

# Registers of the Swedish total population and their use in medical research

Jonas F. Ludvigsson<sup>1,2,3</sup> · Catarina Almqvist<sup>1,4</sup> · Anna-Karin Edstedt Bonamy<sup>5,8</sup> · Rickard Ljung<sup>6</sup> · Karl Michaëlsson<sup>7</sup> · Martin Neovius<sup>5</sup> · Olof Stephansson<sup>5,8</sup> · Weimin Ye<sup>1</sup>

Received: 2 November 2015 / Accepted: 8 January 2016 / Published online: 14 January 2016  
© Springer Science+Business Media Dordrecht 2016

**Abstract** The primary aim of the Swedish national population registration system is to obtain data that (1) reflect the composition, relationship and identities of the Swedish population and (2) can be used as the basis for correct decisions and measures by government and other regulatory authorities. For this purpose, Sweden has established two population registers: (1) The Population Register, maintained by the Swedish National Tax Agency (“Folkbokföringsregistret”); and (2) The Total Population Register (TPR) maintained by the government agency Statistics Sweden (“Registret över totalbefolkningen”). The registers contain data on life events including birth, death, name change, marital status, family relationships and migration within Sweden as well as to and from other countries. Updates are transmitted daily from the Tax Agency to the

TPR. In this paper we describe the two population registers and analyse their strengths and weaknesses. Virtually 100 % of births and deaths, 95 % of immigrations and 91 % of emigrations are reported to the Population Registers within 30 days and with a higher proportion over time. The over-coverage of the TPR, which is primarily due to underreported emigration data, has been estimated at up to 0.5 % of the Swedish population. Through the personal identity number, assigned to all residents staying at least 1 year in Sweden, data from the TPR can be used for medical research purposes, including family design studies since each individual can be linked to his or her parents, siblings and offspring. The TPR also allows for identification of general population controls, participants in cohort studies, as well as calculation of follow-up time.

✉ Jonas F. Ludvigsson  
jonasludvigsson@yahoo.com

<sup>1</sup> Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, 17177 Stockholm, Sweden

<sup>2</sup> Department of Paediatrics, Örebro University Hospital, Örebro, Sweden

<sup>3</sup> Division of Epidemiology and Public Health, School of Medicine, University of Nottingham, Nottingham, UK

<sup>4</sup> Astrid Lindgren Children’s Hospital, Karolinska University Hospital, Stockholm, Sweden

<sup>5</sup> Clinical Epidemiology Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden

<sup>6</sup> Unit of Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

<sup>7</sup> Department of Surgical Sciences, Uppsala University, Uppsala, Sweden

<sup>8</sup> Department of Women’s and Children’s Health, Karolinska Institutet and Hospital, Stockholm, Sweden

**Keywords** Population · Register · Registry · Public health

## Abbreviations

PIN	Personal identity number
PR	Population Register
PR-Tax	Population Register (maintained by the Swedish Tax Agency)
TPR	Swedish Total Population Register (maintained by Statistics Sweden)

## Introduction

In Sweden, there are two Population Registers (PRs) tracking Swedish residents: the Population Register (PR-Tax) and the Total Population Register (TPR). The former

is maintained by the Swedish Tax Agency and the latter by Statistics Sweden. Both registers are instrumental to epidemiological research in Sweden, and can be linked to other health registers through the personal identity number (PIN) [1]. The current paper aims to describe vital aspects of the Swedish PRs, and builds on our previous paper about the PIN [1].

The aim of the Swedish national population registration (Swedish: folkbokföring) and the PR-Tax is to obtain data that (1) reflect the composition, family relationship and identities of the Swedish population and (2) can be used as the basis for correct decisions and measures (e.g., taxation and allowances) by government and other organisations. Major events in an individual's life, such as birth, death, marriage, name changes, family relationships, change of sex, but also place of residence are continuously recorded in the national registration. Being recorded in the Swedish PRs is tied to both rights and duties, including social benefits, paying tax and the right to vote. Figure 1 gives an overview of the Swedish population since 1968 according to year and age group. The population in Sweden was 9,775,572 on April 30, 2015.

The aim of the TPR is to produce statistics and most researchers will only interact with the TPR. Although the PR-Tax and the TPR are two different registers, they share all variables and are therefore jointly described in this paper.

### History of the Population Registers

The Church of Sweden has kept local registers of their parish members since the 17<sup>th</sup> century [2]. This procedure allowed the church, and consequently the Swedish State, to collect taxes and enrol soldiers.

In 1991, the Swedish Tax Agency took over the responsibility for the national registration and the Population Register from the Church of Sweden. The Tax Agency is organised around local units collecting data (Fig. 2). Local data are then stored in the national PR-Tax.

Most people who are born in Sweden or move to the country are recorded in the national PR-Tax (exceptions are described in the section about the PIN). On entering the register, the individual is assigned a PIN [1]. The PR-Tax also distributes population information (including address information) to various government agencies. For these reasons, Swedish residents moving *within* Sweden have to notify the Tax Agency when they change their address.

Five times a week the Tax Agency distributes new data from the PR-Tax to relevant government authorities, among them the Swedish Migration Board, the Swedish Pension Agency and Statistics Sweden (Fig. 2). Each year, approximately three million updates are sent from the Swedish Tax Agency. Transmission of data to other

government agencies is regulated by law (SFS1991:481 and SFS2001:100).

Certain updates (primarily names and addresses) of the PR-Tax are accessible to a number of actors, including private customers through a service called "SPAR" (the Swedish Population and Address Register). This service allows, for instance, mail order companies to contact a large number of potential customers. Government agencies, banks and insurance companies have access to additional data, whereas some highly sensitive data are only distributed to the Police and the Customs Office.

The TPR (Swedish: Registret över TotalBefolkningen, RTB) started in 1968 after a large part of the population data was computerised in 1967. A complete year-specific version of the TPR is available for each year since 1968 (certain data are actually available since 1961).

