Global burden of non-communicable diseases – why prevalence is not a comprehensive measure

Sophie Kaufmann, Annika Hoyer, Ralph Brinks*

German Diabetes Center, Institute for Biometry and Epidemiology, Germany

Received: June 20, 2016
Accepted: September 5, 2016
Online Published: September 9, 2016
DOI: 10.5430/jer.v3n1p10
URL: http://dx.doi.org/10.5430/jer.v3n1p10

ABSTRACT

Background: The prevalence is a common epidemiological measure to describe the morbidity and burden of a non-communicable disease in a population. Often it is used to compare various populations with respect to the disease, e.g. in different countries. However, comparing the prevalence in different populations may lead to erroneous conclusions about the disease specific situation, because the prevalence is a result of a complex interplay between incidence and mortality. In this article, we show how distinctively different population based interventions against type 2 diabetes mellitus (T2DM) lead to the same age-specific prevalence of the disease. By comparing the prevalences of T2DM, the impact of the interventions would be rated as being equivalent – which they are not.

Methods: Based on data from the National Health Interview Surveys (NHIS) and the National Health and Nutrition Examination Surveys (NHANES), we calculate the effects of two hypothetical prevention programs against T2DM in the United States. The first prevention program is based on a pharmacological intervention, which is assumed to affect the incidence of diabetes only. The second program is a lifestyle intervention affecting both, incidence of diabetes and the mortality in the population.

Results: Choosing plausible participation rates, both prevention programs have the same impact on the age-specific prevalence of diabetes. Considering the prevalence alone would erroneously indicate a similar effect of both programs although the lifestyle intervention is superior in terms of the mortality in the population.

Conclusion: A comprehensive description of the burden of a chronic disease in a population should not rely on the prevalence alone, but include further measures, such as the incidence and the mortality rates. This is especially important in setting up surveillance systems for chronic diseases as requested e.g. by the United Nations’ resolution against chronic diseases.

Key Words: Prevalence, Incidence, Intervention, Diabetes, Chronic diseases, Mortality, Illness-death model, National Health Interview Survey, National Health and Nutrition Examination Survey

1. INTRODUCTION

The majority of all global deaths were caused by non-communicable diseases (NCDs), such as cancer, cardiovascular and chronic respiratory diseases, diabetes and neurological disorders.[1] From 1990 to 2010, not only the absolute number of NCD deaths rose from 27 to 35 million, but also the proportion of overall deaths which were attributable to NCDs increased from 57% in 1990 to 65% in 2010. The enormous and increasing burden of NCDs attracted highest economical and political councils. As an example, in 2011 for the first time in history, a United Nations’ General Assembly was solely devoted to NCDs. The passed resolution demands to strengthen country-level surveillance systems and monitoring systems for risk factors, determin-
The prevalence of an NCD is often used for demonstrating and comparing the burden of the disease in different countries, for example in diabetes,[3] cancer[4] or dementia.[5] A prominent example is the Atlas of the International Diabetes Federation, which is updated every two to three years. The IDF Atlas makes extensive usage of the prevalence of type 1 and type 2 diabetes for comparing different countries.

Despite the frequent use of the prevalence of NCDs for making cross-country comparisons, it is very well known in epidemiology that the prevalence of a chronic disease likewise depends on the incidence of the disease and also on the mortality. Although Niels Keiding published a review 25 years ago,[6] the exact quantitative relations between incidence, mortality and prevalence are rarely used in epidemiology and public health. A reason for seldomly using these quantitative relations might be the lacking tradition of compartment models, such as the illness-death model (see below), in NCD epidemiology. This is in contrast to the field of infectious disease epidemiology, where compartment models have been used for decades.[7]

In this article, we will demonstrate that the prevalence is only a weakly informative measure to compare the public health situation between different countries with respect to an NCD. This is done by contrasting the effects of two hypothetical population wide intervention programs against type 2 diabetes mellitus. Although the interventions are very different – one intervention just reduces the incidence of diabetes while the other also reduces the mortality – they have the same impact on the age-specific prevalence. Hence by comparing the prevalence alone, both interventions would be rated to be of equal value.

Moreover, we show how relations between prevalence and incidence can be used to make estimates about the impact of population wide interventions.

