0.6053	0.2549	1.2016	0.6692
Total increase	**1.3000**	**1.5099**	**3.5600**	**3.2300**	**3.8200**	**3.0700**	**8.6800**	**7.8099**

Source: HMD, CoDD and MAC.

Appendix 3 Decomposition of the changes in LE, in years, by mortality chapter – United States

	1970–1984		1985–1998		1999–2013		1970–2013	
Mortality Chapter	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
I – Infectious Diseases	0.0389	0.0309	–0.1876	–0.1514	–0.0851	–0.1764	–0.2338	–0.2969
II – Malignant Neoplasms	–0.0415	–0.1698	0.4333	0.3034	0.8759	–1.5601	1.2677	–1.4265
III – Other Neoplasms	0.0099	0.0139	0.0057	–0.0045	0.0185	–0.0426	0.0340	–0.0332
IV – Diseases of the Blood	–0.0320	0.0009	0.0861	0.0041	–0.0027	–0.0106	0.0514	–0.0056
V – Endocrine/Nutritional	0.0746	0.1528	–0.0899	–0.0889	–0.0111	0.1260	–0.0264	0.1899
VI – Mental Disorders	–0.0044	–0.0243	–0.0142	–0.0661	–0.1060	–0.0502	–0.1246	–0.1406
VII – Nervous System	0.0237	–0.0175	0.0056	–0.0607	–0.0421	–0.0095	–0.0128	–0.0877
VIII – Heart Diseases	1.3615	1.2491	1.3454	0.9152	1.2733	0.8194	3.9802	2.9837
IX – Cerebrovascular Diseases	0.4190	0.7223	0.1371	0.1593	0.2047	0.0960	0.7609	0.9776
X – Other Circulatory Diseases	0.1219	0.1616	0.0547	0.0264	0.0958	0.0757	0.2724	0.2637
XI – Respiratory Diseases	0.3387	0.1490	0.0953	–0.1870	0.1640	0.3647	0.5980	0.3267
XII – Diseases of the Digestive System	0.2718	0.1813	0.1300	0.0959	0.0240	0.5449	0.4258	0.8221
XIII – Diseases of the Skin	–0.0009	0.0022	0.0010	0.0078	–0.0056	–0.0074	–0.0055	0.0026
XIV – Diseases of the Musculoskeletal System	0.0064	0.0059	–0.0036	–0.0126	–0.0034	–0.0522	–0.0006	–0.0589

Appendix 3 **cont.**

	1970–1984		1985–1998		1999–2013		1970–2013	
Mortality Chapter	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
XV – Diseases of the Genitourinary System	0.0665	0.0824	0.0240	–0.0040	0.0218	–0.0214	0.1124	0.0570
XVI – Complications of Pregnancy/Childbirth	0.0000	0.0283	0.0000	0.0029	0.0000	0.0905	0.0000	0.1217
XVII – Conditions of the Perinatal Period	0.5565	0.4316	0.1469	0.1013	0.0630	0.0437	0.7664	0.5766
XVIII – Congenital Malformations	0.0857	0.0898	0.0670	0.0503	0.0177	0.0114	0.1705	0.1515
XIX – Ill–Defined or Unknown	–0.0016	0.0329	0.1195	0.0675	0.0234	0.0062	0.1412	0.1066
XX – External Causes	0.8052	0.3968	0.3737	0.0910	0.1038	1.7420	1.2827	2.2298
Total increase	**4.1000**	**3.5201**	**2.7300**	**1.2499**	**2.6300**	**1.9901**	**9.4600**	**6.7601**

Source: HMD, CoDD and MAC.

POPULATION DEVELOPMENT IN CZECHIA IN 2020

David Morávek[1)] – Jana Koukalová[2)]

Abstract
The population development of Czechia in 2020 was significantly affected by the COVID-19 epidemic. Epidemiological measures or the pandemic itself were reflected in almost all monitored demographic processes, and in many cases long-term trends were interrupted. In addition, existing minimums or maximums were rewritten within the period of the last ten years. The article focuses on the main demographic processes, namely births, deaths, marriages, divorces and migration. The beginning of the examined period is the year 2011, in which the census took place, and then the following years, especially the period 2015–2020.

Keywords: demographic development, population, age structure, nuptiality, divorce, fertility, mortality, migration, Czechia

https://doi.org/10.54694/dem.0295 **Demografie,** 2021, **63 (4): 246–261**

POPULATION BY AGE AND MARITAL STATUS

In 2020, the population in Czechia increased by 7,838 from 10,693,939 (as of 1 January) to 10,701,777 inhabitants (as of 31 December). This was the smallest year-on-year increase in the last decade, apart from the exceptional decrease in 2013. Since the beginning of 2011, when the population was first based on the results of the Census, the Czech population has grown by 215 thousand persons. The population growth was caused by international migration not only in total for the period 2011–2020 but also in the year 2020 (see Table 1). The balance of international migration in 2020 reached 26,927 persons and was thus 17,3 thousand lower than in the previous year, but still higher than in 2011–2016. The natural change in the total caused a decrease in the number of inhabitants. A significant natural decrease of 19,1 thousand persons was recorded in 2020, which was the highest since the beginning of the century. In other years, the natural change caused an increase in population only marginally.

In 2020, the number of people in the elderly, child and adolescent age groups of the population continued to grow, while the number of persons of working age decreased. From the point of view of five-year age groups, most people were aged 40–44 years, for the fifth year in a row. Despite the specificity of 2020, the population in the three main age groups continued to develop in the same trajectories as in previous years. During 2020, the number of persons aged 0–14 years grew by 9,5 thousand to 1,72 million and thus represented 16.1% of the total population. In addition, the number of persons aged 65 and older

1) Czech Statistical Office, Demographic Statistics Unit, Na padesátém 81, 100 82 Prague 10, Czech Republic; Prague University of Economics and Business, Faculty of Informatics and Statistics, Department of Demography, W. Churchill Sq. 1938/4, 130 67 Prague 3 – Žižkov. Contact: david.moravek@czso.cz.

2) Czech Statistical Office, Demographic Statistics Unit, Na padesátém 81, 100 82 Prague 10, Czech Republic; Charles University, Faculty of Science, Department of Demography and Geodemography, Albertov 6, 128 43 Prague – 2, Czech Republic. Contact: krestanova.jana@gmail.com.

Table 1 Vital statistics, 2011 and 2015–2020

Indicator	2011	2015	2016	2017	2018	2019	2020
Natural increase	1,825	–409	4,913	2,962	1,116	–131	–19,089
Net migration	16,889	15,977	20,064	28,273	38,629	44,270	26,927
Total increase	18,714	15,568	24,977	31,235	39,745	44,139	7,838
Per 1,000 population							
Natural increase	0.2	0.0	0.5	0.3	0.1	0.0	–1.8
Net migration	1.6	1.5	1.9	2.7	3.6	4.1	2.5
Total increase	1.8	1.5	2.4	2.9	3.7	4.1	0.7

Source: Czech Statistical Office.

continued to grow by 26,7 thousand to 2,16 million. The share of the elderly population exceeded one-fifth for the first time. However, for both groups of persons, the growth rate in 2020 was the lowest in the decade. The working-age population has been declining for the twelfth year in a row; however, its year-on-year decreases have been more modest in the last three years than at the beginning of the decade. As in previous years, the working-age population represented the largest group of persons (6,82 million, or 63.8% of the total population in 2020). All this, together with the declining number of the working-age population, contributes to ageing of the population, which can be documented by relevant indicators of age distribution (see Table 2). The average age of the population, which has been steadily increasing since the early 1980s, increased by one tenth year-on-year to 42.6 years in 2020. Over the last decade, since the beginning of 2011, it has grown by less than two years, by 1.8 years for men and 1.7 years for women. The difference between the average age of men and women thus decreased slightly from 3.0 to 2.8 years, when in 2020 the average age of men was 41.1 years and the average age of women was 44.0 years. Moreover, the median age increased more than the average age of the population during the years 2011–2020, by 3.5 years from 39.8 years to 43.3 years. Since 2016, the median age has exceeded the average age of the population (by 0.7 years in 2020). The index of ageing, which is expressed as the number of persons

Table 2 Age distribution and characteristics of population, 2011 and 2015–2020 (as at 31 Dec.)