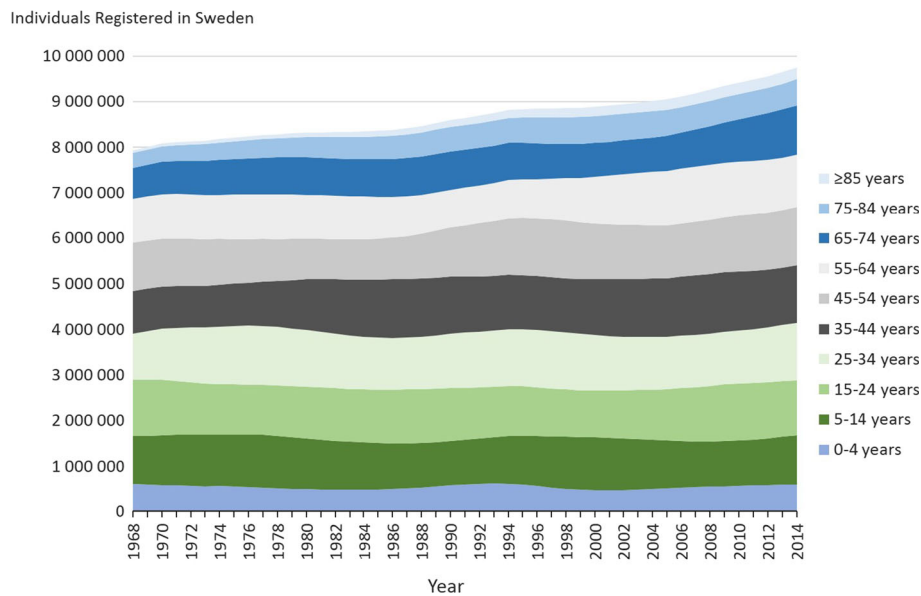
The variables in the PRs have remained relatively constant since 1968 (Table 1), with a few changes. Until 1973, *birth number* was only recorded for births within a marriage while *birth order* signifies the order of births in the same mother. In 1974, regulations on divorce shortened the administrative time, resulting in an increase in the number of registered divorces in 1974 and 1975. Rules for citizenship of children whose parents had different citizenship changed in 1975. Regulations for a widow's pension were changed in 1990, which led to a substantial increase in the number of married couples the year before (1989).

Since 2011, the TPR contains information on apartment (some data available since 2006). Apartment data are collected by the city councils, merged and administered by the Swedish National Land Survey (Swedish: Lantmäteriet), and then transmitted to Statistics Sweden. This allows Statistics Sweden to calculate household-specific information such as size and type of dwelling and number of children per household. It also makes it possible to identify cohabitants.

### Content of the Population Registers (PRs)

#### Birth

At birth, the midwife or the hospital reports the birth to the PR-Tax. The child is automatically assigned a PIN and the Tax Agency records the citizenship of the child as well as the newborn's family name. If the parents (couple) are married, the husband is assumed to be the biological father of the child, and parents will automatically have joint custody. If the couple is not married, paternity is confirmed through a notification to the Social Welfare Committee (Swedish: Socialnämnden) signed by both parents. At that stage, non-married parents are also asked to report whether they want to have joint custody of the child. If, instead,



**Fig. 1** The Swedish population by year and age group from 1968 to 2014. The total population exceeded 8 million in 1979 and 9 million in 2004. Of note, the number of individuals aged 85+ years has increased from 61,000 (0.8 % of the population) in 1968 to 256,000 (2.6 % of the population) in 2014. Meanwhile, the number of children

aged 0–14 years has remained relatively constant (1,658,000 (20.9 % of the population) in 1968 and 1,682,000 (17.3 % of the population) in 2014). Data originate from Statistics Sweden. Accessed on June 2, 2015 <http://www.statistikdatabasen.scb.se>

paternity is confirmed through a court decision, the court will inform the Swedish Tax Agency of its decision. Further, a woman living with the biological mother can be registered as a “parent”, but not as the *biological mother* (for a same-sex parent to be registered as a “parent”, the women have to be married, or live together; similar rules apply to male parents).

## Death

When a person in Sweden dies, the physician confirming the death has to submit a death certificate to the Tax Agency. For Swedish citizens who die abroad, the Swedish embassy or consulate in the relevant country must inform the Tax Agency about the death. However, if a foreign citizen residing in Sweden dies abroad (outside Sweden), the family of the deceased is *not* obliged to inform the Tax Agency about the death (but often the relatives do so).

## Emigration

Individuals who move outside Sweden and plan to stay abroad for  $\geq 1$  year should be *de-registered* from the PRs. This rule also applies to individuals who plan to stay abroad for  $< 1$  year but then decide to stay away longer. Moving abroad does not influence citizenship. Individuals spending more than a year abroad are not removed from the PRs provided that they are working in the service of the Swedish state, or if absence is due to sudden disease or

imprisonment, this also applies to their family members. Swedish priests and missionaries working abroad and their family were previously included in the PRs even if staying abroad for  $\geq 1$  year. However, this exemption has been removed. Swedish individuals who work for the European Union (EU) or United Nations, and Swedish students going abroad for  $\geq 1$  year because of academic studies are regarded as emigrated and will be removed from the PRs.

