Data inventory of health inequalities among adolescents and young adults in the Brussels-Capital Region

HANNELORE DE GRANDE[1]
HADEWIJCH VANDENHEEDE[1]
PATRICK DEBOOSERE[1]

Résumé

Cet article analyse différentes sources de données quant à leur potentiel à étudier et suivre les inégalités de santé parmi les adolescents et les jeunes adultes dans la Région de Bruxelles-Capitale (RBC). Les données du recensement et d'enquêtes sont évaluées quant à leur disponibilité, leur aptitude à se prêter à une analyse statistique, la quantité d'information portant sur des indicateurs-clés, comme la position socio-économique (PSE), l'état de santé, des comportements de santé ou des risques, et les soins de santé. Les certificats de décès couplés avec les données du recensement et l'Enquête de Santé par Interview (HIS) ont été identifiés comme les meilleures sources de données et sont discutées en détail. Les certificats de décès sont couplés aux recensements de 1991 et 2001, ce qui en fait un important outil de monitoring. Le lien entre ces deux bases de données nous permet de suivre la mortalité par cause dans la population jeune de RBC. De plus, le recensement de 2001 fournit de l'information sur le niveau de santé auto-évalué pour l'ensemble de la population. L'information sur les variables de PSE, la structure familiale, l'historique des migrations et la nationalité d'origine sont disponibles et permettent la comparaison de la santé perçue et de la mortalité entre différents sous-groupes, en tenant compte de l'impact spécifique de plusieurs déterminants sociaux. Cependant, le monitoring des inégalités de santé parmi les adolescents et les jeunes adultes ciblé uniquement sur la mortalité et la santé perçue ne suffit pas à en décrire toutes les facettes. D'autres données sont disponibles pour compléter l'information des bases de données couplées, par exemple celles de la HIS ; nous avons analysé la représentativité et les non-réponses dans les données de la HIS (1997, 2001, 2004, 2008) afin d'identifier son potentiel et ses limites pour l'étude des 15-34 ans. Les avantages de cette enquête sont : la taille suffisante de l'échantillon pour le RBC, la représentativité des 15-34 ans pour la plupart des caractéristiques démographi-

---

1.   Interface Demography, Vrije Universiteit Brussel, Belgium.

ques, une information détaillée sur les comportements de santé, la santé mentale et la consommation de soins, ainsi que sur des déterminants sociaux. Il existe aussi certaines limites : comme dans toute enquête se pose le problème des non-répondants, et spécifiquement pour cette enquête, le taux de non-réponses au volet auto-rempli du questionnaire est important et sélectif. De plus, certains modules n'ont pas été utilisés à chaque vague de l'enquête, ce qui limite l'utilisation de données groupées pour certains modules.

Mots clés

Inégalités de la santé, données administratives, enquête de santé, recensement, analyses de sensibilité, méthode de couplage des données, analyse des données manquantes, inventaire des données, métropole, adolescents et jeunes adultes.

Summary

In this paper, various data sources are analysed in terms of their potential for the study and monitoring of health inequalities among adolescents and young adults in the Brussels-Capital Region (BCR). Routine, census and survey data sources are evaluated as to their accessibility, their suitability for use in statistical analysis, and the amount of information they can yield on key indicators such as socioeconomic position (SEP), health outcomes, intermediate factors (i.e. health and risk behaviours), and health care. Death certificates linked to census data and the Belgian Health Interview Survey (BHIS) are identified as the most suitable data sources and are discussed in detail. The linkage of death certificates to the 1991 and 2001 censuses resulted in an important tool for health monitoring because the continuity between these two datasets allows for the tracking of cause-specific mortality rates among the BCR's young population. Additionally, the 2001 census data include information on self-rated health for the entire population. Data on SEP variables, living arrangement indicators and other sociodemographic variables such as migration history and nationality of origin are available, allowing for comparisons between subgroups of self-rated health and mortality and taking into account the specific impact of a number of social determinants. However, it is insufficient to rely solely on cause-specific mortality and self-rated health for the monitoring of health inequalities among young people as this is too narrow an approach to produce a complete picture of the situation. Other data complementing the information obtained through the linking of datasets are available, such as that collected by the HIS. Representativeness and nonresponse within HIS data (1997, 2001, 2004, 2008) are analysed in order to identify the opportunities and limitations of the HIS for the study of young people. The benefits of this survey are its sample size, sufficiently large for the BCR, the representativeness of 15-to-34-year-olds with regard to most demographic characteristics, and the in-depth information it provides on health behaviour, mental health and medical consumption, as well as on social determinants of health. This dataset also has its limitations: as with any survey, there is the problem of nonresponse, and in this specific case, the dropout rate for the self-administered questionnaire is substantial and selective. Also, not all of the modules were

employed every survey year, thus restricting the usefulness of pooled data to only some of the modules.

Keywords

Health inequalities, routine data, Health Interview Survey, census, sensitivity analysis, data linkage procedure, missing data analysis, data inventory, metropolis, adolescents and young adults.

Introduction

> «The errors which arise from the absence of facts are far more numerous and more durable than those which result from unsound reasoning respecting true data», Charles Babbage, 1832.

Monitoring and surveillance are important tools in order to fully understand the complexity of health inequalities (Solar *et al.*, 2007) and setting up effective actions for prevention and intervention (WHO, 2010). Including measures of socioeconomic position (SEP) such as education, occupation and wealth, are indispensible in this respect (Kunst *et al.*, 2001), but are lacking in several data sources. Routine data from population health surveys and surveillance systems are often used to monitor epidemiological changes in health inequalities (Chittleborough *et al.*, 2009). These data can also be supplemented with survey data to obtain information on SEP, health outcomes, behaviours and risks lacking in routine data.

The quality of the data at one's disposal can often make or break a research or monitoring project. As data collection is a time-consuming and expensive matter and respondents are swamped with surveys, the use of existing sources of information is often preferred. A large amount of existing data has the potential to be further exploited, to the benefit of both the researcher and the original data collector. However, there are also some drawbacks to using secondary data which can be related to its origin. Routine data, for example, are not aimed primarily at serving research purposes. Therefore, it needs some preparation in order to render it suitable for analysis with statistical software. It might also be not suited to fully understand the subject under study. In the case of secondary survey data, the researcher cannot control the quality of the data collection or content, leading to further quality checks before analysing the data, especially when the research question differs from the main objectives of the research project for which the data were originally collected. These and other data problems should not be dismissed easily, as

they potentially can bias estimates of the magnitude of health inequalities (Kunst *et al.*, 2001).

To reach an informed decision as to which data to consider for a research or monitoring project, the first step is to prepare an inventory of the available sources. The second step is to select the most appropriate databases and scrutinise these data, including identifying their flaws and opportunities. This paper addresses these two steps in relation to the monitoring of social inequalities in health among adolescents and young adults in the Brussels-Capital Region (BCR). Theoretical considerations are first discussed, followed by a brief overview of the available data sources. The two most suitable data sources are then elaborated upon.

Theoretical background: The health of adolescents and young adults in a metropolitan context

As the majority of the world's population lives in cities nowadays, there is a growing interest in the study of health matters and needs that are specific to cities (Kjellstrom, 2008). While cities offer many opportunities in terms of job availability, schooling options and high housing capacity, they also concentrate health risks and hazards. Air pollution, the lack of green areas and higher crime rates are important factors influencing one's health, but unhealthy lifestyles are of concern as well (WHO, 2010). Unhealthy fast-food diets, sedentary behaviour and substance abuse are often more common in urban than in rural areas.

Young people (15- to 34-year-olds) are an especially vulnerable group. Although most of the young living in (post-)industrialised countries consider themselves to be in good to excellent health (Currie *et al.*, 2012; Ozer, Irwin Jr, 2009; De Grande *et al.*, 2014), their life stage is critical for the development of health and risk behaviours that influence premature morbidity, mortality and health later in life (Brindis *et al.*, 2004). Intervening at this stage is a prerequisite for preventing morbidity and mortality later in life. Thus, public health policies should pay special attention to giving children and adolescents a healthy start in life (Halfon *et al.*, 2000).

There is an uneven distribution of health risks among the young population. Although some health differences are genetic in nature, many inequalities are socially determined by SEP (education, employment and/or income), gender and ethnicity (Solar *et al.*, 2007). Selecting indicators

that are suitable for determining SEP change over the life course is challenging, especially among 15-to-34-year-olds, as it is a transition period during which many different positions are possible: a large share of this group is studying, while others have already settled and started a family or are focusing on their career. Employment situation is not an ideal indicator during this life stage, as a high percentage of this group is still studying and there is high turnover in first jobs. Educational level is thus a more suitable indicator of SEP for this age group, even if a part of the study population is still enrolled in education. Own educational level is considered a valid measure of SEP, at least from the age of maturity, as it captures material, intellectual and other resources of the family of origin (Galobardes *et al*., 2006). Those who are still in secondary education at the age of 19 clearly trail behind on the educational track, indicating disadvantage, while being enrolled in higher education entails social opportunities. For 15- to 19-year-olds, it is necessary to include academic orientation (general, technical, vocational), as it gives a good impression of future social position (Hagquist, 2007). Researchers often choose to include measures of SEP of one or both parents to complement or replace the SEP of adolescents, but there is some discussion as to the contribution of these parents' indicators to the transition period in which adolescents find themselves (Koivusilta *et al*., 2006; West, 1997). Measuring social capital in families (family affluence scales), schools (school performance or academic orientation) or neighbourhoods constitute other interesting avenues of research (Koivusilta *et al*., 2006).

Social inequalities also tend to be larger in cities, where the poor and disadvantaged are often concentrated in marginalised neighbourhoods (Diez Roux, 2001). Brussels, the capital of Belgium, is an interesting metropolis in this context, as it contains sharp socioeconomic discrepancies (Loopmans *et al*., 2009). There is marked differentiation in the spatial distribution of high- and low-income households. This socio-spatial polarisation is largely a reflection of differential access to the labour market, with on the one hand a highly educated, multilingual workforce competing for highly qualified jobs, and on the other, a low-educated, rather monolingual group, of which a high percentage is of foreign origin. As the BCR attracts a young population for both education and employment opportunities, and as out-migration takes place among older age groups, it is the only Belgian region that has rejuvenated in recent decades (Deboosere *et al*., 2009). Strong social inequalities exist among the youth of Brussels: the school dropout rate is high and unemployment rates are soaring (Brussels Observatorium voor de Werkgelegenheid, 2011).