2. MATERIAL AND METHODS

Based on epidemiological data from the National Health Interview Survey (NHIS) and National Health and Nutrition Examination Survey (NHANES), we consider two hypothetical population wide prevention programs against type 2 diabetes mellitus in the United States of America. NHANES and NHIS are representative epidemiological surveys about the health status and health related behaviour of the US population. They are primarily cross-sectional studies, which are conducted in regular waves.[8,9]

To calculate the impact of the interventions on the prevalence of diabetes, we use a recently developed partial differential equation (PDE).[10] The PDE describes the age-specific prevalence $\pi$ of those persons aged $a$ at time $t$ in terms of the illness-death model shown in Figure 1.[11] The illness-death model splits the considered population into the relevant disease states healthy ($H$), ill ($I$) and dead ($D$). The transition rates between these states are the incidence ($\lambda$), the mortality of the healthy ($\mu_0$) and the mortality of the diseased ($\mu_1$). These rates depend on the calendar time $t$ and age $a$.

![Figure 1. Illness-death compartment model. The transition rates $\lambda, \mu_0, \mu_1$ between the compartments $H$ (healthy), $I$ (ill), and $D$ (dead) depend on calendar time $t$ and age $a$.](image)

Let the numbers $H, I, D$ denote the numbers of persons in the associated states. Then, Brinks and Landwehr have shown that the age-specific prevalence $\pi(t, a) = \frac{I(t, a)}{H(t, a)+I(t, a)}$ can be expressed by the following PDE:[12]

$$\frac{\partial}{\partial t} + \frac{\partial}{\partial a} \pi = (1 - \pi) \left( \lambda - \pi (\mu_1 - \mu_0) \right).$$

(1)

The PDE relates the temporal change of the age-specific prevalence $\pi$ with the incidence $\lambda$ and the mortality rates $\mu_0, \mu_1$. Although an extension of Eq. (1) exists for the case of migration,[10] for brevity we assume that the magnitude of migration is negligible compared to the resident population. This is the only assumption we have to make.

Together with an appropriate initial condition (e.g., $\pi(t_0, a) = \pi_0(a)$ for all $a$), calculating the prevalence $\pi(t, a)$ can be achieved by numerical integration of the PDE (1). For this, the PDE is transformed into an ordinary differential equation (ODE) by the Method of Characteristics.[13] Then, the classical Runge-Kutta method of fourth order is used to solve the resulting ODE.[14]

In addition to the age-specific prevalence, we calculate the effect of the interventions on the 20-year survival rate in the general population for the ages 35, 45, 55, and 65 years. For this, we determine the overall mortality $\mu$ in the base case and each of the interventions by the relation...
\( \mu(t, a) = (1 - \pi(t, a)) \mu_0(t, a) + \pi(t, a) \mu_1(t, a). \)

Then, we calculate the 20-year survival probability \( S_{20}(a) \) for a person aged \( a \) in 1995 by

\[
S_{20}(a) = \exp \left( - \int_{0}^{20} \mu(1995 + \tau, a + \tau) d\tau \right). \tag{2}
\]

Hence, \( S_{20}(a) \) describes the probability that a person aged \( a \) (in years) at \( t = 1995.0 \) survives at least 20 years until \( t = 2015.0 \).

In the following sections we describe the situation when no intervention is enrolled in the population (base case) and vary the base case by considering the effects of the two hypothetical interventions.

### 2.1 Base case
In case no intervention is enrolled in the population, we use the incidence rate \( \lambda \) of type 2 diabetes in men. Similarly, the mortality rates \( \mu_0 \) and \( \mu_1 \) of men are applied. All rates stem from NHIS.\(^{[15]}\) We use the first day of the year 1995 as start for the integration and accordingly choose the prevalence of diagnosed type 2 diabetes of men at \( t = 1995.0 \) as initial condition for solving the PDE. For this, we used data from NHANES.\(^{[16]}\) Details about the rates and initial conditions are provided in the Technical Appendix at the end of this article.

