Administrative Data Research Centre Wales
Data Science Building
Swansea University
Singleton Park
Swansea
SA2 8PP

For general enquiries please call +44 (0) 1792 606918
or to speak to a member of the Research Support Team call +44 (0) 2920 870928

ADRCWales@Swansea.ac.uk www.adrn.ac.uk @ADRCWales

Administrative Data Research Centre Wales

An ESRC Data Investment

Better Knowledge Better Society

Data Brief 6: Welsh Demographic Service

Funded by

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

June 2016

Available in Welsh on request

The Welsh Demographic Service (WDS) is part of a set of services that manage administrative information and demographic data for NHS patients in Wales. This Data Brief summarises the data items that are available and particular issues surrounding the use of the WDS.

Mae Gwasanaeth Demograffeg Cymru (WDS) yn rhan o gyfres o wasanaethau sy'n rheoli gwybodaeth weinyddol a data demograffeg ar gyfer cleifion GIG Cymru. Mae'r Briff Data hwn yn crynhoi'r eitemau o ddata sydd ar gael a'r materion neilltuol sydd ynghlwm â'r defnydd o'r WDS.

Introduction

The Welsh Demographic Service (WDS) dataset contains demographic information for all NHS patients in Wales, including NHS number, name, address, date of birth, their GP, and practice registration history. The WDS was introduced in 2009 to replace the NHS Wales Administrative Register (NHSAR), and is predominantly used for the purpose of transferring medical records between practices and for managing payments to GPs. The register is updated daily, with information from local authorities to ensure that address changes and death notices are included.

The NHS in Wales uses the service to manage patients' medical records, to call women for cervical and breast screening, to allocate NHS numbers, to pay contractors such as pharmacists and GPs, to trace patient details, to ensure blood and organ donor details are maintained, for tracing patient details to facilitate payment for non-contract activity in secondary care and as part of child protection processes.

Whilst the WDS is mainly used for administrative purposes, it also provides an invaluable resource for research as it provides a full registration history for the population of Wales since 1990, including change of address and GP registration changes.

Structure and Content

The WDS dataset contains three main sets of information relating to (1) individual attributes, (2) address history, and (3) GP practice history.

Attendance Data

The individual attributes data contains key information about the people included on the WDS register. The table contains one record for each individual with the ALF used as the unique person identifier. Gender is also supplied along with date of death, if known.

Date of birth is deemed as a sensitive and potentially identifiable piece of information, therefore this information will not be available through ADRC-W. In order to overcome this issue and to enable a person's age to be calculated, the week of birth is supplied. This is determined using the Monday prior to the actual date of birth, meaning that the week of birth date will be accurate to within seven days.

Address History

The address history data contains a record for each address to which an ALF has been associated. Each time a person records a change of address or registers with a new GP practice, the WDS is updated and a new RALF is allocated to that person based on their new address. The data also includes 'from' and 'to' dates based on GP registrations, which indicates the duration of residency at each address. A row status field has been generated to indicate which address is currently valid for an individual, where 'A' represents an active address and 'D' represents a non-active address. Individuals who have no active address can be assumed to have moved out of Wales, which can be used to identify participants who have been 'lost to follow up' in cohort studies that utilise ADRC-W data sets.

A Lower Super Output Area (LSOA) code is provided based on the individual's postcode of residency, which can be linked to deprivation scores such as the Welsh Index of Multiple Deprivation (WIMD) and Townsend. LSOA codes can also be aggregated up to Middle (MSOA) or Upper Super Output Area (USOA) codes to enable analyses at a higher geographical level.

GP Practice History

The GP practice history data holds a record for every GP practice to which a person has been registered. In order to protect their anonymity and to prevent practice performance comparisons from being made, an encrypted code is assigned to each GP practice. This code is used to identify all practices that a person has been registered with. As with the address history data, a 'from' and 'to' date is supplied to identify which registrations are historical and which ones are current.

Using the WDS within ADRC-W

Data Generation Process

Prior to being brought into the ADRC-W secure setting for analysis, the WDS dataset is anonymised and encrypted by the NHS Wales Informatics Service (NWIS), who act as the Trusted Third Party (TTP) on behalf of ADRC-W.

Anonymous Linking Field

Probabilistic matching techniques are used to find the NHS number of each individual contained in the WDS register, which is later encrypted into an Anonymous Linking Field (ALF). As ALF's are based on NHS numbers, they are unique to each individual on the WDS register and are used across all ADRC-W data sets. This allows researchers to track all contact a patient has had with the healthcare system and other administrative services, from birth to death.

For data sets that do not include NHS numbers, a series of probabilistic matching techniques are used to assign an ALF to an individual based on other identifiable information, such as name, age, gender and postcode. Each ADRC-W data set relating to individual people contains two variables indicating the matching method used and the confidence that a true match has been obtained.