In a metropolis such as the BCR, high-quality data on the health situation of the population are needed in order to identify health problems and opportunities. This information is essential in order to be able to take preventive actions and target specific subpopulations. The next section sums up the available data sources that could be used to monitor and research the health and well-being of young people (15-to-34-year-olds) in the BCR. The paper then focuses on the data deemed most suitable for the study of this subject.

Data sources

The Belgian Health Care Knowledge Centre (KCE) has published an overview of all available administrative data and surveys concerning health care (Van de Sande *et al.*, 2006). In this paper we will focus solely on data sources providing information on health inequalities among young people in the BCR. These data can be classified into three groups: routine data, census data and survey data. Each group is summarised below. The merits and shortcomings of these data sources are detailed in Table 1.

Routine data

More and more databases cover the entire population, with their information obtained through systems of administrative records (Ingenbleek *et al.,* 2010). Data for the BCR can be extracted from several administrative data sources collected at the federal level. This includes information on health-care consumption expenditures (National Institute for Health and Disability Insurance, NIHDI), hospital statistics (Federal Public Service Health) and cancer (Belgian Cancer Registry), among others. Some of these data sources are linked to population data and have become an important monitoring tool, while others have the potential of becoming one but currently lack conformity. Most of these datasets can be used for scientific purposes, although one has to keep in mind that this was not necessarily the primary aim of the data collection.

Intermutualist Agency (IMA-AIM)

IMA-AIM was founded in 2002 to integrate and analyse the demographic characteristics of the health insurance beneficiaries and the information collected by the various official health insurance agencies. Its database consists of two main pillars: (1) information on medical visits (to general practitioners, specialists, nurses, etc.), and (2) information on prescribed medication and health interventions.

This agency centralises information collected on all persons insured by the mandatory Belgian health insurance programme, in accordance with the Law of 14 July 1994. Membership becomes compulsory if one starts working, reaches 25 years of age and is still studying, or receives a living allowance. Children are automatically joined with the head of the family until the age of 25, unless they become fiscally independent before that time. All ages are therefore represented in this database, allowing for the analysis of adolescents and young adults specifically. General health indicators are published on IMA-AIM's website. Indicators on the level of census tracts and according to age groups are available on demand (IMA-AIM, 2013). Depending on the level of detail requested for, an application to the Commission for Protection of Privacy (CPP) may also be compulsory.

Because the IMA-AIM database was not primarily designed for scientific use, a feasibility study was set up to test its usefulness for the construction of health indicators on a local level within the BCR. This feasibility study was a cooperative effort between the Brussels-Capital Health and Social Observatory, Interface Demography (Vrije Universiteit Brussel, VUB) and the École de Santé Publique (Université libre de Bruxelles, ULB). Interface Demography examined the sociodemographic indicators available in the database while the École de Santé Publique identified useful health indicators.

The first study was carried out by Lammens *et al.* (2009). This study examined to what extent the sociodemographic distribution within the IMA-AIM data reflects that of the entire population, using several data sources – including the National Registry (NR), the 2001 census, and the Labour Force Survey – for comparison. They found that the IMA-AIM data are of fairly good quality and comparability with the other sources. However, information on the census tracts was often lacking, which distorts information that otherwise could be useful as a health monitoring tool at the local level. For the purposes of our research project, we are mainly interested in information concerning SEP and health indicators. Only limited information exists on labour market indicators that could

be used as a proxy for SEP ((self-)employment, activity status and receiving welfare benefits). It is necessary to conduct a series of transformations of the data in order to calculate these indicators. In general, IMA-AIM distributions of these indicators are quite comparable, both in absolute numbers and proportionally, with external data sources such as the Labour Force Survey (Lammens *et al.*, 2009).

The second study of Moureaux *et al.* (2010) described the available health indicators in IMA-AIM. Diseases and other medical conditions entailing a hospital admission can be identified. In terms of adolescent health, these include abortions, type 1 diabetes, chronic pneumonia and asthma, among others. Not all of this information is easily transformed into an applicable health indicator. With regard to abortions, for example, the hospitals do not elaborate on the cause of an intrauterine operation leading to an interruption of a pregnancy between 4 and 6 months. It is thus impossible to identify the difference between miscarriages and abortions within the IMA-AIM data. For a detailed description of each outcome and its associated difficulties, see Moureaux *et al.* (2010). This study was however more focused on diseases and health circumstances common at older age rather than the age range in our study. In addition to hospital admissions and medical visits outside the hospital, the IMA-AIM database also provides information on prescribed medication. However, this accounts for only a small part of the total amount of medicine sold by pharmacists, and diagnostic information is lacking.

Death certificates

Cause-of-death statistics are widely used as health indicators. Their popularity lies in their availability and comparability between regions and countries (Jougla *et al.*, 2001). The information involves data on the entire population instead of sample data, and consists of a detailed chain of diseases leading to death. Until the end of 1997, this information was coded separately in each province and then collected at a national level in Belgium. It was used to produce national vital statistics, and strongly deviated from the international standard (Aelvoet, 2008). Only immediate and underlying causes of death were certified and coded. Socio-demographic information was scant or out of date because of delays in the Belgian National Registry (Gadeyne, 2006). Since 1998, causes of death

have been collected in more detail and coded at the community[2] level and then distributed regionally. BCR death certificates are gathered by the Health and Social Observatory of the BCR and coded by the Flemish Agency for Care and Health (Vlaams Agentschap Zorg en Gezondheid, VAZG), which is also in charge of Flemish death certificates. These recent death certificates also include more detailed information on the background of the deceased as well as the chain of diseases leading to the death (in accordance with the WHO model) (Aelvoet, 2008). The quality of the sociodemographic information remains doubtful for some indicators (Houttekier *et al.*, 2009). This was one of the chief motives for linking death certificates with census data, the latter of which contains detailed and high-quality information on the sociodemographics of the entire population. These linked data are an important tool for health monitoring.

TABLE 1   Overview of available data sources

	Age range	Population	Accessibility	Information on...			
				Social determinants	Health outcomes	Intermediate health behaviours	Health care
Administrative data							
Intermutualist Agency	All ages	Comprehensive	IMA-AIM	+/-	+	-	+
Death certificates	All ages	Comprehensive	BCR Health and Social Observatory	+/-	++	-	-
Medical inspection of children at school	Until 15 yr	Comprehensive	No access	+	++	+	++
	All ages	Comprehensive	Statistics Belgium	++	+	-	-
Census data/Survey data							
Health Interview Survey	All ages	Sample +	WIV-ISP	++	++	++	++
Health Behaviour of School-Aged Children	11, 13, 15 yr	Sample -	HBSC team	++	++	++	+
JOP-Monitor	12-20 yr	Sample -	JOP	++	-	+	-
Sentinel general practitioner	All ages	Sample -	WIV-ISP	-	++	+	++

Age range = Age range of the data. Population = Whether information is available for the entire population (comprehensive) or based on a sample («sample -» = Sample but not representative for youth in BCR, «sample +» = Representative sample for youth in BCR). Accessibility = Through which channel it is possible to gain access to these data. Social determinants = Socioeconomic position, nationality (of origin), migration background, gender. - = Restricted or no availability. + = Information is present in the data. ++ = Information is abundantly present (more than three indicators).

---

2.   Belgium is a federal state, consisting of three language Communities (French, Flemish and German) and three regions. When referring to «community level», we refer to the institutional language community level, not the geographical unit.

Medical inspection of children at school

Information obtained through medical inspection at schools can give us insight into the health situation of the total population of school-aged children. These are the only measurements and tests at a population level that have been available since the suspension of military service in Belgium[3]. When military service was compulsory, each Belgian man had to complete these tests. Not only medical information but also simple measures such as height can be highly valuable to public health: on the basis of such measurements, researchers found consistent differences in health between manual workers and employees, constituting one of the first pieces of evidence for persistent social inequalities in health since modernisation (see also Szreter, 2004).

Medical inspection in Belgian schools is performed at six different points for each child. During these inspections, height and weight are measured and an eye test and vaccinations are performed. The children also answer a list of questions concerning their health, school well-being and household situation. The development of each child through its school career can be followed longitudinally for ten years. Not only is this data source interesting for the amount of information that it contains, it also includes valuable information that can be used as a starting point for observing the study population of 15- to 34-year-olds. Social- and health-related adversities in childhood contribute to social inequalities in adolescence and young adulthood (Koivusilta *et al.*, 2003), and are thus important indicators to be taken into account. This population database could also be followed up over a longer period of time through linkage with other data sources, such as death certificates. As census data are no longer gathered in the traditional way (see section 4), other potential sources need to be considered to guarantee further health monitoring.

In order for school medical inspection data to fulfil its potential, several barriers need to be overcome. There are problems due to a lack of standardisation in procedures for collecting and coding these data between and within communities. Education is organised at the community level and is done differently within the Dutch- and French-language communities, including different procedures for medical inspection. The BCR includes both Dutch and French communities, which complicates the matter further, not least in terms of data collection. Several feasibility

3.   Date of commencement of this Law: 18/01/2013. The last ones liable to military service were discharged on 5/2/1995.

studies were executed to identify and eliminate obstacles, one in the BCR (Decant, 2007) and two in the French community (Decant *et al*., 2009; Godin, 2000). The most substantial problems are summarised below.