### 2.2 Interventions
We study the effect of two hypothetical interventions. As in the base case, we restrict ourselves to the male population. Using \( t = 1995.0 \) as initial condition for integrating the PDE (1), we assume the interventions to start at \( t = 2000.0 \). The interventions have an impact on the rates \( \lambda, \mu_0 \) and \( \mu_1 \) in Eq. (1) as described below. We allow three years for the interventions to take full effect. Between \( t = 2000.0 \) and \( t = 2003.0 \), we interpolate the effect of the interventions on the rates linearly. From beginning of 2003 the interventions have the full impact on the rates until \( t = 2015.0 \), which is chosen as the date of comparing the age-specific prevalences of both interventions.

#### 2.2.1 Pharmacological intervention
The first considered intervention is a pharmacological one. We assume that oral diabetes drugs are prescribed to a considerable part of the population. Two types of interventions are possible: either prescribing the drugs to persons with a highly elevated risk of contracting diabetes (high-risk intervention) or by prescription to persons with an increased risk of contracting diabetes.\(^{[17]}\) The details of how the intervention is accomplished depend on many factors, e.g., the prices of the drugs, the side-effects, the logistics etc.\(^{[17]}\) These details are beyond the scope of this article.

The effect of oral antidiabetics on reducing the incidence of type 2 diabetes has been studied in a meta-analysis.\(^{[18]}\) According to the results of the meta-analysis, we model the impact of the pharmacological intervention on the incidence rate \( \lambda \) by a hazard ratio of \( \text{HR}^{(\lambda)} = 0.70 \). A hazard ratio of 70% means that once the full effect of the intervention is reached, the incidence of diabetes is reduced to 70% of the incidence without the intervention.

If a prevention program is enrolled in a population, typically not all subjects in the population take part. On the one hand there are practical reasons against complete participation such as monetary or logistical restrictions. On the other hand, any pharmacological intervention has side effects and, of course, participation is voluntary. For the overall effect of the pharmacological intervention, we assume that the resulting incidence \( \lambda' \) of type 2 diabetes is a combination of the incidence \( \lambda \) from NHANES without intervention and the reduced incidence \( \text{HR}^{(\lambda)} \lambda \):

\[
\lambda' = \gamma \text{HR}^{(\lambda)} \lambda + (1 - \gamma) \lambda = \lambda \left[ \gamma \text{HR}^{(\lambda)} + (1 - \gamma) \right].
\]

For our simulation we choose \( \gamma = 0.33 \), which implies that the overall incidence \( \lambda \) of the base case is reduced to an incidence \( \lambda' = 0.89 \lambda \) in case of the pharmacological intervention. This does not seem too optimistic or unrealistic for an intervention targeted to the population with high or increased risk of contracting diabetes.

As there is controversial evidence if oral diabetes drugs have an impact on the all-cause mortality,\(^{[19]}\) we assume that the mortality rates \( \mu_0 \) and \( \mu_1 \) remain unchanged.

#### 2.2.2 Lifestyle intervention
For comparison, we study the effects of a lifestyle intervention (diet and exercise). Lifestyle interventions affect the incidence of type 2 diabetes and the mortality rates. It reduces the incidence by about 51%, i.e., the hazard ratio is \( \text{HR}^{(\lambda)} = 0.49 \).\(^{[18]}\) The associated hazard ratios for the mortality rates are \( \text{HR}^{(\mu_0)} = 0.79 \) and \( \text{HR}^{(\mu_1)} = 0.795 \).\(^{[21]}\)

Again, the effect of the participation rate on the lifestyle intervention is modelled by an additional factor \( \theta_0 \) in the population without diabetes and by a factor \( \theta_1 \) in the population with type 2 diabetes. Thus, we have following rates in the lifestyle intervention:
We choose \( \theta_0 = 0.2 \) and \( \theta_1 = 0.1 \), which implies that the participation rate in the lifestyle intervention is lower in the population that already contracted diabetes. This choice is motivated by the fact that persons with diabetes have a higher chance of being physically disabled,\(^{[22]}\) which makes it more difficult to exercise.

### 3. RESULTS

#### 3.1 Base case

Figure 2 shows the calculated age-specific prevalence of type 2 diabetes in different years from 1995 to 2015 if the base case scenario (no intervention) is assumed. We see that the prevalence is increasing between 1995 and 2015 for every age group. The mean of the calculated prevalence rises by 89\% during 1995 and 2015. Indeed, the observed prevalence of diabetes has grown by a similar magnitude: The age-standardized prevalence of diabetes in the male US population increases from 3.3\% in 1995 to 6.6\% in 2014.\(^{[23]}\) Unfortunately, the CDC does not present exact age courses of the prevalence over time, such that a detailed comparison is not possible. However, both of the presented figures indicate the growing burden of type 2 diabetes in the US over the past two decades.