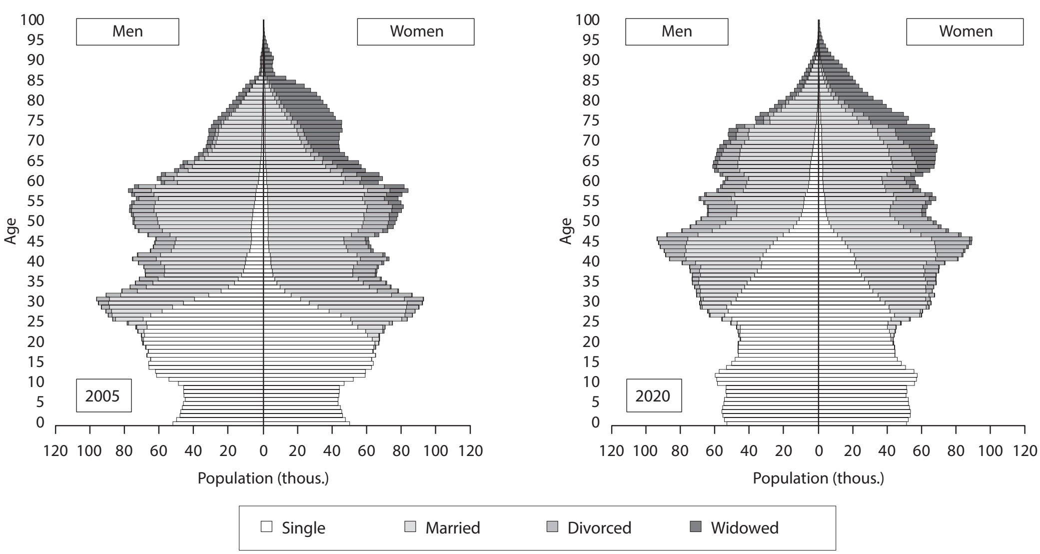
Indicator	2011	2015	2016	2017	2018	2019	2020
Total	10,486.7	10,553.8	10,578.8	10,610.1	10,649.8	10,693.9	10,701.8
0–14	1,521.8	1,623.7	1,647.3	1,670.7	1,693.1	1,710.2	1,719.7
15–64	7,328.0	6,997.7	6,942.6	6,899.2	6,870.1	6,852.1	6,823.7
65+	1,637.0	1,932.4	1,988.9	2,040.2	2,086.6	2,131.6	2,158.3
Percentage of total population							
0–14	14.5	15.4	15.6	15.7	15.9	16.0	16.1
15–64	69.9	66.3	65.6	65.0	64.5	64.1	63.8
65+	15.6	18.3	18.8	19.2	19.6	19.9	20.2
Characteristics of age distribution							
Average age	40.9	41.9	42.0	42.2	42.3	42.5	42.6
Median age	39.8	41.5	41.9	42.3	42.6	43.0	43.3
Index of ageing[1)]	107.6	119.0	120.7	122.1	123.2	124.6	125.5
Total age dependency ratio[2)]	55.4	61.4	63.2	64.8	66.3	67.8	69.0

Source: Czech Statistical Office; authors' calculations.
Note: 1) The number of persons aged 65+ years per hundred persons aged 0–14 years.
2) The number of children (0–14 years old) and older persons (65 years or more) per hundred persons in economic activity (15–64 years old).

aged 65+ per hundred persons aged 0–14, increased within the period from 108 to 126 seniors. The total age dependency ratio – the number of persons aged 0–14 years old and persons aged 65 years and older per 100 population aged 15–64 years – has also been growing steadily in the last decade. The growth of the index was mainly a reflection of the development of the number of the elderly population, which grew significantly in relation to the number of people of working age.

The age structure of the population is formed by unequally numerous generations of persons of individual years of birth. Apart from the influence of mortality in older age, it is mainly a consequence of the fluctuating development of the birth rate, which subsequently affects its development. The change in the age-distribution of the population by marital status continued in 2020 in the direction of long-term trends. Although after 2013 (with the exception of 2020) the level of marriage increased, its decline in the previous two decades and the persistently high level of divorce and decline in mortality and changes in age-distribution are behind the increased proportion (absolute and relative) of single and divorced persons in the population. The share of married persons, including widowed, decreased (see Figure 3). Measures against the spread of the COVID-19 epidemic led to the postponement of marriages and the trend of an increase in the number and share of single persons and a decrease in the share (absolute and relative) of married persons in 2020. The structure of persons by marital status has changed over time not only in the overall representation of individual categories of marital status in the population aged 15+, but also in terms of five-year age groups, with different degrees of intensity and different trajectories. The structure of the population has changed the most for persons between the ages of 30 and 44 over the last decade as generations born in the 1970s, who were crucial to changes in demographic behaviour since 1989, have passed through these age groups. The shares of married persons decreased the most because of a reduction the level of marriage and the postponement of marriage to a later age. The results of the population balance by marital status at the end of 2020 showed a more significant increase in the share of single persons and, conversely, a decrease in the share of married persons in the population than in previous years. The largest part of the population of persons aged 15 and over was represented

Figure 1 Population by age, sex and marital status, 2005 and 2020 (as at 31 Dec.)



Source: Czech Statistical Office.

by married persons, namely 46.1%. Additionally, 32.1% of persons were single, 13.7% were divorced and 8.1% were widowed.

NUPTIALITY

While in the previous six years the number of marriages increased every year, in 2020 there were 45,415 couples who entered into marriage, which was 9.5 thousand less than a year earlier (see Table 3). A significant year-on-year decrease in marriages (by 17%) was caused mainly by epidemiological measures, namely regulations on the possibility and size of wedding ceremonies in terms of the number of participants. The annual total of marriages had a declining trend from the 1990s until 2013, when it reached a historical minimum of 43,5 thousand. This was followed by a six-year period when, on the contrary, there was an increase in the number of marriages. By 2019, the annual total of marriages had risen to 54,870, the highest since 2008. The last decrease was by more than 10% in 1994; since 2011 the year-on-year changes have ranged from −4% to +6%. The largest group consisted of so-called protogamous marriages, i.e. of two single persons, which in the year 2020 amounted to 29,694 marriages or 65.4% of the total number of marriages. Additionally, 11,601 men (25.5% of all grooms) and 11,441 women (25.2% of all brides) entered into their second or further marriage. A year-on-year decrease was recorded in both groups of marriages, respectively. In relative terms, it was more noticeable in first-order marriages, whose number decreased by 19% for both men and women, while the number of higher-order marriages decreased by 13% for both sexes. As in previous years, so-called remarriages for men and women belonged to the second order. From the point of view of marital status, divorced persons dominate in marriages of a higher order. In general, widows rarely enter into a new marriage. The decrease in marriages was also reflected in the level of marriage, which decreased for both single and divorced persons. While keeping the intensity of marriage of singles at the level of 2020, only 51.9% of men and 60.8% of women would enter into their first marriage before the age of 50, which was less by seven p.p. than in the situation of 2019. The mean age of men and women at the beginning of their first marriage, which did not change significantly in previous years, increased year-on-year by 0.5 to 32.6 years for men and by 0.6 to 30.4 years for women.

The seasonal profile of marriage was specific in 2020, when a historical minimum of monthly marriages was recorded in March. On the contrary, attractive data in the calendar supported stronger numbers of marriages in February and October compared to previous years. However, marriages are not evenly distributed throughout a calendar year. Couples most often enter into marriage from June to September (64–69% in 2011–2019) and least often from December to February (6–8% in 2011–2019). In 2020, a slightly higher concentration of marriages was within the four most popular months, which reached 70.9%, up 6.2 p.p. year-on-year, but there was also a higher share of marriages entered into during the winter months from December to February. These three least popular months for marriages in 2020 accounted for a whole tenth of the total of marriages,

Table 3 Marriages, 2011 and 2015–2020

Indicator	2011	2015	2016	2017	2018	2019	2020
Total marriages	45,137	48,191	50,768	52,567	54,470	54,870	45,415
Protogamous marriages (%)	64.3	67.8	67.5	67.7	67.2	66.9	65.4
Remarriages (%) – men	26.1	23.5	24.0	23.8	24.1	24.2	25.5
– women	25.9	23.2	23.2	23.3	23.6	24.0	25.2
Total firtst marriage rate – men (%)	53.5	55.1	56.2	57.6	58.8	59.0	51.9
– women (%)	61.0	62.4	64.3	65.4	66.9	67.5	60.8
Mean age at first marriage – men	32.2	32.4	32.2	32.2	32.2	32.1	32.6
– women	29.6	29.8	29.9	29.8	29.8	29.8	30.4

Source: Czech Statistical Office; authors' calculations.

while a year earlier it was only 6%. Overall, March (647) and April (742) were recorded as the weakest calendar months in terms of the number of marriages, and conversely August (10,084) and September (8,629) as the strongest months. The number of marriages entered into in March 2020 represented the lowest monthly total in the history of the independent Czech state spanning more than a century. The decrease in marriages during the spring months was not compensated by the end of the year.

DIVORCE

According to data obtained from the Ministry of Justice of the Czech Republic, a total of 21,734 divorces were registered in 2020, which was 2,407 less than in the previous year and the least (not only) in the last decade (see Table 4). As in the case of marriages, the epidemic situation contributed to the lower number of divorces, however the number of divorces had a slightly declining trend in previous years. This was the lowest annual number of divorces since 1970. The irregular decrease in divorce rates over the last decade mainly reflects the declining number of marriages in previous decades and changes in the intensity of divorce rates over the duration of the marriage. However, the lower number of marriages terminated by divorce could also be due to epidemiological measures, namely limitation of the activity of courts in times of emergency, as indicated by the very specific distribution of divorces into individual months of the year, or postponement of divorces to a later time. Most men and women are divorced for the first time. In the year 2020 it was the first divorce for 17,612 men (81.0% of all divorced persons) and 17,685 women (81.4% of all divorced persons). The remaining one-fifth (4,122 divorced men and 4,049 divorced women) have already undergone a repeated divorce (or divorce of a higher order). A total of 12,719 marriages with minor children and 9,015 marriages without minor children were divorced. The year-on-year decrease in the number of divorces was 10.7%, respectively 9.0%. Since the beginning of the decade, the percentage of divorces with minor children on the total number of divorces has fluctuated between 56% and 59% (58.5% in 2020). A total of 20,187 children were affected by divorce, which was 2,457 fewer children than in the previous year thanks to a significant year-on-year decrease in all divorces. The average number of minors per divorce with minors has thus increased, from 1.50 to 1.59 children since the beginning of the decade.