The Dutch-language schools regulate medical inspection through pupil guidance centres (Centrum voor Leerlingenbegeleiding, CLB). The information gathered is monitored using an electronic system (Leerling Administratie en Registratie Systeem, LARS – Pupil Administration and Registration System). The French-language educational system organises medical inspections through centres promoting health at school (Promotion de la Santé à l'École, PSE). Before information from both educational systems can be joined, it needs to be standardised within each system. LARS was developed within the Dutch-language system for this purpose, as part of the broader responsibilities of the pupil guidance centres (CLB, 2012). There is no such standardisation procedure within the French-language educational system. The feasibility study conducted by Decant *et al*. (2009) was undertaken as a first modest step towards establishing a standardised procedure for the French community (Brussels and the Walloon Region). However, combining objective indicators such as height, weight, eye tests and vaccinations led to incorrect counts because of a lack of standardisation and harmonisation between the coding schemes in the different centres. Inadequate communication between members of the centres and the central administration further hampered this process.

Combining subjective information on the health status of a child is even more complex, as the content of the face-to-face interview at the end of the medical inspection differs strongly between and within communities. Copies of lists of questions are documented in the Decant's thesis (2007). One way of standardising these questions could be the introduction of a uniform question on self-reported health. Standardisation of one or more questions for all medical tests would also simplify the harmonisation of the data from the different centres. Further research is needed to maximise on this database's potential for future research and monitoring purposes.

Census data

Belgium has a long tradition of gathering census data, its oldest census dating back to the mid-19th century. The purpose of collecting census data is to obtain in-depth knowledge on various sociodemographic characteristics for the entire population. A unique feature of the 2001 cen-

sus is the inclusion of health questions (self-reported health, long-term illness), making it possible to assess the health situation of the entire population. Self-rated health is a powerful indicator that is widely used by researchers and policy makers (Chittleborough *et al.*, 2009; Idler *et al.*, 1997). In addition to this, the data were complemented by Statistics Belgium with register data on mortality and migration as well as death certificates (this procedure will be detailed in section 4). In this way, all the information necessary to calculate standardised mortality rates and ratios is available.

Survey data

Health Interview Survey (HIS)

The Scientific Institute of Public Health (WIV-ISP) carries out an HIS every four to five years. The BCR is oversampled within the HIS, resulting in an average sample of 3'000 respondents per survey year, with around 800 respondents aged 15-34 years. A sample of 3'319 adolescents and young adults can be extracted when combining all survey years. Pooled data can only be used to assess the overall health situation for a given period, and not in a study that focuses on evolutions over time.

The health interview consists of three parts:

– a short household questionnaire, including questions on the household's composition, income, health care expenditure and housing;
– a face-to-face interview, with questions on medical consumption, education, employment, nutrition and chronic conditions;
– a self-administered questionnaire, with personal and subjective questions on lifestyle behaviours (e.g. smoking habits and substance (ab)use), self-rated health and mental health.

This dataset represents a unique tool for the study of health inequalities within the BCR, as its sample is sufficiently large, its list of health indicators is ample, and measures of social position are available at the individual as well as at the household level.

Health Behaviour of School-Aged Children (HBSC)

The HBSC survey specifically addresses the health of young adolescents. It is a collaborative cross-national survey coordinated every four years

by the WHO Regional Office for Europe and carried out in 43 different countries (Currie *et al.*, 2012). The Walloon Region (from 1985 onwards) and the Flemish Region (from 1993) are both adequately represented in this survey. This is not the case for the BCR. The Walloon Region oversampled the BCR's French-language schools in 2002 and 2006 at the request of the Brussels-Capital Health and Social Observatory (2010), while in the Dutch-language part only one school participated. The number of respondents is too small to be of use for health monitoring in the BCR. The age range (11 to 15 years old) is also too limited for our field of study. A great opportunity is missed here, since HBSC surveys gather ample information on SEP measures – with the family affluence scale acting as an important proxy for SEP at young ages (Currie *et al.*, 2012) – health outcomes (self-rated health, injuries, body weight) and intermediate health and risk behaviours (eating behaviour, physical activity, fighting, substance use, smoking, sexual intercourse).

Youth Research Platform (JOP) Survey

The JOP, an interdisciplinary collaboration between three research groups based at VUB, UGent and KU Leuven, organised a survey on the youth of Brussels in 2008. A representative sample of 12- to 20-year-olds in the Dutch-language education system in the BCR was obtained. Since only 17% of the entire school population in the BCR is attaining the Dutch educational system (BRIO, 2013), this survey misses a significant part of the target population. A follow-up study was run in 2013. Information gathered through these surveys is available on social background characteristics (educational orientation, parental employment/education, nationality of origin), but there is only limited information on health indicators (general well-being, satisfaction with life/health) and intermediate health behaviours (violence, safety, substance use), and information on health care is missing altogether.

Sentinel general practitioners

In 1978, WIV-ISP launched an epidemiological surveillance system to monitor specific health problems through a network of volunteer sentinel general practitioners (Van Casteren, 2011). Through the work of this surveillance network, selected diseases can be monitored, their incidence measured and their most important epidemiological characteristics studied. Information on selected themes is available for a number of

years (such as measles, mumps, HIV tests, diabetes and suicide attempts), while other themes are registered more infrequently (such as heart failure, alcohol-related problems and Lyme disease). The sample of sentinel general practitioners is representative at a national level but not at a regional level. Separate information is not available for the BCR. Socioeconomic information is also limited.

Conclusion

It is clear that two data sources stand out for the study of different aspects of health and health inequalities among young persons in the metropolitan context of the BCR: the census linked to death certificates and the HIS. We now provide further information on the procedure used for linking census data with death certificates and point out opportunities and limitations presented by this linked database for analysing our study population. We also probe into the HIS, examining the methodological background behind its survey collection, the representativeness of the data for 15- to 34-year-olds, and nonresponse at various stages in the survey process.

Death certificates linked to census data

Death certificates

Cause-specific mortality has a long tradition as health indicator and as a monitoring tool for public health policy. This information is collected through death certificates, and is needed to develop vital statistics at the national level (which is done by Statistics Belgium). There are several steps between diagnosing the cause of death and producing the vital statistics. This procedure differs in the various Belgian communities. In this paper we will detail the steps for the BCR specifically.

There are two different types of death certificates: (1) for deaths under the age of 1, and (2) for deaths above the age of 1. Given our research population, we will restrict our focus to the latter. A physician determines the cause of death and fills out this information on the death certificate. Since 1998 a death certificate has consisted of four parts. Part A contains strictly confidential information on the deceased person (name, name of spouse, sex, address, date of death). This is stored in the municipality's administrative system and is not accessible for data collection.

Part B repeats some of the information from Part A that is to be forwarded to the National Registry (date/place of death and sex). Part C contains detailed information on the nature of the decease (and specification of the accident, if applicable), the chain of causes of death, and additional medical information (if the deceased was pregnant, whether the physician was the medical attendant, whether an autopsy was necessary, among others). Part D records other personal characteristics of the deceased which are provided by the municipal administration (education level, last/current employment situation, marital status, birthdate, nationality, municipality of domicile/decease). The administration does not have access to part C and is therefore not informed of the cause of death. Upon final completion, parts B, C and D go to a centralised unit responsible for the quality control, data input and encoding of the causes of death according to the International Classification of Diseases (ICD), of which the tenthrevision (ICD-10) is currently being used. Another unit is responsible for statistical processing and further data controls. As already mentioned, VAZG applies ICD codes to the BCR death certificates, while the Brussels-Capital Health and Social Observatory is in charge of the final data checks and the processing of the data.

The overall quality of Belgian vital statistics up to the 1990s was rather poor (De Henauw *et al.*, 1998; Aelvoet, 2008), as in other European countries (Jougla *et al*., 2008). Thanks to the standardisation and harmonisation procedures implemented by the Working Group on Public Health Statistics (Pearce, 2004), quality has since improved considerably. Furthermore, ISTAT (Italian National Institute for Statistics) developed online tutorials to instruct physicians on how to certify causes of death using ICD-10. Nonetheless, some limitations remain in terms of the availability and quality of sociodemographic information recorded on death certificates (Houttekier *et al*., 2009). This applies to Belgium as well as to other European countries (Cohen *et al*., 2007). More information is available for the BCR, the Flemish Region and Italy than in other countries (Cohen *et al*., 2007) because the municipal administration provides the sociodemographic information for the death certificates. Still, this information is mostly either out of date (e.g. marital status) or missing (e.g. educational level) (Houttekier *et al*., 2009). In order to perform meaningful statistical analyses on inequalities in mortality, the gaps must be filled through the use of other registers or of ecological measures of SEP (Cohen *et al*., 2007).

It is compulsory to submit an application to the CPP in order to link these data to other registers. The possible identification of individuals and the exhaustive compulsory character of administrative data render it

more sensitive than anonymous sample data. Because these data are primarily collected for administrative purposes, national laws and regulations largely preclude their use for other purposes. Anything that falls within the definition of «personal details», i.e. any data concerning an identified or identifiable natural person, directly or indirectly, by means of an identification number or one or more components that characterise their physical, physiological, mental, economic, cultural or social identity, is subject to the regulations governing the protection of individuals. Linking this kind of information could thus constitute an intrusion of one's privacy. As Nilstun and colleagues (2006) recommend, each application should be dealt with carefully, applying thorough controls in order to protect confidentiality and anonymity; yet, the requirements should be proportional to the potential harm that may be caused to the persons under study.

Assuming that approval is obtained, in many cases no ready-made identification code exists to link different data. An identification key needs to be created containing information that uniquely identifies each deceased person in both the death certificates and the census. The next section will explain this linkage procedure in detail.

Linkage procedure

The linkage consists of combining census information with National Register data on all-cause mortality and migration and with the death certificates. Summed up, the procedure is as follows (see also Deboosere *et al.*, 1999). In a first step, census data are linked to register data on all-cause mortality and migration using the National Register (NR) number. This direct link is made by Statistics Belgium and allows for an analysis of socioeconomic and ethnic inequalities in *all-cause* mortality. In a second step, cause-of-death information is added from the death certificates database, allowing for analysis of socioeconomic and ethnic inequalities in *cause-specific* mortality.