#### 3.2 Interventions

Figure 3 presents the age-specific prevalence of diabetes in 2015 with and without interventions. Compared to the base case (thin blue line), the age-specific prevalences with interventions are considerably lower. Both intervention programs yield to very similar age courses of the prevalence in 2015.

For better comparison the age-specific prevalences in 2015 are shown in Table 1. The age courses of the prevalences in the two different interventions are very similar.

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Without</th>
<th>Pharmacologic</th>
<th>Lifestyle</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>4.0</td>
<td>3.7</td>
<td>3.7</td>
</tr>
<tr>
<td>45</td>
<td>10.1</td>
<td>9.3</td>
<td>9.3</td>
</tr>
<tr>
<td>55</td>
<td>18.1</td>
<td>16.7</td>
<td>16.8</td>
</tr>
<tr>
<td>65</td>
<td>25.8</td>
<td>24.2</td>
<td>24.2</td>
</tr>
<tr>
<td>75</td>
<td>30.8</td>
<td>29.3</td>
<td>29.2</td>
</tr>
<tr>
<td>85</td>
<td>29.5</td>
<td>28.2</td>
<td>27.9</td>
</tr>
</tbody>
</table>

The 20-year survival rates \( S_{20}(a) \) for several ages in the base case and the two interventions are presented in Table 2.
Table 2. 20-year survival rates in 1995 for several ages without (base case) and with interventions

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Without</th>
<th>Pharmacologic</th>
<th>Lifestyle</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>86.88</td>
<td>86.73</td>
<td>87.37</td>
</tr>
<tr>
<td>45</td>
<td>72.63</td>
<td>72.76</td>
<td>73.87</td>
</tr>
<tr>
<td>55</td>
<td>49.14</td>
<td>49.33</td>
<td>50.59</td>
</tr>
<tr>
<td>65</td>
<td>21.36</td>
<td>21.53</td>
<td>23.13</td>
</tr>
</tbody>
</table>

From Table 2 we see that at all ages the 20-year survival is the nearly the same for the base case and the pharmacological intervention. At the age of 65, the base case has the lowest 20-years survival rate (21.4%). It increases slightly to 21.5% in the pharmacological intervention and by additional 1.6% in the lifestyle intervention.

Note that we are referring to the survival in the overall (general) population. The increased survival probability at the age of 65 of only 1.7% in the lifestyle intervention (compared to the base case) may seem low, but it holds for the overall population and not only for those who are taking part in the intervention programs.

4. DISCUSSION

In this article, we have calculated the effect of two hypothetical interventions against diabetes in the male population of the United States. Based on data from epidemiological surveys and reported incidence reductions from a meta-analysis, we compared the resulting age-specific prevalence in 2015 and the 20-years survival rates of the overall population in 1995.

Although both considered interventions are substantially different with respect to the type (pharmacological versus behavioral), it turned out that the resulting age-specific prevalences in 2015 are essentially the same whereas the general mortality differs considerably. Assuming non-overly optimistic participation rates in both interventions, the overall 20-years survival rate for a person aged 65 years in 1995 differs by 1.6% between both interventions.

Hence, if only the prevalence was considered in comparing these two interventions, no difference would have been observed. A policy or decision maker could erroneously draw the conclusion that the interventions are equivalent, although the lifestyle intervention has the advantageous effect of a lowered mortality in the considered population. It is also well-known that physical activity as part of the lifestyle intervention has a variety of further positive effects on the outcomes of other diseases.\(^24\)

In making cross-country comparisons of the burden of diabetes the prevalence is used frequently, e.g., in the Diabetes Atlas.\(^25\) This may be misleading, too. Due to the dependency of the prevalence on both, incidence and mortality, prevalence alone is not a comprehensive measure to assess the public health situation of a country with respect to a disease. For this reason, additional figures are necessary in evaluating the impact of interventions or in comparing different countries. Additional measures may be quality adjusted life-years or costs\(^26\) years of life lost\(^1\) or healthy life years.\(^27\)

