Social inequalities in Healthy Life Expectancy

ALTERNATIVE METHODS OF ESTIMATION IN THE ABSENCE OF THE NATIONAL CENSUS

Public Health and Surveillance
Rue Juliette Wytsman 14
1050 Bruxelles – Belgique
www.iph.fgov.be

D/2011/2505/41

Santé Publique et Surveillance | Octobre 2011 | Bruxelles, Belgique
N° de dépôt légal**: D/2011/2505/41**

Authors
Rana Charafeddine* (WIV-ISP)
Sylvie Gadeyne (VUB)
Patrick Deboosere (VUB)
Nicolas Berger (WIV-ISP)
Stefaan Demarest (WIV-ISP)
Herman Van Oyen (WIV-ISP)

*** rue Juliette Wytsman 14, B-1050 Bruxelles, rana.charafeddine@wiv-isp.be**

Le projet est financièrement soutenu par le Service Public Fédéral de Programmation Politique Scientifique (Contrat # AG/00/155)

Table of contents

1. General introduction

1.1 Background

Life expectancy has long been used as an indicator of population health. This indicator represents the average number of years a person can expect to live, if in the future they experience the current age-specific mortality rates in the population. It has often been used to compare population health across countries and across different points of time. This is largely because on one hand reliable and comparable mortality data are available in many countries, and because on the other hand mortality was thought to reasonably describe the state of population (ill-)health. This is a sensible assumption while infectious diseases were the main cause of death. However, now that chronic diseases have replaced, or are progressively replacing, infectious diseases, and the risk of becoming ill is not solely linked to the risk of dying, monitoring the increase in life expectancy is no longer sufficient to infer population health (Sihvonen et al., 2009).

Using morbidity data solely to describe the health of the population is also problematic for two reasons. First, there are problems associated with the availability, reliability and comparability of data on the prevalence of diseases or disabilities. However, this issue is being seriously tackled, especially in the context of the development of European level surveys such as the Health Interview Survey (EU-HIS). The other problem is that morbidity data describe those people who are alive at a certain age. The older the age group concerned the smaller and more select that group will be as mortality increases as age increases.

To overcome the above mentioned limitations, a number of summary measures of population health have been developed to take into account mortality and morbidity simultaneously. Among these measures are the indicators of health expectancies that have became widely used recently. Health expectancies are a measure of population health that combine length and quality of life into a single measure.

1.2 What are health expectancies?

The concept of health expectancies as health indicators was proposed by Sanders (1964) and the first example was published in a report of the US Department of Health Education and Welfare (Sullivan, 1971). This report contained preliminary estimates of "Disability-Free Life Expectancy" calculated using a method devised by Sullivan and applicable to any state of health definition.

Health expectancy reflects the current health of a real population adjusted for mortality levels and independent of the age structure. It is defined as the number of remaining years spent in a health state from a particular age assuming current rates of mortality and morbidity. For example in 2004 the female life expectancy at birth in Belgium was 81.4 years, so a baby girl born in 2004 could expect to live to age 81 years, assuming the conditions of 2004 prevailed over her whole life. By considering not only mortality but also ill-health at particular ages we can divide this remaining number of years into years spent in good and bad health – these are then health expectancies. Health expectancies add a quality dimension to the quantity of life lived (Robine, 2006).

Health expectancy is a generic term referring to an entire class of indictors expressed in terms of life expectancy in a given state of health. Therefore, as there are many dimensions of health, there are many health expectancies. Healthy Life Years (HLY) is the HE based on limitations in daily activities and is therefore a disability-free life expectancy. This indicator is one of the most common health expectancies reported (Robine et al., 2003). Others examples of health expectancies indicators include healthy life expectancy (based on the self-rated health question: ‘How is your health in general?) and life expectancy free of specific diseases, for instance dementia-free life expectancy (Roelands et al., 1994).

Health expectancies have many advantages compared to other morbidity and mortality measures. Health expectancy indicators are considered as important population health outcome measures (Stiefel et al., 2010). They are more intuitive and meaningful measure of health to which policy makers can relate due to the measurement in expected years of life. Also, the indicator is already age standardized so comparison are possible without further adjustment between population with different age distributions and among subgroups that make up populations such as the different genders, socio-economic categories, regions and countries.

1.3 Use of the health expectancies indicators

Health expectancies are becoming a standard summary measure of population health that is used for different purposes at the international and regional levels. For instance, it has been used to study health inequalities, to target resources for health promotion, to evaluate the impact of health policies, and to plan health, social, and fiscal policy (Stiefel et al., 2010).

Health expectancies have also been widely used to explore the future course of mortality and morbidity in developed countries. Three scenarios have been examined in different countries (Doblhammer & Kytir, 2001; Graham et al., 2004). The "compression of morbidity" theory purports that the increase in life expectancy is accompanied by a decrease in unhealthy life years (Fries, 1980). In contrast, the "expansion of morbidity" scenario suggests that the increase in life expectancy is achieved mainly through improvements in medical care and secondary prevention strategies extending the life of people with illness and disability (Gruenberg, 1977). The dynamic equilibrium is an intermediate scenario where an increase in unhealthy life years is offset by a decrease in the severity of illness and disability (Manton, 1982). It is important for the planning of future medical and care services to determine which scenario is unfolding as these three scenarios imply different pressures on health services and systems.

A large and growing number of European Member States estimate health expectancies periodically. Moreover, in 2004 the European Union has selected the health expectancy and more specifically, the disability free life expectancy (DFLE) at birth (also termed Healthy Life Years) and at age 50 years as one of the structural indicators to be monitored annually as a key economic outcome measure for social policies related to retirement age and spending for health and long term care for the aging population (EC 2009). It has to be noted that the health expectancy is the only health indicator among the structural indicators. The importance given by the EU to the health expectancy indicator is evident by the fact that it figures on the short list of the European Community Health Indicators (ECHI), that it is included in the list of sustainable development indicators (DFLE at birth and at age 65) and that it is one of the 18 so-called Laeken indicators for social inclusion, adopted at the European Council in Laeken, Belgium in 2001.

1.4 Social inequalities in health expectancies

The inverse association between health expectancy and socioeconomic status has been widely documented internationally (Bronnum-Hansen & Baadsgaard, 2008; Matthews et al., 2006). Also, social inequalities in HE have been found in Belgium. Bossuyt and colleagues have found that in 1997 those with a low level of education have shorter lives than people with a higher level of education. (Bossuyt et al., 2004) They also have fewer years in good health, and can expect more years in poor health in their shorter lives. For instance, in 1990s, among males aged 25 years the difference in health expectancy between the highest and lowest levels of education can be up to 17 years. This difference is substantially greater than the difference in life expectancy, which is a maximum of 5.23 years. Among females this difference is 11.42 years, while the difference in life expectancy is, at most, 3.22 years (Deboosere et al., 2008; Van Oyen et al., 2011).

These inequalities are not static overtime. Studying the evolution of inequalities in healthy life years from 1997 to 2004, Van Oyen and colleagues (2011) found that the social gap has increased during this period. For instance, at 25 years, the difference in HLY between men in the lowest and highest educational categories was 17 years in 1997 and became 18.58 years in 2004. For women, this difference was 11.42 years in 1997 and became 18.18 years in 2004. This highlights the importance of identifying the appropriate databases and methods to monitor inequalities in such an important indicator.

1.5 Estimation of health expectancies by social position in Belgium

To estimate social inequalities in health expectancy, two different types of data are used. First, information is needed about the prevalence of health status by socioeconomic category. This information is usually extracted from surveys. For instance the Belgian studies mentioned above have used data from the Belgian Health Interview Survey, while the EU uses the Statistics on Income and Living Conditions Survey (EU-SILC). Second, information is needed about mortality rate by SES category. In Belgium, this information has been extracted from the mortality follow up of the census. For instance, the above studies linked the 1991 and the 2001 censuses to the National Register data on mortality and emigration in 1991–1994 and 2001–2004. The record linkage was based on a unique identifier present in both the census and register files. As a consequence, socio-demographic characteristics in the census were unambiguously matched to migration and mortality data in the population register in a follow-up period of three years.

The method described above is based on what is called “linked data”. Another method to estimate HLY by SES is to use “unlinked data”. This entails that mortality data, including the SES of the deceased, are extracted from death certificates while population data, including their SES, are coming from the census or census based population estimates. Such data can be potentially biased due to the lack of comparability of the SES information on the death certificate (reported by a proxy informant usually a relative of the deceased) with that of the census (self reported by the person); this is the so called numerator denominator bias. That is why it has been recognized that the most reliable data for group-specific (e.g. SES groups) estimations are provided by linked record studies as individual data linkage avoids inconsistencies between the numerator and the denominator (Valkonen, 1993).

This makes the census and its mortality follow up an important tool for estimating health expectancies by SES in Belgium. However, as there is no planned census after the 2001 census, other approaches must be identified to estimate HE by SES.

1.6 Objectives of the project

In this context, the overall objective of this project is the development of a methodology to estimate and update healthy life expectancy by socio-economic status (SES). This is an important endeavor as this project will contribute through its recommendations to the establishment of a system to monitor the HLY by SES in Belgium.

The aim of this report is to explore the different possible methods as alternative to the census to be used in Belgium to estimate HLY by SES. As described above, one possible method is the use of cross-sectional, unlinked studies. This method involves the use of mortality rates by SES generated from two different cross-sectional datasets: one providing the number of death by SES and the other the distribution of the population by SES. Chapter 2 describes this approach and its limitations in the Belgian context. The other possible method is the use of linked record studies other than the census. Using this approach, health expectancies by SES can be estimated based on mortality rates generated from the follow up of a number of surveys such as the Health Interview Survey or the EU-SILC that include information on the participants' socioeconomic status, health and functional limitations. Chapter 3 describes the criteria used to select the appropriate surveys to be used in this second approach. Chapter 4 presents the estimations done based on the surveys selected. Chapter 5 examines the approaches used in other European member states to estimate mortality rate by SES. Finally, chapter 6 presents the recommendations for the establishment of a system in Belgium to estimate and update healthy life expectancy by socio-economic status.

This project is funded by the Belgian Science Policy (Contract # AG/00/155), in the context of the AGORA program. This project was initiated in January 2010 and has ended in June 2011. Researchers from two institutions collaborate on this project: the Scientific Institute of Public Health (IPH) that is also the coordinator, and the Vrije Universiteit Brussel (VUB).

2. Sterfte en sociale ongelijkheid. De gelinkte en de niet-gelinkte methode vergeleken

Sylvie Gadeyne, Patrick Deboosere

2.1 Inleiding

De gezonde levensverwachting kan op verschillende manieren berekend worden. Een van de mogelijkheden is de Sullivan-methode waarbij de prevalentie van mensen in goede gezondheid toegepast wordt op de sterftetafel. Deze techniek kan uiteraard gedifferentieerd worden in functie van het onderwijsniveau om de gezonde levensverwachting naar opleiding te berekenen. Een sterftetafel wordt klassiek opgesteld op basis van leeftijdsspecifieke sterftekansen. Deze worden doorgaans niet direct geobserveerd, maar afgeleid van leeftijdsspecifieke sterftecijfers. De teller van deze cijfers bestaat uit het aantal sterfgevallen naar opleidingsniveau, de noemer uit het aantal geleefde persoonsjaren in elke opleidingsgroep.

In België kan de gezonde levensverwachting op basis van verschillende databronnen berekend worden. De prevalentie van mensen in goede gezondheid (of in minder dan goede gezondheid) kan afgeleid worden uit de opeenvolgende gezondheidsenquêtes of uit de census 2001, die beide ook informatie bevatten over het opleidingsniveau van de respondenten.
De andere component van de gezonde levensverwachting, de sterftetafel en de daaruit volgende levensverwachting naar onderwijs, kan op verschillende manieren berekend worden, althans voor Brussel Hoofdstedelijk Gewest (H.G.). De data beschikbaar op Interface Demography laten voor Brussel H.G. toe om sterftecijfers naar onderwijsniveau te berekenen op basis van een ‘gekoppelde’ of ‘gelinkte’ dataset en op basis van een ‘niet-koppelde’ of ‘niet-gelinkte’ dataset. Dit stelt ons in staat om na te gaan hoe groot de fout zou zijn indien gebruik gemaakt zou worden van niet-gelinkte data in plaats van gelinkte data om sterftecijfers te berekenen. Omdat noemers en tellers afkomstig zijn uit twee onafhankelijke, niet-gelinkte databronnen, stijgt het risico op fouten bij de berekening van sterftecijfers naar opleiding.

2.2 Gekoppelde en niet-gekoppelde data

De koppeling van de data van de volkstelling van 1/10/2001 en de registerdata over sterfte en emigratie tijdens de periode 1/10/2001 - 31/12/2004 laat toe om rechtstreeks sterftecijfers naar opleiding te berekenen voor de betrokken periode. Het gaat hier om een deterministische koppeling, d.w.z. met zekerheid dat de gekoppelde informatie betrekking heeft op dezelfde persoon in beide bestanden. Dit is mogelijk dankzij de aanwezigheid van het rijksregisternummer in zowel het censusbestand als het registerbestand. Het registernummer fungeert m.a.w. als een unieke identificatiesleutel op basis waarvan met zekerheid een juiste koppeling doorgevoerd kan worden. Het gekoppelde bestand bevat censusinformatie over de demografische en socio-economische kenmerken van elk individu en daarnaast ook registerinformatie over de datum van overlijden en emigratie. De databank laat toe om een correcte inschatting te maken van de sterftecijfers naar opleiding, gegeven dat sterfte en opleidingsniveau correct zijn opgemeten.

Naast de gelinkte dataset is voor Brussel H.G. ook een dataset beschikbaar die de oorzaak van overlijden aangeeft voor elk sterftegeval dat geregistreerd werd tijdens de follow-up periode 2001-2004 (de certificaten). Tijdens de registratie van de overlijdensoorzaak wordt ook het beroep en de opleiding van de overledene genoteerd. De databank laat dus toe om de tellers van de sterftecijfers – het aantal sterfgevallen naar opleidingsniveau – te berekenen. Deze kunnen dan gedeeld worden door het aantal geleefde persoonsjaren naar opleidingsniveau zoals geregistreerd in de census. De sterftecijfers gebaseerd op een niet-gelinkte dataset kunnen beïnvloed worden door een zogenaamd teller-noemer probleem, aangezien beide niet uit dezelfde databron afkomstig zijn.

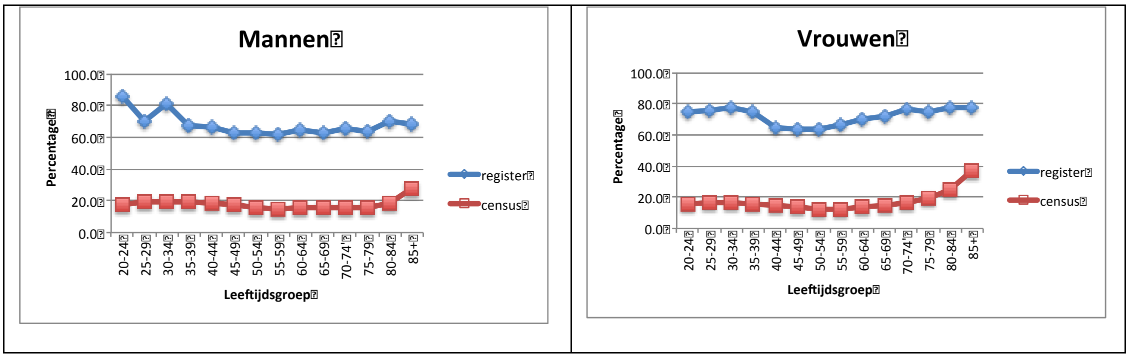
2.3 Sterftecijfers op basis van gelinkte en niet-gelinkte data: grote verschillen?

Om de fout te ramen die voortvloeit uit het gebruik van niet-gelinkte data worden voor Brussel H.G. sterftecijfers berekend op basis van de niet-gelinkte dataset en vergeleken met de cijfers berekend op basis van de gelinkte dataset.

Eerst wordt de kwaliteit van de variabelen in beeld gebracht. De registratie van demografische variabelen in het register is van zeer goede kwaliteit. Er kan dus aangenomen worden dat geslacht, leeftijd en leeftijd van overlijden correct ingeschat worden in het register. De informatie over geboortedatum en geslacht wordt overgenomen op het censusformulier zodat op dit vlak sprake is van een perfecte overeenkomst tussen beide bestanden. De koppeling van registerdata over de datum van overlijden is zoals gezegd gebaseerd op een uniek identificatienummer – het rijksregisternummer – zodat ook hier sprake is van een perfecte overeenkomst.

Een belangrijk verschil tussen de census en de certificaten heeft betrekking op de kwaliteit van de socio-economische variabelen zoals opleiding. Voor de drie jongste leeftijdsgroepen is in de census geen informatie over het onderwijsniveau opgenomen. Bovendien kan het uiteindelijke onderwijsniveau van een individu ook maar vanaf de leeftijd van 20-24 jaar bepaald worden. De analyses worden bijgevolg beperkt tot de bevolking van 20-24 jaar en ouder. In deze populatie bedraagt het percentage ontbrekende waarden voor de opleidingsvariabele in de census 17,1%, met een iets lager percentage bij vrouwen dan bij mannen (16,6% tegenover 17,6%). In de certificaten bedragen deze percentages niet minder dan 71,1% in de totale bevolking, 75,5% bij vrouwen en 65,9% bij mannen. Dit betekent dat bij drie kwart van de vrouwen het onderwijsniveau niet gekend is in de certificaten, wat ongetwijfeld repercussies zal hebben voor de kwaliteit van de tellers van de sterftecijfers (aantal sterfgevallen naar opleiding) berekend op basis van de registerdata.
In figuur 1 wordt het percentage ontbrekende waarden naar leeftijd gegeven. In de censusdata kan zowel bij mannen als bij vrouwen een trend van een stijgend percentage naar leeftijd geobserveerd worden. Voor de certificaten differentieert het patroon meer in functie van geslacht. Bij mannen wordt een hoger percentage waargenomen voor de jongere generaties, vervolgens een stabilisatie en een lichte toename vanaf 75-79 jaar. Bij vrouwen ligt het percentage ook het hoogst in de jongste leeftijdsgroepen, maar treedt na een eerste daling een geleidelijke stijging in van het aantal ontbrekende waarden tot ongeveer 80% in de oudste leeftijdsgroep.

Figuur 2.1: Percentage ontbrekende waarden voor het onderwijsniveau naar leeftijd en geslacht: certificaten- en censusdata vergeleken (eigen berekeningen)



Het spreekt voor zich dat dergelijke hoge niveaus van ontbrekende waarden een belangrijke impact zullen hebben op de sterftecijfers. Bij de niet-gelinkte methode wordt de teller van het sterftecijfer berekend op basis van deze certificatendata, de noemer op basis van de census. Bij de gelinkte methode worden zowel teller als noemer berekend op basis van de censusdata. In annex worden de tabellen gegeven met de sterftecijfers naar opleidingsniveau berekend volgens beide methodes. Deze sterftecijfers vormen de basis voor de aanmaak van sterftetafels. In tabel 1 wordt de levensverwachting op 20-24 jaar gegeven naar onderwijsniveau volgens beide methodes en het verschil tussen beide. Verschillende conclusies dringen zich op.