The number of divorces in a given duration of marriage relative to the number of marriages concluded before a given number of years is regularly the highest between three and six years after marriage and then gradually decreases with increasing length of marriage (see Table 5). In the period 2011–2020, the most significant changes in the intensity of divorce rates over the last decade were recorded in the shortest period, in the interval 0–4 years, where it decreased from 2.12 to 1.49 divorces per 100 marriages (or by 30%). The level of divorce rates also decreased slightly and fluctuated over time for marriages lasting between five to nine years (from 2.20 to 1.94 divorces), 15–19 years (from 1.32 to 1.18 divorces) and 20–24 years (from 1.00 to 0.86 divorces).

Table 4 Divorces, 2011 and 2015–2020

Indicator	2011	2015	2016	2017	2018	2019	2020
Total divorces	28,113	26,083	24,996	25,755	24,313	24,141	21,734
Percentage of repeated divorces – men	19.4	19.3	19.7	19.3	19.1	19.2	19.0
– women	19.1	18.8	19.2	18.6	18.7	18.0	18.6
Divorces without minors	12,282	11,090	10,270	10,559	10,120	9,905	9,015
Divorces with minors	15,831	14,993	14,726	15,196	14,193	14,236	12,719
– percentage of total	56.3	57.5	58.9	59.0	58.4	59.0	58.5
Number of minors in divorced marriages	23,716	23,187	22,855	23,752	22,294	22,644	20,187
Average number of minors per divorce with minors	1.50	1.55	1.55	1.56	1.57	1.59	1.59

Source: Czech Statistical Office; authors' calculations.

Table 5 Divorces by duration of marriage, 2011 and 2015–2020

Duration of marriage (years)	2011	2015	2016	2017	2018	2019	2020
0–4	2.12	1.94	1.82	1.82	1.70	1.74	1.49
5–9	2.20	2.34	2.26	2.36	2.19	2.11	1.94
10–14	1.62	1.70	1.69	1.78	1.72	1.72	1.60
15–19	1.32	1.29	1.24	1.33	1.26	1.32	1.18
20–24	1.00	0.96	0.94	0.97	0.96	0.93	0.86
25–29	0.57	0.55	0.57	0.61	0.59	0.56	0.54
30+	0.22	0.25	0.24	0.26	0.26	0.28	0.25
Total divorce rate (%)	46.2	46.5	45.2	47.2	44.8	44.8	40.6
Mean duration of marriage at divorce (years)	12.9	13.0	13.1	13.2	13.4	13.5	13.7

Source: Czech Statistical Office; authors' calculations.

In addition, in the last three years the intensity of divorce decreased in the interval of 25–29 years (from 0.61 in 2017 to 0.54 divorces in 2020). Conversely, the number of divorces per 100 marriages has increased over the last decade for marriages lasting 30 years or more (from 0.22 to 0.25 divorces). If the intensity of divorce in individual lengths of marriage remained at the level of 2020, 40.6% of marriages would end in divorce, which was by 4.2 p.p. less year-on-year and the lowest since the beginning of the century. In the period 2011–2019, on the other hand, the values of the indicator ranged from 44.5% to 47.8%, with the maximum recorded in 2013. Compared to 2019, the intensity of divorce rates decreased in all durations of marriage. The mean duration of marriage at divorce has been increasing for more than two decades with smaller fluctuations; between the years 2011 and 2020 it increased by 0.8 years from 12.9 years to 13.7 years, and the last year-on-year increase was 0.2 years.

FERTILITY

A total of 110,200 live births were recorded as new inhabitants in Czechia, which was 2,000 less than in the previous year but 1,527 more children than in 2011 (see Table 6). In 2012 and 2013, the number of births decreased slightly, for the next four years it increased to 114,405 in 2017 and since then it has decreased again. The number of live births was lower for both first-born and second-born children and children born in the third or higher order. However, the structure of live births by birth order has been relatively stable over the last ten years. Of all live births, the first-born children accounted for 46.9% to 48.7% (47.6% in 2020), which was the highest percentage of all live births. Then, the second-born children accounted for 36.6% to 38.9% (37.6% in 2020) and live births of the third and higher order accounted for 14.1% to 15.1% (14.8% in 2020) of all live births. The absolute number of live births of all orders decreased year-on-year in 2020. A total of 431 children were born dead in 2020, which was 29 more than a year earlier. On average, a total of 387 children were born dead each year between 2011 and 2020, with higher numbers since 2012 mainly due to a change in the definition of stillbirths. From the point of view of the mother's marital status, children born to married women have long predominated, and this was no different in 2020, when 56,792 children were born to married women, more than half (51.5%) of all live births. Over the last ten years, the proportion of children born to married women has gradually decreased from 58.2% in 2011 to 51.0% in 2017. Since then, it has remained at almost 52%. While in the last five years the share of single mothers has ranged between 43% and 44%, until ten years ago their representation was only around one third (35.6% in 2011). A relative and absolute year-on-year decrease in 2020 was recorded in the case of divorced mothers, whose representation decreased from 4.3% to 4.1% compared to the previous year. The proportion of live births to divorced women has been declining continuously over the past ten years; in 2011 they gave birth to 6.0% of live births. Widowed women contribute only marginally to the number of live births; in 2020 they gave birth to 127 children (or 0.1% of all live births).

Table 6 Births, 2011, 2015–2020

Indicator	2011	2015	2016	2017	2018	2019	2020
Births	108,990	111,162	113,083	114,789	114,419	112,633	110,631
Live births	108,673	110,764	112,663	114,405	114,036	112,231	110,200
– first order	46.9	48.1	48.7	48.7	48.0	47.8	47.6
– second order	38.8	37.3	36.7	36.6	37.2	37.6	37.6
– third and higher order	14.3	14.7	14.6	14.7	14.7	14.6	14.8
Marital status of mother – single	38,666	46,887	48,807	50,379	49,956	49,137	48,799
– married	63,252	57,788	57,930	58,314	58,698	58,138	56,792
– divorced	6,514	5,911	5,730	5,539	5,227	4,818	4,482
– widowed	241	178	196	173	155	138	127
Stillbirths	317	398	420	384	383	402	431

Source: Czech Statistical Office.

The structure of live births by birth-order differs depending on whether the children were born in or outside a marriage. Firstly, of the total number of live births in marriage, 38.7% were born in the first order, 44.6 % in the second order and 16.7 % in the third and higher order. The structure has not changed much in the last ten years. Second order children dominate children born in marriage, while children born outside marriage have the highest share of first order children. Secondly, at the beginning of the past decade, the proportion of children born outside marriage was 41.8%, then increased until 2017, when it reached 49.0%, then decreased slightly in 2018 and 2019 for the first time in 30 years of continuous growth. In 2020, after a two-year decline, the percentage of live births outside marriage increased by 0.3 p.p. to 48.5%. The increase was mainly due to a higher percentage of live births outside marriage among firstborns (from 57.3% to 58.1%). Overall, almost half of live births are currently born outside marriage. In 2020, 58.1% of children of the first order, 38.9% of children of the second order and 41.9% of children of the third and higher order were born outside marriage.

The total fertility rate remained unchanged at 1.71 children per woman for the third year in a row in 2020 (Table 8). Ten years ago, in 2011, it decreased year-on-year to 1.43 (from 1.49), but then continued to grow to the current level. The highest relative year-on-year increase was recorded between 2013 and 2014, when total fertility increased by almost 5% (from 1.46 to 1.53 children per woman). In the next three years the total fertility rate grew at an average year-on-year rate of 3.7%, then growth slowed, resp. stopped. The mean age of mothers at childbirth has continued to rise over the last decade, but the growth has been smaller than in the previous two decades. Between 2011 and 2020, it increased by 0.5 years from 29.7 years to 30.2 years. The mean age of mothers at the childbirth remained at the same level as in the previous year. Compared to 2011, the mean age of mothers at childbirth in the first order increased by 0.7 years to 28.5 years. The mean age of mothers at childbirth in the second order increased by 0.4 years to 31.3 years in the same period, while the mean age of mothers at childbirth in the third and higher order fluctuated between 33.2 – 33.4 years (33.3 years

Table 7 Live births outside marriage by birth order, 2011, 2015–2020

Indicator	2011	2015	2016	2017	2018	2019	2020
Percentage of live births outside marriage	41.8	47.8	48.6	49.0	48.5	48.2	48.5
– first order	53.1	58.0	58.5	58.6	57.9	57.3	58.1
– second order	29.9	37.5	38.2	39.0	39.0	39.2	38.9
– third and higher order	37.0	40.8	41.7	42.3	42.0	41.5	41.9

Source: Czech Statistical Office.

Table 8 Fertility indicators, 2011, 2015–2020

Indicator	2011	2015	2016	2017	2018	2019	2020
Total fertility rate	1.427	1.570	1.630	1.687	1.708	1.709	1.707
– first order	0.699	0.787	0.829	0.858	0.856	0.852	0.849
– second order	0.535	0.570	0.582	0.600	0.619	0.624	0.622
– third and higher order	0.192	0.212	0.219	0.230	0.234	0.233	0.236
Net reproduction rate	0.689	0.759	0.787	0.816	0.829	0.826	0.830
Mean age of mother at childbirth	29.7	30.0	30.0	30.0	30.1	30.2	30.2
– first order	27.8	28.2	28.2	28.2	28.4	28.5	28.5
– second order	30.9	31.2	31.2	31.3	31.3	31.3	31.3
– third and higher order	33.3	33.4	33.3	33.4	33.4	33.4	33.3

Source: Czech Statistical Office; authors' calculations.

in 2020). The net reproduction rate has also remained at the same level for the last three years (2018–2020), namely 0.83 girls per woman.