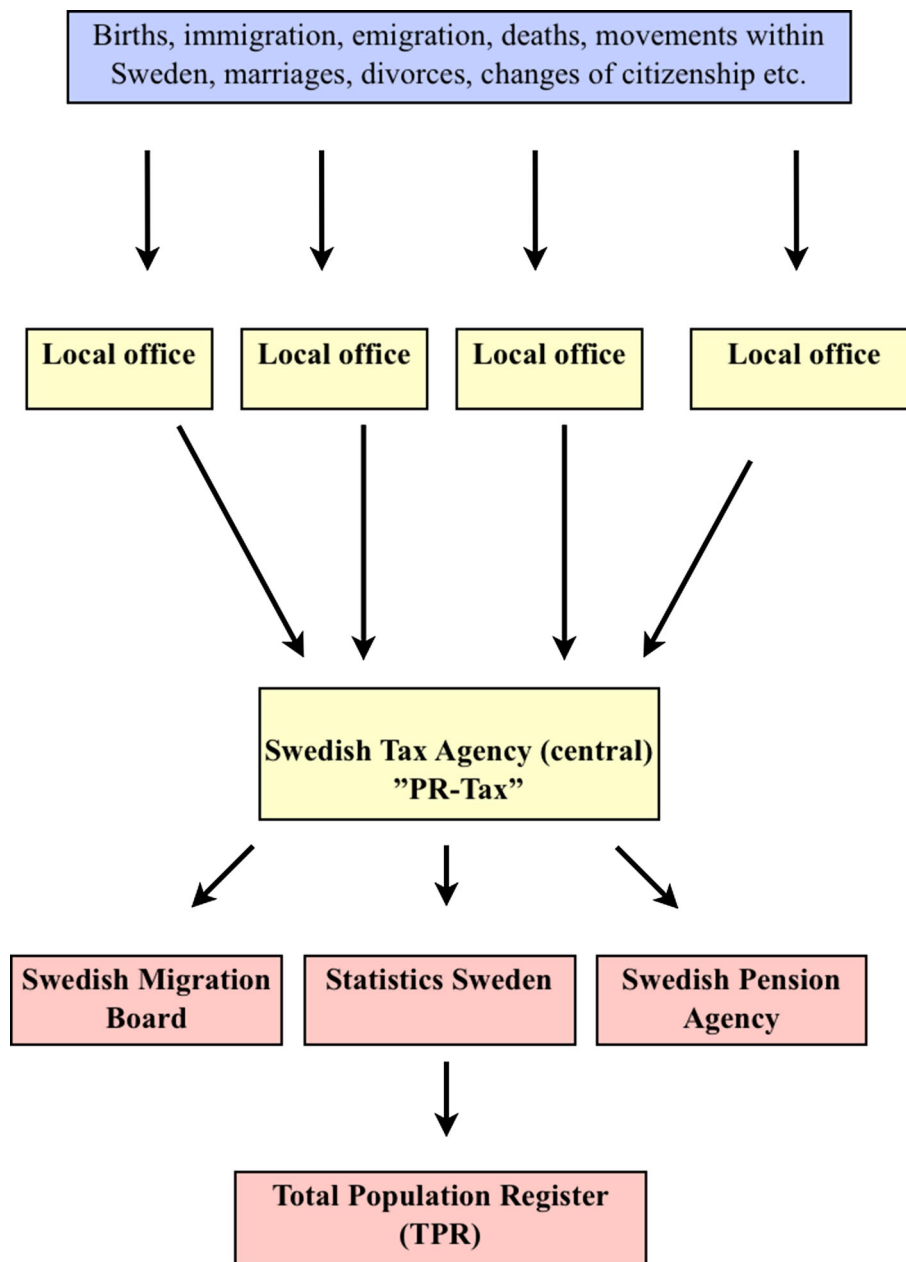
Special rules for emigration apply to Swedish individuals moving to Denmark, Finland, Iceland or Norway. These individuals will remain in the PRs until the receiving country notifies the Swedish individuals that from now on they will be included in *their* PR (e.g., in Norway this takes 6 months). Individuals living both abroad and in Sweden can choose to remain in the Swedish PRs provided that they spend at least  $1/7$  of each year in Sweden (1 day per week or equivalent). Emigration and re-immigration do not affect one’s PIN (all emigrants retain their PIN).

## Immigration to Sweden

Individuals moving to Sweden and planning to stay there for  $\geq 1$  year will in most cases be recorded in the PRs, provided that they fulfil certain criteria for work, education or self-support.

Until 1998, only the year of immigration was recorded for individuals who had immigrated before 1968 (sometimes not even the year of immigration was recorded).

**Fig. 2** Flowchart: The Swedish Tax Agency delivers data to Statistics Sweden and its Total Population Register (TPR). In 2015, the Swedish Tax Agency had 103 local offices



### Migration within Sweden

All individuals who move within Sweden must notify the Swedish Tax Agency. The place of residency is the place where the individual lives most of the time and where the person sleeps for at least one out of seven nights. Each individual can only have one registered address. This also applies to children with two guardians at different addresses. In 2012, one out of four children ( $n = 480,000$ ) had separated parents (including those whose parents had

never lived together) [3]. Of these children, some 35 % had two residences [3].

Children who must relocate because of academic studies will remain being registered with their parents until the age of 18 years or up until 21 years of age if they need to complete elementary school or high school, i.e. non-university studies. A registered address will not be changed because of hospital admission or imprisonment or because an elderly person has been moved for caring reasons to a retirement home in a different city.

**Table 1** Selected variables in the Total Population Register (TPR) at Statistics Sweden

Variable	Additional data stored in the TPR on each individual	EU variables <sup>a</sup>	Family variables <sup>b</sup>
Personal identity number (PIN)	PIN data on husband/wife, partner, biological mother, biological father, parent, child, adopted children (as well as children “adopted away”), adoptive mother, adoptive father and “other persons”, including custodians (all the above individuals are sometimes called “relation persons”)		Family identity per household
Confidential/protected identity			
Sex	Sex of partner and of relation person(s) are also recorded		
Name: Given name, middle name, family name			
Civil status	Duration, dates and changes in civil status (see text for categories)	Civil status	Type of family. Number of individuals in the family. Number of children
Duration of marriage			
Date entering the population registers	Data on register entry also for adoptive fathers, adoptive mothers and custodians		
Date of deregistration	Date when an individual is removed from the population registers (usually on emigration)		
<i>Birth-related data</i>			
Date of birth	Date of adoption and date of custody are also registered		
Parish at birth <sup>d</sup>	From January 2016, district/city at birth will be recorded for all newborns instead of parish		
Country of birth	Of the individual and of his/her parents	Country and region of residence	
County of birth	For individuals born in Sweden		
Age in years and days on Dec 31 (current year)		Age	Number of children below the ages 6, 18, 22 years. Number of children born. Age of youngest child in family
Birth order <sup>c</sup>	Birth order of the father and mother is also recorded		
<i>Address data</i>			
Address including postal area	Up to four addresses in Sweden or abroad can be registered	Degree of urbanization	
Housing/residence	Includes information on household since 2011		
<i>Other</i>			
Income			
<i>Origin</i>			
Date of immigration, country of immigration, municipality			
County code	Name of county		
County			
Citizenship			
Date of citizenship			
Emigration: date and country	If the individual has emigrated to another Nordic country, the city in the new country is recorded		

TPR Total Population Register, EU European Union

<sup>a</sup> Statistics Sweden also stores a set of EU variables

<sup>b</sup> Data are usually registered for “family” and separately for “household” (the latter may include more individuals than the family)

<sup>c</sup> Birth order refers to all children of the same mother. Until 1973, “birth number” only applied to children born within a marriage

<sup>d</sup> Until 1999, it was possible to be born in a non-territorial parish in Sweden. There used to be five non-territorial parishes in the Church of Sweden: the German congregation in Stockholm, the Finnish congregation in Stockholm, the congregation of the Royal Court, the admiralty in Karlskrona and the German congregation in Gothenburg. None of these parishes formally exists since 1999. From 2016, national registration in newborns will consider “city” (Swedish *kommun*) and no longer parish

## Marriage and divorce

The Swedish PRs record marriages. Marriage of children <18 years are however only recognised if both individuals were non-Swedish and  $\geq 15$  years at time of marriage (and only then will the marriage be treated on a case-by-case basis).