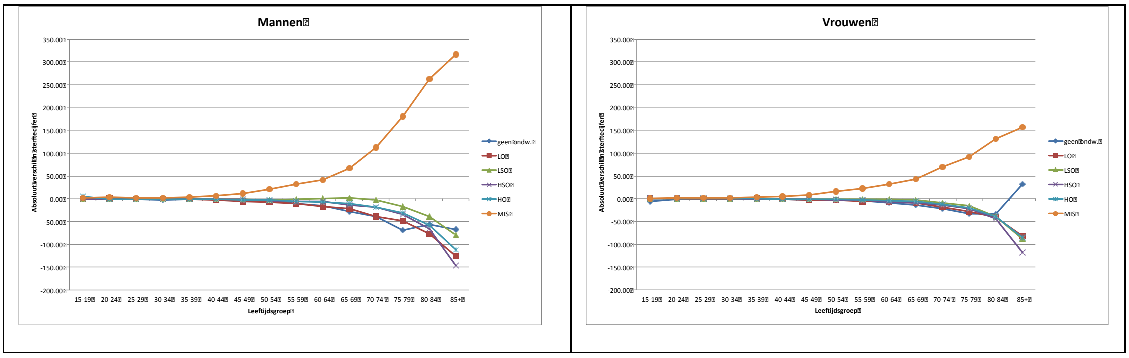
De sterfte wordt duidelijk onderschat aan de hand van de niet-gelinkte methode, met een aanzienlijk hogere levensverwachting op 20 jaar als gevolg. De afwijking lijkt het grootst voor de hogere onderwijsniveaus, hoewel de relatie niet lineair verloopt en afhankelijk is van geslacht. Bij vrouwen wordt het kleinste verschil genoteerd voor vrouwen zonder diploma, waar de levensverwachting op 20 jaar 'maar' 8 jaar hoger is bij de niet-gelinkte methode. Het grootste verschil wordt geobserveerd voor vrouwen met een diploma hoger secundair onderwijs, met een verschil van niet minder dan 38 jaar! Bij mannen zijn de verschillen lager en wordt de kleinste afwijking geobserveerd bij de lager secundair opgeleiden van bijna 2,5 jaar. De onderschatting van de mortaliteit bij de verschillende onderwijsgroepen wordt volledig verklaard door de veel hogere sterfte van respondenten met een ontbrekende waarde voor onderwijsniveau in de certificaten. Zij hebben een levensverwachting die 32 lager ligt bij de mannen en 21 jaar lager bij de vrouwen in vergelijking met de categorie met een ontbrekende waarde voor het onderwijsniveau in de census. Een deel van de onderschatting van de sterfte vloeit dus voort uit het ontbreken van informatie over het onderwijsniveau van de overleden personen.

Tabel 2.1: Levensverwachting op 20 jaar naar onderwijsniveau en geslacht volgens de niet-gelinkte en de gelinkte methode (eigen berekeningen)

	Totaal	Mannen	Vrouwen
Geen ondw.			
niet-gelinkt	68.9	69.6	68.7
gelinkt	58.3	55.4	60.7
verschil	10.5	14.1	8.0
LO			
niet-gelinkt	71.5	67.3	73.6
gelinkt	56.7	53.0	59.9
verschil	14.8	14.4	13.7
LSO			
niet-gelinkt	65.4	56.3	74.6
gelinkt	58.0	53.9	61.4
verschil	7.3	2.4	13.2
HSO			
niet-gelinkt	84.4	69.2	99.6
gelinkt	59.1	55.6	61.9
verschil	25.3	13.6	37.7
HO			
niet-gelinkt	80.2	74.1	92.4
gelinkt	61.3	59.1	64.6
verschil	18.9	15.1	27.8
Ontbrekend			
niet-gelinkt	22.5	17.0	34.0
gelinkt	52.3	49.4	55.3
verschil	-29.8	-32.4	-21.3

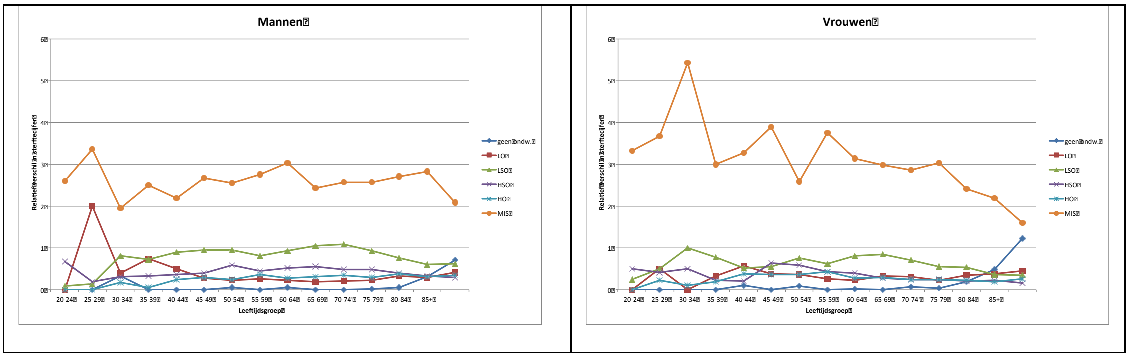
Om het leeftijdspatroon van de verschillen tussen beide methodes in kaart te brengen, worden verschillen berekend tussen de series leeftijdsspecifieke sterftecijfers. Figuur 2 geeft de absolute verschillen weer voor mannen en vrouwen afzonderlijk. Voor de volledigheid worden in annex de tabellen gegeven met de sterftecijfers berekend volgens beide manieren.

Figuur 2.2: Absolute verschillen tussen de leeftijdsspecifieke sterftecijfers berekend op basis van de gelinkte data en de leeftijdsspecifieke sterftecijfers berekend op basis van de niet-gelinkte data, mannen en vrouwen (eigen berekeningen)



Op jonge leeftijd blijven de absolute verschillen betrekkelijk laag, wat enigszins logisch is aangezien sterfte laag is op jonge leeftijd. Naarmate de leeftijd stijgt, neemt de afwijking tussen beide reeksen toe en vooral in de oudste leeftijdsgroepen worden grote verschillen genoteerd. De figuur illustreert duidelijk de onderschatting van de mortaliteit in de verschillende onderwijscategorieën, zowel bij mannen als bij vrouwen. De patronen bij mannen en vrouwen zijn vergelijkbaar. Voor de onderwijsgroepen zijn ze veelal iets groter bij vrouwen, maar voor de groep met ontbrekende informatie is de afwijking groter bij mannen vooral voor de oudste leeftijdsgroepen. De afwijking tussen beide series is voor deze groep overigens al op relatief jonge leeftijd zichtbaar, zowel bij mannen als bij vrouwen.

Figuur 2.3: Relatieve verschillen tussen de leeftijdsspecifieke sterftecijfers berekend op basis van de gelinkte data en de leeftijdsspecifieke sterftecijfers berekend op basis van de niet-gelinkte data, mannen en vrouwen (eigen berekeningen)

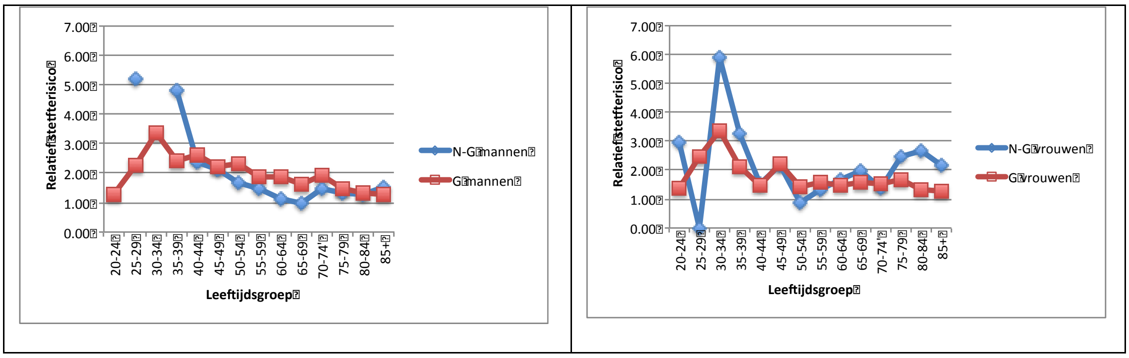


Dit blijkt ook uit figuur 3 die de relatieve verschillen tussen de sterftecijfers berekend volgens beide methodes weergeeft. Van jongs af aan ligt het sterftecijfer van individuen met een ontbrekende waarde voor het opleidingsniveau veel hoger bij de niet-gelinkte data: tot 3 keer hoger bij mannen en tot meer dan 5 keer hoger bij vrouwen. De sterfte van personen met een gekend onderwijsniveau is zoals gezegd lager in het niet-gelinkte bestand en dit resulteert in relatieve verschillen kleiner dan de eenheidswaarde. De laagste relatieve verschillen worden opgetekend voor het lager secundair onderwijs, zowel bij mannen als bij vrouwen. Voor de individuen met een diploma lager onderwijs liggen de cijfers berekend op basis van het gelinkte bestand 1,9 keer hoger op de leeftijd van 30-34 jaar tot 4,5 keer hoger op de leeftijd van 75-79 jaar (totale bevolking). Voor het hoger secundair wordt een groot verschil op de jongste leeftijden genoteerd (7,75 hoger op 20-24 jaar), maar op de andere leeftijden varieert het verschil van 1,1 tot 1,5. Het hoger secundair en hoger onderwijs toont opnieuw waarden die veelal drie tot vier keer hoger liggen in het gelinkte bestand dan in het niet-gelinkte bestand.

2.4 Relatieve sterfteverschillen naar opleiding op basis van gelinkte en niet-gelinkte data

De vraag stelt zich ook welke fout er optreedt bij de berekening van de sociale ongelijkheid in sterfte wanneer gebruik gemaakt wordt van niet-gelinkte data. Figuur 4 geeft de ratio tussen het sterftecijfer van de lager opgeleiden en dat van de hoger opgeleiden volgens de gelinkte (G) en de niet-gelinkte (N-G) methode bij mannen en vrouwen.

Figuur 2.4: Relatieve sterfteverschillen naar opleidingsniveau (LO/HO) berekend op basis van de gelinkte data en op basis van de niet-gelinkte data, mannen en vrouwen



Bij mannen is de curve berekend op basis van de niet-gelinkte data ‘onderbroken’. Op 20-24 jaar kan het relatieve verschil tussen de lager en de hoger opgeleiden niet berekend worden omdat de mortaliteit van de hoger opgeleiden gelijk is aan nul. Op 30-34 jaar bedraagt het relatieve verschil niet minder dan 45,8, wat het gevolg is van de extreem lage sterfte van de hoger opgeleide mannen (0,02 pro mille).

Abstractie gemaakt van deze ‘outliers’, laat figuur 4 een duidelijke trend zien voor mannen. Op jonge leeftijd wordt de sociale ongelijkheid tussen de lager en de hoger opgeleiden overschat met een factor van ongeveer 2,5 (relatief verschil van ongeveer 2 tegenover ongeveer 5) wanneer gebruik gemaakt wordt van niet-gelinkte data. Vanaf de leeftijd van 45-49 jaar is niet langer sprake van een overschatting, maar van een onderschatting tot de leeftijd van 70-74 jaar. Vanaf de leeftijd van 75-79 jaar komen de resultaten van beide methodes relatief goed overeen. De vergelijking van het sterfteverschil tussen de andere onderwijsniveaus en het hoger onderwijs, geeft aan dat de verschillen doorgaans overschat worden bij de niet-gelinkte methode. Voor het lager secundair is sprake van een overschatting over gans de lijn. De verschillen tussen beide methodes zijn het grootst op jonge leeftijd (4,7 keer grotere verschillen bij de niet-gelinkte methode), maar blijven ook op oudere leeftijd aanzienlijk (factor 3 op 70-74 jaar en vervolgens factor 2). Het verschil tussen beide methodes verloopt iets grilliger voor de vergelijking hoger secundair en hoger onderwijs, hoewel de tendens van een overschatting op jonge leeftijd en een onderschatting op oudere leeftijd ook wel aanwezig is (net zoals voor het lager onderwijs).

Voor vrouwen toont figuur 4 een overschatting van de relatieve sterfteverschillen naar onderwijsniveau in de jongste leeftijdsgroepen, een relatief goede overeenkomst in de groepen 40-44 tot en met 70-74 jaar en opnieuw een overschatting van de verschillen in de oudste leeftijdsgroepen. Voor het lager secundair is net zoals bij mannen sprake van een systematische overschatting van de verschillen bij de niet-gelinkte methode, veelal in de grootteorde van factor 1,5 tot 2,5. Voor het hoger secundair onderwijs is vanaf de leeftijd van 55-59 jaar sprake van een relatief goede overeenkomst tussen beide reeksen, voor die leeftijd zien we een overschatting van de verschillen (veelal rond factor 1,5).

2.5 Conclusie

Het is duidelijk dat het werken met niet-gelinkte data een leidt tot een aanzienlijke afwijking bij de berekening van de sterftecijfers en afgeleide indicatoren, zoals sterftekansen en levensverwachting. Zowel bij vrouwen als bij mannen leidt het gebruik van niet-gelinkte data tot een belangrijke onderschatting van de mortaliteit in de verschillende onderwijsgroepen. De fout blijkt voor een deel gerelateerd aan het groot aantal ontbrekende waarden voor de onderwijsvariabele in het certificatenbestand (op basis van dewelke de tellers – het aantal sterfgevallen naar onderwijsniveau – berekend worden) in vergelijking met het censusbestand (op basis van dewelke de noemers – het aantal geleefde persoonsjaren naar onderwijsniveau berekend worden). De groep met een ontbrekende waarde heeft duidelijk een veel groter sterftecijfer bij de niet-gelinkte methode dan bij de gelinkte methode.
Op vlak van sociale ongelijkheid in sterfte tussen de opleidingsniveaus (relatieve verschillen), is duidelijk dat de niet-gelinkte methode doorgaans grotere verschillen suggereert dan de gelinkte methode.

Deze conclusies hebben een aantal belangrijke implicaties. Indien sterfte naar onderwijsniveau en sociale ongelijkheid bestudeerd dienen te worden aan de hand van niet-gelinkte data, dan moet het onderwijsniveau op een correcte en exhaustieve manier geregistreerd worden in de certificaten. In elk geval kunnen sterfte en sterfteverschillen niet correct ingeschat worden indien het onderwijsniveau voor bijna drie vierde van de bevolking niet gekend is. Eventueel zouden respondenten waarvoor het onderwijsniveau onbekend is in de certificaten herverdeeld kunnen worden volgens de frequentieverdeling van onderwijs zoals geobserveerd in de census, of het onderwijsniveau zou geïmputeerd kunnen worden op basis van gekende informatie (zoals leeftijd of beroep). Deze werkwijzen bieden natuurlijk geen enkele zekerheid omtrent de juistheid van het onderwijsniveau dat op deze manier aan de respondenten wordt toegekend.

Zelfs indien het onderwijsniveau volledig correct geregistreerd zou worden in de certificaten dan nog kunnen fouten optreden bij het gebruik van twee verschillende databronnen voor de berekening van de teller en de noemer van de sterftecijfers naar onderwijsniveau (verschillende registratie van onderwijs, technische problemen, registratiefouten in sterfte). De beste optie is dat beide datasets – de registratie van algemene sterfte in het rijksregister of van de doodsoorzaak in de certificaten die de tellers voorziet en de census die de noemers voorziet – gekoppeld kunnen worden. Via een uniek identificatienummer – zoals het rijksregisternummer – aanwezig in beide bestanden of via de aanmaak van een identificatiesleutel op basis van een aantal variabelen kan met relatief veel zekerheid aangenomen worden dat de gekoppelde informatie betrekking heeft op dezelfde persoon zodat ook de berekening van de sterftecijfers naar onderwijsniveau relatief correct kan verlopen. De census van 2001 was de laatste ‘klassieke census’ die in België is doorgegaan. Vanaf 2011 wordt deze vervangen door een administratieve census. In deze is de registratie van onderwijs nog niet voor iedereen verzekerd. Dus ook hier zal de opvolging van sociale ongelijkheid in sterfte niet volledig of exhaustief kunnen verlopen.

Annex

Totale bevolking

	geen onderwijs		**LO**		**LSO**		**HSO**		**HO**		**Ontbrekende waarde**	
	niet-gelinkt	gelinkt	niet-gelinkt	gelinkt	niet-gelinkt	gelinkt	niet-gelinkt	gelinkt	niet-gelinkt	gelinkt	niet-gelinkt	gelinkt
20-24	0.00	0.62	0.80	0.64	0.16	0.84	0.11	0.44	0.06	0.48	3.46	1.00
25-29	0.16	0.63	0.21	0.94	0.64	0.76	0.24	0.61	0.06	0.40	2.76	1.11
30-34	0.00	1.46	0.70	1.32	0.47	0.62	0.20	0.68	0.05	0.39	3.91	1.48
35-39	0.09	1.34	0.64	1.21	1.15	1.54	0.41	1.34	0.15	0.52	5.35	2.12
40-44	0.00	1.70	0.84	2.78	2.02	2.59	0.95	1.94	0.43	1.30	9.41	3.12
45-49	0.21	2.97	1.39	5.19	3.93	4.45	1.87	3.18	0.65	2.31	16.92	6.61
50-54	0.00	3.82	1.81	6.86	5.31	7.16	2.38	5.39	1.44	3.60	27.64	9.02
55-59	0.30	7.44	2.23	9.85	7.84	8.87	3.84	8.16	1.61	5.66	40.75	13.26
60-64	0.00	12.20	3.35	14.08	11.32	11.62	4.75	11.31	2.60	8.39	59.52	22.61
65-69	0.63	20.94	4.63	18.95	16.28	17.63	6.50	17.31	4.08	12.88	86.11	32.03
70-74'	0.86	30.00	7.37	33.23	20.57	27.23	9.20	24.87	5.97	21.20	134.77	48.63
75-79	6.29	51.83	17.44	51.62	28.55	43.96	11.99	38.28	12.67	38.06	196.37	77.53
80-84	30.35	70.64	27.19	77.60	32.60	71.69	18.82	69.92	17.96	67.52	284.31	118.70
85+	175.31	162.45	70.03	159.21	64.21	150.78	31.83	156.63	44.86	143.18	449.38	264.78

Mannen

	geen onderwijs		**LO**		**LSO**		**HSO**		**HO**		**Ontbrekende waarde**	
	niet-gelinkt	gelinkt	niet-gelinkt	gelinkt	niet-gelinkt	gelinkt	niet-gelinkt	gelinkt	niet-gelinkt	gelinkt	niet-gelinkt	gelinkt
20-24	0.00	0.80	1.31	0.66	0.18	1.29	0.12	0.62	0.00	0.52	4.70	1.40
25-29	0.35	1.05	0.44	1.09	1.01	1.24	0.22	0.70	0.08	0.48	3.49	1.79
30-34	0.00	2.04	0.94	1.25	0.54	0.73	0.32	0.98	0.02	0.37	4.83	1.93
35-39	0.00	0.97	0.82	1.64	1.58	1.77	0.59	1.61	0.17	0.68	5.84	2.67
40-44	0.00	1.53	1.11	4.04	2.79	2.94	0.97	2.46	0.47	1.55	10.94	4.09
45-49	0.18	3.38	1.53	6.90	5.75	6.08	2.37	4.02	0.73	3.08	19.58	7.68
50-54	0.00	4.78	2.35	8.96	7.75	9.53	3.40	7.62	1.42	3.89	32.18	11.69
55-59	0.53	11.02	2.91	13.11	11.69	12.59	5.82	11.13	1.98	6.93	48.53	16.01
60-64	0.00	14.82	3.76	19.90	17.72	16.72	7.88	14.10	3.32	10.50	70.37	28.89
65-69	0.30	28.47	5.58	27.60	27.77	25.68	11.87	24.75	5.83	16.83	109.55	42.53
70-74'	1.13	40.18	10.85	48.94	38.79	41.82	16.88	35.29	7.46	25.28	184.77	72.09
75-79	3.31	72.43	24.51	72.96	53.00	68.99	22.83	56.36	18.57	49.01	286.81	106.22
80-84	25.89	82.48	32.55	109.78	59.11	98.20	30.90	95.97	25.59	83.80	405.35	143.18
85+	161.15	228.61	91.12	216.41	128.65	207.71	59.72	206.02	59.74	170.98	606.77	290.40

Vrouwen

	geen onderwijs		**LO**		**LSO**		**HSO**		**HO**		**Ontbrekende waarde**	
	niet-gelinkt	gelinkt	niet-gelinkt	gelinkt	niet-gelinkt	gelinkt	niet-gelinkt	gelinkt	niet-gelinkt	gelinkt	niet-gelinkt	gelinkt
20-24	0.00	0.51	0.32	0.63	0.144	0.29	0.112	0.27	0.105	0.46	2.18	0.59
25-29	0.00	0.29	0.00	0.81	0.196	0.20	0.249	0.50	0.035	0.33	1.95	0.36
30-34	0.00	0.88	0.46	1.39	0.392	0.50	0.077	0.34	0.078	0.41	2.78	0.93
35-39	0.17	1.66	0.46	0.80	0.661	1.27	0.218	1.05	0.139	0.37	4.73	1.44
40-44	0.00	1.82	0.59	1.56	1.213	2.22	0.922	1.46	0.389	1.06	7.57	1.94
45-49	0.23	2.70	1.25	3.47	2.218	2.90	1.444	2.48	0.577	1.57	13.71	5.32
50-54	0.00	3.09	1.26	4.78	3.169	5.09	1.568	3.62	1.456	3.31	22.46	5.97
55-59	0.13	4.77	1.57	6.75	4.700	5.83	2.338	5.91	1.194	4.26	32.48	10.34
60-64	0.00	9.79	2.98	8.93	6.668	7.92	2.572	9.38	1.785	5.98	48.81	16.42
65-69	0.92	14.30	3.95	12.81	8.794	12.39	2.982	12.43	1.981	8.12	67.03	23.49
70-74'	0.66	22.57	5.38	24.25	10.837	19.44	4.563	18.56	3.943	15.64	104.55	34.45
75-79	7.98	40.15	14.34	42.27	16.981	32.11	5.707	27.79	5.843	25.41	156.02	64.73
80-84	32.15	65.86	25.30	66.22	21.655	60.73	12.421	56.13	9.445	49.32	241.43	110.02
85+	178.77	146.31	65.74	147.58	47.030	135.60	22.692	140.46	30.144	115.69	415.69	259.30

3. Inventory and evaluation of surveys as data sources to estimate HLE by SES

3.1 Evaluation criteria

To use surveys as source of data to estimate HLE by SES, it is necessary for the survey to have a number of essential characteristics. In the context of this study, these characteristics are: the health indicators included in the survey, the SES indicators included in the survey, the mortality follow up of the survey, and the design criteria of the survey.