In contrast to previous years, when fertility peaked in women at the age of 30, the maximum fertility rate was recorded at the age of 29 in 2020 (see Figure 2). In the last decade, its value has ranged between 117 children born per 1,000 women of a given age (in 2011 and 2013) and 129 (in 2017). In 2020 it was at a similar level as in the previous year – 127 children per 1,000 women of a given age. The rate of first-order fertility has increased over the last ten years in almost the entire reproductive age range of women, with the exception of women aged 16–20 years. From a relative point of view, the fertility rate increased the most in women aged 35 and over (to 1.5 times or more). In second-born children, there was a decrease in the fertility rate in women aged 17, 20 and 30 years. The fertility rate increased, relatively mostly again in the oldest age categories, starting at the age of 37. Fertility rates by age in children of the third and higher order were higher in 2020 compared to 2011 across all ages.

Figure 2 Age-specific fertility rates of women by age and by birth order, 2011 and 2020

Source: Czech Statistical Office; authors' calculations.

The age with the maximum fertility rate in the period between 2011 and 2020 was 28 or 29 years for the first order; the only exception was the year 2020, when the highest fertility rate was reached at the age of 27. The highest fertility rate of the second order was recorded in all years at the age of 31 or 32 years, in the third order at the age of 34 or 35 years, with the exception of 2013, in which it was already at the age of 33 years. The structure of total fertility by birth order does not change significantly.

MORTALITY

In 2020, a total of 129,289 inhabitants in Czechia died, which was 16,927 (or 15.1%) more than in the previous year and at the same time 18,159 (or 16.3%) more than the 2015–2019 average (Table 9). The last time the total number of deaths exceeded 129 thousand was in 1990, and the last time the number of deaths exceeded that of 2020 occurred 34 years ago, in 1986. Throughout the decade 2011–2020, the number of deaths due to ageing of the population had an increasing tendency with two fluctuations in 2014 and 2016, when there was a year-on-year decrease. Compared to 2019, in 2020 there was a decrease in the number of deaths of children under one year of age – a total of 249 of them died, which was 39 less year-on-year and the least in the last decade. Infant mortality in 2020 fell to 2.3‰ and was the lowest recorded level of infant mortality. In 2011–2019, its values oscillated around 2.6‰. The decline in the past year was driven more by neonatal mortality, which affects children over four weeks of age. Boys tend to have higher infant mortality rates than girls; in 2020 it reached 2.7‰, and 1.8‰ for girls.

Worsened mortality conditions led to a significant decrease in life expectancy at birth in 2020, by 1.0 year for men and by 0.7 years for women year-on-year, when life expectancy reached 75.3 years for men and 81.4 years for women. While in 2011–2019 life expectancy at birth had a growing trend, it decreased between 2019 and 2020, which grossly corresponded to the value from 2013 for men and the value from 2015 for women. In 2020, life expectancy turned out to be declining for all ages. If, for example, we look at the age of 65, which is generally considered the retirement age, men aged 65 had an average of 15.2 years expected to live in 2020, which was 1.1 year less than in the previous year. The life expectancy of women aged 65 was lower by 0.8 years, namely 19.2 years of age.

As a result of the COVID-19 epidemic, the year 2020 deviated from the typical mortality seasonality profile. The number of deaths in the individual months and their year-on-year development were significantly affected. Record high numbers of deaths were concentrated in the last three months of the year, when the "autumn wave" of the coronavirus epidemic broke out in Czechia. While in the first eight months the number of deaths was still around the average of recent years with a normal variance, in September it started to deviate more from the five-year average and in October it already exceeded record values. In 2020, most people died in November (15,751), which represented a year-on-year increase of 70.5%, then in October (14,189) and December (14,165). The monthly number of deaths exceeded 14,000 for the first time since December 1995, and the November maximum was similar to January 1970. The months with the lowest number of deaths were May (8,795)

Table 9 Deaths, 2011, 2015–2020

Indicator	2011	2015	2016	2017	2018	2019	2020
Deaths	106,848	111,173	107,750	111,443	112,920	112,362	129,289
Deaths under one year of age	298	272	317	304	292	288	249
Infant mortality rate (‰)	2.7	2.5	2.8	2.7	2.6	2.6	2.3
Life expectancy at birth – men	74.7	75.6	76.0	76.0	76.1	76.3	75.3
– women	80.8	81.5	81.8	81.8	81.9	82.1	81.4
Life expectancy at 65 – men	15.5	15.8	16.1	16.1	16.1	16.3	15.2
– women	18.9	19.3	19.7	19.6	19.7	19.9	19.2

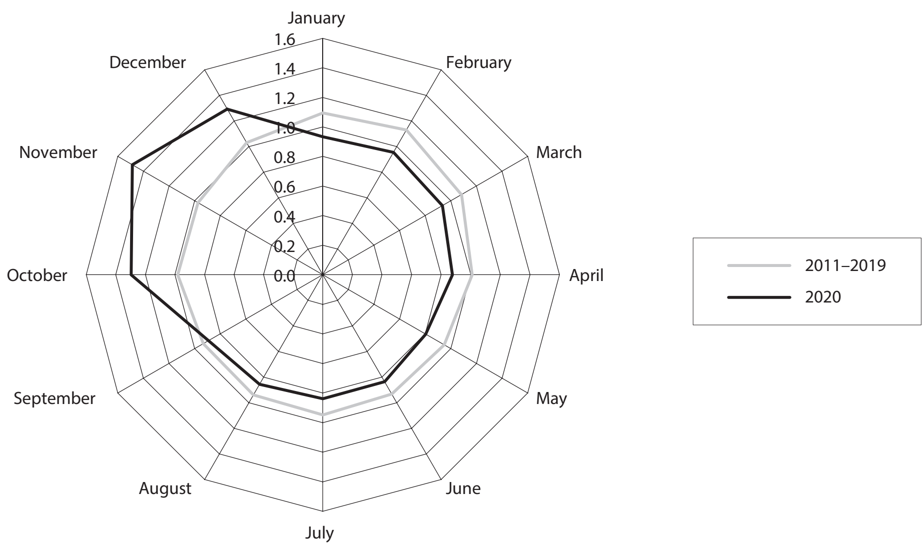
Source: Czech Statistical Office.

and June (8,847). When adjusted for the same number of days in each month of the year, most deaths were in February, resp. generally in the winter months along with March, and the least number of deaths was in the period from May to September (see Figure 3).

The development of the age structure of deaths over time, in addition to the mortality, also influences the development of the age structure of the population itself, especially the shift of numerically stronger or weaker generations to older age. Indicators of life tables can eliminate the influence of the age structure of the population on the number of deaths. In comparison with the male part of the population in the last decade, the deceased women were concentrated in a narrower interval at an older age (see Figure 4), in correspondence with the situation of empirical deaths and similarly as in previous years. Between 2011 and 2019, there was a relatively smooth shift in the curve of the life table deaths (not in childhood and infancy) towards an increase in the life table deaths at a very old age with its decrease in earlier and middle senior age, when the turning point was approximately the modal age. The epidemic situation in 2020 and changes in the probability of death caused a different development between 2019 and 2020. The curve of the life table deaths of women de facto returned to the state of 2011 (only slightly shifted to the right). For men it returned to the level of 2011 at most ages above the modal age, and for ages 67–88 years an even higher number of life table deaths in 2020 than in 2011 was recorded. The age with the highest life table deaths in 2020 was 87 years for women and 82 years for men. In comparison with 2011 and 2020, the modal age at death in women increased by only one year, in men it remained the same (between 2011 and 2019, however, it increased by two years for both sexes).

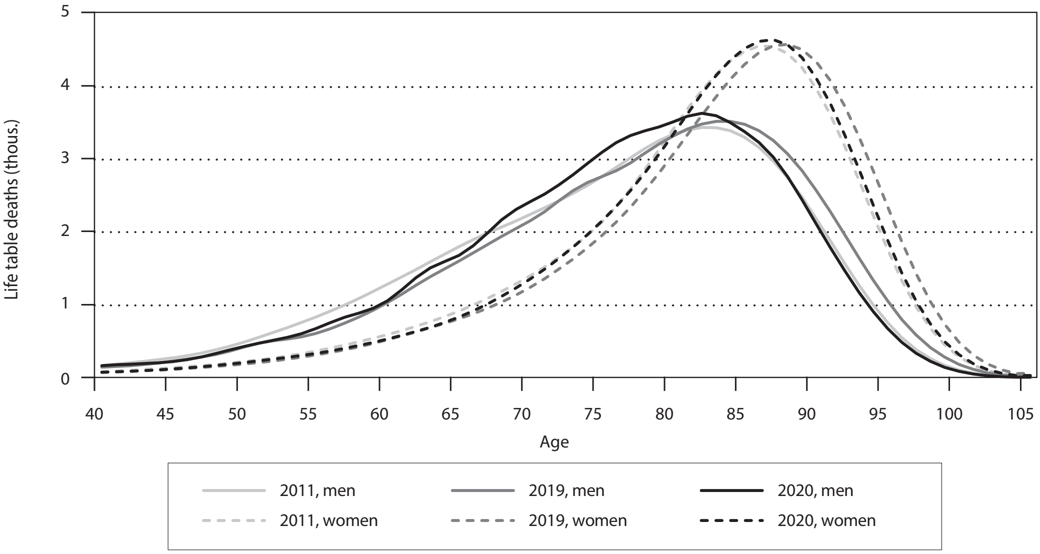
As seen in Figure 5, the increase in the life expectancy of men at birth between 2011 and 2019 was mostly due to a reduction in mortality in the 55–59 age group. The wider age range from 50 to 69 years then included an increase in life expectancy by 0.9 years, while overall their life expectancy at birth increased by 1.6 years. For women, the older age groups contributed more to the increase in life expectancy at birth in the same time period – the decrease in mortality between the ages of 75 and 89 ensured an increase of 0.6 years out of a total increase of 1.3 years between 2011 and 2019, improving the mortality of women aged 80–84 years. The decrease