As death certificates are anonymised and do not contain NR numbers, an identification key (ID key) must be constructed as part of the second step. This ID key uniquely identifies each individual and is constructed in both the census-register and death certificates databases. It consists of variables included in both datasets: place of residence, sex, date of birth and date of decease. Using this ID key, on average 85% of deceased individuals can be linked, implying an exact match between the ID keys in the census-register and death certificates databases. For the remaining 15%, for which no exact equivalent ID key exists in both databases,

another linkage procedure is applied. Tests show that most of the non-linked records are due to errors, mostly clerical, in one or more variables of the ID key. Hence, in a probabilistic linkage, one part of the ID key, such as day of birth, is left out, while controlling for the probability of false positives. For example, dropping place of residence could link 60% of the remaining deceased persons. Most of these persons would have migrated to another municipality before they died, leaving a discrepancy between the information in the census-register database and the death certificates database. Probabilistic linkage enables linkage of such records.

Nationally, this procedure resulted in 98.0% of the records being linked for the 1991-1997 period (the linkage of the 1991 census with death certificates covering the period 01/03/1991-31/12/1997). A slightly smaller percentage (95.5%) was linked for the BCR due to greater out-migration from the capital to other countries, a group for which we do not have cause-of-death information. As for the 2001 census-death certificates linkage, the follow-up period is much shorter nationally because of the unavailability of death certificate data from the French community for the period 2001-2003. Hence, for the French community, only the years 2004 and 2005 are covered, while for the BCR and the Flemish Region, cause-of-death information is available for the period 01/10/2001-31/12/2009[4]. The final linkage percentages are similar to those of 1991: 98.8% in the Flemish Region and 95.5% in the BCR.

Statistics Belgium has been working on a more sophisticated linkage procedure since the last traditional census was performed in 2001. The 2011 census approach will be completely different from the previous censuses, which have always consisted of exhaustive surveys. In the past considerable parts of the necessary data were obtained by asking the person in question. A vast amount of information can be recovered from administrative databases. It is therefore more desirable to use these administrative data than to burden citizens with a survey (Statistics Belgium, 2014). The 2011 census will contain information from several administrative data sources, including the Crossroads Bank for Enterprises, the Labour Market and Social Protection data warehouse under the Crossroads Bank for Social Security, the General Administration of Patrimonial Documentation, and educational information from the communities, among others. Researchers can access the data by following the proper procedures, which include lodging an application with the Privacy Commission.

---

4. At time of writing this paper.

Opportunities and limitations

A major advantage of this linked database is its detailed and high-quality information on SEP, living arrangements, context variables and other sociodemographic information (from the census and the register) (Table 2). As can be observed, this list is abundant, and all sorts of research questions and public health concerns relating to inequalities can be addressed with these indicators. Also, different levels of analysis are possible as information is available at the individual, census-tract, neighbourhood and municipal levels.

TABLE 2   Sociodemographic information
in the 1991 and 2001 census
and register data

SEP indicators
Education
Current educational level enrolled in
Highest educational level obtained
Age at finalising studies
Highest education within household
Employment
Activity status (student, unemployed, employed)
Occupational status
Income
Type of income
Housing status (renting, owning)
Housing comfort
Living arrangements
Marital status
Household composition
Household position
Number of persons in household
Other sociodemographic information
Sex
Nationality
Nationality of birth
Place of birth
Context
Municipality + duration of residence
Housing conditions
Neighbourhood
Degree of urbanisation

Concerning health outcomes, several measures are available in the linked census and death certificate data. Already mentioned is the self-rated health variable available in the 2001 census, which has been shown to be an important tool for health monitoring at young ages (Breidablik *et al.*, 2008). This variable assesses how a person feels about their overall state of health, including mental and physical health, and is a good predictor of premature mortality (Idler *et al.*, 1997). All-cause and cause-specific mortality are the other health outcomes of interest. Particularly at young ages – before the age of forty – the cost to society of premature mortality is high. As adolescents and young adults die mostly of external causes such as (traffic) accidents and suicides (Patton *et al.*, 2009), information on causes of death is an important instrument for public-health policy. Information on cause-specific mortality makes it possible to identify risks among the young population and evaluate efforts towards effective prevention and intervention policies by comparing mortality trends over time.

Another asset of this linked dataset is the in-depth information it provides on causes of death, with follow-up for a long period of time. Until 1998, the causes of death were coded according to the Ninth Revision of the ICD (ICD-9); since then, the Tenth Revision (ICD-10) has been used. The use of different coding schemes in the 1990s compared to the 2000s can complicate comparisons over time, depending on the level of detail on specific diseases/medical conditions that is needed for analysis. ICD-10 is much more detailed, with around nine thousand categories, compared to five thousand for ICD-9. Chapters and cause-of-death titles have been added and rearranged. In general, selection and modification rules are similar, but in some instances these can lead to a shift in classification (Anderson *et al.*, 2001). An important shift concerns deaths due to motor-vehicle accidents (MVA), a major cause of death among young persons. Under ICD-10, there is a 15%-reduction of MVA deaths because new selection rules ask for the involvement of a «motor» (Anderson *et al.*, 2001). In contrast, under ICD-9, an accident was coded as an MVA death even if it did not involve a «motor» (for example, collision with stationary object or animal-drawn vehicle). In terms of other major causes of death among young persons, such as suicides and homicides, comparability is nearly 100%. Anderson *et al.* (2001) demonstrated that comparability between most causes of death over time is sufficient for large groups of causes. For deaths among adolescents and young adults, we propose a categorisation into major cause groups, shown in Table 3, a slight adaptation of the one used in De Grande *et al.* (2013). Eight broad categories can be discerned; these can be further broken down

according to specific research or policy evaluation needs, as long as possible differences in coding are accounted for.

In addition to the abundant sociodemographic and health information this linked data provides, it also has the advantage of overcoming numerator-denominator problems. While it is possible to produce mortality statistics on risk factors (such as age and occupation) without the link, calculating the prevalence or incidence of specific causes of death is only possible in combination with aggregate information on the total population extracted from other sources. There is no certainty that this population remains in accordance with the risk population over a longer period of time, as migration flows change the composition of the population. When using aggregate information, there is no control over the mutual relations of different factors that influence the mortality risk. The census provides us with information on the denominator, while the NR and the death certificates provide the information on the numerator, which opens up the possibility of answering such questions as whether the risk of premature death is higher among lower-educated young persons than among higher-educated ones in a given period, or whether there are differences in young-adult mortality according to nationality of origin, and for which causes. Survival curves can furthermore be calculated for specific causes of death, and social inequalities in mortality can be identified (e.g. Gadeyne, 2006) and monitored over time (e.g. De Grande *et al.,* 2013).

TABLE 3   Categorisation of major causes of death
among adolescents and young adults

ICD-9	Description	ICD-10
001-139	Infections	A00-B99
140-239	Neoplasms	C00-D48
240-291; 293-302; 306-799	Other medical conditions	D49-F9; F17; F20-R99
E950-E959	Suicide	X60-X84; Y87.0
E810-E829	Road accidents	V01-V89
E960-969	Homicide	X85-Y09
291, 292, 303-305, 4255, 5710-5713, 5770-5771, E850-855, E858, E860, E935, E937-940	Drug- or alcohol-related deaths	F10-F16; F18-F19; I426, K70, K85, X40-X49; Y10-Y14; Y16
Rest E800-E999	Other external causes	Rest V01-Y98

One limitation of these data revolves around health outcomes. Social inequalities in premature mortality and self-rated health only partly identify health differentials. Furthermore, self-rated health is only mea-

sured at a single point in time (2001). It is also important to note that the 2001 census was the last census based on exhaustive questionnaires. The 2011 census, which at time of writing was still to be published, dispensed with the use of a questionnaire, instead linking various administrative data sources to obtain information that is as close as possible to that which was previously collected by traditional censuses (Statistics Belgium, 2014). It does therefore no longer contain information on self-rated health. The 2011 census also does not contain some other information useful for the monitoring of health inequalities which was available in the 2001 census, such as housing quality, satisfaction with neighbourhoods, and country of birth. The quality of the indicators available, such as educational level and activity status, remains uncertain. Nonetheless, the advantages of this procedure are obvious: it is much less costly, citizens are not burdened by a questionnaire, and more information can be followed up longitudinally (such as housing and education) and updated regularly. As Statistics Belgium is also in charge of linking these data with information from death certificates, the linkage procedure detailed in section 4.2 will no longer be necessary. The linked data can be requested and obtained, provided that there is CPP approval.

Health Interview Survey

The second database that is particularly apt for the study of health (inequalities) among adolescents and young adults in the BCR is the HIS. This database will be discussed exhaustively below. More specifically, we will go into detail about the sampling procedure, the content relevant to our field of study, and the nonresponse over several survey phases, concluding with a missing data analysis on 15- to 34-year olds.

General sampling frame and content

The HIS was developed for the first time in 1997 as a cooperative effort between the WIV-ISP, the designated ministers and their administrations, the three language communities, the BCR and the Walloon Region. There was an urgent need for additional bottom-up health information showing health in all its facets as perceived by individuals rather than by health care professionals. So far, the HIS has been collected five times over a time span of less than twenty years (1997, 2001, 2004, 2008 and 2013). The data collected in 2013 are coded and controlled for accuracy

and missingness during 2014 but are not available at the time of writing. The sampling frame for each survey year consists of all households listed in the National Register (NR). All members of these households are part of the sampling frame. The households are selected on the basis of a stratified clustered multi-stage design. Detailed information on this selection procedure can be found in the research protocols for each survey year (WIV-ISP, 2005).