3.1.1 Health indicators

To estimate HLE, it is important for the survey to include a number of health indicators. Examples of such indicators may include general self perceived health or reported chronic diseases or disabilities. As HLE are used to compare health of a particular country over time or across countries, it is important to use the same definitions of health states over time and across countries. For example, differences between health expectancies calculated for different countries can often be explained by differences in the measurement instruments used to collect the prevalence data (Boshuizen & van de Water, 1994).

In order to allow comparable calculations of health expectancies across Europe, the Minimum European Health Module (MEHM) was designed. The MEHM has been implemented in several national and European surveys such as EU-SILC, SHARE, and EHIS (Robine et al, 2003). The objective of the MEHM is to obtain, with a short instrument, information comparable across Europe on three health domains: self perceived health, chronic long standing conditions, and long-term activity limitations (Cox et al., 2009).

Self perceived health (SPH) has become a conventional method for measuring health in populations in the last two decades. The utility of this measure derives from a number of factors. First, it is easily administered as it consists of asking individuals to rate their own health on a scale ranging from poor to excellent. Second, regardless of the semantic variations in the questions, this measure has been shown to be strongly and consistently predictive of mortality and of functional limitations. For instance, in a review of 27 studies comparing SRH and mortality, the authors concluded that in all but four of the 27 studies, general SRH reliably predicted survival in populations even after accounting for known risk factors (Idler

& Benyamini, 1997). More recently, a meta-analysis between SRH and mortality was conducted (DeSalvo et al., 2006). The authors found that persons with poor SRH had a 2-fold higher mortality risk compared with persons with excellent SRH. Third, this measure provides a succinct way of addressing health from a broad perspective rather than assessing mortality or the absence of disease (Krause and Jay, 1994). It is suggested that this measure includes information about: diagnosed diseases and their severity as well as undiagnosed or early stage illness that might be missed by a physician, personal observation about functional status and performance in everyday life, sensation and perceptions about one's own body and mind (e.g. pain, tiredness), and individual understanding of their social and family history (Idler & Benyamini, 1997). The self perceived health item as part of the MEHM is based on the existing recommendations of the WHO.

The second item is a measure of the prevalence of chronic conditions. Reported longstanding diseases or conditions affect the health-related quality of life and are one of the major causes of utilisation of health services.

The third item is the global activity limitation indicator (GALI). Disability is a multifaceted phenomenon that can occur in a number of role and activities such as working, bathing, shopping, and socializing. To reflect this diversity, health surveys contain detailed items about limitations in specific activities. These multiple items are very informative but they also frustrate professionals and policy makers who often need brief summaries of population functioning and health (Verbrugge, 1997). Just as the SPH is one global generic health indicator, there has been a recent effort to create a single indicator of global disability: the global activity limitation indicator (GALI). It is an indicator that is able to identify subjects, in both general and/or specific populations, who perceive themselves as having long-standing, health related restrictions or limitations in their usual activities. In a French survey, it was found that the GALI is highly predictive of functional problems and that it captures more systematically severe levels of disability as it is most sensitive to personal care activities restrictions and less sensitive to functional limitations that are not associated with activity restrictions (Cambois et al., 2007). It is important to highlight that the healthy life years (HLY) item, which is a structural indicator of the European community is calculated based on the GALI from SILC data.

Box 1: The Minimum European Health Module

I would like now to talk about your health:
1. Self perceived health: How is your health in general?
1. Very Good
2. Good
3. Fair
4. Bad
5. Very Bad

2. Chronic conditions: Do you have any long standing illness or health problem?
1. Yes
2. No

3. Global activity limitation indicator: For at least the past 6 months, to what extent have you been limited because of a health problem in activities people usually do?
1. Severely limited
2. Limited but not severely
3. Not limited at all

3.1.2 SES indicators

The availability of a range of indicators reflecting individual or household social status is essential to monitor social inequalities in HLE using surveys. The position of an individual within the social hierarchy may be determined by the dimensions of occupation, income and education. Each dimension relates to a specific aspect of social stratification, and therefore it is preferable to use all three instead of only one (Kunst & Mackenbach, 1994). Educational level differentiates people in terms of access to information and aptitude to benefit from new technologies. Income creates differences in access to scarce materials. Occupational status includes both these aspects and adds to them benefits derived from the exercise of specific jobs such as prestige, power and social and technical skills.

Among those indicators, education has a number of advantages. First, it applies to the adult population regardless of labour market position. Second, it affects potential earnings and access to material resources that influence health, and therefore shares some of the health effects with other indicators (occupation, income and wealth) and is strongly correlated with them (Galobardes et al., 2006). Third, as formal education is normally completed in young adulthood and partly reflects the characteristics of the family and community of origin, it is an indicator that measures early life socioeconomic position and that remains relatively stable over the life course from early adulthood onwards (Davey Smith et al., 1998).

3.1.3 Mortality follow-up

To study HLE, we need mortality data in addition to health data. One way to secure such data is to request information on the mortality follow up of the survey participants on the individual level. The duration of such mortality follow up depends on the size of the survey sample; the smaller the sample the longest the follow up period must be. In any case, the follow-up period must not be less than 2 to 3 years.

Obtaining individual mortality follow up data is a resource and time consuming process (see Appendix 1). Yet, the dataset obtained is very rich as it links health and socio-economic information with the vital status of each individual participating in the survey.

3.1.4 Representative sample

The survey sample should be representative of the Belgian population. An important tool to have a representative sample is the frame that refers to the list of units (eg, persons, households, etc) in the survey population. The selection of the sample is directly based on this list, and therefore the frame determines how well a target population is covered. In Belgium, the frame for national surveys is often the National Register. The National Register is the best and most complete population list in Belgium, yet it excludes some groups as the homeless who have no official address and those living illegally in the country.

As we are interested in deriving estimates for the general population, we will exclude surveys that are concerned with a specific subgroup of the population. For instance, the Survey of Health, Ageing and Retirement in Europe (SHARE) that includes a wide range of health and SES indicators will not be considered in this project as it pertains to people 50 years and older.

3.1.5 Accounting for design aspects

A main interest of surveys is the estimation of population parameters using the sample, often selected by complex schemes such as stratified multi-stage cluster sampling. Statistical methods for estimating population parameters and their associated variances are based on assumptions about the characteristics of the underlying distribution of the observations. Among these are that the observations were selected independently and have the same probability of being selected (Tibaldi et al., 2003). However surveys often violate these assumptions. Procedures need to be developed to account for complex sampling design in order to have correct estimations and variances. Therefore, surveys to be used need to include the sample weights and provide any other directives to adjust for the study design.

3.1.6 Sample size

Determining an appropriate sample size is an important task in study design. Sample size decision is primarily based on the desired analysis to be undertaken as the sample size influences the margin of error or degree of precision in estimation. Also, the size of the sample is based on the available budget and on efficiency that is about obtaining the same amount of information with as small a sample as possible. The surveys to be included must have a sample with sufficient size to allow estimation with an acceptable degree of precision.

3.1.7 Response rate

The quality and reliability of survey data can be significantly affected by the degree of response to a survey. A low participation rate is of concern as it may bias the results. That is why, it is important to compare respondents and non respondents on the variable of interest. Previous studies have shown that people in higher SES categories are more likely to participate in health surveys than those in lower SES categories (Goldberg et al., 2001; Purdie et al., 2002). Still, if the difference in participation rate between socio-economic groups is unrelated to the health outcome, then participation will not affect results in terms of socio-economic inequalities (Galea & Tracy, 2007; Lorant et al., 2007). A study conducted by Lorant and colleagues (2007) investigated if the SES difference in non-participation is related to health outcome by comparing the results of the 2001 Belgian census with the 2001 Belgian Health Interview Survey. They found that people in lower SES categories were less likely to participate in the survey, when they had a poor health status compared to people in higher educational categories. The authors hypothesized that this phenomenon may be related partly to an issue of privacy. People with low SES status have been found to engage in riskier health behaviors such as having a more sedentary lifestyle or being a smoker. For fear of stigma, these individuals may decline to participate in health surveys. If this is the case then health inequalities may be underestimated in surveys.

3.1.8 Periodicity

An important issue to consider when looking for a database to produce regular indicators is the periodicity of the database. For instance, the Labour Force Survey is a large survey with an interesting range of SES indicators. In 2002, the Labour Force Survey included a module on health. However, it would be problematic to recommend using this survey for regular HLE indicators as health module are not regular in the survey.

3.2 Evaluation of surveys in relation to theoretical criteria

As proposed in Section 3.1, we will compare the surveys based on a number of criteria. These are: availability of health indicators, availability of SES indicators, availability of an appropriate mortality follow up period, representative sample, survey design aspects accounted for, sample size, response rate, periodicity. Initially, we considered 4 surveys to be potentially interesting to estimate HLE by SES. These are: the Health Interview Survey (HIS), the Statistics on Income and Living Condition (SILC), Survey of Health, Ageing and Retirement in Europe (SHARE), and the Labor Force Survey (LFS). However, as SHARE pertains only to those 50 years and older and the LFS include health question only periodically, we have finally chosen to explore further only two national surveys: HIS and SILC.

The Health Interview Survey (HIS) provides information on the health status of the total Belgian population by means of interviews in a representative sample of the population. The specific aims of the HIS are: 1) identification of health problems, 2) description of the health status and health needs of the population, 3) estimation of prevalence and distribution of health indicators, 4) analysis of social (in)equality in health and access to the health services, 5) study of health consumption and its determinants, 6) study of possible trends in the health status of the population.

The SILC aims at studying poverty, social exclusion and living conditions on the basis of indicators that can be compared at the European level. The objective of this survey is to provide a framework for the production of updated data concerning the evolution of income and living conditions for Belgium. It is collecting on an annual basis timely and comparable multidimensional micro-data on income, housing condition, labour, education and health.

Health indicators

In the HIS, the individuals are questioned about a wide variety of health domains including general health perceptions, morbidity and functional status, use of health services, lifestyle and socio-economic factors. Also the questions of the MEHM are included in the survey.

Although the SILC studies essentially poverty, social exclusion and living conditions, it includes a number of health questions. These are the three questions of the Minimum European Health Module (MEHM). In both surveys there is an effort to provide indicators and concepts that are agreed upon on the European level and are therefore comparable with European level data. For the current study, we are interested to use the Global Activity Limitation Indicator (GALI). This will allow the estimation of the disability free life expectancy or Healthy Life Years.

SES indicators

The SILC collects a wide range of timely and comparable multidimensional data on income, housing condition, labour, and education. The choice of SES indicators is more restricted in the HIS, still this survey includes detailed SES information that is appropriate in this context. Especially, we are interested in using the educational level to reflect socioeconomic status and such information is widely available in both surveys.

Mortality follow-up

The process of obtaining mortality follow up information for both surveys is regulated by the Privacy Commission. A request was sent to the Commission and we were able to obtain the data. For the HIS, we have mortality follow up information for the waves of the years 1997 and 2001 until 31, December 2007. For the SILC, we obtained follow up information for the SILC 2004 until 31/12/2009. In the case of the SILC, the request process is described in Annex 1. For the His, the request was submitted 7 years ago, and meanwhile the process of submission has been modified. Therefore, the HIS follow up process will not be described here.

Representative sample

In both surveys the target population are all people with a residency in Belgium; that is all those registered in the National Register. There is no restriction on nationality or age. However the questions of interest in this study have been asked for those 15 years and older in the HIS and for those 16 years and older in the SILC. The sampling frame is the National Register, which allows identifying the administrative reference person of a household, the first study unit. The second study unit, the individual is identified at the level of the household. Data from the HIS are representative for the three regions, however this is not the case of the SILC where data are only representative of the Belgian population.

Design aspects accounted for

The participants of the HIS are selected from the National Register through a stratified multistage clustered sampling. Correct estimates can be obtained by re-weighting the data and accounting for clustering and stratification. Procedures have been developed to accurately estimate population parameters using HIS data (Tibaldi et al., 2003). The SILC has a stratified multistaged clustered sampling. To draw correct estimations we used the sample weight provided.

Sample size

For the HIS, the sample size is about 10.000 effective interviews per cycle, divided between the regions: 3500 in the Flemish community, 3500 for the Walloon Region and 3000 in the Brussels Region. For the SILC, the sample size is approximately 6000 households (about 11.000 individuals). So, the sample size is comparable in the examined surveys.

Response rate

For the HIS, the response rate is approximately 60%, and for the SILC it is about 64%.

Periodicity

The HIS takes place every 4-5 years. Already four waves have been completed: 1997, 2001, 2004, and 2008. The SILC is an annual survey.

3.3 Conclusion

The HIS and the SILC have the required characteristics to correctly estimate social inequalities in healthy life years. Both surveys have the required variables and are of high quality concerning the study design. The advantage of the HIS is that the data are representative on the regional level, while this is not the case of the SILC. The SILC has the main advantage of being collected annually while the HIS is undertaken every 4 to 5 years.

4. Estimation of Healthy Life Years by educational level

4.1 Methods to estimate Healthy Life Years

In the last thirty years, there has been a significant increase in the number of Health Expectancies' calculations (Stiefel et al., 2010). Most of these calculations have been undertaken using the Sullivan method. This method starts from the mortality and survival life table that is based on age group specific mortality data for the population and time period concerned. The principle in calculating Sullivan health expectancies is that for each age range in the life table, total life years are divided between healthy and unhealthy years. For this, the data required are the age-specific prevalence (proportions) of the population in healthy and unhealthy states (often obtained from cross-sectional surveys), and age-specific mortality information taken from a period life table. The age-specific prevalence is directly applied to the person years of the life table: it provides the total number of years lived with disability; the total number of years lived without disability, and summing both, the total number of years lived. The concern about the Sullivan method is that it is based on current morbidity prevalence rates, and not current incidence rates. Also, it assumes the current mortality rates will prevail in future cohorts as they reach the same age. Therefore, it has been claimed that the Sullivan method produces biased estimates as it cannot reflect sudden changes in disability transition rates.

More recently, the multi-state method has been developed for estimating HE based on incidence morbidity rates instead of prevalence rates allowing the detection of changes in disability incidence over time (Khoman et al., 2008). However, this method is not used very frequently compared to the Sullivan method as its implementation requires longitudinal data which are not readily available n many countries including Belgium. Therefore, even with the drawback mentioned above, the Sullivan method is still the method of choice for estimating population health due to its simplicity, relative accuracy and ease of interpretation. Also, it has been suggested that the Sullivan method and multi-state method may produce similar results if transition rates are smooth and regular over time (Jagger et al., 2007).

In this context, we will use the Sullivan method for estimating HLY by SES based on an abridged life table where HLY are estimated by 5 year age groups. In this report we will not describe the methodology to develop a life table and apply to it health prevalence rates as it has been described in details elsewhere. Examples of such publications are “The life table and its applications” of Chiang (Chiang, 1984) and a number of reports and manuals developed in the context of the REVES program (Bossuyt & Van Oyen, 2000; Jagger et al., 2007). Based on these publications, the estimation of HLY was done by educational level using a program developed in the SAS software. In Appendix 2, you will find the SAS program used for the HLE estimation.

The prevalence of disability and of mortality show considerable fluctuation due to sampling variation and therefore HLY will be estimated with their standard errors. When mortality rates are extracted from the census then the variation resulting from the mortality rates is negligible and this part for the variance can be ignored. In our case, as mortality rates are estimated from a survey, we shall estimate the standard errors taking into account the variances of both mortality and morbidity rates.

To compare the results from the HIS and the SILC, we will compare mainly the educational gradient in both surveys. For this, we will estimate if the difference in the HLY of those in the lowest educational category and the HLY of those in the highest educational category is statistically significant. The test will be done by age group using the z-statistic. In Appendix 2, you will find the SAS program used for this test of difference.

4.2 Data

As mentioned in the previous section, we will use the HIS 2001 and SILC 2004 to estimate HLY by educational level. For a detailed description of these surveys, please refer to chapter 3.

The initial HIS survey has about 12 111 individuals, however after the linkage with the National Register, this number dropped to 10 093. The number of unlinked observation is quite high, but this is related to a problem in this specific linkage between HIS 2001 and the National Register. About 2018 people were not matched because there was a problem in their identification number. The problem of the mismatch was identified but due to the time constraint in this project, we were not able to another linkage. The initial sample of SILC survey is 10 146 individuals, and the linked dataset is 9 775 individuals. Only 371 individuals were not matched with the national register and therefore dropped from the dataset.

The ages of the subjects belonging to these two cohorts change during the follow up time, the longer the follow-up period is, the larger will be the difference between the ages of entry and those "acquired" during the follow up. To account for this, Lexis expansions of the original data is often used. For this, we divided the follow up period of each subject into 1-year age bands(Kirkwood & Sterne, 2003).

As mentioned in the previous section, both surveys include information on the Global Activity Limitation Indicator (GALI) that is the health outcome of interest in this report. As this question is based on the Minimum European Health Module, the questions are comparable in both surveys. For the HIS, the formulation in Fench is: Etes-vous limité(e) depuis au moins 6 mois à cause d'un problème de santé, dans les activités que les gens font habituellement? Oui, sévèrement limité(e), Oui, limité(e), Non, pas du tout. For the SILC, it is : Durant les six derniers mois ou plus, un problème de santé vous a-t-il limité dans vos activités ? Oui, de façon importante, Oui, un peu, Non.

As to educational level, its definition differs slightly by survey. In the HIS, the level of education is a household variable and not an individual variable. It reflects the highest educational level achieved within the household by the person of reference or his/her partner. In the SILC, the level of education is individual. Although these two educational indicators are slightly different, we opted to use the indicator developed and standardized by each survey team. For both surveys, the educational indicator was recoded as: primary educational level or less, lower secondary, higher secondary and higher education.

The health prevalence data are weighted to account for the study design of the specific survey.

4.3 HLY by SES

This section presents the results of the estimation of HLY by educational level using the HIS survey and the SILC survey.

Table 4.1 shows the prevalence of disability by educational level using both surveys. This table shows higher prevalence of disability in the SILC compared to the HIS. For instance, 19% of the males in the HIS reported having limitations compared to 25% in the SILC. Similarly, among females 21% reported limitations in the HIS compared to 29% in the SILC. This pattern of higher disability in the SILC applies for all educational categories. Concerning the SES gradient, we find substantial inequalities in both surveys in the prevalence of disability between those having attained higher education and the other categories. For instance, among males in the HIS the difference in the prevalence between the highest and lowest educational categories is approximately 9%, and for the SILC it is about 22%. Among females, these figures are 15% for the HIS and 21% for the SILC.