Information about divorce (dissolution of marriage) is submitted by the Swedish district court to the Tax Agency where it is recorded. When the divorce has taken place abroad, those involved in the divorce must inform the Tax Agency.

Until 2009 Swedish government agencies also registered “partnership” (and “separated after partnership”). Data on partnership can still be found in the Swedish Population Registers but no new “partnership data” are entered into the registers after 2009 when marriage in Sweden became gender neutral.

## Country of birth

Between 1947 and 2013, children born abroad were regarded as born in Sweden provided that certain conditions of their parents were met. The child was then assigned the same county and parish of birth as that recorded in the mother at the time of birth. For non-Swedes, country of birth is registered with the name of the country at year of birth of each individual (even if that country has ceased to exist, e.g. Soviet Union). Before 1947, country of birth other than Sweden was usually recorded as “foreign-born” without further clarification.

## Sex

The PRs record biological sex. Men are assigned “1” and women “2”. The code for sex is generated from an individual’s PIN (the second last digit is odd in men and even in women).

Since 1972, the sex designation of an individual can be changed after application to the National Board of Health and Welfare. Upon approval of that application, the PIN and the sex in the PRs are replaced.

## Citizenship

Among Swedish citizens, only Swedish citizenship is listed. In non-Swedish residents the PRs will record up to three citizenships per person, with the citizenship obtained first in time listed first. Citizenship from states that are not yet acknowledged by the Swedish government is not registered. Since April 1, 2015, the Swedish Tax Agency records a Swedish citizenship for every child born to a Swedish mother or father independent of the country of

birth. Before that date, children to Swedish mothers were always given Swedish citizenship but only some children to Swedish fathers (provided the father was married to the mother or if the non-Swedish mother lived in Sweden).

If both parents lack Swedish citizenship, a child will be assigned citizenship according to the laws of the country of the parents (the child will not automatically become a Swedish citizen even if born in Sweden). Citizenship does not automatically imply that the child is registered in the PRs. Most children born abroad to Swedish parents will not enter the PRs until moving to Sweden, the exception being children born to certain diplomats or individuals working abroad on behalf of the Swedish State.

When a child is adopted in Sweden, both the adoptive parents and the biological parents are recorded in the PRs. After a Swedish court has decided on an adoption, the court will inform the Tax Agency about its decision. In the event an adoption has taken place outside Sweden the Swedish adoptive parents have an obligation to inform the Tax Agency when the child enters Sweden in order for the child to obtain Swedish citizenship.

## Special registers based on the TPR

Using the TPR, Statistics Sweden each year produces a number of special registers. Among them are the (1) Multi-generation register (contains individuals born since 1932 and any relationship to biological and adoptive parents) [4], (2) the Migration register (the most important data in this register are grounds for residence and the reason to immigrate) and (iii) the PIN register [1], which contains all PINs in the TPR since 1968 as well as any PIN recorded in the annual censuses since 1960 or later. This register also contains PIN *changes*. A detailed description of circumstances when a PIN may be changed has been published elsewhere [1].

## Updates

Statistics Sweden provides a monthly, quarterly, and annual version of the TPR. The monthly version of the TPR is used during 1 month and then replaced by the next month’s TPR. The *annual TPR* refers to the population on December 31 of each year. Only the annual register data (updates), available since 1968, are used as the basis for individual statistical registers constructed by Statistics Sweden and for the selection of controls in research. Statistics Sweden also constructs a family register. Previously, the TPR underestimated the number of cohabitants in Sweden (some unmarried individuals living together without common children lacked a formal relation in the register). Through apartment data it is now possible to identify cohabitants living in apartments.



## Unique identifiers

### Personal identity number (PIN)

Linking data in the PRs with other data is possible through the use of the PIN [1]. The PIN consists of birthdate, birth number and a check digit. A detailed description of its construction has been published elsewhere [1]. The vast majority of children born in Sweden receive a PIN at birth. Almost all Swedish delivery units have an “e-birth” system. This system, initiated in 2009, generates a PIN when the delivery nurse enters the sex and time of birth of the newborn. Before 2009, a newborn would receive his or her PIN only 2 weeks after birth; prior to that time, identities were tracked using reserve numbers (in some instances also called “patient numbers”). Healthcare contacts during the first 2 weeks were sometimes incompletely linked to the correct PIN, especially for same-sex twins. There is no key between reserve numbers and PINs on a national level.

Children born outside Sweden will be assigned a PIN in special circumstances. These circumstances include a child being born to a Swedish diplomat or Swedish State employee or according to special arrangements. For example, some Swedish women may choose to deliver their babies in nearby countries such as Norway or Finland because of the long distance to Swedish hospital care facilities. Children born in Sweden to foreign diplomats are not assigned a PIN and do not enter the PRs.

Immigrants can be assigned a PIN provided they plan to stay in Sweden for  $\geq 1$  year AND (1) are employed, (2) have been admitted to  $\geq 1$  year of studies or (3) provide evidence that they can support themselves. Similar rules apply to individuals from the European Economic Area (EEA) and individuals outside the EEA.

### Coordination number

An individual planning to stay in Sweden for  $< 1$  year can be assigned a unique identifier that is permanently attached to him/her: the coordination number [1]. This unique identifier is similar in structure to the PIN, except that the number 60 is added to the birthday. This means that an individual born on the May 21, 1968 is assigned a coordination number beginning 19680581. The number of individuals who have ever had a coordination number is currently estimated at 1,000,000. This figure includes individuals with the older “total population number” (“TP number”; in use until 1999), which was a *temporary* PIN preceding the use of the coordination number. Individuals with a coordination number may later receive a PIN.

The purpose of the coordination number is to help government agencies and society to keep track of individuals without a PIN, as well as to facilitate communication within and between government agencies. For an individual to receive a coordination number, the identity (name, date of birth, place of birth and citizenship) of that individual must be established. Coordination numbers are administered by the Swedish Police, the Swedish Tax Agency and universities (for temporary foreign students).