Figure 3 Deaths – monthly indices, average of the period 2011–2019, 2020



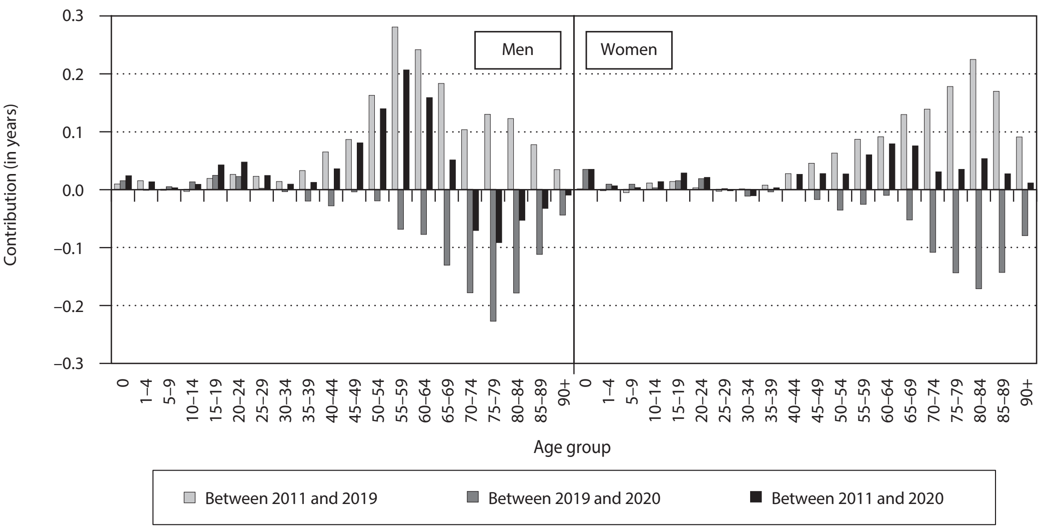
Source: Czech Statistical Office; authors' calculations.

Figure 4 Life-table deaths by sex and age, 2011, 2019 and 2020



Source: Czech Statistical Office.

Figure 5 Contributions of age groups to the difference in life expectancy by sex, 2011, 2019 and 2020



Source: Czech Statistical Office; authors' calculations.
Note: Method of calculating according to Pressat (1985).
Differences – between 2011 and 2019: +1.62 years (men), +1.27 years (women); between 2011 and 2020: +0.59 years (men), +0.54 years (women); between 2019 and 2020: –1.03 years (men) and –0.72 years (women).

in the life expectancy of men between 2019 and 2020 was due to a worsening of mortality in all age groups with relevant mortality rates (except the infancy age), with the deepest decline recorded in the 75–79 age group (by 0.2 years) and adjacent five-year age categories (all three together caused a decrease of 0.6 years). From the point of view of the decade defined by 2011 and 2020, the life expectancy at birth for men increased by 0.6 years between these extreme time points, as the positive development of mortality in the 65–69 age group contributed to growth. However, for older men their higher mortality in 2020 than in 2011 had a negative effect on the overall development of life expectancy at birth. The decrease in the life expectancy of women between 2019 and 2020 was mainly due to the change in the mortality rate in the same age groups, which on the contrary contributed the most to its increase between 2011 and 2019 (age groups 75–79, 80–84 and 85–89). The age groups in the range of 75–89 years included 0.5 years out of the overall decrease in the life expectancy of women between 2019 and 2020 (by 0.7 years). When combined for the entire decade (2011–2020), a positive effect of an increase in life expectancy of 0.5 years is evident in all age groups with a relevant mortality rate, mostly in the 55–69 age group.

Mortality by cause of death

Cause of death statistics are not fully comparable over time. Since 2013, there has been a significant update of the International Classification of Diseases (ICD-10) as well as adjustments in the process of data collection and processing in Czechia. From the data for 2018, the statistics on causes of death reflected the transition to a new version of IRIS software (used internationally to select the underlying cause of death) conditioned by the adoption of the ICD-10 update valid on 1 January 2018. In 2020, the ICD-10 (and the IRIS) was operatively updated on the disease COVID-19, which is treated as influenza and for which was assigned the code U07 from Chapter XXII. *Codes for special purposes* reserved in the emergency classification for selecting the underlying cause of death. In 2020, based on the standardized mortality rates (Table 10) the groups of causes of death in both men's and women's populations were ranked in the same way as the absolute number of deaths for these groups of causes, with the only exception being when endocrine, nutritional and metabolic diseases ranked sixth in men, while according to the absolute number of deaths they ranked seventh (they switched places with diseases of the digestive system). The same was the case in the previous year 2019. Between 2019 and 2020, the values of standardized mortality rates for individual groups of causes of death (the ICD chapters) moved in the same trajectories as the absolute numbers of deaths. For men, mortality increased for all common groups of causes except external causes and neoplasms (a decrease of 2.6% and 1.8%, respectively). The largest increase in mortality (excluding COVID-19) was recorded for endocrine, nutritional and metabolic diseases (by 8.3%) and circulatory diseases (by 7.2%), while for respiratory and digestive diseases the increase in mortality was lower (by 2.7% and 1.9%, respectively). In terms of the development of the whole decade, the negative development of men in 2020 did not outweigh the trend of improving mortality rates for circulatory system diseases, when in 2020 men had a lower mortality rate than in 2011 (by 19.7%). There was also a decrease in standardized men's mortality between 2011 and 2020 for external causes (by 13.5%) and neoplasms (by 11.9%). For women, mortality increased in 2020 (excluding COVID-19) from nervous system diseases (by 14.7%), endocrine, nutritional and metabolic diseases (by 14.4%) and circulatory system diseases (by 6.4%), other causes of death decreased (by 2–5%). In addition, for women, the negative development in 2020 did not reverse the long-term positive trend in the case of mortality from diseases of the circulatory system (a decrease of 22.6% between 2011 and 2020). The long-term development of standardized mortality from the last year-on-year (between 2019 and 2020) for women also differed for respiratory diseases, which generally increased over the decade (between 2011 and 2020 by 20.3%), but decreased year-on-year by 4.9%.

Table 10 Standardised mortality rates*) by selected causes of death (per 100,000), 2011 and 2020

Underlying cause of death (code according ICD-10)	Men		Women	
	2011	2020	2011	2020
Deaths – total	1 682.1	1 718.8	1 090.9	1 084.3
Neoplasms (C00–D48)	412.9	363.8	236.0	213.2
– Malignant neoplasm of colon, rectum and anus (C18–C21)	60.1	47.2	30.1	23.8
– Malignant neoplasm of pancreas (C25)	26.7	25.6	18.9	18.8
– Malignant neoplasm of trachea, bronchus and lung (C33–C34)	98.7	73.4	31.6	30.7
– Malignant neoplasm of prostate (C61)	45.0	40.0	33.1	29.0
Endocrine, nutritional and metabolic diseases (E00–E90)	37.0	70.5	32.1	55.3
– Diabetes mellitus (E10–E14)	31.4	60.3	26.7	45.7
Diseases of the nervous system (G00–G99)	30.0	46.7	22.5	41.5
– Alzheimer disease (G30)	14.1	26.3	13.5	29.4
Diseases of the circulatory system (I00–I99)	850.2	682.5	615.7	476.6
– Ischaemic heart diseases (I20–I25)	456.8	336.8	302.7	204.7
– Acute myocardial infarction (I21–I22**)	117.0	54.1	60.0	23.5
– Heart failure (I50)	69.8	87.6	46.8	64.5
– Cerebrovascular diseases (I60–I69)	158.1	93.7	138.2	72.9
– Atherosclerosis (I70)	56.1	21.0	45.2	15.9
Diseases of the respiratory system (J00–J99)	109.9	123.6	51.3	61.7
Diseases of the digestive system (K00–K93)	63.9	64.0	39.0	36.3
External causes of morbidity and mortality (V01–Y98)	96.9	83.9	37.2	32.4
– Transport accidents (V01–V99, Y85)	12.8	9.9	4.2	2.9
– Intentional self-harm (X60–X84, Y870)	27.0	20.4	4.8	3.9
COVID-19 (U07)	-	160.8	-	78.5
Other	81.2	123.0	57.0	88.8

Note: *) The European population standard issued by Eurostat (2013) was used for standardization.
**) Since 2018, subsequent myocardial infarction I22 has used the acute form I21 instead as the underlying cause of death.
Source: Czech Statistical Office; authors' calculations.