Table 4.1: Crude (but weighted) and age adjusted prevalence of disability and their 95% confidence intervals by educational level and sex for the HIS 2001 and SILC 2004

		Males			**Females**		
HIS	**Education**	N	Crude %	Age adjusted % (95% CI)	N	Crude %	Age adjusted % (95% CI)
	Primary education	656	35.52	23.54 (20.29-26.80)	768	42.84	31.48 (27.42-35.55)
	Lower secondary	823	28.07	23.53 (20.83-26.23)	872	27.64	21.03 (18.53-23.54)
	Higher secondary	1319	18.57	18.38 (16.32-20.44)	1365	20.29	20.15 (18.08-22.22)
	Higher education	1537	13.99	14.94 (13.21-16.68)	1580	14.11	16.20 (14.31-18.09)
	Total	4335	21.31	19.26 (18.18-20.35)	4585	23.34	20.58 (19.49-21.66)
		Males			**Females**		
SILC	**Education**	N	Crude %	Age adjusted % (95% CI)	N	Crude %	Age adjusted % (95% CI)
	Primary education	895	42.46	38.71 (34.54-42.87)	1199	49.71	41.49 (37.70-45.27)
	Lower secondary	841	27.47	29.70 (26.46-32.94)	849	34.75	33.33 (29.95-36.71)
	Higher secondary	1837	22.37	24.10 (22.13-26.08)	1656	24.88	26.88 (24.70-29.05)
	Higher education	1296	17.05	16.93 (14.96-18.90)	1413	19.04	20.58 (18.37-22.78)
	Total	4869	25.53	24.99 (23.82-26.16)	5117	30.72	28.79 (27.60-29.98)

Table 4.2: Vital status by educational level and sex for the follow up of the HIS 2001 and SILC 2004

		Males				Females				Total			
HIS	Education	Death registration		No change in status		Death registration		No change in status		Death registration		No change in status	
		N	%	N	%	N	%	N	%	N	%	N	%
	Primary education	111	18.20	498	81.80	91	17.67	592	82.33	202	17.91	1.090	82.09
	Lower secondary	77	10.23	683	89.77	51	5.92	750	94.08	128	8.05	1.433	91.95
	Higher secondary	64	4.36	1.202	95.64	41	2.63	1.233	97.37	105	3.48	2.435	96.52
	Higher eductation	53	2.69	1.394	97.31	24	1.81	1.441	98.19	77	2.26	2.835	97.74
	Total	305	6.64	3.777	93.36	207	5.27	4.016	94.73	512	5.95	7.793	94.05

		Males				Females				Total			
SILC	Education	Death registration		No change in status		Death registration		No change in status		Death registration		No change in status	
		N	%	N	%	N	%	N	%	N	%	N	%
	Primary education	131	15.67	729	84.33	123	12.08	1.026	87.92	254	13.58	1.755	86.42
	Lower secondary	49	6.27	755	93.73	30	4.00	792	96.00	79	5.12	1.547	94.88
	Higher secondary	74	4.45	1.705	95.55	39	2.59	1.562	97.41	113	3.57	3.267	96.43
	Higher eductation	39	3.04	1.212	96.96	20	1.44	1339	98.56	59	2.21	2.551	97.79
	Total	293	6.46	4.401	93.54	212	4.81	4.719	95.19	505	5.61	9.120	94.39

Table 4.2 presents the distribution of the survey participants by vital status and educational level for both surveys. We see that the prevalence of mortality is comparable in both surveys. In the HIS sample, 5.95% of the respondents died compared to 5.61% of the SILC sample. Concerning the SES gradient, we find significant inequalities in mortality by educational categories in both surveys. For instance, in the HIS almost 18% of the lowest SES category died compared to 2% in the highest category. In the SILC, these figures are respectively 14% and 2%.

In Appendix 3, you will find the tables showing HLY and their variances by educational level for all age groups. To highlight some results, the HLY by SES and gender at two selected ages (25 and 65 years) are presented in this section.

Figure 4.1 shows life expectancy broken down by HLY and years in disability for males aged 25 years in the HIS and the SILC. This figure shows very comparable LE in both surveys, but higher HLY in the HIS compared to the SILC. For instance, LE at 25 years for those with the lowest educational achievement is almost 50 years in the HIS and the SILC. Concerning HLY for the same group, it is 35 years in the HIS (28 % of LE spent in disability) compared to 28 years in the SILC (42% of LE spent in disability). Such higher figures for the years spent in disability in the SILC is expected as in table 1 the prevalence rates of disability are systematically higher in the SILC compared to the HIS. In both surveys important educational inequalities are observed in LE and in HLY. For instance, the difference in LE between the highest and lowest educational group is almost 6 years in the HIS and 9 years in the SILC. For HLY, these figures are respectively 7 years and 14 years.

Figure 4.2 shows life expectancy broken down by HLY and years in disability for females aged 25 years in the HIS and the SILC. As expected LE are higher for females compared to males, and years spent in disability are more important for females compared to males. Here also, we find more comparability between the SILC and the HIS in LE but not in HLY. Again, the SILC shows higher years in disability compared to the HIS. For instance, at 25 years women with the lowest educational achievement are expected to live 41% in disability according to the HIS, while this figure is 48% in the SILC. Important educational inequalities are detected in both surveys for LE and HLY.

Figure 4.1: HLY and years in disability for males aged 25 years

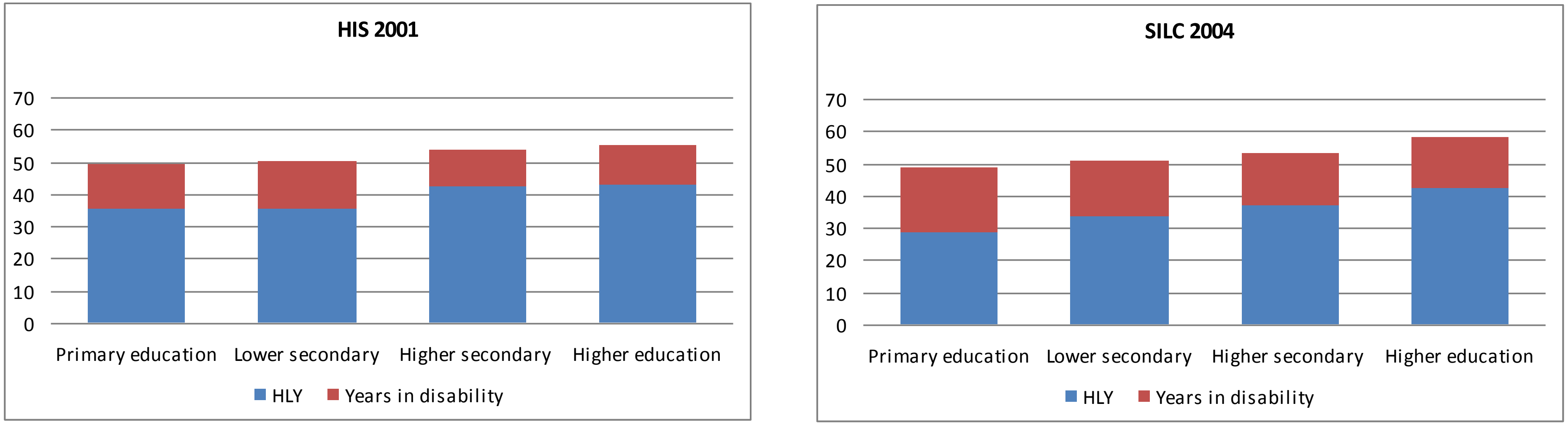


Figure 4.2: HLY and years in disability for females aged 25 years

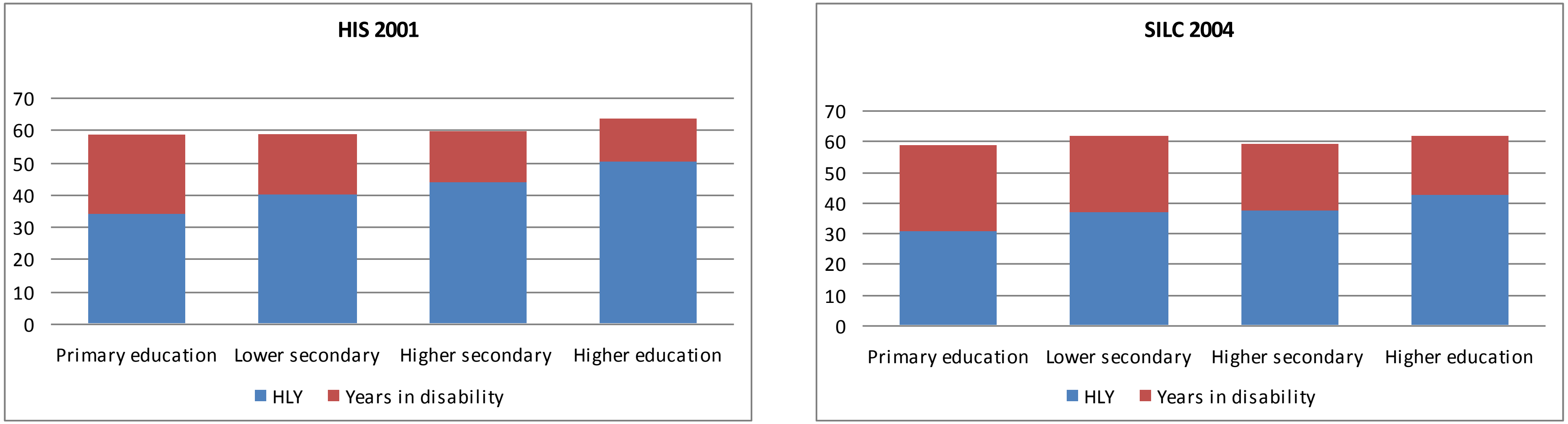


Figure 4.3: HLY and years in disability for males aged 65 years

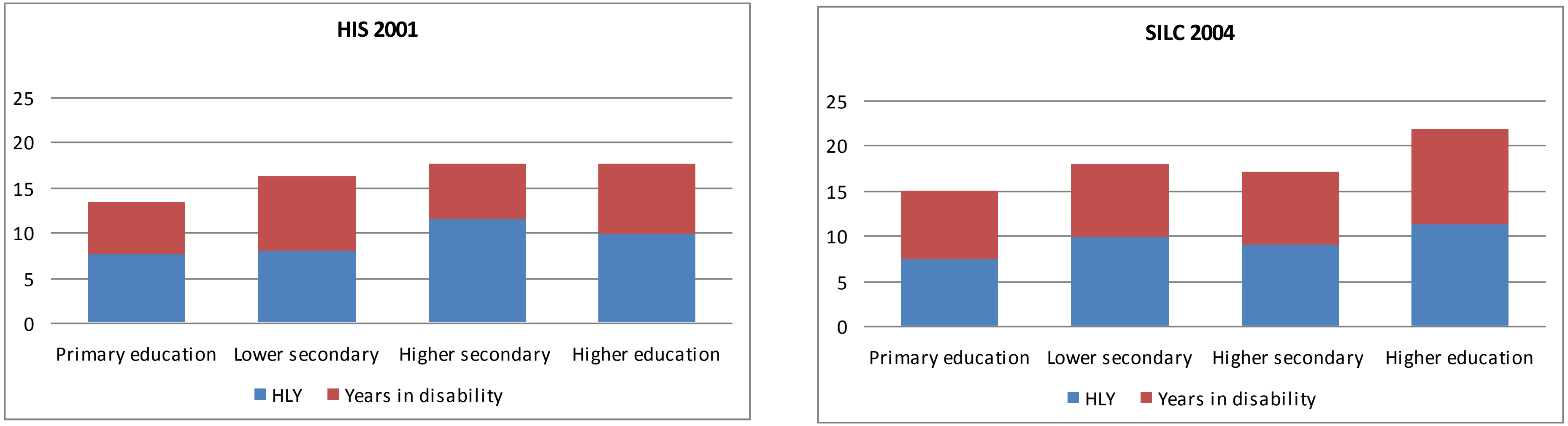
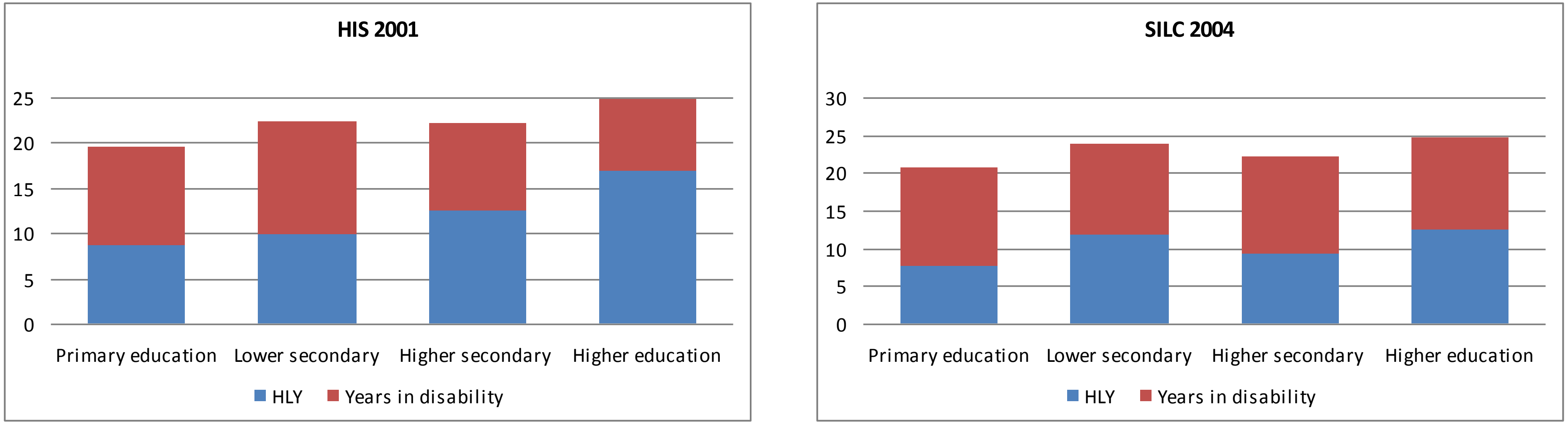


Figure 4.4: HLY and years in disability for males aged 65 years



Figures 4.3 and 4.4 show life expectancy broken down by HLY and years in disability for males and females aged 65 years in the HIS and the SILC. Among males aged 65 years old, we no longer see comparable LE between surveys; here differences are detected in both LE and HLY between the HIS and the SILC. Among females of 65 years old, we see similar pattern as in younger age concerning comparability between surveys in term of LE and HLY. For males and females, we see the same pattern as in the younger age group for higher years spent in disability in the SILC compared to the HIS. For instance, males with the lowest educational achievement in the HIS are expected to live 44% of their life with disability compared to 51% in the SILC. For females, these figures are 56% and 64%. As well, inequalities in LE and HLY are observed in these higher age groups.

Table 4.3 shows the result of the z-statistic for the difference between HLY in the lowest and highest educational categories. Statistically significant inequalities in HLY are detected in both surveys at in the 25 years old category. However, at 65 years old, significant inequalities are only found in the HIS but not in the SILC.

Table 4.3: Difference between HLY in the lowest and highest educational categories

Males

Survey	Age	HLY (variance)		Difference	z	p
		Primary education	Higher education	High-Low		
HIS	25	35,49 (2,03)	42,65 (1,02)	7,16 (3,05)	2,35	<0.02
	65	7,49 (0,27)	9,82 (0,77)	2,33 (1,04)	2,25	<0.05
SILC	25	28,44 (1,58)	42,28 (2,05)	13,84 (3,64)	3,80	<0.001
	65	7,38 (0,32)	11,28 (1,96)	3,89 (2,28)	1,71	<0.10

Females

Survey	Age	HLY (variance)		Difference	z	p
		Primary education	Higher education	High-Low		
HIS	25	34,25 (1,68)	50,48 (2,58)	16,24 (4,26)	3,81	<0.001
	65	8,62 (0,36)	16,95 (2,93)	8,33 (2,76)	3,02	<0.01
SILC	25	30,37 (1,16)	42,73 (3,07)	12,36 (4,23)	2,92	<0.01
	65	7,56 (0,23)	12,57 (3,05)	5,01 (3,28)	1,53	<0.20

5. Estimation of mortality by socioeconomic status in European countries

A growing number of European countries are involved in estimating healthy life expectancies by SES. Although the source for morbidity data is often from cross sectional surveys, the source of mortality data differs. This chapter describes the approaches used by a number of these countries to estimate mortality rates by socioeconomic status. See Table 5.1 for a summary.

Austria

Mortality rates by SES in Austria are generated through the linkage of individual records of the censuses 1981, 1991, and 2001 with death certificates in a 12-month follow-up period. No ‘direct’ matching through the national register number could be used. Instead, data was matched using primary matching variables: sex, date of birth, and last residential address of the deceased. The overall merging rate was about 90% for the 1981 and the 1991/92 death records, compared to almost 94% for the 2001/2002 death records. Using this dataset a study examined inequalities by educational level (Klotz & Doblhammer, 2008). In 2001/2002, life expectancy at age 35 for males was 40.1 years for the lowest educated group and 46.2 for the highest educated group. For women, these figures were 46.6 years and 49.4 years, respectively.

Bulgaria

In Bulgaria, data for estimating mortality by SES come from a linkage between the 1992 population census and death certificates. The follow-up period stopped in 1998.The link between census and mortality data was performed on the basis of a personal identification number that is uniquely assigned to each Bulgarian citizen and is included on death and census records. Approximately 93 % of all death certificates for the study period were linked to the 1992 census records.

Using such data, a study obtained age and sex specific death rates from a parametric specification of the mortality hazard based on a Gompertz model (Kohler et al., 2008). Life expectancy lost and probability of dying between age 40 and 70 are calculated. It is estimated that men with high education can expect to live 2.56 years longer than men with low education between ages 40 and 70. For women, the difference in life expectancy lost between low and high education is 0.64 years.

Denmark

In Denmark, data regarding education and occupation is collected in national registers from the 1970 census. Since 1980, education data have been reported and updated annually. To calculate mortality rate by social status, SES data from the registers is linked to mortality records for all inhabitants. A study examined the inequalities for the 30 to 74 year olds in the period 1996-2005. In this age group and period about 2.3% of the mortality data failed to be linked (due to unclassified or missing data on education). The authors calculated age-standardized death rates for every calendar year by education and sex. For the period 1996-2005, the analysis reveals an increase in social inequality in life expectancy: for the lower educational level the increase in life expectancy at 30 years was 0.73 years for men and 0.42 years for women whereas for the higher educational level the increase was 1.06 years and 0.75 years, respectively (Bronnum-Hansen & Baadsgaard, 2007). Using the same source of mortality data, another study examined health expectancy by occupational status (Bronnum-Hansen, 2005). The authors found that among 30-year-old men, high-level salaried employees had the longest expected lifetime in perceived good health, 41 years, which amounts to 89% of life expectancy, compared to 34 years (73%) for farmers, 32 years (73%) for unskilled workers, and 19 years (56%) for economically inactive men. Expected lifetime in perceived good health for high-level salaried female employees from age 30 was 46 years (91% of life expectancy). The lowest was found for assisting spouses, 36 years (71%) and economically inactive women, 25 years (56%). Large differences were also found when data on long-standing illness were used.

Finland

A study analyzed Finish data together with Bulgarian data in a comparative fashion (Kohler et al., 2008). The approach for calculating mortality by socioeconomic status was thus fundamentally the same. Like for Bulgarian data, census and mortality data were linked using a unique personal identification code. Non-linkage of death records to census records is less than 0.5%. The difference with Bulgarian method is that an 11% random sample of longitudinal census data file was studied in the case of Finland. Analysis reveals that health inequality by education is less pronounced than in Bulgaria, at least for males: men with high education can expect to live 1.79 years longer than men with low education between ages 40 and 70. For women, the difference in life expectancy lost between low and high education is 0.62 years.

France

The data used in France comes from a permanent demographic sample currently including about one million people (approximately 1% of the population), randomly selected. Socio-demographic data is updated at each census and information regarding mortality is permanently updated by the means of vital status forms. In other words, a follow-up of the representative sample is constantly ensured.

Using these data, a study used Cox models to estimate relative risks with age as the time variable, occupational group or educational level being treated as quantitative variables (Saurel-Cubizolles et al., 2009). Relative index of inequality (RII) were also estimated to obtain a global measure of the magnitude of inequalities in mortality. Using data from the period 1990-1999 of 30 to 64 years old the analysis reveals wider social inequalities among men than among women. Based on education, RII is estimated to vary between 2.72 and 3.23 for men and between 2.29 and 2.99 for women. From an occupational perspective, the difference between men and women is more evident: RII is estimated to vary between 5.54 and 6.68 for men and between 2.96 and 3.96 for women.

Lithuania

Recently, linkage procedures have been used in Lithuania (Shkolnikov et al., 2007). Linkage between the 2001 census, death and migration records was accomplished using personal identification numbers as unique identifiers for the same individuals. The follow-up ended in 2004. The method allowed matching about 95% of the death records to the corresponding census records. A special redistribution procedure was applied to the remaining census-unlinked deaths.