TABLE 4 Health and sociodemographic indicators of the HIS according to type of questionnaire, for all survey years

Indicator	Type of questionnaire			Survey year			
	HH	FtF	SA	1997	2001	2004	2008
Self-reported health			X	X	X	X	X
Mental health							
Self-reported mental health		X		X	X	X	X
GHQ-12			X		X	X	X
SF-36 vitality subscale			X		X	X	X
SCL-90-R			X		X	X	X
Suicide ideation and attempt			X			X	X
Nutrition & physical activity							
Underweight/overweight/obesity		X		X	X	X	X
Consumption of fruit/brown bread/fish daily		X			X	X	X
Eating vegetables daily		X				X	X
Consuming soft drinks daily		X					X
Adequate level of physical activity			X	X	X	X	X
Health behaviour & knowledge							
Smoking habits			X	X	X	X	X
Alcohol consumption over the last year			X	X	X	X	X
Lifetime cannabis use			X		X	X	X
Accidents		X					X
Violence			X				X
Knowledge of and attitudes towards HIV*			X	X		X	X
Medical consumption							
Contact with doctor (GP, specialist)		X		X	X	X	X
Use of medicines		X				X	X
Problems with paying health expenses	X			X	X	X	X
Postponement of health expenses	X			X	X	X	X
Sociodemographic characteristics							
Maximum household education completed	X	X		X	X	X	X
Equivalent household income	X	X		X	X	X	X
Perception of financial hardship	X	X		X	X		X
Sex	X			X	X	X	X
Household position	X			X	X	X	X

HH = Household questionnaire. FtF = Face-to-face questionnaire. SA = Self-administered questionnaire. GHQ = General Health Questionnaire consisting of 12 questions. SF-36 vitality subscale = Positive mental health scale. SCL-90-R = Mental health subscales for somatisation, depression, anxiety and sleeplessness.

Caution is needed when using combined information from all survey years («pooled data»). Although the pooling increases the statistical power of the results by enlarging the sample and makes it possible to study specific subgroups, changes may be obscured over time. Furthermore, not all questions are included every survey year. Restricting research questions to information available for all survey years may limit research possibilities heavily.

As mentioned above, the main way that the HIS adds value is through its horizontal approach to data collection: several types of information (such as health status, health determinants, personal characteristics and medical consumption) are collected simultaneously for the same person. A detailed overview of each indicator that could be of interest for the monitoring of health inequalities in young persons is given in Table 4. While most sociodemographic indicators are available for each survey year, there is considerable variation in health information over the years. In addition, some health information is gathered through face-to-face interviews (e.g. information on medical consumption and nutrition), while questions concerning mental health and health behaviours are generally reserved for the self-administered questionnaire.

As mentioned, around 800 respondents aged 15-34 years (living in the BCR) complete the survey each year. A sample of 3'319 adolescents and young adults can be extracted when combining all survey years. The sample for the BCR is sufficiently large to be used as a tool for health monitoring, but a missing data analysis of this sample for the young population in the BCR is needed, as this subgroup is not specifically targeted in the survey's sampling frame.

Methods in nonresponse analysis

The problem of missing data is ubiquitous in both social and health research (Allison, 2002). Standard statistical procedures presume that every case has information on all the variables included in the analysis, but the reality differs strongly from this premise. There are two broad categories of missing information: unit nonresponse and item nonresponse. Unit nonresponse arises when, for any of a number of reasons, selected respondents did not participate in the survey. Item nonresponse refers to missing information for specific questions or modules.

Nonresponse in the HIS occurs at different stages of the interview process. First, there is non-participation of households: failure to identify or locate the sample unit, failure to make contact with the household, and

household refusal are the most common reasons for this. In order to restrict selective dropout, a replacement household is sought when a household does not participate. The replacement household has three features in common with the non-participating household: census tract, household size and age of the reference person. Burzykowski *et al.* (1999) analysed household non-participation in the 1997 HIS and found that the probability of non-participation increased with smaller household size and/or when the reference person is young (15-24 years), female and non-Belgian. Furthermore, non-participation in the BCR was higher than in the other regions. The characteristics of the participating households need to be compared to the total population to ensure the representativeness of the sample for certain characteristics (such as age and sex distribution or nationality).

A second stage of missingness is dropout of members within households. In this case, a proxy can be asked to complete the interview. This can be a family member or another person present in the household. The most common reasons for nonresponse at this stage are that members cannot be contacted, are not able to fill in the questionnaire (e.g. due to physical or mental impairment), or are too young to complete the questionnaire (< 15 years).

A third stage of missingness is nonresponse in the self-administered questionnaire. In the majority of cases, this is because a proxy completed the face-to-face interview but was not entitled to respond to the subjective questions on behalf of the other person so as to not distort the results.

For each of these stages of missingness, it is important to verify the characteristics of the respondents. We analyse the first and second stage of nonresponse through a depiction of all the survey years separately for the study population of 15-34-year-olds in the BCR and by comparing this sample with the distribution for the total population. We make use of census data and NR data on the demographic characteristics of Brussels' population, available on Statistics Belgium's interactive website[5], and of individual information from the NR concerning household composition in 2006. Using these data, we are able to compare the distribution of the HIS sample according to age, sex, nationality and household composition with the distribution of these characteristics among the general population. We analyse the third stage of missingness and present results from a logistic regression, estimating nonresponse to the self-administered questionnaire (0 = questionnaire completed; 1 = dropout)

---

5.   http://statbel.fgov.be/nl/statistieken/webinterface/beSTAT_home.

according to various background characteristics (sex, age, nationality, household composition, highest education in the household, equivalent household income).

This paper will not go into detail in terms of specific item nonresponse, as this is already discussed at length in Burzykowski *et al*. (1999) and in the metadata on WIV-ISP's website[6].

Unit nonresponse

In this section we explore to what extent the non-participation of households impacts on the distribution of the young population in the sample. Pooled data are presented, and where population statistics are available, data per survey year are presented as well.

The distribution of the age structure per five-year age group in the four survey years is shown in Table 5. Percentages of the share of the young population in the total population (NR) and in the sample (HIS) (15-34 years) were calculated. There is a slight but significant underrepresentation of the young population in the sample (25.8%) compared to the total population (29.5%), especially among the youngest age groups (15-24) (X^2 = 17.9, df = 3, p < 0.001). In general the age structure is quite comparable over the survey years and does not differ significantly between the sample and the population in 2004 and 2008.

TABLE 5   *Age structure* in the NR and HIS, totals (N) and percentages of distribution of five-year age groups

Age group	1997		2001		2004		2008		Total	
	NR	HIS	NR	HIS	NR	HIS	NR	HIS	NR	HIS
15-19	18.7	15.9	19.0	18.2	18.5	17.5	18.7	17.0	18.8	17.1
20-24	23.7	22.8	23.0	17.4	22.9	22.4	22.4	24.4	23.0	21.9
25-29	29.3	28.7	29.4	27.7	28.9	28.3	29.6	30.4	29.3	28.8
30-34	28.3	32.7	28.6	36.6	29.7	31.8	29.2	28.2	29.0	32.1
15-34										
%	29.5	26.2	29.3	25.2	29.5	24.5	29.6	27.0	29.5	25.8
N	280'604	813	280'959	757	294'667	844	310'208	905		3'319
X^2	7.81*		27.8***		1.87		3.55		17.9***	

Source: National Register (NR) by Statistics Belgium and Health Interview Survey (HIS), own calculations.
* = p < 0.050; *** = p < 0.001.

___

6.   https://www.wiv-isp.be/epidemio/epinl/index4.htm.

The sex distribution is quite stable over time. The sample accurately reflects the distribution of men and women within the BCR's young population (not in Table[7]).

Concerning current nationality, the comparison of the percentage of Belgians by five-year age group is restricted to 2001, 2004 and 2008 (Table 6). In 2004 and 2008 the percentage of Belgians is fairly comparable between the sample and the total population. Because of an overrepresentation of Belgians in the 2001 sample, that sample and the pooled sample (X^2 = 20.2, df = 3, p < 0.001) differ slightly from the total population. As information on non-Belgian nationality groups or on nationality of origin is available at Statistics Belgium, comparisons cannot be made.

TABLE 6 Percentage of *Belgians (current nationality)* in the NR and HIS, totals (N) and percentages of distribution of five-year age groups

Age group	1997	2001		2004		2008		Total	
	HIS	NR	HIS	NR	HIS	NR	HIS	NR	HIS
15-19	66.7	73.4	79.0	78.3	79.5	79.1	80.4	77.0	76.7
20-24	66.5	68.4	80.3	69.4	76.2	67.7	67.3	68.5	71.8
25-29	63.1	63.1	73.8	63.2	62.3	60.4	61.1	62.2	64.7
30-34	60.9	59.9	69.7	60.6	65.3	57.8	59.6	59.4	64.0
15-34									
%	63.7	65.3	74.4	66.7	69.4	64.8	65.4	65.6	68.1
N	518	183'910	563	196'438	582	200'976	591	581'324	2'254
X²		21.4***		4.1		1.9		20.2***	

Source: Statistics Belgium (http://statbel.fgov.be/nl/statistieken/webinterface/beSTAT_home) and Health interview Survey (HIS), distribution in NR not available for 1997.
*** = p < 0.001.

Table 7 summarises the differences in household composition according to survey year, comparing these with both the 2001 census data and National Register data, for which we have age-specific information on household composition[8] in 2006. The distribution of household composition for the total population does not differ substantially in 2006 compared with 2001. One-person households are underrepresented in the HIS. Another difference occurs in the «other household composition» rest

---

7.	There is a fifty-fifty sex distribution for most of the years, with minimal differences between routine data and the surveys of 2001 (NR M/V = 50/50; HIS = 52/48) and 2004 (NR M/V = 50/50; HIS = 48/52).

8.	Household composition gives information on the type of household in which the respondent resides. It does not specify the position of the respondent in this household type. For example, a respondent living in a single-parent family can be either the parent or the child.

group, which consists of persons cohabiting with non-kin or family members other than parents. While this is a rare household composition form in the census (2.7%) and in the National Register (3.0%), there is a higher percentage present in the HIS (11.4% on average).

TABLE 7   Percentage for each *household composition* according to survey year and compared to the distribution in the 2001 census and the National Register (NR) in 2006

	Survey year				Total	Census	NR
	1997	2001	2004	2008		2001	2006
One person	19.9	21.1	19.8	20.2	20.2	25.0	25.0
Single parent	11.1	10.4	17.5	17.6	14.3	14.6	15.6
Couple -	13.5	13.2	14.5	16.7	14.6	12.9	13.7
Couple +	43.2	42.0	37.7	35.6	39.4	44.8	42.8
Other	12.3	13.2	10.5	9.9	11.4	2.7	3.0
N	813	757	844	905	3'319	283'688	299'276
Census 2001 X²					924***		
NR 2006 X²					1,415***		

Source: Census 2001, National Register 2006 and Health Interview Survey (HIS), own calculations.
Single parent = One-parent household with one (or more) child(ren). Couple - = Household type consisting of a couple without children. Couple + = Household type consisting of a couple with one (or more) child(ren). Other = Other household composition, not specified above. *** = $p < 0.001$.