INTERNATIONAL MIGRATION

The volume and structure of migration flows were affected by epidemiological measures restricting movement across the country's borders and amendments to the Act on the Residence of Foreigners in the Czech Republic. A total of 55,661 persons immigrated to Czechia from abroad, of which 32,914 were men and 22,747 were women (see Table 11). Although there were 9,910 (or 15%) fewer immigrants year-on-year, their number was still above average compared to the average from 2015–2019 and the entire decade. Most immigrants in the last intercensal period (2011–2020) came from abroad to Czechia in 2019 (65.6 thousand), the least in 2011 (22.6 thousand). The number of emigrants in 2020 reached 28,734, of which 19,484 were men and 9,250 were women, and increased for the fourth year in a row. While in 2017 the annual increase was 1%, and about 10% in 2018 and 2019, in 2020 the increase was 35% when there were absolutely 7,433 thousand more emigrants than the year before. A higher number of emigrants than in 2020 was last recorded in 2013 (30.9 thousand). The balance of international migration in 2020 reached 26,927 persons and was almost equally distributed between men (13,430) and women (13,497). Compared to 2019 (a balance of 44.3 thousand), the increase in population by international migration was 17.3 thousand lower and the lowest in the last four years. However, it was higher compared to 2011–2016. During the last decade of 2011–2020, the largest increase in persons due to international migration was in the previous year 2019. Once, in 2013, a decrease in the population due to international migration was registered (−1.3 thousand).

From the point of view of age-distribution (in five-year age groups), people aged 25–29 have come to Czechia from 2012 on a regular basis. The other groups were immigrants aged 20–24 years and 30–34 years, with more persons in the last three years being 30–34 years old. Together, these three groups, 20–34 years, accounted for almost half of all immigrants (46% in 2020, 43–50% in other years of the decade). The whole group of people of working age (15–64 years) then included 82–90% of all immigrants. Their share was the highest in 2020, in which at the same time the lowest proportion of children under the age of 15 was registered in the whole decade. These accounted for less than 9% of immigrants (9.3% to 15.7% in other years). Seniors aged 65 and over traditionally represent a very small group of immigrants, amounting to 1–2%. The situation was no different in 2020 (1.7%), although the absolute number of immigrants of this age increased slightly. On the other hand, the number of immigrant children under the age of 15, as well as people of working age between 15–64 years, decreased in 2020. The average age of immigrants increased year-on-year for the sixth year in a row, exceeding 32 years for the first time since 2001. As in previous years, the age categories of 25–29, 30–34 and 20–24 contributed the most to the increase in international migration in 2020. The number of persons of this age increased by almost 16,0 thousand during 2020 but was 26% less than in 2019. The amount of children under the age of 15 (3.5 thousand) who immigrated did not deviate too much from the average of the previous five years; the year-on-year decrease of 18% was mainly due to its above-average amount in 2019.

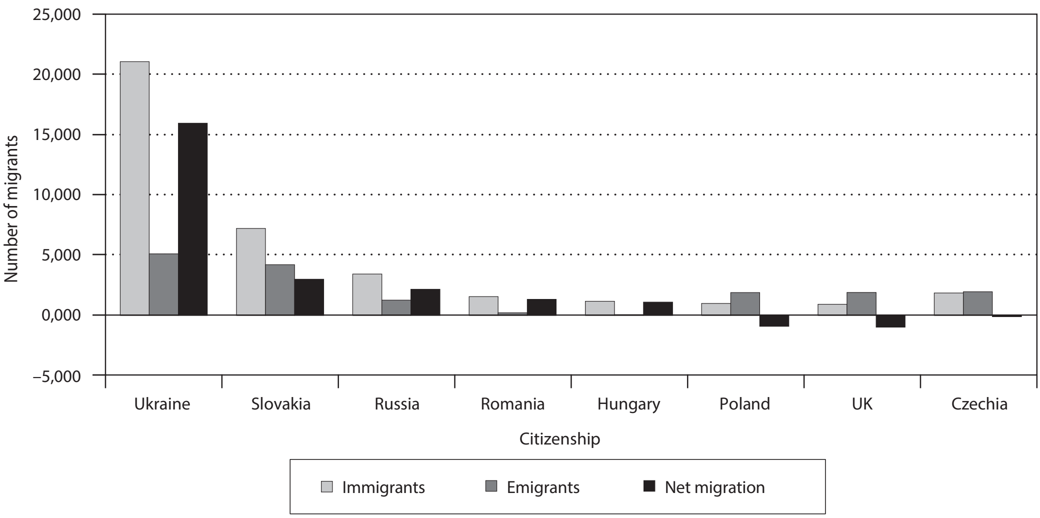
The structure of migrants according to their citizenship did not change much year-on-year in 2020, as the migration balance was, as in 2019, the highest among Ukrainian citizens, the second highest among Slovak citizens and the third among Russian citizens. Romanians again recorded the fourth highest balance, while Hungarians made the top five instead of Bulgarians. Similarly, there was no year-on-year change in immigrants, when the five most numerous groups of immigrants were Ukrainians, Slovaks, Russians, Czechs and Romanians. However, it was different when it came to emigrants, where Slovaks came in second to Ukrainians in 2020, Czechs descended from second place to third, and specifically the fourth and fifth most numerous groups consisted of emigrants with British and Polish citizenship. The change in the structure is mainly due to the implementation of the amendment to the Act on the Residence of Foreigners, where citizens of Great Britain and Poland, but also Slovakia, Germany, Austria and the Netherlands, were the ones who dominated among completed temporary stays lasting more than ten years. With the exception of Slovaks in all these named groups

Table 11 International migration, 2011, 2015–2020

Indicator	2011	2015	2016	2017	2018	2019	2020
Immigrants	22,590	34,922	37,503	45,957	58,148	65,571	55,661
– men	12,440	19,022	20,817	26,839	34,621	39,904	32,914
Emigrants	5,701	18,945	17,439	17,684	19,519	21,301	28,734
– men	3,109	10,502	9,417	9,964	11,201	12,348	19,484
Volume of migration	28,291	53,867	54,942	63,641	77,667	86,872	84,395
Net migration	16,889	15,977	20,064	28,273	38,629	44,270	26,927
0 –14	2,214	3,406	3,270	3,328	3,684	4,241	3,498
15–64	14,357	12,443	16,581	24,748	34,758	39,805	24,166
65+	318	128	213	197	187	224	–737

Source: Czech Statistical Office.

Figure 6 International migration by selected citizenship*), 2020



Note: *) Citizenships whose number of immigrants, emigrants or net migration was among the top five in 2020.
Source: Czech Statistical Office.

of EU citizens, the number of registered emigrants exceeded the number of immigrants, and the balance of international migration in 2020 was negative for foreigners with these citizenships.

CONCLUSION

Demographic events in 2020 in Czechia, as in other countries of the world, were affected by worsened epidemiological conditions due to the COVID-19 pandemic, especially after the introduction of the necessary measures to prevent its spread in the population. When comparing the years 2019 and 2020, the most unfavourable epidemiological situation resulted in a worsening of the mortality of the Czech population. The life expectancy of both men and women has fallen by about one year, while in the past its values have had a long-term upward trend. Although the number of children born decreased year-on-year, on the other hand the total fertility rate remained unchanged. The number of marriages decreased significantly because of the measures in place against the spread of COVID-19, namely regulations on the possibility and size of wedding ceremonies in terms of the number of participants. As in the case of marriages, the epidemic situation contributed to a lower number of divorces, however the number of divorces had a slightly declining trend in previous years. The volume and structure of migration flows were affected by epidemiological measures restricting movement across the country's borders and amendments to the Act on the Residence of Foreigners in the Czech Republic. Because the pandemic has not disappeared from our lives, it can be expected that demographic events will be affected in the coming years. In this context, some of the conclusions drawn from this article may be premature and it is therefore necessary to observe the further development of the population. In addition, the question arises as to what other factors, known or unknown, may influence population development from various perspectives, such as economic, social or political, etc.

References (sources of data)

- Demographic Yearbook of the Czech Republic. 2021. Prague: Czech Statistical Office. Available at: https://www.czso.cz/csu/czso/demographic-yearbook-of-the-czech-republic-2020.
- European Commission. 2013. *Revision of the European Standard Population-Report of Eurostat's task force.* Luxembourg: Publications Office of the European Union.
- Internal Database of Demographic Evidence. Prague: Czech Statistical Office.
- Life Tables for the Czech Republic. 2021. Prague: Czech Statistical Office. Available at: https://www.czso.cz/csu/czso/obyut_ts.
- Population – Methodology. 2021. Prague: Czech Statistical Office. Available at: https://www.czso.cz/csu/czso/population-methodology.
- Population change – Methodology. 2021. Prague: Czech Statistical Office. Available at: https://www.czso.cz/csu/czso/population-change-methodology.
- Population of the Czech Republic – year 2020. 2021. Prague: Czech Statistical Office. Available at: https://www.czso.cz/csu/czso/population-of-the-czech-republic-year-2020.
- Population Trends in the Czech Republic – 2005 … 2020 (Czech only). 2021. Prague: Czech Statistical Office.
- Pressat, R. 1985. Contribution des écarts de mortalité par âge à la différence des vies moyennes [Contribution of age-specific mortality differentials to mean lifetimes]. *Population (French edition)*, pp. 766–770. https://doi.org/10.2307/1532986.

DAVID MORÁVEK

Graduated from the Faculty of Informatics and Statistics at the Prague University of Economics and Business, where he is currently enrolled in the PhD Statistics program. Since 2018, he has been working at the Department of the Demographic Statistics Unit of the Czech Statistical Office. He is interested in the study of mortality and population forecasting.