Education level served as a proxy for socio-economic position. Data regarding education were available both from the census and the death records. Although the two sources of information did not use the same categories to describe education levels, it was possible to classify individuals in one of the 3 following broad categories of educations: *higher education* (at least 14 years of schooling), *secondary education* (10-13 years of schooling), *lower than secondary* *or unknown education* (up to 9 years of schooling). Specific research was undertaken to determine whether the use of one source of information rather that another would matter (differences might arise from actual changes during the interval between the census and registration, from the attitude of declarants or to the definition of the categories and the formulation of the questions used for collecting information). Analysis revealed significant disagreement between the two sources of information (death and population records), but no systematic overstatement in the census records. It was decided to use census data to measure education level.

Researchers used this data to estimate life expectancies at age 30 (Shkolnikov et al., 2007). Standard errors were calculated using the Chiang method. Poisson regressions were also used to compute mortality rate ratios by education (controlling for other factors). The analysis shows that life expectancy at age 30 differs by 11.3 years between men with the highest and those with the lowest educational levels. For women, the difference was estimated to equal 6.9 years.

Netherlands

Socio-demographic data were collected within the framework of the prospective GLOBE study (Gezondheid en LevensOmstandigheden Bevolking Eindhoven en omstreken). In 1991 a random sample of 27 070 non-institutionalised Dutch persons (aged 15-74 years) drawn from municipal population registers from Eindhoven and its surrounding was sent a postal questionnaire. The response rate was 70.1%. Two sub-samples of the respondents were approached for a structured interview[1].

In the study a total 4087 respondents were eligible for the analyses (van Oort et al., 2005). Education is chosen as a proxy for socio-economic position. The follow-up of all subjected that started in 1991 ended in 1998. Information about mortality (and changes of address) was collected annually via municipal population registers and could therefore be easily linked to socio-demographic data. No information is provided regarding the data that failed to be linked.

[1] One sub-sample overrepresented people with chronic lung disease, severe heart disease, diabetes, and persistent back trouble (response 72.3%), n = 2865). The other sub-sample was a random sample of respondents to the postal questionnaire (response 79.4%, n = 2802).

All analyses were weighted to make the sample representative for the original GLOBE respondents. Hazard ratios for levels of education adjusted for age, gender and other covariates (were calculated material factors, behavioural factors, and for psychosociological factors). It was shown that hazards are between 1.43 and 4.46 times higher for the lowest educated level than for the highest educated group.

Spain

Mortality rates by SES were estimated in the two major cities of the country: Madrid and Barcelona. The population at risk comes from the 1991 Municipal Censuses. The censuses information is revised every 5 years through the active collection of data by the statistical office of each municipality. Between these major revisions, the information is continually updated to incorporate data on births and deaths. Death register and municipal censuses are thus linked. Using 1993 and 1994 mortality data, linkage problems (due to problems of record linkage or missing data) appeared in 1.5% of deaths in Madrid and 9% in Barcelona.

Education levels served as an indicator of socio-economic position. Three levels of education were distinguished: *no education* (0-4 years of schooling), *primary education* (5-11 years) and *secondary or higher education* (more than 11 years).

This data has been used to estimate age- and sex-specific mortality rates and life expectancy at 25 years for each educational level using the life-table method (Borrell et al., 1999). Rate ratios were also calculated. Results showed that the mortality risk for the 25 to 34 year old males in Madrid is between 6.08 and 8.27 times higher for low educated than for higher educated. For females in Madrid, the same risk is estimated to vary between 4.91 and 8.17. The analyses show that the differences in mortality by education are more important for the younger population than for the middle-aged and the elderly population.

Switzerland

In Switzerland, recent mortality studies by SES are based on the Swiss National Cohort, a national longitudinal research platform. Initial data comes from the 1990 census. Although no unique personal identifier is available in Switzerland, the exact data of birth was included in that census for the first time, making it possible to link census data and mortality data from that period of time onwards. Accordingly, the 1990 census was linked with the 2000 census, mortality and emigration records using both deterministic and probabilistic methods of matching. In the first step, pairs of records that were matched on sex, date of birth, martial status, nationality, religion and place of residence were identified. In subsequent steps, probabilistic record linkage was used, which estimates the probability that a pair of records from different datasets relates to the same person. Additional variables, in particular information on the spouse and family structure were used in the process. Based on probability weights, possible matches were accepted or rejected. Linkage was automated using the Generalized Record Linkage System (GRLS) packaged developed by Statistics Canada. Of all individuals registered in the 1990 census, 81.9% could be linked to a 2000 census record, and during the period 1990-2000, 2.6% to an emigration record and 8.6% to a mortality record. For the remaining individuals with a 1990 census record (6.9%) no satisfactory link could be found. Note also that of all the death recorded from 1990 to the end of 2000, 95.3% could be linked successfully to a census record.

This dataset was used to analyse inequalities for those aged between 40 and 79 (Faeh & Bopp, 2010). Socio-economic position was assessed by the education level as measured in the 1990 census. Two approaches were used. For regression analysis (i.e. the calculation of relative change of hazard ratios by the means of Cox model), the estimated years of education was used. For descriptive mortality and prevalence rates (calculated using lexis expansion for 5-year age classes), the denotations used in the censuses were reclassified into 3 categories of education. The study shows mortality per additional year of education decrease in the youngest age group by about 7 % (in French Switzerland) and 9% (in German Switzerland) in men an by about 6% (in French Switzerland) and 4% (in German Switzerland) in women.

The United Kingdom

In the UK, mortality rates by socio-economic position are preferably estimated by combining death records (numerator) with mid-year population estimates (denominator) obtained at every census (i.e. every 10 years). The method involves the correction for biases resulting from the use of unlinked data. In order to contribute to the monitoring of health inequalities over time, the Office for National Statistics investigated the feasibility of using survey data – namely, the Labour Force Survey (LFS) – to provide population denominators for the estimation of mortality rates by socio-economic status. The LFS collects detailed data on occupation for a sample of approximately 60,000 households in the UK quarterly and annually (annual dataset being boosted). The survey measures socio-economic position by the means of the National Statistics Socio-economic Classification (NS-SEC), which is derived from an individual's occupation and employment status and the size of their organization. Since size of organisation is not collected on the death register, a version of NS-SEC is used which is derived from occupation and employment status alone. This is known as 'reduced NS-SEC'. A specific methodology was developed for estimating mi-year populations by age (5-years classes) and NS-SEC using the LFS. LFS data were standardized by age to ONS mi-year population estimates, and adjusted for potential health selecting using the same proportional adjustments as were used for a previous census-based denominator study. The age-specific rates for each NS-SEC class were then standardised to the European standard population to produce age-standardised mortality rates for each class.

The study using the described methodology focused on men aged 25-64. For instance, for the higher managerial class age-standardised mortality rate was estimated between 162 and 175 (per 100,000). For the Routine class it is estimated between 525 and 543. (Johnson & Langford, 2010)Also, in the UK inequality in mortality has been examined by geographical area (Wells & Gordon, 2008).

EUROSTAT recommendations

Eurostat has been requested to develop comparable information on mortality by socio-economic status on a regular basis for all EU Member States. In fact, the data described above are the results mainly of ad hoc studies. Given the heterogeneity of the possible methods to estimate such indicators, Eurostat is currently working on guidelines regarding the preferable methodologies for the estimation of mortality by SES. This is a work in progress, but some recommendations are already available:

1. It is suggested to use educational attainment as proxy for socio-economic position: provided a standard classification is used (i.e. ISCED), education level can determined for all individuals, it is more stable and it is in line with countries' practices.
2. The preferable measure of mortality is life expectancy. Life expectancy has a ready interpretation and it is comparable across populations with different age structure.
3. Regarding the source of information used to calculate mortality rates, it is recommended to link death certificates with census information from the censuses broken down by educational attainment, using a personal identity number. Indeed, the procedure avoids inconsistencies between the numerator and the denominator. Nevertheless, a bias is to be expected because the SES is measured at the time of the

census and not at the time of death (but it should be limited if the follow-up period is short). If no unique personal number is available, deterministic methods or probabilistic methods should be used.

In parallel, Eurostat is encouraging the National Statistical Institutes of Member States to carry out the "census records linkage" and follow-up on the occasion of the 2011 census round. Therefore, a list of guidelines and of related methodological issues is being developed on order to help the NISs interested in doing this exercise. These guidelines refer to: linkage procedure and software;handling of non-matched deaths; calculation of deaths rates; follow-up period; follow up of migration and emigration and possible sources of bias.

Table 5.1: Estimation of mortality by SES in a number of European countries

Country	SES indicator	Period	Source of data	Linkage approach
Austria REF	Education	1981-1982 1991-1992 2001-2002	Mortality follow up of the national census	Individual linkage , no unique personal identifier
Bulgaria REF	Education	1992-1998	Mortality follow up of national census	Individual linkage, unique identifier
Denmark	Education Occupation	1996-2005 1986-1990	Mortality follow up of the national census	Individual linkage, unique identifier
Finland	Education	1994-1998	Mortality follow up mortality of a sample of the national census	Individual linkage, unique identifier
France	Education Occupation	1990-1999	Mortality follow up of the permanent demographic sample	Individual linkage, unique identifier
Lithuania	Education	2001-2004	Mortality follow up of the national census	Individual linkage, unique identifier
Netherlands (Eindhoven)	Education	1991-1998	Mortality follow up of the GLOBE longitudinal survey	Individual linkage, unique identifier
Spain (Madrid, Barcelona)	Education	1993-1994	Mortality follow up of the municipal census	Individual linkage, unique identifier
Switzerland	Education	1990-2000	Mortality follow up of the Swiss National Cohort longitudinal study No unique personal identifier	Individual linkage , no unique personal identifier
United Kingdom	Reduced NS-SEC	2001-2007	Mortality follow up of the national Labor Force Survey	Unlinked approach

6. Evaluation of the alternatives and recommendations

The aim of this project is to propose an approach to estimate HE by SES using databases other than the national census. Using the census and its mortality follow up is the golden standard for such estimations but this approach is not feasible anymore in Belgium as the 2001 census was the last ‘classical’ census. We examined two alternative approaches: an unlinked approach using two different databases and a linked approach using the follow up mortality of survey data.

The comparison of the linked and unlinked approaches to estimate mortality rates undertaken in chapter 2 illustrates that the use of non-linked data results in a considerable bias, mortality rates being significantly underestimated both among men and women and for quasi all educational levels. The bias results from the quality of the educational information in the death certificates, with a very high percentage of missing information. More than 70% of the deaths in the certificates have a missing value for education. This group clearly has a much higher mortality rate in the non-linked data. Social inequalities in mortality between the educational groups generally seem overestimated when using non-linked data.

These conclusions have important implications. If mortality and social inequalities in mortality need to be investigated through non-linked data, educational level should be registered correctly and exhaustively in the certificates. Mortality and inequalities cannot be investigated correctly when information on educational level is missing for more than three quarters of the population. Respondents for whom educational data are missing could eventually be distributed in such a way that the educational distribution of the total population in the census 2001 is respected. Or educational information could be derived from other socioeconomic characteristics, such as professional class.

However, even if educational level is registered correctly in the certificates, errors can occur when using non-linked datasets to compute mortality rates by educational level (differing registration of education in both datasets, technical problems, registration errors in mortality). The best option is that data can be linked. The availability of a unique identification number allows for a relatively correct linkage of data, ensuring that the linked information applies to the one and same person. This is the best approach to calculate mortality rates by educational level. The 2001 census was the last ‘classical’ census in Belgium. It will be replaced by an administrative census. In this administrative census the registration of education is not ensured for the total population in Belgium. This definitely has repercussions for the follow-up of social inequalities in mortality.

As an alternative, using the follow up mortality of surveys is an interesting approach. In this approach, we do not have the problem of numerator denominator bias, but other issues have to be considered. The main issue is the selection bias generated due to the difference between the census and the survey. The surveys are based on a sample while the census aims at enumerating the whole population. Also the surveys are not compulsory and so respondents may chose not to participate. This implies that a bias will be generated due to the selection of survey participants. In fact, a study has shown that people in lower SES categories were less likely to participate in a survey when they had a poor health status compared to people in higher educational categories (Lorant et al., 2007). In this case, health inequalities may be underestimated. Although in both the HIS and the SILC an effort is deployed to increase the response rate and to account for participation rate, still selection bias may be an important issue using this approach. Further the size of the survey may be too small and surveys may not cover all age groups or may be not designed to representative at levels other than the national level.

Based on this, we issue a number of recommendations to monitor HE by SES in Belgium:

1-The golden standard for estimating the mortality part of the HE is to use a unique identifier to link the census with the National Register. In the absence of such source of data other approaches will be used.

2-If an unlinked approach is to be used, it is important to highlight that there is a significant nominator denominator bias that has to be acknowledged when using education level as the SES variable and death certificates as the data source. Other SES variables and data sources need to be explored.

3-If survey follow up is to be used, a choice has to be made on the survey. The surveys that were evaluated (HIS and SILC) provide estimation of HE with comparable variances.

This is because the sample size is comparable in both surveys. Therefore, the decision of using one survey over the other to monitor HE by SES should not be based on the statistical stability of the estimation but rather on other criteria.

If one is interested in monitoring HE by SES by region then the HIS is a better choice. In fact, the SILC data is representative at the national level, but there is no additional effort to generate estimate that are accurate at the regional level. Still, estimations for the Walloon and the Flemish regions may be accurate, but it is not the case for Brussels. The HIS, however, is designed to be representative of the three regions and therefore theoretically to derive acute estimators. In the present report, no estimations have been undertaken at the regional level, but we expect that the estimation is feasible even if the sample size is smaller (3500 in Flanders, 3500 in Wallonia, and 3000 in Brussels) and therefore the statistical precision would be lower. Methods can be used to overcome this limitation, such as using the bootstrap method to estimate the variance of HE.

As seen in Chapter 4 and Appendix 3 the difference in SES is significant for younger ages in both surveys, however, for older ages, the gradient exists only in the HIS. This may be another advantage of the HIS where there is an oversampling of elderly people. Therefore, if the aim of the monitoring is to estimate inequalities at older ages, then the HIS may be a better option. This is especially true that a number of studies have found significant social inequalities in HLY at older ages.

The main advantage of the SILC is its annual periodicity. The data of the SILC are issued yearly, however the HIS is collected every 4 to 5 years. Therefore, if the aim is to monitor HLY by SES frequently, then the SILC may be the best alternative.

4-It is important to highlight that the HIS and the SILC can both be used to estimate inequalities in HLY in Belgium but the results are not interchangeable. In other words, it is possible to compare HLY estimates within surveys but not between surveys. Therefore, a decision has to be made on the survey to be used at this point and the future estimations must also be based on the same survey. Otherwise, comparison would be meaningless.

5-Concerning the SES stratification, we used in this report educational level. As seen in the chapter reviewing the situation in a number of European countries and the Eurosat recommendations, education is the variable of choice. Therefore, we recommend its use for the monitoring system.

6-Concering the health outcome to be used, we propose the use the indicator we used in this report that is the Global Activity Limitation Indicator (GALI). As mentioned previously in this report, the GALI allows the estimation of disability free life expectancy that is a structural indicator of the EU. This indicator is based on the Minimum European Health Module (MEHM) that is included in the SILC since 2004 and in the HIS since 2001. It is also interesting for potential European comparison as the MEHM has been implemented in several European countries and European surveys.

7-Concerning the methods for estimating health expectancies, we recommend the use of the Sullivan method. This is the method that is predominately used in the literature. This method has a number of drawbacks mainly that it is based on current morbidity prevalence rates, and not current incidence rates. Also, it assumes the current mortality rates will prevail in future cohorts as they reach the same age. Still, this method is viewed as the method of choice to estimation HE due to its simplicity, relative accuracy and ease of interpretation. Also, we recommend calculating the variances with the HE. This will allow a better understanding of the stability of the estimates.

We also recommend to account for the fact that the ages of the subjects in the surveys change during the follow up time, the longer the follow-up period is, the larger will be the difference between the ages of entry and those "acquired" during the follow up. To account for this, Lexis expansions of the original data is often used. For this, we divided the follow up period of each subject into 1-year age bands(Kirkwood & Sterne, 2003). The steps and programs used to estimate HE are found in Appendix 2.

Practical recommendations

To set up a monitoring of HLY by SES in Belgium using the mortality follow-up of surveys, it is recommended to:

1- Decide upon the survey to be used. We recommend the use of the SILC because it is a yearly survey. Also by pooling the data of several years, we may have enough data to make regional comparisons.

In addition, if at the EU level a decision is made with respect to the monitoring of HLY by SES, Belgium would have to adapt the proposed monitoring to follow the EU recommendations. As the SILC is already used at the EU level to monitor HLY, there is great chance that the same survey will be used to monitor HLY by SES. However, the proposed duration of mortality follow-up may be different from the one proposed below.

2- Consider the mortality follow-up for at least three years to have enough cases. So, for instance, if the SILC 2004 is used, then take the mortality follow-up until at least 2006.

3- Submit a request to the privacy commission to use the SILC data and its mortality follow-up. Refer to Appendix 1 for a sample request. The cost of having access to such data is less than 1500 Euro. But this may vary between years.

4- Expect a waiting time of about 6 months to receive the data after submitting the forms to the privacy commission. If the request was regular (each year) then a shorter waiting period may be expected.

5- Use the Global Activity Limitation Indicator (GALI) as the health indicator. This indicator allows the estimation of the disability free life expectancy that is a structural indicator of the EU. This indicator is based on the Minimum European Health Module (MEHM) that is included in the SILC since 2004.

5- Use the educational level as a proxy for socio-economic status. In this report, we categorized educational level into 4 categories: primary educational level or less, lower secondary, higher secondary and higher education. We recommend the use of this categorization.

6- Use the Sullivan method to estimate health expectancies and their variances. This will allow a better understanding of the stability of the estimates. To do this, please refer to the statistical programs in Appendix 2 and to the booklet: "Health expectancy calculation by the Sullivan Method: A Practical Guide" (Jagger et al., 2007). This booklet walks the user through the estimation of the HLY and its variance using Excel, SAS or SPSS. In our analysis we used SAS.

Therefore, to replicate the work undertaken in the context of this project, the following tasks are necessary:

- Request the SILC data and its mortality follow-up (3 years) from the privacy commission. This step is not very challenging as a sample request is included in this report. Still, the request needs to be adapted every year and most importantly need to be followed up. For estimations at the regional level, the SILC data need to be pooled over 3 consecutive years.

- The data need to be explored and simple descriptive analysis undertaken for the main variables of interest.

- Adapt and run the statistical programs that allow the estimation of HLY by SES.

- Finally, the findings need to be interpreted and a report written that would include a synthetic review of the updated literature on this subject.

We estimate that such a project may be implemented by a 2 person-months per year for the estimation at the Belgian level. This researcher must have a good knowledge in statistical programming and in HLY estimations. However, the first year a longer period may be needed depending on the experience of the researcher.

Appendix 1: Forms submitted to the Privacy Commission to obtain SILC data

Comité de surveillance statistique
Formulaire de demande de communication de données d'étude.

Nom de l'institution/organisation demanderesse : Institut Scientifique de Santé Publique
Nom et prénom du responsable (directeur, administrateur délégué, recteur, fonctionnaire dirigeant,…)*: Herman Van Oyen, chef de Direction Opérationnelle Santé Publique et Surveillance
Adresse : J. Wytsmanstraat 14, 1050 Bruxelles

Tél : 04 78 440 379, 02 642 5029
Fax : 02 642 5001
E-mail : herman.vanoyen@iph.fgov.be

Responsable des données (personne physique responsable du respect de toutes les obligations concernant l'exécution du contrat de confidentialité, de la loi du 4 juillet 1962 relative à la statistique publique et de ses arrêtés d'exécution) :

Nom et prénom : Rana Charafeddine
Fonction : Collaborateur Scientifique
Adresse :J. Wytsmanstraat 14, 1050 Bruxelles

Tél : 02 642 5739
Fax : 02 642 5001
E-mail : rana.charafeddine@iph.fgov.be

A compléter en cas de sous-traitance

Nom du sous-traitant :
Adresse du sous-traitant :
Durée de la sous-traitance :
Travaille sous la responsabilité et le contrôle du demandeur : oui/non*
Preuve de l'engagement du sous-traitant : jointe/non-jointe*

Tél :
Fax :
E-mail :

*Biffer les mentions inutiles

Description/situation de l'organisation :

L'Institut scientifique de santé publique (ISP) est un établissement scientifique de l'Etat fédéral belge.

Sa mission principale est d'apporter un soutien scientifique à la politique de santé.

Il fournit également de l'expertise et des prestations de service public dans le domaine de la santé publique.