Coordination numbers are stored in the Reference Register maintained by the Swedish Tax Agency. Besides the coordination number this register stores data on name, citizenship, city and country of birth, and allows for the linkage between a coordination number and a PIN in the same individual. Individuals with coordination numbers are not included in the PRs.

## Coverage and quality

### Aspects of coverage

Although theoretically the PRs should cover everyone in Sweden, in reality this may not be the case. Individuals who should be in the PRs or have been in the PRs in Sweden make up the target population of the register. The total population in a certain year consists of all individuals who have been in the PRs at some time during that year. For this, the individuals must have lived in Sweden during at least 1/7th (i.e. 52 days) of the time during that year.

When newborns (e.g., born at home) or immigrants are not reported to the PRs, coverage of the registers is decreased. It is estimated that  $< 0.1$  % of newborns are not reported to the PRs within 30 days. Because of the comprehensive administrative process at immigration, the time from entry to Sweden to registration may be longer than that of newborns. Naturally, the PRs fail to take illegal immigrants into account.

The PRs may also overestimate the number of individuals in Sweden. Some emigrants estimate that they may benefit from remaining in the PRs and therefore fail to report emigration. Individuals who truly stay abroad for  $\geq 1$  year (and who should be removed from the PRs) contribute to over-coverage of the PRs. The over-coverage has been estimated to 0.1 % for Nordic citizens but substantially higher for individuals born outside the Nordic countries (potentially 4–8 %)(personal communication: Karin Wegfors, Statistics Sweden, May 29, 2015). One way of analysing over-coverage is to examine the relative mortality among individuals born outside Sweden versus

those born in Sweden. Death taking place among individuals who have spent a long time abroad may not be reported to the PRs and hence people leaving Sweden without reporting their emigration may seem to live longer than feasible. The Swedish Tax Agency regularly audits all individuals aged >100 years in Sweden; the last audit was in 2012.

In a report published by Statistics Sweden in 2015 [5], the agency writes that low mortality in foreign-born individuals in the PRs suggest that a substantial number of individuals (from the age of 1 year or older) are not residing in Sweden. Figures vary over time. Over-coverage of individuals born in the other Nordic countries has been estimated at 5–20,000 individuals. Because of the vast discrepancies in numbers, Statistics Sweden has recently suggested a regression model based on a number of indicators (commuting to a neighbouring country, academic studies, inactivity followed by death, whether events in the same family occur, income and migration data). According to this model, the over-coverage in the PRs was between 39,000 and 60,000 individuals in 2000–2010. Statistics Sweden currently estimates that over-coverage is equal to 0.25–0.5 % of the Swedish population (personal communication, Karin Wegfors, Statistics Sweden, May 29, 2015). A consequence of over-coverage is that individuals who have truly emigrated contribute risk time when selected for registry-based research, but no data on outcome measures.

### Quality of the PRs

The quality of PR data is generally regarded as high. Most data on births, deaths and civil status are reported by professional and administrative personnel such as midwives, physicians, court officials and wedding officials. However, quality may be lower when it comes to change of residence within Sweden or migration in that reporting of data in these areas depends on the individual.

One way to increase the quality of data is through regular audits. Statistics Sweden uses the return of tax declarations (every year) and voting cards (on Election day, every 4th year) because of missing addressees as a means to correct errors.

Another aspect of quality is timeliness. Swedish government agencies try to avoid delay of notifications. It has been estimated that 94 % of all births are reported to the Swedish Tax Agency within 2 days and another 6 % (totalling almost 100 %) within 30 days. Missing data are estimated at 0.1 % of all births. Corresponding figures for death, immigration, emigration and change of residence within Sweden are presented in Fig. 3. The proportion of immigrants registered within 30 days may vary with circumstances and is likely lower when the Swedish customs is overstrained because of high demands from asylum

seekers. On rare occasions, the delay in reporting means that a death or a birth is actually reported during the wrong year (almost always reported in the following year), thereby incorrectly influencing the estimated total population of Sweden in that year. Approximately 0.2 % of all births are reported in the year next to the actual birth year (most of these births likely took place on Dec 31).

## Use of the TPR

### Outside research

The PRs are used both internally and externally. Internal use consists of the official population statistics of Sweden, where the population is presented according to sex, age, civil status and place of residence. Vital status, divorce, immigration and emigration are also presented (several times each year). From the TPR, Statistics Sweden randomly selects a small number of individuals who are interviewed in detail in order to survey the workforce, household expenses, and living standard.

The TPR serves external users such as decision makers (e.g., those planning for child care, care of the elderly, schools, healthcare and housing construction), government departments and agencies (e.g. the National Board of Health and Welfare (Swedish: Socialstyrelsen) uses TPR data to calculate various subsidies and allowances), but also private companies, organisations and individuals (including media and journalists).

### Research use

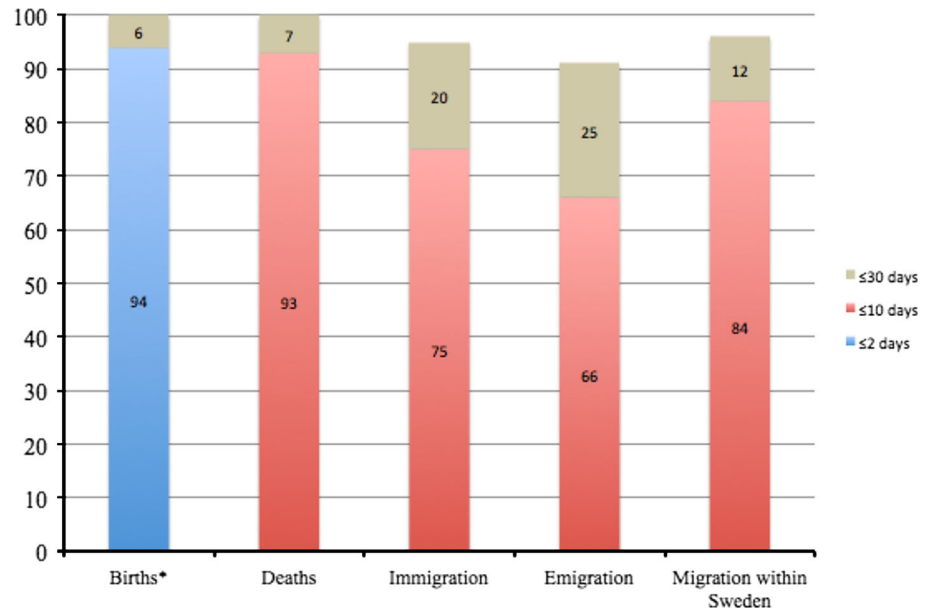
The TPR is a fundamental part of epidemiological research in Sweden. Knowledge of the number of individuals in a certain year or in a certain part of Sweden allows researchers to calculate both *prevalence and incidence* of diseases [6, 7].