JANA KOUKALOVÁ (KŘESŤANOVÁ)

Graduated from a master's program in demography and since 2015 has been a PhD student of demography at the Faculty of Science, Charles University. Since 2016, she has been working in the Department of Demographic Statistics at the Czech Statistical Office. She specialises in the analysis of demographic development and especially nuptiality and divorce rates.

Population and vital statistics of the Czech Republic: 2020, cohesion regions and regions

Cohesion region (NUTS 2), region (NUTS 3)	Population 1 July	Population 31 December	Marriages	Divorces	Live births	Abortions	Deaths			Increase (decrease)			Mar riages	Divorces	Live births	Deaths	Total increase
							Total	Within 1 years	Within 28 days	Natural	Net migration	Total	per 1,000 inhabitants				
Česká republika	**10,700,155**	**10,701,777**	**45,415**	**21,734**	**110,200**	**30,368**	**129,289**	**249**	**172**	**−19,089**	**26,927**	**7,838**	**4.2**	**2.0**	**10.3**	**12.1**	**0.7**
Praha	1,327,272	1,335,084	5,455	2,633	14,713	3,587	13,621	26	17	1,092	9,715	10,807	4.1	2.0	11.1	10.3	8.1
Střední Čechy	1,392,407	1,397,997	5,631	3,209	14,437	3,974	15,302	21	17	−865	13,721	12,856	4.0	2.3	10.4	11.0	9.2
Jihozápad	1,234,648	1,234,592	5,271	2,552	12,428	3,467	14,866	28	20	−2,438	3,048	610	4.3	2.1	10.1	12.0	0.5
Severozápad	1,113,663	1,110,315	4,740	2,417	10,629	4,256	14,819	31	18	−4,190	−1,124	−5,314	4.3	2.2	9.5	13.3	−4.8
Severovýchod	1,518,116	1,516,135	6,318	3,127	15,537	4,332	18,329	44	32	−2,792	928	−1,864	4.2	2.1	10.2	12.1	−1.2
Jihovýchod	1,703,839	1,704,179	7,426	3,156	18,208	4,241	20,465	43	29	−2,257	4,634	2,377	4.4	1.9	10.7	12.0	1.4
Střední Morava	1,213,141	1,210,641	5,178	2,288	12,418	3,325	15,920	23	16	−3,502	−427	−3,929	4.3	1.9	10.2	13.1	−3.2
Moravskoslezsko	1,197,069	1,192,834	5,396	2,352	11,830	3,186	15,967	33	23	−4,137	−3,568	−7,705	4.5	2.0	9.9	13.3	−6.4
Hlavní město Praha	1,327,272	1,335,084	5,455	2,633	14,713	3,587	13,621	26	17	1,092	9,715	10,807	4.1	2.0	11.1	10.3	8.1
Středočeský kraj	1,392,407	1,397,997	5,631	3,209	14,437	3,974	15,302	21	17	−865	13,721	12,856	4.0	2.3	10.4	11.0	9.2
Jihočeský kraj	643,759	643,551	2,792	1,345	6,552	1,954	7,800	15	10	−1,248	716	−532	4.3	2.1	10.2	12.1	−0.8
Plzeňský kraj	590,889	591,041	2,479	1,207	5,876	1,513	7,066	13	10	−1,190	2,332	1,142	4.2	2.0	9.9	12.0	1.9
Karlovarský kraj	294,187	293,311	1,232	672	2,682	948	4,026	6	3	−1,344	−9	−1,353	4.2	2.3	9.1	13.7	−4.6
Ústecký kraj	819,476	817,004	3,508	1,745	7,947	3,308	10,793	25	15	−2,846	−1,115	−3,961	4.3	2.1	9.7	13.2	−4.8
Liberecký kraj	443,161	442,476	1,906	991	4,557	1,621	5,252	10	6	−695	−519	−1,214	4.3	2.2	10.3	11.9	−2.7
Královéhradecký kraj	551,605	550,803	2,228	1,124	5,526	1,489	6,694	15	14	−1,168	324	−844	4.0	2.0	10.0	12.1	−1.5
Pardubický kraj	523,350	522,856	2,184	1,012	5,454	1,222	6,383	19	12	−929	1,123	194	4.2	1.9	10.4	12.2	0.4
Kraj Vysočina	509,855	508,852	2,154	890	5,349	1,307	6,450	6	5	−1,101	140	−961	4.2	1.7	10.5	12.7	−1.9
Jihomoravský kraj	1,193,984	1,195,327	5,272	2,266	12,859	2,934	14,015	37	24	−1,156	4,494	3,338	4.4	1.9	10.8	11.7	2.8
Olomoucký kraj	631,767	630,522	2,665	1,255	6,584	1,799	8,126	10	4	−1,542	49	−1,493	4.2	2.0	10.4	12.9	−2.4
Zlínský kraj	581,374	580,119	2,513	1,033	5,834	1,526	7,794	13	12	−1,960	−476	−2,436	4.3	1.8	10.0	13.4	−4.2
Moravskoslezský kraj	1,197,069	1,192,834	5,396	2,352	11,830	3,186	15,967	33	23	−4,137	−3,568	−7,705	4.5	2.0	9.9	13.3	−6.4

Radek Havel

Population and vital statistics of the Czech Republic in towns with population above 50 thousands: 2020

Town	Population 1 July	Population 31 December	Marriages	Divorces	Live births	Abortions	Deaths	Increase (decrease)			Marriages	Divorces	Live births	Deaths	Total increase
								Natural	Net migration	Total	per 1,000 inhabitants				
Praha	1,327,272	1,335,084	5,455	2,633	14,713	3,587	13,621	1,092	9,715	10,807	4.1	2.0	11.1	10.3	8.1
Brno	381,702	382,405	1,717	680	4,501	1,019	4,652	–151	1,210	1,059	4.5	1.8	11.8	12.2	2.8
Ostrava	286,329	284,982	1,283	538	2,945	832	3,984	–1,039	–1,947	–2,986	4.5	1.9	10.3	13.9	–10.4
Plzeň	175,253	175,219	706	333	1,779	450	2,034	–255	632	377	4.0	1.9	10.2	11.6	2.2
Liberec	104,508	104,261	434	239	1,098	360	1,241	–143	–398	–541	4.2	2.3	10.5	11.9	–5.2
Olomouc	100,696	100,514	416	201	1,102	329	1,196	–94	–55	–149	4.1	2.0	10.9	11.9	–1.5
České Budějovice	94,350	94,229	445	193	1,061	325	1,104	–43	–191	–234	4.7	2.0	11.2	11.7	–2.5
Hradec Králové	92,800	92,683	362	186	1,012	182	1,218	–206	–50	–256	3.9	2.0	10.9	13.1	–2.8
Ústí nad Labem	92,414	91,982	385	184	983	368	1,239	–256	–478	–734	4.2	2.0	10.6	13.4	–7.9
Pardubice	91,965	91,755	386	185	945	236	1,130	–185	213	28	4.2	2.0	10.3	12.3	0.3
Zlín	74,676	74,478	339	138	702	196	943	–241	–216	–457	4.5	1.8	9.4	12.6	–6.1
Havířov	70,887	70,165	331	142	647	221	1,024	–377	–658	–1,035	4.7	2.0	9.1	14.4	–14.6
Kladno	69,216	68,896	274	152	677	281	866	–189	–252	–441	4.0	2.2	9.8	12.5	–6.4
Most	65,905	65,341	277	123	616	283	877	–261	–432	–693	4.2	1.9	9.3	13.3	–10.5
Opava	56,197	55,996	254	121	552	174	712	–160	–294	–454	4.5	2.2	9.8	12.7	–8.1
Frýdek–Místek	55,194	55,006	205	120	575	115	732	–157	–394	–551	3.7	2.2	10.4	13.3	–10.0
Jihlava	51,270	51,125	212	121	542	165	660	–118	27	–91	4.1	2.4	10.6	12.9	–1.8
Karviná	51,822	50,902	220	97	452	178	868	–416	–810	–1,226	4.2	1.9	8.7	16.7	–23.7

Radek Havel

Abstracts of Articles Published in the Journal Demografie in 2021 (Nos. 1–3)

Sylva Höhne – Jana Paloncyová

CHILD CUSTODY AND MAINTENANCE AFTER PARENTAL SEPARATION ACCORDING TO THE RESULTS OF COURT PROCEEDINGS

This article examines child custody in 2018 after parental separation. In most cases mothers are given sole custody of the children and fathers are responsible for the payment of child maintenance. However, the share of fathers who are awarded sole custody increases slightly with the age of the child and shared custody is more common arrangement in the case of children aged 4 to 15 years. An agreement between parents significantly reduces the duration of court proceedings on this matter.

Keywords: child custody, maintenance, parental separation, court decision, Czech Republic

Demografie, 2021, **63: 3–21**

Albína Malinová

AN ANALYSIS OF CASES OF WORK INCAPACITY AND OF PERSONS WHO EXPERIENCED A TERM OF WORK INCAPACITY IN THE CZECH REPUBLIC IN 2009–2019

This article analyses the level of work incapacity in the Czech Republic in 2009–2019. The analysis is focused on the number of cases of work incapacity, the number of persons who experienced a term of work incapacity, and the most common causes of work incapacity. The level of work incapacity increased and it was higher among women than men, and the most common health reason for work incapacity was an unspecified acute upper respiratory infection.