Statut du demandeur / base légale :

Services publics :

<u>0 Etat fédéral (article 15, alinéa 1, 1°, de la loi du 4 juillet 1962)</u>
0 Régions et communautés (article 15, alinéa 1, 2°, de la loi du 4 juillet 1962)
0 Provinces et communes (article 15, alinéa 1; 3°, de la loi du 4 juillet 1962)

Personnes physiques ou morales (article 15, alinéa 1, 4°, de la loi du 4 juillet 1962) :

0 Institution scientifique
0 Particulier
0 Entreprise
0 Autre:
………………………………………………………………………………………………..

Les données demandées seront utilisées pour :

<u>0 Recherche scientifique</u>
0 Recherche statistique
<u>0 Aide à la décision politique</u>
0 Intérêt général
0 Autre :
……………………………………………………………………………………………..

But de la recherche/étude (description précise et explicite des finalités et objectifs statistiques poursuivis) :

L’objectif de ce projet est d’identifier une base de données alternative au recensement national pour estimer les années de vie en bonne santé (AVBS) selon le statut socio-économique.

Pour estimer les AVBS nous avons besoin de données de mortalité et de données de morbidité. Actuellement, en Belgique, les taux de mortalité pour estimer les AVBS par statut socio-économique sont générés à partir des recensements de 1991 et

2001. Etant donné que le recensement national ne va plus exister sous sa forme actuelle, nous devons trouver des méthodes alternatives pour estimer les taux de mortalité par statut socio-économique. Dans ce projet, nous évaluerons 2 méthodes:

Méthode 1 : Les taux de mortalité par statut socio-économique sont calculés à partir du suivi mortalité d'enquêtes tel que l'enquête santé par interview (HIS) ou EU-SILC qui comprennent des informations sur le statut socio-économique des participants, leur santé et leurs limitations fonctionnelles. La durée de ce suivie dépend de la taille de l'échantillon, mais ne doit pas être inférieur à 2ou 3 ans.

Méthode 2 : Les taux de mortalité par statut socio-économique sont estimés à partir de deux bases de données distinctes : l'une fournissant le nombre de décès par statut socio-économique et l'autre fournissant la répartition de la population par statut socio-économique.

Dans les 2 cas, ces taux de mortalité seront associés à des données de morbidité provenant de diverses enquêtes comme l'enquête santé par interview (HIS) ou de l'EU-SILC pour estimer les AVBS par statut socio-économique.

Estimation de la durée de la recherche : 1 an de 01/01/2010 à 31/12/2010

Description suffisamment détaillée des données demandées (type d'enquête, année de référence, détail des données + période) :

Nous souhaitons utiliser les données de l'enquête sur les revenus et les conditions de vie (SILC) de 2004 pour évaluer les questions décrites ci-haut. Pour cela, nous demandons les données suivantes :

<u>1-Les données SILC au niveau du ménage :</u>
Situation de logement actuelle: H5, H7, H8, H9, H10, H11, H12, H50
Situation financière: H51, H52, H53, H54, H55, H56, H57, H58, H59, H60, H61, H62, H63
Revenus perçus par votre ménage: H64, H65, H66, H67, H68, H69, H70, H71

<u>2- Les données SILC au niveau individuel:</u>
Activités professionnelles: I5, I 6, I7, I8, I9, I10, I11, I19, I 20, I 21, I22, I25, I26, I27, I28, I29, I30, I31, I32, I33, I39, I40
Santé : I134, I135, I136, I137, I138, I139, I140
Informations biographiques: I149, I150, I151, I152, I153, I154, I155, I156, I157, I158

Formation scolaire: I159, I160, I161, I163, I164, I165, I166, I 167, I168, I 169, I 170, I171 **Indicateurs SILC**: Risque de pauvreté, difficultés à joindre les deux bouts **Autres :** âge, sexe, région, nombre de personne par ménage, âge et sexe de chaque personne du ménage, pondération, date de l'interview <u>3-Le suivi mortalité des données SILC de 2004:</u> De plus, nous désirerons avoir une variable supplémentaire avec les données décrites ci-haut qui précise la date de décès ou d'émigration des répondants qui sont décédés ou ont émigré avant le 31/12/2009. Cette variable est obtenue par la vérification du statut vital des répondants du SILC dans le registre national.

Expliquer en quoi les données demandées sont nécessaires à la réalisation des finalités et objectifs décrits ci-dessus (par catégorie de données, apporter la preuve de la proportionnalité). Il convient d'apporter la preuve que la communication des données demandées fait partie intégrante des objectifs statistiques poursuivis.		
……..	Besoin exact	Comme décrit ci-haut, les données SILC au niveau individuel seront nécessaires pour ce projet de 2 façon: -Les données du suivi de mortalité du SILC seront utilisées pour estimer les taux de mortalité par catégorie d'âge, sexe, et statut socio-économique (méthode 1). Nous avons déjà obtenu l'autorisation pour les données de suivi de mortalité pour les participants des Health Interview Surveys de 1997 et 2001 jusqu'en décembre 2008. Nous voulons comparer les résultats obtenus des 2 enquêtes et identifier la base de données optimale pour recommander son utilisation pour la surveillance des espérances de vie sans limitation par statut socio-économique au niveau Belge. -Les données de morbidité se trouvant dans SILC seront associées aux taux de mortalité pour estimer les AVBS selon le statut socio-économique.
	Preuve de proportionnalité	Les données de EU-SILC sont utilisés au niveau européen pour calculer des AVBS. Nous voulons tester la faisabilité d'utiliser cette même base de donnés pour estimer les AVBS par statut socio-économique en Belgique. Nous utiliserons

		plusieurs indicateurs pour refléter le statut socio-économique.
		………..
		………..

Durée de conservation des données[2]:

1 an

Mesures de sécurité
Quelles sont les mesures mises en place afin d'éviter que des données ne puissent être identifiées via la publication de résultats ?

Nous allons présenter tous nos résultats de manière agrégée. De cette façon, il n'est pas possible d'identifié les données à la publication des résultats.

Quelles sont les mesures mises en place afin de veiller à la protection et à la sécurité des données ?

-Les données seront stockées sur le réseau de l'institut qui est un réseau protégé.
-Ces données ne seront accessibles qu'aux chercheurs membres du projet et non pas à la totalité des chercheurs de l'institut.

[2] A l'échéance du contrat, les données et backups doivent être détruits. Un usage prolongé de ces données, pour les mêmes finalités, n'est pas autorisé, à moins que le contrat ne soit reconduit. Si les objectifs statistiques sont atteints avant l'expiration du terme initialement prévu, les données et backups devront être détruits avant l'expiration du terme, soit au moment où les objectifs statistiques sont atteints (voir article 3 du contrat de confidentialité).

Uniquement, pour les personnes physiques ou morales poursuivant un but de recherche scientifique (article 15, alinéa 1, 4°, de la loi du 4 juillet 1962) :

Description des méthodes d'analyse qui seront utilisées :

Expliquer en quoi le projet de recherche répond aux normes scientifiques en vigueur :

Fait à :Bruxelles Date :

Signature :

Nom :Herman Van Oyen
Fonction :Chef de Direction Opérationnelle Santé Publique et Surveillance

Numéro de contrat : ……….

CONTRAT DE CONFIDENTIALITE
ENTRE
LA DIRECTION GENERALE STATISTIQUE ET INFORMATION ECONOMIQUE ET ………………………………………..

Entre

La Direction générale de la Statistique et de l'Information économique, Rue de Louvain 44, 1000 Bruxelles, représenté par Monsieur Lambert VERJUS, Président du Comité de Direction, ci-après dénommé "Direction générale Statistique et Information économique", d'une part,

Et

Herman Van Oyen ci-après dénommé "Chercheur", d'autre part,

IL EST CONVENU :

Article 1 - Objet

La Direction générale Statistique et Information économique communique, en exécution des articles 15 et 15bis de la loi du 4 juillet 1962 relative à la statistique publique, au Chercheur les données d'étude codées indiquées en annexe 1 pour l'exécution du projet ……. Les objectifs et la durée de la recherche sont définis limitativement en annexe 2.

Le Chercheur s'engage à respecter toutes les obligations découlant de la loi du 4 juillet 1962 relative à la statistique publique, de ses arrêtés d'exécution et du présent contrat. Le présent contrat ne peut en aucun cas porter préjudice aux dispositions de la décision du Comité de surveillance statistique (délibération nr. … du …).

Article 2 - Exécutants de la recherche

La recherche sera exécutée par ………

Article 3 - Utilisation des données

Le Chercheur peut exclusivement utiliser les données d'étude codées communiquées pour les objectifs décrits en annexe 2. Il ne peut pas prendre plus de copies que ce qui est exigé pour ces objectifs.

Il est interdit au Chercheur de transmettre les données d'étude codées communiquées ou une partie de celles-ci à des tiers, sauf avec l'accord du Comité de surveillance statistique et de la Direction générale Statistique et Information économique qui prendra contact avec ce nouvel utilisateur avec qui un contrat de confidentialité sera établi. Cette communication ne peut avoir lieu qu'après autorisation et approbation du contrat de confidentialité précité par le Comité de surveillance statistique.

Le Chercheur peut seulement utiliser les données d'étude codées communiquées pendant la durée de la recherche mentionnée en annexe 2. Après cette période, les données et backups seront entièrement détruits par le Chercheur. Un usage prolongé de ces données d'étude codées, pour les mêmes finalités, n'est pas autorisé, à moins que le présent contrat ne soit reconduit. Si les objectifs statistiques décrits en annexe 2 sont atteints avant l'expiration du terme prévu dans le présent contrat, le Chercheur détruira les données et backups avant l'expiration du terme du présent contrat, soit au moment où les objectifs statistiques sont atteints.

Le Chercheur n'est autorisé à utiliser les données d'étude codées communiquées qu'en vue d'en faire des analyses, d'effectuer des études et d'établir des statistiques globales et anonymes. En aucun cas, les données d'étude codées communiquées ne peuvent être utilisées à des fins de contrôle ou de répression.

Le Chercheur s'engage à respecter la confidentialité des données d'étude et à veiller à ce que les données d'étude soient utilisées exclusivement par les membres de son personnel, en vue de l'exécution de la recherche visée en annexe 2.

Le Chercheur s'engage à veiller à la protection et à la sécurité des données d'étude et à ce que les données individuelles ne puissent pas être identifiées directement ou indirectement par le biais des résultats diffusés.

Article 4 - Obligations de la Direction générale Statistique et Information économique

La Direction générale Statistique et Information économique s'engage à mettre à disposition les données d'étude codées indiquées en annexe 1, pour les objectifs et pendant la période spécifiés en annexe 2, pour autant que les données nécessaires soient disponibles au sein de la Direction générale Statistique et Information économique.

Ces données d'étude codées seront mises à disposition du Chercheur par la Direction générale Statistique et Information économique :

- dans les 30 jours qui suivent la conclusion du présent contrat de confidentialité*
- en cas de contrat pluriannuel : dans les 30 jours qui suivent la diffusion de ces données par la Direction générale Statistique et Information économique*

(*biffer la mention inutile).

Article 5 - Prix

Les données d'étude codées indiquées en annexe 1 sont mises à la disposition du Chercheur par la Direction générale Statistique et Information économique contre le paiement de … € en tant que participation aux frais de la Direction générale Statistique et Information économique pour l'établissement et l'envoi des données. Le paiement est à effectuer, dans le mois qui suit la réception des données, sur le compte 679-2005886-23 ouvert au nom de la Direction générale Statistique et Information économique, avec référence au présent contrat.

Toutes les amendes et frais éventuels qui résulteraient du non-respect des obligations légales qui sont attachées au présent contrat sont exclusivement à charge du Chercheur.

Article 6 - Responsabilité de la Direction générale Statistique et Information économique

Les parties conviennent expressément que la Direction générale Statistique et Information économique n'est pas responsable des erreurs portant sur le contenu des données communiquées.

Article 7 - Dispositions spécifiques

Le Chercheur s'engage à mettre gratuitement les analyses, études et statistiques globales et anonymes à la disposition de la Direction générale Statistique et Information économique, qui pourra les utiliser librement.

Le Chercheur s'engage à signaler préalablement au directeur général de la Direction générale Statistique et Information économique toute situation qui, au regard des stipulations du présent contrat de confidentialité, pourrait donner lieu à doute ou ambiguïté : un arrangement serait alors recherché, tout en restant dans le cadre du présent contrat de confidentialité.

Les résultats ne peuvent être diffusés que sous une forme globale et anonyme. Au moins quinze jours avant leur diffusion, le Chercheur doit les soumettre à la Direction générale Statistique et Information économique, et cette dernière peut éventuellement en interdire la diffusion.
Le terme « diffusion » doit être entendu dans un sens large qui tienne compte de l'évolution de la société de l'information. Il couvre toute diffusion qu'elle se fasse de manière écrite, orale ou en ligne.

A chaque diffusion de ces données d'étude globales et anonymes, quelle que soit la forme de cette diffusion, la Direction générale Statistique et Information économique doit être citée comme source : « Source : SPF Economie – DG Statistique et Information économique ».

Article 8 - Responsable des données

Le Chercheur s'engage également à indiquer dans l'annexe 3 une personne physique qui est personnellement responsable du respect de toutes les obligations concernant l'exécution du présent contrat, de la loi du 4 juillet 1962 relative à la statistique publique et de ses arrêtés d'exécution.

Cette personne s'engage à contrôler effectivement l'utilisation légitime des données communiquées.

Article 9 - Contrôle par la Direction générale Statistique et Information économique et/ou le Comité de surveillance statistique

Le Chercheur accepte expressément que les représentants de la Direction générale Statistique et Information économique ou du Comité de surveillance statistique aient, à chaque instant et sans mise en demeure préalable, accès aux locaux et à l'infrastructure informatique où les données communiquées sont conservées, pour contrôler l'exécution des dispositions de la décision du Comité de surveillance statistique, de la loi du 4 juillet 1962 relative à la statistique publique, de ses arrêtés d'exécution et du présent contrat.
Sur simple demande, la Direction générale Statistique et Information économique ou le Comité de surveillance statistique peut obtenir accès aux autres systèmes ICT et locaux afin de contrôler si aucune violation au présent contrat n'est commise.

Article 10 - Sanctions

En cas de non-respect des dispositions du présent contrat de confidentialité, la Direction générale Statistique et Information économique se réserve le droit de résilier aussitôt le contrat de confidentialité au moyen d'une lettre recommandée, de réclamer au Chercheur des dommages et intérêts et de refuser de conclure tout autre contrat de confidentialité de ce type avec le Chercheur que ce soit pour l'année en cours ou pour les années suivantes.

La Direction générale Statistique et Information économique a le droit, sans être redevable d'aucune indemnité, de mettre fin au présent contrat de confidentialité à tout moment si pour des raisons techniques ou d'opportunité, la mise à disposition des données d'étude codée spécifiées en annexe 1 n'est plus possible, à titre provisoire ou définitif.

Le Chercheur a pris connaissance des articles 22 et 23 de la loi du 4 juillet 1962 relative à la statistique publique, dont une copie est jointe en annexe 4 du présent contrat de confidentialité.

Article 11 - Droit applicable – tribunal compétent

Seul le droit belge s'applique à ce contrat. En cas de différend, les tribunaux de Bruxelles sont exclusivement compétents.

Etabli à Bruxelles le en autant d'exemplaires que de parties au contrat, chacune reconnaissant en avoir reçu un exemplaire original.

Pour la Direction générale Statistique et Information économique,

Pour le Chercheur,

Lambert VERJUS
Président du Comité de Direction

Herman Van Oyen
Chef, Direction Opérationnelle,
Institut Scientifique de Santé Publique

Annexe 1 :

Définition des données demandées

Annexe 2 :

Description du thème de la recherche
Description des objectifs de la recherche
Indication de la durée de la recherche
Durée de conservation des données par le Chercheur
Indication de l'exécutant de la recherche
Fréquence de la recherche

Annexe 3 :

Données d'identification de la personne physique responsable
Signature de la personne physique responsable

Annexe 4 :

Extrait de la loi du 4 juillet 1962 relative à la statistique publique

ANNEXE 1:

Définition des données demandées :

Nous souhaitons utiliser les données de l'enquête sur les revenus et les conditions de vie (SILC) de 2004 pour évaluer les questions décrites ci-haut. Pour cela, nous demandons les données suivantes :

1-Les données SILC au niveau du ménage :
Situation de logement actuelle: H5, H7, H8, H9, H10, H11, H12, H50
Situation financière: H51, H52, H53, H55, H56, H57, H58, H59, H60, H61, H62, H63
Revenus perçus par votre ménage: H64, H65, H66, H67, H68, H69, H70, H71

2- Les données SILC au niveau individuel:
Activités professionnelles: I5, I 6, I7, I8, I9, I10, I11, I19, I 20, I 21, I22, I25, I26, I27, I28, I29, I30, I31, I32, I33, I39, I40
Santé : I134, I135, I136, I137, I138, I193, I140
Informations biographiques: I149, I150, I151, I152, I153, I154, I155, I156, I157, I158
Formation scolaire: I159, I160, I161, I163, I164, I165, I166, I 167, I168, I 169, I 170, I171
Indicateurs SILC: Risque de pauvreté, difficultés à joindre les deux bouts
Autres : âge, sexe, région, nombre de personne par ménage, âge de chaque personne du ménage, pondération

3-Le suivi mortalité des données SILC de 2004:
De plus, nous désirerons avoir une variable supplémentaire avec les données décrites ci-haut qui précise la date de décès ou d'émigration des répondants qui sont décédés ou ont émigré avant le 31/12/2009. Cette variable est obtenue par la vérification du statut vital des répondants du SILC dans le registre national.

(Date et signature)

ANNEXE 2:

Thème de la recherche	Estimation des espérances de vie en bonne santé par statut socioéconomique
Objectifs de la recherche	L'objectif de ce projet est d'identifier une base de données alternative au recensement national pour estimer les années de vie en bonne santé (AVBS) selon le statut socio-économique.
Durée de la recherche	1 an de 01/01/2010 à 31/12/2010
Durée de conservation des données (par le Chercheur)	1 an
Exécutants de la recherche	Herman Van Oyen, Rana Charafeddine, Stefaan Demarest, Patrick Deboosere
Fréquence de la recherche	Ponctuelle

(Date et signature)

ANNEXE 3:

Madame : Rana Charafeddine

Collaborateur Scientifique
Rana.charafeddine@iph.fgov.be
Tél : 02-642 5739

Numéro de registre national : XXXXXXXXXXX

Signature de la personne physique responsable,

ANNEXE 4:

LOI DU 4 JUILLET 1962
RELATIVE A LA STATISTIQUE PUBLIQUE (*M.B.* 20-07-1962)
(extrait)

Dispositions pénales.

Article 22.- Est puni d'une amende de 26 € à 10.000 € :

1° Celui qui, étant tenu de fournir des renseignements en vertu de la présente loi et des arrêtés pris pour l'exécution de celle-ci, ne remplit pas les obligations qui lui sont imposées;

2° Celui qui s'oppose aux recherches et constatations visées à l'article 19 ou à l'exécution d'office prévue à l'article 20 ou entrave l'activité des personnes chargées des recherches et constatations ou de l'exécution d'office;

3° Celui qui utilise à des fins non admises par la présente loi les données individuelles recueillies en vertu de la présente loi ou les données globales mais confidentielles visées à l'article 2, *littera* c, deuxième alinéa;

4° Celui qui viole les obligations de faire ou de ne pas faire imposées, en matière de collecte de données statistiques, par un acte juridique directement applicable émanant d'un organe de l'Union européenne.

La peine est doublée et un emprisonnement de huit jours à un mois peut en outre être prononcé, si l'infraction a été commise dans les cinq ans à compter du jour où une condamnation antérieure, du chef de l'une des infractions prévues par le présent article, est devenue irrévocable.

Article 23.- Les dispositions du livre I du Code pénal, sans exception du chapitre VII et de l'article 85, sont applicables aux infractions prévues par l'article 22.