Data on immigration allow for research on health among *immigrants*, [8] as well as on the aetiology of disease [9]. Lindblad et al. [10] found that ADHD medication was more common among international adoptees than in a comparison population, but *adoption* has also been linked to and risk of certain psychiatric diseases in childhood [11].

Bjorkenstam et al. [12] recently showed that *divorce* increases the risk of newly onset psychiatric disease. Wallby et al. [13] examined the healthcare contact pattern and immunisation status in children to single versus cohabiting mothers. The TPR contains extensive information on familial relations and is the basis for the Multi-generation Register [4]. The Multi-generation Register has been extensively used in research. For instance, Swedish data suggest that crime aggregates in certain families [14]



**Fig. 3** Coverage (%) of vital status and migration data in Swedish population registers. Data refer to % within 10 days (red) and within 30 days (brown), except for birth data\* where the “94 %” refers to  $\leq 2$  days (blue). About 75 % of all immigrations are reported to the Swedish Tax Agency within 10 days and another 20 % are reported within 30 days (totalling 95 %). (Color figure online)



and that maternal bereavement in the year preceding pregnancy is linked to the risk of stillbirth [15].

The TPR also contains geographical information, including street address and apartment data (the latter is the basis for information on households). This information allows for studies on socioeconomic gradients and health [16] using Statistics Sweden’s “Small Area Marketing Statistics” (SAMS) classification system.

*Migration* within Sweden has implications for the selection of controls in register-based research in that controls are often matched on residence (county, city and/or parish). A girl born in Uppsala County in 1974 would have had her complete inpatient care recorded in the National Patient Register from birth [17]. If this girl then moved to another county in 1977, she could no longer serve as a matched control to other girls in Uppsala county because her comorbidity may not have been completely recorded in the new county (the National Patient Register only became nationwide in 1987).

*Date of death* is useful for researchers examining overall mortality [18]. As opposed to the Cause of Death Register, the TPR does not contain any data on cause of death and only allows the identification of deceased individuals. Furthermore, TPR data on date of death, as well as *date of emigration* are often used to calculate a correct follow-up time in cohort studies [19–26].

Importantly, the TPR is also used to identify controls for cohort and case–control studies [27, 28]. Some Swedish epidemiological researchers will—after relevant approval

by a Regional Ethical Review Board—ask the National Board of Health and Welfare to identify individuals with a certain disease and then ask Statistics Sweden to identify matched controls [27, 29–32].

### Confidentiality, data access and ethical aspects

Statistics Sweden, and indirectly the TPR, is regulated by the personal data act (1998:204), as well as by additional regulations and laws on the treatment of personal data. As opposed to data in the PR-Tax where most data are public, data in the TPR are confidential and protected by the Secrecy Act (chapter 9, §4). This act stipulates that data *become confidential* on transmission to Statistics Sweden. However, personal data from Statistics Sweden may still be delivered to researchers after appropriate approval by an Ethical Review Board and by an additional review by Statistics Sweden [33]. This means that even if there is an approval from the (Regional) Ethical Review Board, Statistics Sweden may still deny researchers access to TPR data.

Statistics Sweden usually delivers de-identified data for research. There is currently no EU regulation for any of the Population Registers in Sweden, but an agreement on a unified data protection plan has been reached in the so-called Trilogue (the General Data Protection Regulation, GDPR) [34].

Confidentiality/secrecy in the TPR is of special relevance if access to data may harm an individual (see also

our paper on the PIN [1]). For instance, certain individuals may have a protected address or a change of PINs for security reasons. Certain personal data are regarded as especially sensitive (e.g., those that relate to adoption and citizenship). Special safety precautions should also protect e.g. (1) women who have divorced their husband and when there is a credible threat to the woman or her family, (2) individuals who are persistently harassed by political opponents (especially immigrants harassed by their former regime), (3) some judicial personnel (e.g., policemen working with serious crime) and (4) participants in witness protection programs. In October 2014, some 13,000 individuals (0.1 % of the Swedish population) had a protected identity status. Only for certain purposes will Statistics Sweden deliver data on individuals with a security indicator signalling protected identity. However, individuals with protected identity are included in the housing census. When individuals are persecuted and at high risk, the Tax Agency may allow these persons to keep their recorded place of residence even if they move to another part in the country (“kvarskrivning”). This exception is currently relevant to about 2000 people in Sweden [35].

Several ethical aspects are relevant in epidemiological research using the TPR [33], and include the use of large-scale data without prior informed consent of study participants, and the risk of breaching the identity of study participants should the data be accessed by non-researchers or for inappropriate purposes [33]. Meanwhile, the use of already collected data with national coverage allows for the creation of large and unbiased study populations, minimises costs for data collection, and eliminates certain selection bias.

## Discussion

The PRs in Sweden have a long history and are an important component of medical research. While the primary aim of national registration is to provide complete population-based data to help the government and other agencies make correct decisions, the data can also be used for medical research. The content of the PR-Tax and the TPR is very similar. Statistics Sweden does however link TPR data to additional individual-based data, and serves medical researchers with data. Similar PRs are also available in other Nordic countries [33]. Given the comparable legal frameworks and handling of ethics applications, it is natural that there is already a number of Nordic research collaborations involving the Swedish TPR [36–39].

As mentioned above, the TPR can be used for a range of research purposes, including follow-up of study participants as it contains data on date of death and emigration. It

can also be used for selection of controls and study participants in cohort studies. Of utmost importance in epidemiological research is that controls and cases are selected from the same source population [40]. Data on residence of living is also important given the different coverage of the National Patient Register up until 1987 when this register became nationwide [17].

The main strengths of the national PRs of Sweden are their completeness and the link to the PIN [1] allowing data linkage and consistency of definitions (few variables have changed in the past 40 years). The registers have some limitations. One minor limitation is that children with separated parents can have only one registered address, although they may alternate between several addresses [41]. Some students moving for academic studies fail to report their move: a study from the Uppsala County in 2011 found that 15 % of full-time students in Uppsala (a city with a large student population) had a registered address other than Uppsala [5]. Furthermore, the PRs do not record more than one citizenship when an individual is Swedish.