Proxy interviews and dropout in the self-administered questionnaire

Proxy interviews are more common in the BCR than in other Belgian regions (Table 8). In the pooled dataset we observe that 16.4% of adolescents and young adults did not complete the face-to-face interview, instead using a proxy. In two-thirds of these cases, a proxy was used because the young member of the household was not present at the moment of the interview. In addition, 17.6% were not able to complete the survey at all, while 10.8% refused to participate. Those who were interviewed as proxy are primarily members who do not speak the survey language(s) (61.2%) or have a mental impairment (28.6%). When comparing the survey years it is striking that starting in 2001, the use of proxy interviews increased over time, peaking in 2008 at 22.4%.

The use of proxy interviews has important implications for the further progress of the survey, as proxies are not entitled to fill in the self-administered questionnaire. More than half of the nonresponse in the self-administered questionnaire can be accounted for by a proxy having been used for the face-to-face interview. While the increase in nonresponse among 15- to 34-year-olds is observed in all regions, it is markedly more

problematic in the BCR (Figure 1). Between 2001 and 2004, nonrespon-
se more than doubled in the BCR. More than one-fourth of the Brussels
respondents who completed the face-to-face interview (or used a proxy
for that part) did not fill in the self-administered questionnaire in 2008.
This difference with the Flemish Region is striking, which had 10% non-
response. The Walloon Region also exhibited high dropout in 2008 com-
pared with 2004, but its nonresponse is still smaller than that of the
BCR.

TABLE 8   Percentage *proxy interviews* according to region
           and survey year for 15- to 34-year-olds,
           with 95% confidence intervals

Survey year	Region					
	FR		BCR		WR	
1997	9.8	[8.0-11.6]	15.7	[13.2-18.3]	9.9	[8.0-11.8]
2001	5.7	[4.3-7.1]	7.8	[5.9-9.7]	6.7	[5.3-8.1]
2004	9.5	[7.6-11.4]	18.4	[15.7-21.0]	10.6	[8.7-12.5]
2008	11.1	[9.0-13.3]	22.4	[19.6-25.1]	20.0	[17.3-22.7]
Pooled	8.9	[8.0-9.8]	16.4	[15.1-17.7]	11.2	[10.2-12.2]

Source: HIS, own calculations.
FR = Flemish Region; BCR = Brussels-Capital Region; WR = Walloon Region.

As the respondents who dropped out at this stage of the interview pro-
cess had completed the face-to-face interview, it is possible to analyse
some of the characteristics of the non-respondents. There is a concern
that representation according to several social determinants is not met
in the self-administered questionnaire. Therefore, Table 9 shows the
odds ratios for dropping out of the interview process after completing
the face-to-face interview for several of the young population's back-
ground characteristics: sex, age, nationality of origin, educational level,
equivalent income and household composition.

FIGURE 1Nonresponse in the self-administered questionnaire according to region for 15- to 34-year-olds (FR = Flemish Region, BCR = Brussels-Capital Region, WR = Walloon Region)
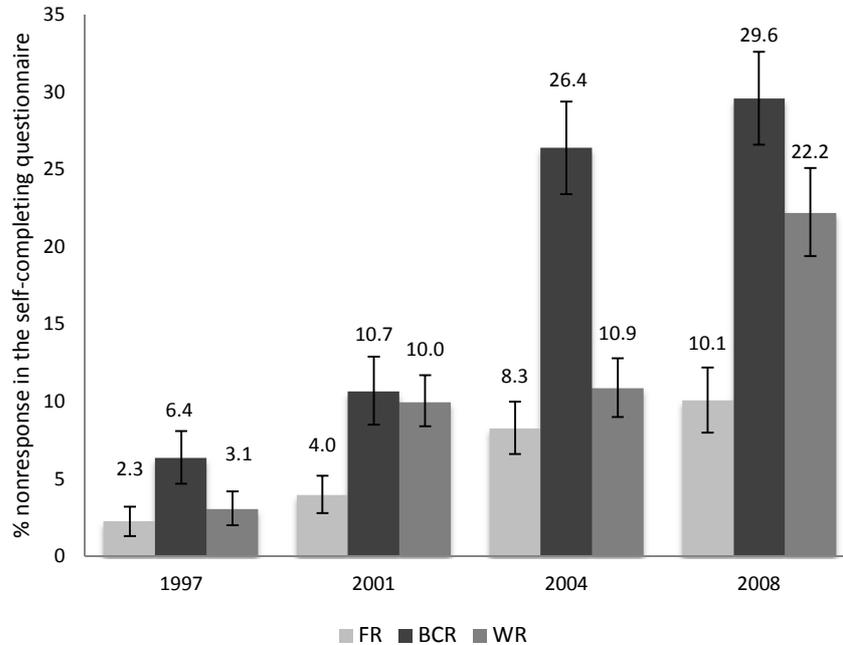
Source: HIS, own calculations.

We restrict the analysis to these variables, as these are key independent variables for social inequalities in adolescent and young adult health. The first column shows bivariate models adjusted for age, and the second column shows the multivariate logistic model. Dropouts differ significantly from respondents who completed the questionnaire in terms of most characteristics included in the model, except for equivalent income. Nonresponse is more likely for men, people of non-Belgian origin, people who live in lower-educated households, and adolescents living with one or both parent(s). There is strong variation according to household position. One-person households are more likely to complete the interview than couples with children, while children living with one or both parent(s) are the most likely to drop out. The likelihood of dropout also increases with household size (not in Table). The relation with age changes significantly after controlling for household composition: while young adults 25-34 years had lower odds of dropping out in the bivariate analysis, these odds are higher after controlling for household com-

position. Especially one-person households have lower odds of dropping
out.

TABLE 9   Odds ratios (OR) and confidence intervals (CI) of dropout
in the self-administered questionnaire for 15- to 34-year-olds,
pooled data (1997, 2001, 2004, 2008)

Characteristics	Bivariate OR[a] [CI]		Multivariate OR [CI]		N
Sex					
Men (ref)	1.00		1.00		1'649
Women	0.73	[0.61-0.87]***	0.72	[0.59-0.86]**	1'670
Nationality of origin					
Belgian (ref)	1.00		1.00		1'901
Turkish/Moroccan	1.83	[1.42-2.36]***	1.98	[1.48-2.65]**	434
European	1.52	[1.21-1.91]***	1.58	[1.24-2.02]*	639
Other	1.80	[1.36-2.38]***	2.16	[1.60-2.91]***	341
Age group					
15-19 (ref)	1.00		1.00		569
20-24	0.96	[0.74-1.25]	1.33	[1.00-1.77]*	727
25-29	0.73	[0.57-0.95]*	1.44	[1.04-2.00]*	957
30-34	0.59	[0.45-0.76]***	1.12	[0.79-1.60]	1'066
Maximum HH education					
Low	1.61	[1.30-2.00]***	1.40	[1.08-1.81]**	794
Middle	1.21	[0.98-1.15]	1.20	[0.94-1.52]	909
High (ref)	1.00		1.00		1'523
Missing	2.01	[1.25-3.24]**	1.82	[1.08-3.06]*	93
Equivalent income					
Less than 750 eur	1.27	[0.97-1.66]	0.76	[0.55-1.05]	995
750'999 eur	1.09	[0.80-1.50]	0.80	[0.57-1.14]	511
1'000-1'499 eur	1.12	[0.84-1.51]	1.00	[0.74-1.36]	800
1'500 eur or higher (ref)	1.00		1.00		686
Missing	2.88	[2.11-3.94]***	2.25	[1.61-3.15]***	327
Household composition					
One-person household	0.38	[0.27-0.54]***	0.51	[0.35-0.76]**	672
Child in single-parent or other family	1.94	[1.34-2.74]***	3.04	[2.06-4.48]***	444
Child in two-parent family	1.78	[1.28-2.49]**	2.65	[1.85-3.84]***	608
Partner in couple without children	0.71	[0.51-0.98]*	1.43	[1.01-2.02]*	482
Single parent	1.21	[0.79-1.84]	1.88	[1.15-3.06]*	164
Other	1.57	[1.09-2.24]*	1.87	[1.25-2.81]**	243
Partner in couple with child(ren) (ref)	1.00		1.00		697

[a] = Each model is age-adjusted. Equivalent income = Weighted according to the OECD's modification sca-
le (WIV-ISP 2012). (ref) = Reference category. *** = p < 0.001. ** = p < 0.010. * = p < 0.050.

Discussion

The analysis of unit nonresponse shows that although the young population is not specifically targeted, most of the demographic indicators are similarly distributed in the sample and in the total population. The WIV-ISP produced some weights that correct for province, age group, sex and household size. Impact of these weights on the estimates is negligible when focusing on the young population of the BCR alone.

The dropout rates during the interview process are of greater concern than the initial nonresponse as they are increasing per survey year and are highly selective. There is a higher dropout frequency among non-Belgians, lower-educated persons and men. Another specific group of dropouts are young persons in their twenties living with their parents. A possible explanation is that a large share of them do not reside with their parents on weekdays because they are residing elsewhere pursuing higher education. Almost one-third of the non-contacted young persons who used a proxy were enrolled in higher education. The use of proxies has important implications for the completion of the survey. In the BCR, interviewers resort to the use of a proxy – instead of re-contacting the household – more often than in other regions. While earlier research (Burzykowski *et al.*, 1999) found that participation is higher for larger households, dropout for the self-administered questionnaire increases with household size. The completion of the whole interview process for all members of the household should be the target of the interviewer. This is an issue for not only young respondents but also older adults.