Keywords: work incapacity, age, gender, development trend, Poisson regression, Czech Republic

Demografie, 2021, **63: 22–38**

Branislav Šprocha

THE CONTINUING TRANSFORMATION OF NUPTIALITY AND DIVORCE IN CZECHIA AND SLOVAKIA AFTER 1989 IN A COHORT PERSPECTIVE

Nuptiality and divorce are processes that have undergone several important changes in Czechia and Slovakia in the last three decades. The main aim of this paper is a cohort analysis of the quantum and tempo of nuptiality among single persons and of divorce rates among the selected marriage cohorts that have been most affected by the transformation of marriage behaviour in Czechia and Slovakia since 1989. The results show a significant and, among selected cohorts, gradually steeper decline in the rate of first marriages. At the same time, there has been a continuous and inter-cohort intensification of the process of postponing entry into the first marriage in both countries and in both sexes. These changes are occurring more dynamically in Czechia. This means that in the birth cohorts from the late 1970s, more than a third of men and almost 30% of women never marry. In Slovakia, the figure is about 30% of men and a quarter of women. The main reason for this is the significant drop in the probability of marriage at a younger age and insufficient recuperation in older ages. The steadily increasing cross-sectional level of divorce rates in both countries was also reflected in the development of the cohort probability of divorce. The highest risk of divorce (47%–48%) was identified in Czechia among the marriage cohorts from the second half of the 1980s and the early 1990s. In Slovakia, the highest risk was slightly lower (33–34%) and was observed among the marriage cohorts from the first half of the 1990s. Younger cohorts were affected by the faster increase in cross-sectional divorce rates. The result of the differences in the dynamics of the divorce rate trends between Czechia and Slovakia was thus a certain equalisation of the intensity of divorce among younger cohorts in both countries.

Keywords: nuptiality, divorce, cohort approach, Czechia, Slovakia **Demografie,** 2021, **63: 91–104**

Branislav Šprocha – Branislav Bleha

TRANSFORMATION OF FERTILITY IN URBAN AND RURAL AREAS IN SLOVAKIA AFTER 1989

Fertility in Slovakia has undergone several important changes since 1989. The main goal of this paper is to analyse the most important shifts in the rate and timing of fertility and of fertility in relation to marital status in urban and rural communities. At the same time, we also try to identify how the transformation period affected existing differences in selected aspects of fertility between urban and rural communities. In addition, we try to estimate how the current transformation has affected the cohort fertility and parity structure of women in those cohorts that were most affected by the transformation process after 1989. Our results confirmed the long-term higher fertility of women in rural communities as well as the obvious differences in the structure of women by parity between urban and rural communities. After 1989 transformational changes in the process of fertility occurred more dynamically in urban communities. Because of the faster recovery from the postponement of childbirth in urban communities, there has been some convergence in terms of fertility rate numbers. On the other hand, the more significant postponement of childbirth in urban areas has contributed to a deepening of the urban/rural differences in terms of fertility tempo.

Keywords: fertility, urban and rural communities, Slovakia **Demografie,** 2021, **63: 139–157**

Mohammad S. Zahangir – Mohammed Chowdhury – Mosammat Z. Nahar – Hafiz Khan – Mohammad Masum

TRENDS AND DETERMINANTS OF KNOWLEDGE AND AWARENESS OF HIV/AIDS AMONG MARRIED WOMEN IN BANGLADESH: AN URBAN–RURAL COMPARISON

The aim of this study was to examine the trends and determinants of knowledge and awareness of HIV/AIDS among women in urban and rural areas of Bangladesh. This study used data from the 2014 Bangladesh Demographic and Health Survey (BDHS). A binary logistic regression model was employed to detect potential risk factors (covariates) associated the outcome variable. While women's HIV/AIDS knowledge in rural areas has shown an increasing trend over the years in which the survey has been conducted, a narrow decline in HIV/AIDS knowledge was observed among women in urban areas in recent years. Education and mass media have played the major role in spreading knowledge about HIV/AIDS among women in both urban and rural areas. Since rural women's HIV/AIDS knowledge score was relatively lower than that of urban women, awareness raising through mass media should focus especially on rural areas.

Keywords: HIV/AIDS infection, knowledge score, logistic regression, odds ratio **Demografie,** 2021, **63: 158–171**

MANUSCRIPT SUBMISSION

The journal accepts manuscripts in print or electronic format accompanied by a cover letter with complete contact information, including an email address.

MANUSCRIPT FORMAT

Articles should not exceed 20 standard pages (1 SP = 1800 characters with spaces) or 36 000 characters with spaces. Submissions to the Discussion and Summaries sections should be a maximum of 8 standard pages and reviews a maximum of 4 standard pages. Along with the manuscript please submit a five-line abstract and a twenty-line summary in English, a complete list of references in alphabetical order, and a brief (five-line) biography about the author with information on the author's specialisation and most important publications. The preferred format for manuscripts is Word, Excel for the source data for tables and figures, and *.tif, *.jpg, or *.eps for illustrations and maps. Manuscript pages should be numbered and tables, figures and illustrations should be numbered consecutively and their placement indicated in the text.

The peer-review procedure is double-blind. Decisions whether to publish are made by the Editorial Board and the author is informed of their decision within fourteen days of the Editorial Board's meeting.

MANUSCRIPT PREPARATION

A. ARTICLES (IN MS WORD)

1. Use the 'align left' attribute to align paragraphs.
2. In-text highlighting (italics, bold) and the use of indexes will be formatted as submitted.

B. FIGURES, ILLUSTRATIONS

1. Figures must be submitted with source data in Excel along with the preferred type.
2. All illustrations – e.g. maps, photographs – should be submitted as a separate file in *.tif, *.jpg, *.eps, with a reference in the text to its placement .

C. REFERENCES AND HEADINGS

Examples of reference formats:

Books

- Roubíček, V. 1997. *Úvod do demografie*. Prague: Codex Bohemia.
- Hantrais, L. (ed.). 2000. *Gendered Policies in Europe. Reconciling Employment and Family Life*. London: Macmillan Press.
- *Potraty*. 2005. Prague: Ústav zdravotnických informací a statistiky.

Articles in periodicals

- Bakalář, E. and Kovařík, J. 2000. 'Fathers, Fatherhood in the Czech Republic.' *Demografie*, 42, pp. 266–272.

For periodicals that use consecutive page numbering within a volume it is not necessary to indicate the issue number.

Chapter contributions

Daly, M. 2004. 'Family Policy in European Countries.' In *Perspectives on Family Policy in the Czech Republic*, pp. 62–71. Prague: MPSV ČR.

Electronic sources

1. Specify the medium (on-line, CD ROM, database, data set, CD)
2. date downloaded [cit. 29. 10. 2005]
3. web address (<http://www.czso.cz>)

Conference papers

Maur, E. '*Problems with the Study of Migration in the Czech Lands in Early Modern History*.' Paper presented at the conference 'The History of Migration in the Czech Lands during the Early Modern Period. Prague, 14. 10. 2005.

References

The list of references must be in alphabetical order. References to two or more works by the same author are listed in order of the year of publication. If two or more publications by one author date from the same year, distinguish between them adding the letter a, b, c, etc. to the year of publication. Example:

Syrovátka, A. 1962a. 'Injuries in the Household.' *Czech Paediatrics*, 17, pp. 750–753.

Syrovátka, A. 1962b. 'Child Mortality from Automobile Accidents in the Czech Lands.' *Czech Medical Journal*, 101, pp. 1513–1517.

In-text references

(*Srb*, 2004); (*Srb*, 2004: pp. 36–37); (*Syrovátka et al.*, 1984).

Table and figure headings

Table 1: Population and vital statistics, 1990–2010

Figure 1: Relative age distribution of foreigners and total population of CR, 31 Dec 2009

ADDRESS OF THE EDITORIAL OFFICE:

Czech Statistical Office
Editorial Office of Demografie
Na padesátém 81, 100 82 Prague 10-Strašnice
Phone: 274 052 834, e-mail: redakce@czso.cz
https://www.czso.cz/csu/czso/demografie_review_for_population_research

CZECH
STATISTICAL
OFFICE

Demografie
Review
for Population Research

www.czso.cz

Demography, Review for Population Research
Demografie, revue pro výzkum populačního vývoje

Published by the Czech Statistical Office
Vydává Český statistický úřad

Adress of the Editorial Office: Na padesátém 81, 100 82 Prague 10 – Strašnice, Czech Republic
Phone: +420 274 054 240
E-mail: redakce@czso.cz
www.czso.cz

The journal is published quarterly. As of 2011, three issues are published in Czech (articles are accompanied by abstracts and summaries in English) and one issue in English. *Demografie* is published in print format and in electronic format on the website of the Czech Statistical Office (https://www.czso.cz/csu/czso/demografie_review_for_population_research).

Subscriptions and orders: Information Services, Czech Statistical Ofce, Na padesátém 81, 100 82 Prague 10 – Strašnice, Czech Republic, e-mail: objednavky@czso.cz.

Typesetting: Družstvo TISKOGRAF, David Hošek
Design and layout: Ondřej Pazdera
Print: Czech Statistical Office
Issues of the journal can be bought from the CZSO Publications Shop (58 CZK per copy).
The price of an annual subscription is 4 x 58 CZK + postage.
Index number 46 465, ISSN 0011-8265 (Print), ISSN 1805-2991 (Online),
Reg. Zn. MK ČR E 4781

Number 4/2021, volume 63
This issue was published in December 2021