Another minor limitation is incorrect data. A comparison of the PR-Tax and the TPR in 2011 found major discrepancies in about 500 individuals (this however amounts to <6/100,000 registered individuals).

The major limitation of the Population Registers is probably incorrect coverage. Undercoverage is usually of a temporary nature (for example immigrants and deaths not yet recorded), whereas over-coverage may be a more serious issue, even if equivalent only to 0.25–0.5 % of the registered Swedish total population. This figure includes individuals who have died abroad where death has not been reported to Swedish government agencies. Researchers using the family database (the multi-generation part [4]) of the TPR should also be aware of the left-truncation of data. In Sweden, the PIN was initiated in 1947. At that time, all living parents to individuals aged  $\leq 15$  years were recorded, but parents of individuals born in 1931 or earlier were not recorded. Parental data are also missing to a larger extent when parents have been born abroad [4]. While the TPR registers *de jure* parenthood, a small proportion of fathers are not the *de facto* fathers. In most large-scale epidemiological studies, this misclassification is likely to have little influence on research findings as non-paternity (while varying between countries) seem to occur in <5 % of births in most populations [42].

Data on PIN changes are important for medical research in the sense that without knowledge of PIN changes, the medical history of one individual might mistakenly be merged and assigned to two individuals. Throughout time, PINs have been re-used ( $n > 15,000$  individuals) [1]. This could potentially result in morbidity from two individuals

(or births in two women) being incorrectly assigned, as happening in the same individual had not Statistics Sweden kept a record on re-used PINs.

When delivering data to researchers, Statistics Sweden will assign *one* serial number to an individual with two PINs, and all of his or her healthcare history will be linked to that serial number. If two different individuals have been given the same PIN (re-use of the PIN), Statistics Sweden will nevertheless assign two serial numbers to the two individuals so that, e.g., death or “first offspring” will not occur twice in the same individual.

Hence, we conclude that both quality and timeliness of the Swedish population registers are high.

**Acknowledgments** Jonas F. Ludvigsson would like to dedicate this paper to Ingvar Johannesson, former civil servant at Statistics Sweden, neighbour and friend. Johannesson authored the report Total Population Register 2002 BEO1O2. Ludvigsson would also like to acknowledge the help of secretary Angelica Netterhall, Karin Wegfors (Statistics Sweden), Jesper Brodin (Statistics Sweden) and Gunnilla Autio (Swedish Tax Agency). We acknowledge financial support from the Swedish Research Council through the Swedish Initiative for Research on Microdata in the Social And Medical Sciences (SIMSAM) framework Grant No. 340-2013-5867.

## References

- Ludvigsson JF, Otterblad-Olausson P, Pettersson BU, Ekblom A. The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research. *Eur J Epidemiol*. 2009;24(11):659–67.
- Swedish\_National\_Tax\_Agency. Den svenska folkbokföringens historia under tre sekler [English: The Swedish national population registration - history of the last 300 years], 2015.
- Statistics\_Sweden (Swedish: Statistiska Centralbyrån S. Olika familjer lever på olika sätt - om barns boende och försörjning efter en separation. Demografiska rapporter 2014:01 [English: Different families have different ways of living—about children’s living and support after a separation], 2014.
- Ekblom A. The Swedish Multi-generation Register. *Methods Mol Biol*. 2011;675:215–20.
- Statistics\_Sweden (Swedish: Statistiska Centralbyrån S. Over-coverage in the Total Population Register—a register study. Background facts. Population and Welfare 2015;1.
- Eriksson JK, Neovius M, Ernestam S, Lindblad S, Simard JF, Askling J. Incidence of rheumatoid arthritis in Sweden: a nationwide population-based assessment of incidence, its determinants, and treatment penetration. *Arthritis Care Res (Hoboken)*. 2013;65(6):870–8.
- Busch K, Ludvigsson JF, Ekstrom-Smedby K, Ekblom A, Askling J, Neovius M. Nationwide prevalence of inflammatory bowel disease in Sweden: a population-based register study. *Aliment Pharmacol Ther*. 2014;39(1):57–68.
- Johansson B, Helgesson M, Lundberg I, Nordquist T, Leijon O, Lindberg P, Vingard E. Work and health among immigrants and native Swedes 1990–2008: a register-based study on hospitalization for common potentially work-related disorders, disability pension and mortality. *BMC Public Health*. 2012;12:845.
- Ji J, Ludvigsson JF, Sundquist K, Sundquist J, Hemminki K. Incidence of celiac disease among second-generation immigrants and adoptees from abroad in Sweden: evidence for ethnic differences in susceptibility. *Scand J Gastroenterol*. 2011;46(7–8):844–8.
- Lindblad F, Weitoft GR, Hjern A. ADHD in international adoptees: a national cohort study. *Eur Child Adolesc Psychiatry*. 2010;19(1):37–44.
- Wicks S, Hjern A, Dalman C. Social risk or genetic liability for psychosis? A study of children born in Sweden and reared by adoptive parents. *Am J Psychiatry*. 2010;167(10):1240–6.
- Bjorkenstam E, Hallqvist J, Dalman C, Ljung R. Risk of new psychiatric episodes in the year following divorce in midlife: cause or selection? A nationwide register-based study of 703,960 individuals. *Int J Soc Psychiatry*. 2013;59(8):801–4.
- Wallby T, Modin B, Hjern A. Child health care utilisation in families with young or single mothers in a Swedish county. *J Child Health Care*. 2013;17:17–29.
- Frisell T, Lichtenstein P, Langstrom N. Violent crime runs in families: a total population study of 12.5 million individuals. *Psychol Med*. 2011;41(1):97–105.
- Laszlo KD, Svensson T, Li J, Obel C, Vestergaard M, Olsen J, Cnattingius S. Maternal bereavement during pregnancy and the risk of stillbirth: a nationwide cohort study in Sweden. *Am J Epidemiol*. 2013;177(3):219–27.
- Sariaslan A, Larsson H, D’Onofrio B, Langstrom N, Fazel S, Lichtenstein P. Does population density and neighborhood deprivation predict schizophrenia? A nationwide Swedish family-based study of 2.4 million individuals. *Schizophr Bull*. 2015;41(2):494–502.
- Ludvigsson JF, Andersson E, Ekblom A, Feychting M, Kim JL, Reuterwall C, Heurgren M, Olausson PO. External review and validation of the Swedish national inpatient register. *BMC Public Health*. 2011;11(1):450.
- Ludvigsson JF, Montgomery SM, Ekblom A, Brandt L, Granath F. Small-intestinal histopathology and mortality risk in celiac disease. *JAMA*. 2009;302(11):1171–8.
- Herweijer E, Leval A, Ploner A, Eloranta S, Simard JF, Dillner J, Netterlid E, Sparen P, Arnheim-Dahlstrom L. Association of varying number of doses of quadrivalent human papillomavirus vaccine with incidence of condyloma. *JAMA*. 2014;311(6):597–603.
- Melander O, Maisel AS, Almgren P, Manjer J, Belting M, Hedblad B, Engstrom G, Kilger U, Nilsson P, Bergmann A, Orholm M. Plasma proneurotensin and incidence of diabetes, cardiovascular disease, breast cancer, and mortality. *JAMA*. 2012;308(14):1469–75.
- Melander O, Newton-Cheh C, Almgren P, Hedblad B, Berglund G, Engstrom G, Persson M, Smith JG, Magnusson M, Christensson A, Struck J, Morgenthaler NG, Bergmann A, Pencina MJ, Wang TJ. Novel and conventional biomarkers for prediction of incident cardiovascular events in the community. *JAMA*. 2009;302(1):49–57.
- Ludvigsson JF, Welander A, Lassila R, Ekblom A, Montgomery SM. Risk of thromboembolism in 14,000 individuals with coeliac disease. *Br J Haematol*. 2007;139(1):121–7.
- Ludvigsson JF, Olsson T, Ekblom A, Montgomery SM. A population-based study of coeliac disease, neurodegenerative and neuroinflammatory diseases. *Aliment Pharmacol Ther*. 2007;25(11):1317–27.
- Song H, Ekheden IG, Zheng Z, Ericsson J, Nyren O, Ye W. Incidence of gastric cancer among patients with gastric precancerous lesions: observational cohort study in a low risk Western population. *BMJ*. 2015;351:h3867.
- Holmqvist ME, Neovius M, Eriksson J, Mantel A, Wallberg-Jonsson S, Jacobsson LT, Askling J. Risk of venous thromboembolism in patients with rheumatoid arthritis and association with disease duration and hospitalization. *JAMA*. 2012;308(13):1350–6.