Declining response rates in survey research are a worrying international trend (Stoop *et al.*, 2010), especially in urban contexts (Pickery *et al.*, 2008). Characteristics of both the target persons (e.g. survey fatigue, perceived invasion of privacy) and the survey (e.g. length, sensitive questions) can affect willingness to cooperate. However, lessons can be learned from methods used to increase response rates and minimise introduced bias. Stoop *et al.* (2010) summarise several strategies. Extended interviewer efforts, through intensive training of interviewers and the use of standardised contact forms, enhance the response rate but do not necessarily decrease bias. Incentives are most effective when they are prepaid and unconditional. Refusal conversion – the attempt to convince target persons to cooperate after an initial refusal – also increases the final response rate obtained. This strategy, in combination with asking interviewers for feedback on the response process, can identify and minimise bias. The use of incentives and intensified interviewer training

are possible solutions that were introduced for the 2013 HIS (J. Van der Heyden, personal communication, 27/10/ 2012).

Conclusion

This paper has presented various possible data sources for the study of health inequalities among young people in the BCR and has identified the opportunities and limitations for each of these sources. From an inventory of routine data, census data and survey data of interest, two sources stood out as the most apt for the study of health inequalities among young people 15 to 34 years of age: death certificates linked to census data, and the HIS.

Census data linked to death certificate data are a valuable research tool because it contains exhaustive information on the entire BCR population. Health indicators can be analysed for different subgroups and covariates, and different levels of analysis (individual, neighbourhood, municipality) are possible. Health outcomes include self-rated health (only for the 2001 census) and cause-specific mortality. The latter is obtained through linkage of census data with death certificates and provides insight into the root causes of health problems. In adolescence and young adulthood, these are mainly external causes (e.g. suicide, unintentional injury, road accident, drug overdose). While important inequalities can be identified in this way, other non-fatal health outcomes also need to be addressed. The HIS are particularly apt to fill this gap, as these surveys collect information on a wealth of health indicators and information on lifestyle behaviour.

We elaborated at length on the use of HIS data because this data source provides information on various health indicators and background characteristics, is performed every four to five years, and oversamples the BCR. We analysed nonresponse at different stages of the interview process. The reliability of the data collected through face-to-face interviews is quite good, but selective dropout exists for the self-administered questionnaire. This may bias the results, especially in more recent survey years, as nonresponse to this part of the survey has increased substantially in the BCR. Decreasing the use of proxy interviews should be used as a strategy to minimise the probability of bias in future waves. New strategies should be introduced to increase the initial response rates as well as ensure completion of the entire survey. The use of incentives and intensified interviewer training are strategies that will be intro-

duced in the next wave. When using the data for specific analyses of the youth of Brussels, the results should be treated with caution. Working with the pooled data for all survey years may minimise bias. If questions are not available for one or several survey years further weighting and sensitivity analyses, such as those introduced by Burzykowski *et al.* (1999), should be used. Although there are some limitations to using the HIS, it is important as a complement to other data sources, such as the census linked to death certificates.

Ensuring a healthy start in life should remain a priority for public health policy, and should intersect with other policy domains such as education, employment, housing and infrastructure. One of the key challenges to providing prevention and intervention strategies adapted to adolescents and young adults is the prioritising of the collection and maintenance of high-quality data sources. These data sources need to be kept up to date, and, where possible, new resources need to be attracted to improve available sources or collect supplementary data.

Acknowledgements

References

AELVOET W. (2008), *Ambtelijke Registraties voor een Evaluatief Gezondheidsbeleid. Aspecten van Beleidsonderbouwing op Landelijk, Intermediair en Lokaal Niveau,* Doctoral dissertation of medical sciences at Antwerp University, pp. 1-404.

ALLISON P. D. (2002), *Quantitative Applications in the Social Sciences: Missing Data,* Thousand Oaks, Sage Publications.

ANDERSON R. N., MINIÑO A. M., HOYERT D. L., ROSENBERG H. M. (2001), «Comparability of Cause of Death between ICD-9 and ICD-10: Preliminary Estimates», *National Vital Statistics Reports*, 49 (2), pp. 1-32.

BREIDABLIK H.-J., MELAND E., LYDERSEN S. (2008), «Self-Rated Health during Adolescence: Stability and Predictors of Change (Young-Hunt Study, Norway)», *European Journal of Public Health*, 19 (1), pp. 73-78, http://dx.doi.org/10.1093/eurpub/ckn111.

BRINDIS C. D., PARK J. M., VALDERRAME T. L., LEE C. M., MARGOLIS R., KOLBE L. J., ACHREKAR A. P., HANNAN C., ANGLIN T. M. (2004), *Improving the Health of Adolescents and Young Adults: A Guide for States and Communities*, Atlanta, GA, Centre for Disease Control and Prevention.

BRIO (2013), *Het Nederlandstalig Onderwijs in Brussel – 19*, Retrieved on 27/03/2015 from http://www.briobrussel.be/ned/webpage.asp?WebpageId=1115.

BRUSSELS-CAPITAL HEALTH AND SOCIAL OBSERVATORY (2010), *Gezondheidsindicatoren van het Brussels Gewest*, Brussel, Gemeenschappelijke Gemeenschapscommissie.

BRUSSELS OBSERVATORIUM VOOR DE WERKGELEGENHEID (2011), *Statistisch Portret van de Brusselse Jongeren bij hun Toetreding tot de Arbeidsmarkt*, Brussel, ACTIRIS.

BUBBAGE C. (1832), «Of Price as Measured by Money. Chapter 17», C. BUBBAGE, (ed), *On the Economy of Machinery and Manufactures*, London, Charles Knight.

BURZYKOWSKI T., MOLENBERGHS G., TAFFOREAU J., VAN OYEN H., DEMAREST S., BELLAMAMMER L. (1999), «Missing Data in the Health Interview Survey 1997 in Belgium», *Arch Public Health*, 57, pp. 107-129.

CENTRUM VOOR LEERLINGENBEGELEIDING (CLB) (2012), Leerlingenbegeleiding: Administratie en Registratie Systeem (LARS) concept. Retrieved on 22/10/12, from: http://www.clb-lars.be/content/concept.

CHITTLEBOROUGH C. R., TAYLOR A. W., BAUM F. E., HILLER J. E. (2009), «Monitoring Inequities in Self-Rated Health over the Life Course in Population Surveillance Systems», *American Journal of Public Health*, 99 (4), pp. 680-689, http://dx.doi.org/10.2105/AJPH.2008.141713.

COHEN J., BILSEN J., MICCINESI G., LÖFMARK R., ADDINGTON-HALL J., KAASA S., NORUP M., VAN DER WAL G., DELIENS L. (2007), «Using Death Certificate Data to Study Place of Death in 9 European Countries: Opportunities and Weaknesses», *BMC Public Health*, 283 (7).

CURRIE C., ZANOTTI C., MORGAN A., CURRIE D., DE LOOZE M., ROBERTS C., SAMDAL O., SMITH O. R. F., BARNEKOW V. (2012), *Social Determinants of Health and Well-Being Among Young People,* Health Policy for Children and Adolescents, 6, pp. 1-218.

DEBOOSERE P., EGGERICKX T., VAN HECKE E., WAYENS B. (2009), «De Brusselse Bevolking: een Demografische Doorlichting», *Brussels Studies, Synthesenota,* 16 (3), pp. 1-17.

DEBOOSERE P., GADEYNE S. (1999), «De Nationale Databank Mortaliteit. Aanmaak van een Databank voor Onderzoek van Differentiële Sterfte naar Socio-Economische Status en Leefvorm», *Steunpunt Demografie Working Paper,* (7), pp. 1-28.

DECANT P. (2007), *Le développement d'indicateurs sanitaires locaux : recherche sur l'utilisation des données de médecine scolaire dans le contexte de la région Bruxelloise,* Master Thesis in Public Health (Promes), ULB.

DECANT P., **FAVRESSE D.**, **DE SMET P.**, **GODIN I.** (2009), *Recherche exploratoire sur le processus de recueil informatisé de données sanitaires des enfants scolarisés en Communauté Française*, Brussels, Sipes, ULB.

DE GRANDE H., **DEBOOSERE P.**, **VANDENHEEDE H.** (2013), «Evolution of Educational Inequalities in Mortality Among Young Adults in an Urban Setting», *International Journal of Public Health*, 58, pp. 1-11, http://dx.doi.org/10.1007/s00038-013-0478-x.

DE GRANDE H., **VANDENHEEDE H.**, **GADEYNE S.**, **DEBOOSERE P.** (2014), «Health Status and Mortality Rates of Adolescents and Young Adults in the Brussels-Capital Region: Differences According to Region of Origin and Migration History», *Ethnicity & Health*, pp. 1-22.

DE HENAUW S., **DE SMET P.**, **AELVOET W.**, **KORNITZER M.**, **DE BACKER G.** (1998), «Misclassification of Coronary Heart Disease in Mortality Statistics. Evidence from the WHO-MONICA Ghent-Charleroi Study in Belgium», *Journal of Epidemiology and Community Health*, 52 (8), pp. 513-519, http://dx.doi.org/10.1136/jech.52.8.513.

DE SPIEGELAERE M., **CLOSON M.**, **DEBOOSERE P.**, **HUMBLET P.** (2009), «Gezondheid en Levenskwaliteit in Brussel», *Brussels Studies*, Synthesenota, (11), 11 p.

DIEZ ROUX A. V. (2001), «Investigating Neighborhood and Area Effects on Health», *American Journal of Public Health*, 91 (11), pp. 1'783-1'789.

GADEYNE S. (2006), *The Ultimate Inequality. Socio-economic Differences in All-Cause and Cause-Specific Mortality Differences in Belgium in the First Part of the 1990s*, PhD Dissertation, Brussels, Interface Demography, VUB.

GALOBARDES B., **SHAW M.**, **LAWLOR D. A.**, **LYNCH J. W.**, **SMITH G. D.** (2006), «Indicators of Socioeconomic Position (Part. 1)», *Journal of Epidemiology and Community Health*, 60 (1), pp. 7-12, http://dx.doi.org/10.1136/jech.2004.023531.

GODIN I. (2000), *Étude de faisabilité des données sanitaires des centres de médecine scolaire. Rapport sur l'état des lieux de la situation et ébauches de recommandations.* Master Thesis Promes (Public Health Research), ULB.