Pris connaissance dans le cadre du contrat de confidentialité entre la Direction générale Statistique et Information économique et le Chercheur

(Date et signature)

Appendix 2: Statistical programs used

I-Preparation of the mortality data using the STATA software

This example is based on the SILC data.

```
*********************************
merging different SILC datasets : match_rr has the mortality follow up
information, the p_file has the information on the individual level for
those ages 16 years and older, the r_file
********************************

merge  rb030 using "X:\Hlifeexp\HLY calculation\SILC\match_rr.dta"
merge rb030 using "X:\Hlifeexp\HLY calculation\SILC\p-file_wiv.dta"
merge rb030 using "X:\Hlifeexp\HLY calculation\SILC\r-file_wiv.dta"


save "X:\Hlifeexp\HLY calculation\SILC\Original_merged.dta"


***************************************************************************
recodding variables (have 371 unmatched observations due to mismatch SILC
and national register)
These variables are needed to set the survey data as a cohort: DOE (data of
exit from the cohort), DOO (date of origin that is date of interview), and
DOB (date of birth).
***************************************************************************


drop if  geboortedatum ==.

generate status=1
replace status=0 if lft==.

tostring datdec, replace

generate DOE=date(datdec,"DMY")
replace DOE=mdy(12,31,2009) if status==0
format %td DOE

generate DOO = mdy(PB100, PB090, PB110)
format %td DOO

tostring geboortedatum, replace
generate DOB=date( geboortedatum , "YMD")
format %td DOB


generate id2=_n

*********************************************************************
* set data for cohort analysis using current age and not age of entry and
Lexis expansion
*********************************************************************


stset DOE, fail(status) id(id2) origin(DOB) enter(DOO) scale(365.25)

stsplit curage, at(0(1)85)
```

```
**********************************
* Calculate person years
*********************************

/*person years for the split database*/

generate pyears2=_t - _t0


**********************************
* To prepare for HLY calculation
*********************************

generate X=curage

generate Dx=_d

generate Px=pyears2

recode PE040 (0/1=1) (2=2)(3/4=3) (5/6=4), gen(educ4)

save "X:\Hlifeexp\HLY calculation\SILC\Original_merged_V2.dta", replace



********************************************************
* To generate stratified databases to be exported to SAS
********************************************************

keep if sex==1 & (PE040==0 | PE040==1)
save "X:\Hlifeexp\HLY calculation\SILC\men educ1\men_educ1.dta"

keep if sex==1 & PE040==2
save "X:\Hlifeexp\HLY calculation\SILC\men educ2\men_educ2.dta"

keep if sex==1 & (PE040==3 | PE040==4)
save "X:\Hlifeexp\HLY calculation\SILC\men educ3\men_educ3.dta"

keep if sex==1 & (PE040==5 | PE040==6)
save "X:\Hlifeexp\HLY calculation\SILC\men educ4\men_educ4.dta"

keep if sex==2 & (PE040==0 | PE040==1)
save "X:\Hlifeexp\HLY calculation\SILC\women educ1\women_educ1.dta"

keep if sex==2 & PE040==2
save "X:\Hlifeexp\HLY calculation\SILC\women educ2\women_educ2.dta"

keep if sex==2 & (PE040==3 | PE040==4)
save "X:\Hlifeexp\HLY calculation\SILC\women educ3\women_educ3.dta"

keep if sex==2 & (PE040==5 | PE040==6)
save "X:\Hlifeexp\HLY calculation\SILC\women educ4\women_educ4.dta"
```

II-Estimation of Healthy Life Expectancy using the SAS software

For more details concerning this program, please refer to the manual developed by the EHEMU project: "Health Expectancy Calculation by the Sullivan method: A practical guide" found on the following link: http://www.ehemu.eu/pdf/Sullivan_guide_final_jun2007.pdf

Analyses were done by gender and educational level. This is the program used for estimating HE for males in the lowest educational category.

```
****************************************************************************
* Analysis for males in the educational level 1 (Primary education or
less);
****************************************************************************

;
/******** Calculation of Life Expectancy (LE) ********/


DATA hlymen_educ1;
SET mysas.men_educ1;
RETAIN x agegr Px Dx;

IF  x=0                        THEN agegr=0;  IF (x ge 1 and x le 4)
THEN agegr=1;
IF (x ge 5 and x le 9)   THEN agegr=5;  IF (x ge 10 and x le 14) THEN
agegr=10;
IF (x ge 15 and x le 19) THEN agegr=15; IF (x ge 20 and x le 24) THEN
agegr=20;
IF (x ge 25 and x le 29) THEN agegr=25; IF (x ge 30 and x le 34) THEN
agegr=30;
IF (x ge 35 and x le 39) THEN agegr=35; IF (x ge 40 and x le 44) THEN
agegr=40;
IF (x ge 45 and x le 49) THEN agegr=45; IF (x ge 50 and x le 54) THEN
agegr=50;
IF (x ge 55 and x le 59) THEN agegr=55; IF (x ge 60 and x le 64) THEN
agegr=60;
IF (x ge 65 and x le 69) THEN agegr=65; IF (x ge 70 and x le 74) THEN
agegr=70;
IF (x ge 75 and x le 79) THEN agegr=75; IF (x ge 80 and x le 84) THEN
agegr=80;
IF (x ge 85)                  THEN agegr=85;

run;


proc sort data=hlymen_educ1;
by agegr;
run;

PROC MEANS DATA=hlymen_educ1 IDMIN SUM;
VAR Px Dx;
BY agegr;
*ID x Px1 Dx2 B pix;
OUTPUT OUT=mysas.hlymen_educ1 (DROP=_type_ _freq_) SUM(Px)=Px SUM(Dx)=Dx;
RUN;
```

```
DATA mysas.hlymen_educ1;
RETAIN x agegr Px Dx nmx ax nqx nrx;
SET mysas.hlymen_educ1;
nmx=Dx/Px;
if agegr =0 then ax=0.09;
if agegr ne 0 then ax=0.5;

nqx=(5*nmx)/(1+((1-ax)*5*nmx));

IF agegr=1 THEN nqx=(4*nmx)/(1+((1-ax)*4*nmx));
IF agegr=85 THEN nqx=.;
nrx=1-nqx;
FORMAT nmx 8.6 nqx 10.8;
RUN;

PROC EXPAND DATA=mysas.hlymen_educ1 OUT=mysas.hlymen_educ1 (DROP=time)
METHOD=NONE;
CONVERT nrx=nrxrx / TRANSFORM=(PRODUCT);
RUN;

DATA mysas.hlymen_educ1;
RETAIN x agegr Px Dx nmx ax nqx lx pix;
SET mysas.hlymen_educ1;
lagnrxrx=LAG(nrxrx);
IF agegr=15 THEN lx=100000;
ELSE lx=100000*lagnrxrx;
DROP nrx nrxrx lagnrxrx;
FORMAT lx 8.1;
RUN;

PROC SORT DATA=mysas.hlymen_educ1 OUT=mysas.hlymen_educ1;
BY DESCENDING agegr;
RUN;

DATA mysas.hlymen_educ1;
RETAIN x agegr Px Dx nmx ax nqx lx nLx;
SET mysas.hlymen_educ1;
laglx=LAG(lx);
nLx=((lx*5*ax)+(laglx*5*(1-ax)));
IF agegr=0 THEN nLx=((lx*0.2)+(laglx*0.8));
IF agegr=1 THEN nLx=((lx*4*ax)+(laglx*4*(1-ax)));
IF agegr=85 THEN nLx=lx/nmx;
DROP laglx;
FORMAT nLx 9.1;
RUN;

PROC EXPAND DATA=mysas.hlymen_educ1 OUT=mysas.hlymen_educ1 (DROP=time)
METHOD=NONE;
CONVERT nLx=Tx / TRANSFORM=(SUM);
RUN;
PROC SORT DATA=mysas.hlymen_educ1 OUT=mysas.hlymen_educ1;
BY agegr;
RUN;

DATA mysas.hlymen_educ1;
RETAIN x agegr Px Dx nmx ax nqx lx nLx Tx ex;
SET mysas.hlymen_educ1;
ex=Tx/lx;
FORMAT ex 4.1;
RUN;
```

```
******** merging databases to use health prevelance with life
expectancy********/;

PROC SORT DATA=mysas.hlymen_educ1 OUT=mysas.hlymen_educ1;
  BY agegr;
RUN;

PROC SORT DATA=mysas.prev_men_educ1 OUT=mysas.prev_men_educ1;
  BY agegr;
RUN;

DATA mysas.hlymen_educ1_2 ;
  MERGE mysas.hlymen_educ1 mysas.prev_men_educ1;
  BY agegr;
RUN;

DATA mysas.hlymen_educ1_2;
SET mysas.hlymen_educ1_2;
pix=GALI_p2;
if pix=. then pix=0;
Nx=GALI_n;
if Nx=. then Nx=0;
RETAIN agegr Px Dx nmx ax nqx lx nLx Tx ex pix Nx;
KEEP agegr Px Dx nmx ax nqx lx nLx Tx ex pix Nx;
run;



/******** Calculation of Disability-Free Life Expectancy (DFLE) ********/

DATA mysas.hlymen_educ1_2;
SET mysas.hlymen_educ1_2;
nLx_disafr=nLx*(1-pix);
FORMAT nLx_disafr 9.1;
RUN;

PROC SORT DATA=mysas.hlymen_educ1_2 OUT=mysas.hlymen_educ1_2;
BY DESCENDING agegr;
RUN;

PROC EXPAND DATA=mysas.hlymen_educ1_2 OUT=mysas.hlymen_educ1_2 (DROP=time)
METHOD=NONE;
CONVERT nLx_disafr=Tx_disafr / TRANSFORM=(SUM);
RUN;

PROC SORT DATA=mysas.hlymen_educ1_2 OUT=mysas.hlymen_educ1_2;
BY agegr;
RUN;

DATA mysas.hlymen_educ1_2;
RETAIN agegr Px Dx nmx ax nqx lx nLx Tx ex Nx pix nLx_disafr Tx_disafr
DFLEx percDFLEx;
SET mysas.hlymen_educ1_2;
DFLEx=Tx_disafr/lx;
```

```
percDFLEx=100*(DFLEx/ex);
FORMAT DFLEx percDFLEx 4.1 ;
RUN;


/******** Approximate standard errors of the prevalence rates********/



DATA mysas.hlymen_educ1_3;
set mysas.hlymen_educ1_2;
RETAIN agegr Px ax Dx nqx lx nLx ex pix nLx_disafr Tx_disafr DFLEx Nx
S2_pix L2S2_pix;

S2_pix=(pix*(1-pix))/Nx;
L2S2_pix=(nLx**2)*S2_pix;

KEEP agegr Px ax Dx nqx lx nLx ex pix nLx_disafr Tx_disafr DFLEx Nx S2_pix
L2S2_pix;
FORMAT S2_pix 8.6 L2S2_pix 9.0;
RUN;

PROC SORT DATA=mysas.hlymen_educ1_3 OUT=mysas.hlymen_educ1_3;
BY DESCENDING agegr;
RUN;

PROC EXPAND DATA=mysas.hlymen_educ1_3 OUT=mysas.hlymen_educ1_3(DROP=time)
METHOD=NONE;
CONVERT L2S2_pix=T_L2S2_pix/TRANSFORM=(SUM);
RUN;

PROC SORT DATA=mysas.hlymen_educ1_3 OUT=mysas.hlymen_educ1_3;
BY agegr;
RUN;

DATA mysas.hlymen_educ1_3;
RETAIN agegr Px ax Dx nqx lx nLx ex pix nLx_disafr Tx_disafr DFLEx Nx
S2_pix L2S2_pix
         T_L2S2_pix S2_DFLEx S_DFLEx;
SET mysas.hlymen_educ1_3;

S2_DFLEx=T_L2S2_pix/(lx**2);

S_DFLEx=sqrt(S2_DFLEx);

FORMAT S2_pix 8.6 L2S2_pix T_L2S2_pix 10.0 S2_DFLEx 7.5 S_DFLEx 5.3;
RUN;


/******** Standard errors taking into account the variance of the mortality
rates ********/

DATA mysas.hlymen_educ1_4;
SET mysas.hlymen_educ1_3;

S2_DFLEx_1=S2_DFLEx;

S2_qx=((nqx**2)*(1-nqx))/Dx;

IF agegr=85 THEN S2_qx=0;
```

```
KEEP agegr Px ax Dx nqx lx ex Nx pix DFLEx S2_DFLEx_1 S2_qx;
FORMAT S2_DFLEx_1 7.5 S2_qx 10.8;
RUN;

PROC SORT DATA=mysas.hlymen_educ1_4 OUT=mysas.hlymen_educ1_4;
BY DESCENDING agegr;
RUN;

DATA mysas.hlymen_educ1_4;
SET mysas.hlymen_educ1_4;
lagDFLEx=LAG(DFLEx);
RUN;

PROC SORT DATA=mysas.hlymen_educ1_4 OUT=mysas.hlymen_educ1_4;
BY agegr;
RUN;

DATA mysas.hlymen_educ1_4;
SET mysas.hlymen_educ1_4;
y=(((1-ax)*5*(1-pix))+lagDFLEx);
IF agegr=0 THEN y=(((1-ax)*(1-pix))+lagDFLEx);
IF agegr=1 THEN y=(((1-ax)*4*(1-pix))+lagDFLEx);
IF agegr=85 THEN y=0;
z=((y**2)*(lx**2)*(S2_qx));
DROP lagDFLEx;
FORMAT y 7.4 z 12.3;
RUN;

PROC SORT DATA=mysas.hlymen_educ1_4;
BY DESCENDING agegr;
RUN;

PROC EXPAND DATA=mysas.hlymen_educ1_4 OUT=mysas.hlymen_educ1_4 (DROP=time)
METHOD=NONE;
CONVERT z=T_z / TRANSFORM=(SUM);
RUN;

PROC SORT DATA=mysas.hlymen_educ1_4 OUT=mysas.hlymen_educ1_4;
BY agegr;
RUN;

DATA mysas.hlymen_educ1_4;
RETAIN x agegr Px ax Dx nqx lx ex Nx pix DFLEx S2_DFLEx_1 S2_qx y z T_z
S2_DFLEx_2 S2_DFLEx;
SET mysas.hlymen_educ1_4;
S2_DFLEx_2=T_z/(lx**2);
S2_DFLEx=S2_DFLEx_1+S2_DFLEx_2;
FORMAT S2_DFLEx_2 S2_DFLEx 7.5;
RUN;

DATA mysas.hlymen_educ1_5;
SET mysas.hlymen_educ1_4;
DROP ax lx nqx S2_qx y z T_z;
sex=1;
educ=1;
survey=1;
run;
```

Appendix 3: Final tables

Tables of life expectancy and healthy life years by age groups, sex and education using the HIS and the SILC

			HIS			SILC		
Age groups	**Sex**	**Education**	**Life Expectancy**	**HLY**	**Variance of HLY**	**Life Expectancy**	**HLY**	**Variance of HLY**
15	Male	Primary	59,12	44,55	2,1	57,62	35,14	2,28
20	Male	Primary	54,12	40	2,07	52,62	30,74	2,25
25	Male	Primary	49,12	35,49	2,03	48,69	28,44	1,58
30	Male	Primary	45,6	32,4	0,9	43,69	25,11	1,29
35	Male	Primary	40,6	28,23	0,83	38,69	21,41	1,11
40	Male	Primary	35,6	23,82	0,78	33,69	18,46	0,98
45	Male	Primary	30,6	19,95	0,72	29,33	15,7	0,81
50	Male	Primary	25,6	16,06	0,66	24,73	13,33	0,68
55	Male	Primary	21,08	13,03	0,52	21,18	11,07	0,53
60	Male	Primary	17,15	10,51	0,37	17,8	8,9	0,41
65	Male	Primary	13,38	7,49	0,27	15,03	7,38	0,32
70	Male	Primary	10,98	5,81	0,17	12,32	5,57	0,28
75	Male	Primary	8,13	3,75	0,13	10,11	4,05	0,31
80	Male	Primary	6,43	2,47	0,12	6,86	2,53	0,32
85	Male	Primary	4,7	2,27	0,07	5,25	1,42	0,68

Age groups	Sex	Education	HIS Life Expectancy	HLY	Variance of HLY	SILC Life Expectancy	HLY	Variance of HLY
15	Male	Lower secondary	58,35	43,88	2,30	60,69	41,78	2,39
20	Male	Lower secondary	54,29	39,89	1,86	55,69	37,17	2,38
25	Male	Lower secondary	50,20	35,72	1,47	50,69	33,42	2,28
30	Male	Lower secondary	45,20	30,96	1,44	45,69	29,82	2,15
35	Male	Lower secondary	40,20	26,69	1,37	41,78	26,89	1,63
40	Male	Lower secondary	35,93	23,33	1,13	37,50	23,74	1,42
45	Male	Lower secondary	31,51	19,81	0,97	32,89	19,97	1,35
50	Male	Lower secondary	27,35	16,67	0,82	27,89	15,91	1,29
55	Male	Lower secondary	23,49	13,44	0,68	23,98	13,23	1,19
60	Male	Lower secondary	20,00	11,13	0,53	20,73	11,26	1,09
65	Male	Lower secondary	16,22	7,85	0,43	17,93	9,80	1,01
70	Male	Lower secondary	12,66	4,99	0,38	13,69	6,79	0,88
75	Male	Lower secondary	9,60	3,69	0,33	9,63	4,55	0,75
80	Male	Lower secondary	6,54	2,57	0,30	6,22	3,34	0,64
85	Male	Lower secondary	3,53	0,41	0,11	5,52	2,29	1,23

Age groups	Sex	Education	HIS Life Expectancy	HLY	Variance of HLY	SILC Life Expectancy	HLY	Variance of HLY
15	Male	Higher secondary	62,56	50,87	1,76	63,03	46,21	0,98
20	Male	Higher secondary	58,15	46,59	1,55	58,03	41,62	0,96
25	Male	Higher secondary	53,67	42,24	1,40	53,03	37,04	0,95
30	Male	Higher secondary	48,67	37,69	1,39	48,30	32,97	0,91
35	Male	Higher secondary	43,67	33,34	1,36	43,83	29,14	0,85
40	Male	Higher secondary	38,67	28,92	1,35	39,03	25,18	0,82
45	Male	Higher secondary	33,91	25,00	1,31	34,21	21,28	0,80
50	Male	Higher secondary	29,60	21,07	1,26	29,89	18,22	0,76
55	Male	Higher secondary	25,64	17,93	1,17	25,37	15,11	0,71
60	Male	Higher secondary	21,48	14,97	1,08	20,92	11,74	0,67
65	Male	Higher secondary	17,57	11,51	1,02	17,13	9,06	0,63
70	Male	Higher secondary	13,95	9,08	0,97	13,55	6,73	0,57
75	Male	Higher secondary	11,10	6,96	1,02	9,76	4,06	0,51
80	Male	Higher secondary	7,94	5,13	1,15	7,37	3,16	0,63
85	Male	Higher secondary	8,55	6,73	2,04	5,88	2,61	1,07

Age groups	Sex	Education	HIS Life Expectancy	HLY	Variance of HLY	SILC Life Expectancy	HLY	Variance of HLY
15	Male	Higher education	64,42	51,79	1,27	68,09	51,87	2,08
20	Male	Higher education	59,42	47,02	1,26	63,09	46,87	2,08
25	Male	Higher education	55,06	42,65	1,02	58,09	42,28	2,06
30	Male	Higher education	50,06	37,96	1,01	53,09	38,00	2,03
35	Male	Higher education	45,35	33,80	0,96	48,40	33,84	1,99
40	Male	Higher education	40,35	29,33	0,95	43,40	29,35	1,98
45	Male	Higher education	35,57	25,06	0,92	38,40	25,12	1,96
50	Male	Higher education	30,57	20,82	0,90	34,21	21,63	1,92
55	Male	Higher education	26,01	16,62	0,87	29,72	17,85	1,90
60	Male	Higher education	21,60	13,19	0,82	25,79	14,69	1,93
65	Male	Higher education	17,52	9,82	0,77	21,78	11,28	1,96
70	Male	Higher education	13,29	6,48	0,70	17,69	8,20	1,98
75	Male	Higher education	11,11	5,46	0,66	15,07	6,29	2,30
80	Male	Higher education	7,90	4,04	0,59	11,24	4,37	2,31
85	Male	Higher education	4,46	2,79	0,47	6,68	1,86	2,24

Age groups	Sex	Education	HIS Life Expectancy	HLY	Variance of HLY	SILC Life Expectancy	HLY	Variance of HLY
15	Female	Primary	65,54	41,72	1,64	66,98	36,47	2,09
20	Female	Primary	63,50	38,67	1,77	63,56	33,11	1,43
25	Female	Primary	58,50	34,25	1,68	58,56	30,37	1,16
30	Female	Primary	53,50	30,56	1,40	53,56	27,68	0,86
35	Female	Primary	48,50	27,68	1,13	48,56	23,84	0,74
40	Female	Primary	43,50	23,61	1,03	43,56	20,61	0,63
45	Female	Primary	38,50	20,50	0,86	38,56	18,03	0,54
50	Female	Primary	33,50	17,42	0,70	33,56	14,50	0,47
55	Female	Primary	28,50	14,83	0,58	29,76	12,37	0,35
60	Female	Primary	23,50	11,70	0,47	24,76	9,70	0,30
65	Female	Primary	19,54	8,62	0,36	20,80	7,56	0,23
70	Female	Primary	16,13	6,39	0,29	16,42	5,39	0,19
75	Female	Primary	13,15	4,48	0,26	13,33	4,00	0,17
80	Female	Primary	9,71	3,38	0,26	9,78	2,09	0,13
85	Female	Primary	6,85	1,54	0,19	7,77	1,40	0,18