This work has a few limitations. Although the effect sizes of the considered interventions stem from a meta-analysis, it remains open if an impact of the calculated magnitude can indeed be achieved. For example, the effects of these intervention can be confounded or modified by other factors (e.g., genetic or environmental). However, the aim of this work was to compare two interventions that were plausible and not too over-optimistic. Thus, we included additional factors modelling the moderate participation rates in the different interventions. Nevertheless, the interventions are hypothetical and serve only as an example. For the same reason, details about enrolling the intervention programs, such as costs and logistics, are not discussed here. Another limitation lies in the currentness of the data. A recent publication has shown that the trend of diabetes incidence in the past years cannot be extrapolated from the years before. Instead, a decrease of the age-adjusted incidence has been observed.\(^28\) The decrease has not been included in the our calculations, because detailed age- and sex-specific incidence rates of the past years were not published yet. A third limitation can be seen in the lack of including statistical uncertainty into our calculations. The estimated effect sizes of the interventions have statistical uncertainties that will propagate into uncertainty of the prevalence estimates. Although the techniques of how this can be achieved are readily available (see for instance\(^27\)), we have chosen not to include them for clarity of our intended message about the comprehensiveness of the prevalence. We just wanted to show that different public health situations induced by different types of interventions may lead to the same age-specific prevalence but different other epidemiological outcomes such as mortality. Thus, we argue that despite its frequent use in public health, prevalence alone is not a comprehensive measure, which may lead to erroneous conclusions.

Estimating the impact of the interventions on the population level was done by a recently developed relation between prevalence, incidence and mortality rates.\(^10\) Similar relations have a tradition ranging back at least to the early 1990ies,\(^16\) but are rarely used in chronic disease epidemiology, public health or decision making. Apart from the demonstration that prevalence has a limited meaningfulness.
in assessing the health situation of a population, we hope to contribute to the expanded use of mathematical models in these and related settings to oppose the enormous burden of non-communicable diseases.

**DESCRIPTION OF THE DATA FOR THE CALCULATION**

This section presents some details about the calculations and underlying data. The hypothetical interventions start in the year 2000. To allow some lead time, we use the prevalence of type 2 diabetes five years before as initial condition for the numerical integration of the PDE (1). The associated age-specific prevalence of type 2 diabetes for \( t = 1995.0 \) is shown as a black solid line in Figure 2. These data were obtained by interpolation of the values reported from NHANES 1988–94 and NHANES 1999–2002.\[^{[16]}\]

In addition to the age-specific prevalence for \( t = 1995.0 \), the base case scenario uses the incidence of type 2 diabetes in men from the NHIS study.\[^{[15]}\] The article by Gregg and colleagues reports the incidence for three periods (1985–89, 1990–99, 2000–11). The relevant periods for this article are the later two periods. We assumed the reported incidences to be in the mid of these periods (1995.0 and 2006.0). Incidence rates between \( t = 1995.0 \) and \( t = 2006.0 \) were obtained by linear interpolation, rates later than \( t = 2006.0 \) were estimated by linear extrapolation. Figure 4 shows the age-specific incidence rates for several years in the base case.

Similar to the incidence rates, the mortality rates \( \mu_0 \) and \( \mu_1 \) were taken from the three NHIS waves.\[^{[15]}\] Only the later two surveys (1990–99 and 2000–11) are relevant for this article. As in the relevant age range (30-85 years), the mortality rates can be approximated reasonably well by linear functions after making a logarithmic transformation, interpolation (for \( t \) between 1995.0 and 2006.0) and extrapolation (for \( t \) between 2006.0 and 2015.0) were applied after logarithmizing.

For the calculation of \( S_{20}(a) \) we use ages \( a \) not older than 65 in 1995, which allows us to stay within the modelled age range up to 85 years. Note that person aged 65 years in 1995 is 85 at the end of the calculated period 2015.

All calculations have been performed with the statistical software R, version 3.2.1 (The R Foundation for Statistical Computing). The associated R script may be obtained on request from the authors.

**ACKNOWLEDGEMENTS**

All authors contributed to conception and design, paper preparation, read, and approved the final paper.

**CONFLICTS OF INTEREST DISCLOSURE**

The authors declare that there is no conflict of interest regarding the publication of this paper.

**REFERENCES**