HAGQUIST C. E. I. (2007), «Health Inequalities Among Adolescents – The Impact of Academic Orientation and Parents' Education», *European Journal of Public Health*, 17 (1), pp. 21-26, http://dx.doi.org/10.1093/eurpub/ckl087.

HALFON N., **INKELAS M.**, **HOCHSTEIN M.** (2000), «The Health Development Organization: An Organizational Approach to Achieving Child Health Development», *Milbank Quarterly*, 78 (3), pp. 447-497, http://dx.doi.org/10.1111/1468-0009.00180.

HEALTH INTERVIEW SURVEY (HIS), Survey detailed in: https://his.wiv-isp.be/SitePages/Home.aspx by the Scientific Institute for Public Health (WIV/ISP).

HEALTH BEHAVIOUR OF SCHOOL-AGED CHILDREN (HBSC), Survey detailed in: http://
www.hbsc.org/about/index.html by the HBSC Research Team in cooperation with
WHO Regional Office for Europe.

HOUTTEKIER D., COHEN J., BILSEN J., DEBOOSERE P., VERDUYCKT P., DELIENS L. (2009), «De-
terminants of the Place of Death in the Brussels Metropolitan Region», *Journal of
Pain and Symptom Management*, 37 (6), pp. 996-1005, http://dx.doi.org/10.1016/j.
jpainsymman.2008.05.01.

IDLER E. L., BENYAMINI Y. (1997), «Self-Rated Health and Mortality: A Review of
Twenty-seven Community Studies», *Journal of Health and Social Behavior*, 38,
pp. 21-37, http://dx.doi.org/10.2307/2955359.

INGENBLEEK A., COPPIETERS Y., LAMMENS L., COLS F., D'HOORE W., DEBOOSERE P., LEVÊQUE
A. (2010), *Cadre de référence pour le système belge d'information sanitaire à visée
longitudinale et lien avec la plate-forme eHealth,* Working Paper N°1, Project BeL-
HIS, AGORA AG/JJ/13, Bruxelles.

INTERMUTUALIST AGENCY (IMA-AIM) (2013), IMA-AIM Atlas. Retrieved on 03/12/13,
from: http://atlas.ima-aim.be/home.aspx?ReturnUrl=%2f.

JOP-MONITOR, Survey detailed in: http://www.jeugdonderzoeksplatform.be/activi-
teiten/monitor.htm by the Jeugdonderzoeksplatform (Youth Research Platform).

JOUGLA E., ROSSOLLIN F., NIYONSENGA A., CHAPPERT J. L., JOHANSSON L. A., PAVILLON G.
(2001), *Comparability and Quality Improvement of European Causes of Death Sta-
tistics-Final Report*, Brussels, European Commission, DG Sanco & Eurostat. Retriev-
ed from: http://ec.europa.eu/health/ph_projects/1998/monitoring/fp_monitor-
ing_1998_frep_04_en.pdf.

JOUGLA E., ROSSOLIN F., PAVILLON G. (2008), «Cause of Death Statistics: Production
Process, Quality, and International Comparability», P. BOYLE, M. SMANS (eds), *IARC
Scientific Publication: Atlas of Cancer Mortality in Europe,* pp. 1-7. Retrieved from:
http://www.iarc.fr/en/publications/pdfs-online/epi/sp159/.

KJELLSTROM T. (2008), *Our Cities, our Health, our Future. WHO Commission on Social
Determinants of Health from the Knowledge Network on Urban Settings Acting on
Social Determinants for Health Equity in Urban Settings,* Kobe, Japan, WHO Centre
for Health Development.

KOIVUSILTA L., ARJA R., ANDRES V. (2003), «Health Behaviours and Health in Adoles-
cence as Predictors of Educational Level in Adulthood: A Follow-Up Study from Fin-
land», *Social Science & Medicine*, 57 (4), pp. 577-593, http://dx.doi.org/10.1016/
S0277-9536 z(02)00405-7.

KOIVUSILTA L. K., RIMPELA A. H., KAUTIAINEN S. M. (2006), «Health Inequality in Adoles-
cence. Does Stratification Occur by Familial Social Background, Family Affluence, or
Personal Social Position?», *BMC Public Health*, 6 (110).

KUNST A., BOS V., MACKENBACH J. (2001), *Monitoring Socio-Economic Inequalities in
Health in the European Union: Guidelines and Illustrations,* Rotterdam, Department
of Public Health.

LAMMENS L., DE SPIEGELAERE M., DEBOOSERE P. (2009), *Haalbaarheidsstudie naar het Gebruik van IMA Gegevens voor de Constructie van Gezondheidsindicatoren op Wijkniveau*, Internal document VUB/Brussels-Capital Health and Social Observatory.

LOOPMANS L. M., KESTELOOT K. C. (2009), «Sociale Ongelijkheden», *Brussels Studies*, Synthesenota, (15), 12 p.

MOUREAUX C., HEYMANS I., MOREAU N., HUMBLET P. (2010), *Recherche sur l'utilisation des données de l'Agence intermutualiste – 3ème étape*, Internal Document Brussels-Capital Health and Social Observatory.

NILSTUN T., CARTWRIGHT C., LÖFMARK R., DELIENS L., FISCHER S., MICCINESI G., NORUP M., VAN DER HEIDE A., EURELD (EUROPEAN END-OF-LIFE DECISION) CONSORTIUM (2006), «Access to Death Certificates: What Should Research Ethics Committees Require for Approval?», *Annals of Epidemiology,* 16 (4), pp. 281-284, http://dx.doi.org/10.1016/j.annepidem.2005.01.010.

OZER E. M., IRWIN JR C. E. (2009), «Adolescent and Young Adult Health: From Basic Health Status to Clinical Interventions», R. M. LERNER, L. STEINBERG (eds), *Handbook of Adolescent Psychology*, 3rd ed., New Jersey, John Wiley & Sons Inc., pp. 618-641.

PATTON G. C., COFFEY C., SAWYER S. M., VINER R. M., BOSE K., VOS T., FERGUSON J., MATHERS C. (2009), «Global Patterns of Mortality in Young People: A Systematic Analysis of Population Health Data», *The Lancet*, 374, pp. 881-892, http://dx.doi.org/10.1016/S0140-6736(09)60741-8.

PEARCE D. (2004), *Partnership on Health Statistics,* Report on the first year activities. Presentation at the Directors of Social Statistics (DSS), Meeting 29-30/09/04, Luxembourg, European Commission/Eurostat.

PICKERY J., CARTON A. (2008), «Oversampling in Relation to Differential Regional Response Rates», *Survey Research Methods*, 2 (2), pp. 83-92.

SOLAR O., IRWIN A. (2007), *A Conceptual Framework on the Social Determinants of Health*. Retrieved on 09/06/11, from: http://www.who.int/social_determinants/resources/csdh_framework_action_05_07.pdf.

STATISTICS BELGIUM (2014), *Census 2011*. Retrieved on 20/03/14, from: http://statbel.fgov.be/nl/statistieken/gegevensinzameling/volkstelling/Census2011.

STOOP I., BILLIET J., KOCH A., FITZGERALD R. (2010), *Improving Survey Response. Lessons Learned from the European Social Survey,* Chicester, John Wiley & Sons, http://dx.doi.org/10.1002/9780470688335.

SZRETER S. (2004). Industrialization and health. *British Medical Bulletin*, 69, pp. 75-86, http://dx.doi.org/10.1093/bmb/ldh005.

VAN CASTEREN V. (2011), *Sentinel General Practitioners*. Retrieved on 22/10/12, from: https://www.wiv-isp.be/epidemio/epinl/index10.htm.

VAN DE SANDE S., **DE WACHTER D.**, **SWARTENBROECKX N.**, **PEERS J.**, **DEBRUYNE H.**, **MOLDE-NAERS I.**, **LEJEUNE B.**, **VANDAMME V.**, **RAMAEKERS D.**, **LEYS M.** (2006), *Inventaris van Databanken Gezondheidszorg*, KCE Reports, Vol. 30A, Brussel, Federaal Kenniscentrum voor de gezondheidszorg (KCE), Ref. D/2006/10.273/16.

VAN DER HEYDEN J. (2012), Personal communication with WIV-ISP staff member on nonresponse in HIS, 27/10.

WEST P. (1997), «Health Inequalities in the Early Years: Is There Equalisation in Youth?», *Social Science & Medicine (1982)*, 44 (6), pp. 833-858.

WHO (2004), «Rules and Guidelines for Mortality Coding», *International Statistical Classification of Diseases and Related Health Problems. Tenth Revision,* Geneva, WHO, ICD-10, pp 31-128.

WHO (2010), *Why Urban Health Matters.* Retrieved on 22/10/12, from: http://www.who.int/world-health-day/2010/media/whd2010background.pdf.

WIV-ISP (2005), *The Health Interview Survey 2004 Belgium. Research Protocol*. Retrieved on 22/10/12, from: https://www.wiv-isp.be/epidemio/epien/crospen/hisen/his04en/protocol2004.pdf

WIV-ISP (2012), *Manual for the Users of the Health Interview Survey*, Brussels, WIV-ISP.

List of abbreviations (some abbreviations come from Dutch
or French terminology, but only the English translation
is mentioned here)

BCR	Brussels-Capital Region
CLB	Pupil Guidance Centres
CPP	Commission for Protection of Privacy
HBSC	Health Behaviour of School-aged Children Survey
HIS	Health Interview Survey
ICD	International Classification of Diseases
IMA/AIM	Intermutualist Agency
ISTAT	Italian National Institute of Statistics
JOP	Youth Research Platform
KCE	Belgian Health Care Knowledge Centre
LARS	Pupil Administration and Registration System
NIHDI	National Institute for Health and Disability Insurance
NR	National Register
PSE	Centre for the Promotion of Health at School
ULB	University of Brussels (French-language)
VAZG	Flemish Agency for Health and Care
VUB	University of Brussels (Dutch-language)
WIV-ISP	Scientific Institute of Public Health