Age groups	Sex	Education	HIS Life Expectancy	HLY	Variance of HLY	SILC Life Expectancy	HLY	Variance of HLY
15	Female	Lower secondary	68,84	49,68	1,45	69,58	43,78	3,16
20	Female	Lower secondary	63,84	45,04	1,41	64,58	39,59	3,13
25	Female	Lower secondary	58,84	40,14	1,40	61,47	36,75	2,51
30	Female	Lower secondary	53,84	35,46	1,36	56,47	33,58	2,35
35	Female	Lower secondary	48,84	30,72	1,33	51,47	30,12	2,23
40	Female	Lower secondary	43,84	26,69	1,27	46,47	26,94	2,14
45	Female	Lower secondary	38,84	22,61	1,22	41,47	23,66	2,07
50	Female	Lower secondary	34,35	18,58	1,13	36,47	20,40	2,00
55	Female	Lower secondary	31,72	16,29	0,90	31,92	17,87	1,91
60	Female	Lower secondary	26,72	12,79	0,84	28,14	14,98	1,85
65	Female	Lower secondary	22,41	9,91	0,75	23,84	11,90	1,78
70	Female	Lower secondary	18,41	6,66	0,73	19,22	9,01	1,72
75	Female	Lower secondary	15,08	4,85	0,74	15,29	6,65	1,79
80	Female	Lower secondary	11,59	3,28	0,70	11,18	4,67	1,81
85	Female	Lower secondary	7,83	2,02	0,69	8,67	5,38	2,53

Age groups	Sex	Education	HIS Life Expectancy	HLY	Variance of HLY	SILC Life Expectancy	HLY	Variance of HLY
15	Female	Higher secondary	69,67	53,17	1,64	69,10	46,16	1,54
20	Female	Higher secondary	64,67	48,86	1,61	64,10	41,68	1,52
25	Female	Higher secondary	59,67	44,06	1,60	59,10	37,33	1,51
30	Female	Higher secondary	54,67	39,39	1,59	54,10	33,44	1,48
35	Female	Higher secondary	50,08	35,06	1,52	49,10	29,71	1,45
40	Female	Higher secondary	45,08	30,79	1,50	44,10	25,83	1,42
45	Female	Higher secondary	40,08	26,56	1,47	39,34	22,05	1,40
50	Female	Higher secondary	35,36	22,93	1,43	34,77	18,55	1,37
55	Female	Higher secondary	30,36	19,41	1,38	30,40	15,20	1,35
60	Female	Higher secondary	26,56	16,01	1,31	26,31	12,27	1,34
65	Female	Higher secondary	22,18	12,53	1,25	22,24	9,26	1,32
70	Female	Higher secondary	18,07	10,19	1,19	18,74	6,64	1,36
75	Female	Higher secondary	13,60	7,18	1,12	15,12	4,77	1,38
80	Female	Higher secondary	10,14	5,39	1,22	11,21	2,85	1,30
85	Female	Higher secondary	7,20	3,94	1,29	8,03	1,18	1,61

Age groups	Sex	Education	HIS Life Expectancy	HLY	Variance of HLY	SILC Life Expectancy	HLY	Variance of HLY
15	Female	Higher education	73,37	60,07	2,60	-	-	-
20	Female	Higher education	68,37	55,32	2,59	66,85	47,42	3,08
25	Female	Higher education	63,37	50,48	2,58	61,85	42,73	3,07
30	Female	Higher education	58,37	45,75	2,58	56,85	38,29	3,06
35	Female	Higher education	53,37	41,07	2,57	51,85	34,04	3,04
40	Female	Higher education	48,37	36,41	2,56	47,09	29,97	3,03
45	Female	Higher education	43,37	31,79	2,55	42,09	25,83	3,01
50	Female	Higher education	38,64	28,03	2,52	37,57	22,18	3,01
55	Female	Higher education	34,17	24,35	2,49	32,57	18,75	2,97
60	Female	Higher education	29,46	20,67	2,45	28,43	15,37	3,00
65	Female	Higher education	24,81	16,95	2,39	24,62	12,57	3,05
70	Female	Higher education	21,00	14,90	2,30	19,62	9,63	2,90
75	Female	Higher education	16,47	11,77	2,13	16,29	8,15	2,92
80	Female	Higher education	13,55	9,61	2,00	12,59	5,75	2,97
85	Female	Higher education	10,59	8,25	1,75	8,25	3,19	2,69

Difference between HLY in the lowest and highest educational categories among men in the HIS

Age groups	HLY Low	Variance HLY low	HLY high	Variance HLY high	Difference HLY	Difference variance	z	pvalue
15	44,55	2,10	51,79	1,27	7,23	3,37	2,15	0,0319
20	40,00	2,07	47,02	1,26	7,02	3,32	2,11	0,0345
25	35,49	2,03	42,65	1,02	7,16	3,05	2,35	0,0188
30	32,40	0,90	37,96	1,01	5,56	1,91	2,91	0,0036
35	28,23	0,83	33,80	0,96	5,57	1,79	3,11	0,0019
40	23,82	0,78	29,33	0,95	5,51	1,73	3,19	0,0014
45	19,95	0,72	25,06	0,92	5,11	1,64	3,12	0,0018
50	16,06	0,66	20,82	0,90	4,75	1,56	3,05	0,0023
55	13,03	0,52	16,62	0,87	3,60	1,39	2,59	0,0096
60	10,51	0,37	13,19	0,82	2,68	1,19	2,25	0,0244
65	7,49	0,27	9,82	0,77	2,33	1,04	2,25	0,0246
70	5,81	0,17	6,48	0,70	0,67	0,87	0,77	0,4438
75	3,75	0,13	5,46	0,66	1,71	0,79	2,17	0,0297
80	2,47	0,12	4,04	0,59	1,56	0,71	2,20	0,0277
85	2,27	0,07	2,79	0,47	0,52	0,53	0,98	0,3267

Difference between HLY in the lowest and highest educational categories among women in the HIS

Age groups	HLY Low	Variance HLY low	HLY high	Variance HLY high	Difference HLY	Difference variance	z	pvalue
15	41,72	1,64	60,07	2,60	18,36	4,24	4,33	0,0000
20	38,67	1,77	55,32	2,59	16,65	4,37	3,81	0,0001
25	34,25	1,68	50,48	2,58	16,24	4,26	3,81	0,0001
30	30,56	1,40	45,75	2,58	15,20	3,97	3,82	0,0001
35	27,68	1,13	41,07	2,57	13,38	3,70	3,62	0,0003
40	23,61	1,03	36,41	2,56	12,80	3,60	3,56	0,0004
45	20,50	0,86	31,79	2,55	11,28	3,41	3,31	0,0009
50	17,42	0,70	28,03	2,52	10,61	3,22	3,30	0,0010
55	14,83	0,58	24,35	2,49	9,52	3,08	3,09	0,0020
60	11,70	0,47	20,67	2,45	8,97	2,92	3,08	0,0021
65	8,62	0,36	16,95	2,39	8,33	2,76	3,02	0,0025
70	6,39	0,29	14,90	2,30	8,52	2,59	3,29	0,0010
75	4,48	0,26	11,77	2,13	7,29	2,39	3,04	0,0023
80	3,38	0,26	9,61	2,00	6,24	2,26	2,76	0,0058
85	1,54	0,19	8,25	1,75	6,71	1,94	3,46	0,0005

Difference between HLY in the lowest and highest educational categories among men in the SILC

Age groups	HLY Low	Variance HLY low	HLY high	Variance HLY high	Difference HLY	Difference variance	z	pvalue
15	35,14	2,28	51,87	2,08	16,72	4,35	3,84	0,0001
20	30,74	2,25	46,87	2,08	16,12	4,32	3,73	0,0002
25	28,44	1,58	42,28	2,06	13,84	3,64	3,80	0,0001
30	25,11	1,29	38,00	2,03	12,88	3,32	3,88	0,0001
35	21,41	1,11	33,84	1,99	12,43	3,11	4,00	0,0001
40	18,46	0,98	29,35	1,98	10,88	2,96	3,68	0,0002
45	15,70	0,81	25,12	1,96	9,42	2,77	3,40	0,0007
50	13,33	0,68	21,63	1,92	8,30	2,60	3,19	0,0014
55	11,07	0,53	17,85	1,90	6,78	2,43	2,79	0,0053
60	8,90	0,41	14,69	1,93	5,79	2,34	2,47	0,0134
65	7,38	0,32	11,28	1,96	3,89	2,28	1,71	0,0871
70	5,57	0,28	8,20	1,98	2,63	2,26	1,17	0,2432
75	4,05	0,31	6,29	2,30	2,24	2,61	0,86	0,3919
80	2,53	0,32	4,37	2,31	1,84	2,63	0,70	0,4845
85	1,42	0,68	1,86	2,24	0,44	2,92	0,15	0,8813

Difference between HLY in the lowest and highest educational categories among women in the SILC

Age groups	HLY Low	Variance HLY low	HLY high	Variance HLY high	Difference HLY	Difference variance	z	pvalue
15	36,47	2,09	-	-	-	-	-	-
20	33,11	1,43	47,42	3,08	14,31	4,51	3,18	0,0015
25	30,37	1,16	42,73	3,07	12,36	4,23	2,92	0,0035
30	27,68	0,86	38,29	3,06	10,60	3,92	2,71	0,0068
35	23,84	0,74	34,04	3,04	10,19	3,79	2,69	0,0071
40	20,61	0,63	29,97	3,03	9,36	3,66	2,56	0,0105
45	18,03	0,54	25,83	3,01	7,80	3,55	2,20	0,0279
50	14,50	0,47	22,18	3,01	7,67	3,49	2,20	0,0277
55	12,37	0,35	18,75	2,97	6,38	3,32	1,92	0,0544
60	9,70	0,30	15,37	3,00	5,67	3,30	1,72	0,0857
65	7,56	0,23	12,57	3,05	5,01	3,28	1,53	0,1264
70	5,39	0,19	9,63	2,90	4,25	3,09	1,37	0,1693
75	4,00	0,17	8,15	2,92	4,14	3,09	1,34	0,1795
80	2,09	0,13	5,75	2,97	3,66	3,10	1,18	0,2383
85	1,40	0,18	3,19	2,69	1,79	2,87	0,62	0,5329

References

Borrell, C., Regidor, E., Arias, L. C., Navarro, P., Puigpinos, R., Dominguez, V., & Plasencia, A. (1999). Inequalities in mortality according to educational level in two large Southern European cities. *International Journal of Epidemiology, 28*(1), 58-63.

Boshuizen, H., & van de Water, H. (1994). An international comparison of health expectancies. (pp.1-146). Leiden: TNO.

Bossuyt, N., Gadeyne, S., Deboosere, P., & Van Oyen, H. (2004). Socio-economic inequalities in health expectancy in Belgium. *Public Health*, *118*(1), 3-10.

Bossuyt, N., & Van Oyen, H. (2000). [Health expectancy by socio-economic gradient in Belgium]. (pp.1-96). Brussels: Scientific Institute of Public Health.

Bronnum-Hansen, H. (2005). Health expectancy in Denmark, 1987-2000. *The European Journal of Public Health*, *15* 20-25.

Bronnum-Hansen, H., & Baadsgaard, M. (2007). Increasing social inequality in life expectancy in Denmark. *European Journal of Public Health*, *17* 585-586.

Bronnum-Hansen, H., & Baadsgaard, M. (2008). Increase in social inequality in health expectancty in Denmark. *Scand J Public Health*, *36* 44-55.

Cambois, E., Robine, J. M., & Mormiche, P. (2007). Une forte baisse de l'incapacité en France dans les années 1990? Discussion autour des questions de l'enquête santé. *Population, 62*(2), 361-386.

Chiang, C. L. (1984). A new life table - for survival and stages of disease. In Robert E.Krieger Publishing Company Malabar, F. (Ed.), The life table and its applications (pp.245-273). Malabar, Florida: R.E. Krieger Publishing Company.

Cox, B., Van Oyen, H., Cambois, E., Jagger, C., Le Roy, S., Robine, J. M., & Romieu, I. (2009). The reliability of the Minimum European Health Module. *Int.J.Public Health, 54*(2), 55-60.

Davey Smith, G., Hart, C., Hole, D., MacKinnon, P., Gillis, C., Watt, G., Blane, D., & Hawthorne, V. (1998). Education and occupational social class: which is the more important indicator of mortality risk? *J.Epidemiol.Community Health*, *52*(3), 153-160.

Deboosere, P., Gadeyne, S., & Van Oyen, H. (2008). The 1991–2004 Evolution in Life Expectancy by Educational Level in Belgium Based on Linked Census and Population Register Data. *Eur J Population*, *25*(2), 175-196.

DeSalvo, K. B., Bloser, N., Reynold, K., He, J., & Muntner, P. (2006). Mortality prediction with a single general self-rated health question. A meta-analysis. *J Gen Intern Med*, *21* 267-275.

Doblhammer, G., & Kytir, J. (2001). Compression or expansion of morbidity? Trends in healthy-life expectancy in the elderly Austrian population between 1978 and 1998. *Social Science & Medicine*, *52* 385-391.

Faeh, D., & Bopp, M. (2010). Educational inequalities in mortality and associated risk factors: German- versus French-speaking Switzerland
7. *BMC Public Health*, *10*.

Fries, J. F. (1980). Aging, natural death and the compression of morbidity. *New England Journal of Medicine*, *303* 130-135.

Galea, S., & Tracy, M. (2007). Participation rates in epidemiologic studies. *Annals of Epidemiology*, *17*(9), 643-653.

Galobardes, B., Shaw, M., Lawlor, D. A., Lynch, J. W., & Davey, S. G. (2006). Indicators of socioeconomic position (part 1). *J.Epidemiol.Community Health*, *60*(1), 7-12.

Goldberg, M., Chastang, J. F., Leclerc, A., Zins, M., Bonenfant, S., Bugel, I., Kaniewski, N., Schmaus, A., Niedhammer, I., Piciotti, M., Chevalier, A., Godard, C., & Imbernon, E. (2001). Socioeconomic, demographic, occupational, and health factors associated with participation in a long-term epidemiologic survey: A prospective study of the french GAZEL cohort and its target population. *Am.J.Epidemiol.*, *154*(4), 373-384.

Graham, P., Blakely, T., Davis, P., Sporle, A., & Pearce, N. (2004). Compression, expansion, or dynamic equilibrium? The evolution of health expectancy in New Zealand
1. *Journal of Epidemiology and Community Health*, *58*(8), 659-666.

Gruenberg, E. (1977). The failure of success. *The Milbank Memorial Fund Quarterly/ Health and Society*, *55* 3-24.

Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. *J.Health Soc.Behav.*, *38*(1), 21-37.

Jagger, C., Cox, B., Le Roy, S., Clavel, A., Robine, J. M., Romieu, I., & Van Oyen, H. (2007). Health expectancy calculation by the Sullivan Method: A Practical Guide. Third Edition. (pp.1-44). Montpellier, France.

Johnson, B., & Langford, A. (2010). Intercensal denominators-feasibility of using the Labour Force Survey to estilate mortality rates by NS-SEC. *Health Statistics Quarterly*, *45* 3-27.

Khoman, E., Mitchell, J., & Weale, M. (2008). Incidence-based estimates of life expectancy of the healthy for the UK: coherence between transition probabilities and aggregate life-tables
3. *Journal of the Royal Statistical Society Series a-Statistics in Society*, *171* 203-222.

Kirkwood, B., & Sterne, J. (2003). Poisson regression. In Kirkwood, B., & Sterne, J. (Eds.), Essential medical statistics (pp.249-262). Malden, MA: Blackwell Science.

Klotz, J., & Doblhammer, G. (2008). Trends in educational mortality differentials in Austria between 1981/82 and 2001/2002: A study based on a linkage of census data and death certificates. *Demographic Research*, *19* 1759-1780.

Kohler, I. V., Martikainen, P., Smith, K. P., & Elo, I. T. (2008). Educational differences in all-cause mortality by marital status - Evidence from Bulgaria, Finland and the United States
9. *Demographic Research, 19* 2011-2042.

Kunst, A. E., & Mackenbach, J. P. (1994). Measuring socio-economic inequalities in health. Copenhagen: WHO Regional Office for Europe.

Lorant, V., Demarest, S., Miermans, P. J., & Van Oyen, H. (2007). Survey error in measuring socio-economic risk factors of health status: a comparison of a survey and a census. *Int.J.Epidemiol., 36*(6), 1292-1299.

Manton, K. G. (1982). Changing concepts of morbidity and mortality in the elderly population. *Milbank Memorial Fund Quarterly/ Health and Society, 60* 183-244.

Matthews, R. J., Jagger, C., & Hancock, R. M. (2006). Does socio-economic advantage lead to a longer, healthier old age? *Soc Sci Med, 62* 2489-2499.

Purdie, D. M., Dunne, M. P., Boyle, F. M., Cook, M. D., & Najman, J. M. (2002). Health and demographic characteristics of respondents in an Australian national sexuality survey: comparison with population norms. *J.Epidemiol.Community Health, 56*(10), 748-753.

Robine, J. M. (2006). Summarizing health status. In Pencheon, D., Guest, C., Melzer, D., & Jam, G. (Eds.), Oxford Handbook of Public Health (pp.160-168). Oxford: Oxford University Press.

Robine, J. M., Romieu, I., & Michel, J. P. (2003). Trends in health expectancies. In Robine, J. M., Jagger, C., Mathers, C. D., Crimmins, E., & Suzman, R. (Eds.), Determining health expectancies (pp.75-101). Chichester: John Wiley & Sons.

Roelands, M., Van Oyen, H., & Baro, F. (1994). Dementia-free life expectancy in Belgium. *European Journal of Public Health, 4 (1)* 33-37.

Saurel-Cubizolles, M. J., Chastang, J. F., Menvielle, G., Leclerc, A., & Luce, D. (2009). Social inequalities in mortality by cause among men and women in France. *Journal of Epidemiology and Community Health, 63*(3), 197-202.

Shkolnikov, V. M., Jasilionis, D., Andreev, E. M., Jdanov, D. A., Stankuniene, V., & Ambrozaltiene, D. (2007). Linked versus unlinked estimates of mortality and length of life by education and marital status: Evidence from the first record linkage study in Lithuania
4. *Social Science & Medicine, 64*(7), 1392-1406.

Sihvonen, A. P., Koskinen, S., & Martelin, T. (2009). Socio-economic health inequalities and how they have changed: Healthy life expectancy.Trends in socioeconomic health differences 1980-2005 (pp.119-125). Helsinki: University Press.

Stiefel, M. C., Perla, R. J., & Zell, B. L. (2010). A healthy bottom line: healthy life expectancy as an outcome measure for health improvement efforts. *Milbank Q., 88*(1), 30-53.

Sullivan, D. F. (1971). A single index of mortality and morbidity. *HSMHA Health Reports, 86* 347-354.

Tibaldi, F., Bruckers, L., Van Oyen, H., Van der Heyden, J., & Molenberghs, G. (2003). Statistical software for calculating properly weighted estimates for health interview survey data. *Soc.-Präventivmed*, *48* 269-271.

Valkonen, T. (1993). Problems in the measurement and international comparisons of socio-economic differences in mortality. *Soc Sci Med*, *36*(4), 409-418.

van Oort, F. V. A., van Lenthe, F. J., & Mackenbach, J. P. (2005). Material, psychosocial, and behavioural factors in the explanation of educational inequalities in mortality in the Netherlands
11. *Journal of Epidemiology and Community Health*, *59*(3), 214-220.

Van Oyen, H., Charafeddine, R., Deboosere, P., Cox, B., Lorant, V., Nusselder, W., & Demarest, S. (2011). Contribution of mortality and disability to the secular trend in health inequality at the turn of century in Belgium. *The European Journal of Public Health*, *Epub ahead of print*.

Verbrugge, L. M. (1997). A global disability indicator. *Journal of Aging Studies*, *11*(4), 337-362.

Wells, C., & Gordon, E. (2008). Geographical variations in premature mortality in England and Wales, 1981–2006. *Health Statistics Quarterly*, *38* 7-18.