26. Bjorkenstam C, Weitoft GR, Hjern A, Nordstrom P, Hallqvist J, Ljung R. School grades, parental education and suicide—a national register-based cohort study. *J Epidemiol Community Health*. 2011;65(11):993–8.
27. Ljung R, Lagergren J, Bexelius TS, Mattsson F, Lindblad M. Increased risk of acute pancreatitis among tetracycline users in a Swedish population-based case-control study. *Gut*. 2012;61(6):873–6.
28. Andersson RE, Olaison G, Tysk C, Ekblom A. Appendectomy and protection against ulcerative colitis. *N Engl J Med*. 2001;344(11):808–14.
29. Lind M, Svensson AM, Kosiborod M, Gudbjornsdottir S, Pivodic A, Wedel H, Dahlqvist S, Clements M, Rosengren A. Glycemic control and excess mortality in type 1 diabetes. *N Engl J Med*. 2014;371(21):1972–82.
30. Lagergren J, Bergstrom R, Lindgren A, Nyren O. Symptomatic gastroesophageal reflux as a risk factor for esophageal adenocarcinoma. *N Engl J Med*. 1999;340(11):825–31.
31. Huang J, Magnusson M, Torner A, Ye W, Duberg AS. Risk of pancreatic cancer among individuals with hepatitis C or hepatitis B virus infection: a nationwide study in Sweden. *Br J Cancer*. 2013;109(11):2917–23.
32. Eriksson JK, Neovius M, Bratt J, Petersson IF, van Vollenhoven RF, Geborek P, Ernestam S. Biological vs. conventional combination treatment and work loss in early rheumatoid arthritis: a randomized trial. *JAMA Intern Med*. 2013;173(15):1407–14.
33. Ludvigsson JF, Haberg SE, Knudsen GP, Lafolie P, Zoega H, Sarkkola C, von Kraemer S, Weiderpass E, Norgaard M. Ethical aspects of registry-based research in the Nordic countries. *Clin Epidemiol*. 2015;7:491–508.
34. Council\_of\_the\_European\_Union. Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), 2015.
35. Swedish\_National\_Tax\_Agency. Kvarskrivning.
36. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, Pukkala E, Skytthe A, Hemminki K. Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med*. 2000;343(2):78–85.
37. Furu K, Kieler H, Haglund B, Engeland A, Selmer R, Stephansson O, Valdimarsdottir UA, Zoega H, Artama M, Gissler M, Malm H, Norgaard M. Selective serotonin reuptake inhibitors and venlafaxine in early pregnancy and risk of birth defects: population based cohort study and sibling design. *BMJ*. 2015;350:h1798.
38. Stephansson O, Kieler H, Haglund B, Artama M, Engeland A, Furu K, Gissler M, Norgaard M, Nielsen RB, Zoega H, Valdimarsdottir U. Selective serotonin reuptake inhibitors during pregnancy and risk of stillbirth and infant mortality. *JAMA*. 2013;309(1):48–54.
39. Sandin S, Schendel D, Magnusson P, Hultman C, Suren P, Susser E, Gronborg T, Gissler M, Gunnes N, Gross R, Henning M, Bresnahan M, Sourander A, Hornig M, Carter K, Francis R, Parner E, Leonard H, Rosanoff M, Stoltenberg C, Reichenberg A. Autism risk associated with parental age and with increasing difference in age between the parents. *Mol Psychiatry*. 2015. doi:10.1038/mp.2015.70.
40. Fletcher RH, Fletcher SW, Fletcher GS. *Clinical epidemiology. The Essentials*, 5th edn. Baltimore: Lippincott, Williams & Wilkins; 2012.
41. Statistics\_Sweden (Swedish: Statistiska Centralbyrån S. Different families live in different ways—a survey on residence and support of children after a separation. BE51BR1401., 2014.
42. Bellis MA, Hughes K, Hughes S, Ashton JR. Measuring paternal discrepancy and its public health consequences. *J Epidemiol Community Health*. 2005;59(9):749–54.