Residential Anonymous Linking Field

Addresses are recorded in the WDS based on GP registrations. The WDS register is updated each time a person registers with a new GP or notifies the surgery of a change of address. This provides a complete address history for each individual in the WDS.

Address details held in the WDS are matched to the Royal Mail Postcode Address File (PAF) in order to derive and assign a PAF address key to each address[1]. The PAF address key is then encrypted into a Residential Anonymous Linking Field (RALF), in the same way that an NHS number is encrypted into an ALF. Each individual RALF therefore relates to a specific postal address. If an address cannot be found in the PAF, then a RALF cannot be allocated.

Observations

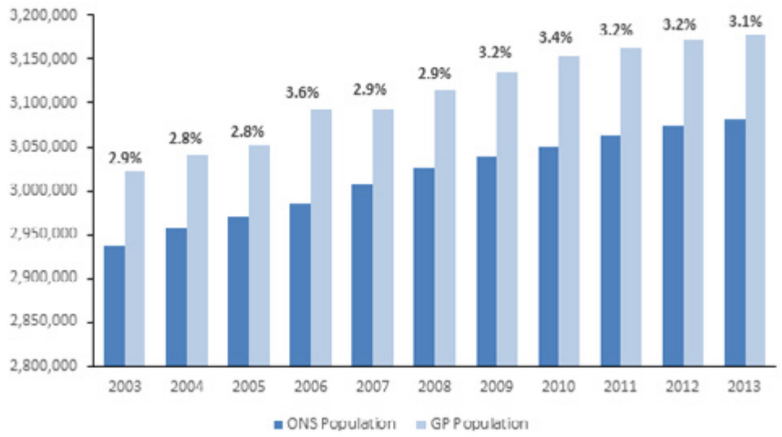
The Welsh Government produces an annual publication that provides workforce data relating to General Medical Practitioners[2]. The statistical release presents data on the number of practitioners, patients per practitioner and the demographic characteristics of practitioners. Information on the number of registered patients per practitioner and the number of practitioners can be used to provide an estimate of the GP population in Wales. These are presented in Figure 1 for the years 2003 to 2013 and are provided with a comparison against ONS Mid-Year Population Estimates.

It can be seen that within Wales, the GP based measure of population is consistently higher than ONS estimates. Furthermore, there is some indication to suggest that the size of this differential has widened between 2003 and 2013. Within Wales, the GP population for 2010 is estimated to be approximately 103,000 more than the Census based estimate. It is noted however that Mid-Year population estimates are based upon the latest available Census data (2011) combined with annual estimates of population change arising from births, deaths and migration and errors in these estimates of population change will be compounded over time.

Both GP workforce data and ONS population estimates are also provided for Health Boards in Wales. As indicated in Figure 2, comparisons of this data reveal that the scale of this differential varies across Wales. In both relative and absolute terms, the size of this differential is largest within Cardiff and Vale University Health Board, with the size of the GP population being roughly 21.5 thousand times larger than that provided by ONS estimates.

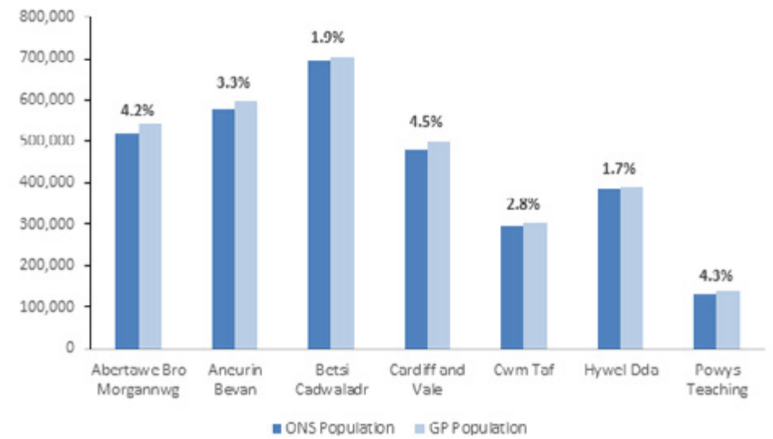
The discrepancy between population estimates based on Census figures and the registered lists of GPs is commonly referred to as list inflation. There are several reasons why some patients on GP lists no longer exist, including death, emigration, migrants returning home or individuals moving house. List inflation has been shown to vary by age, gender and area characteristics. It is noted to be particularly problematic among young males living in relatively deprived inner city areas[3]. To avoid miss-allocation of resources, administrative processes are in place for people moving between UK territories to ensure that their medical records are up to date.

Figure 1: GP and ONS Population Trends in Wales



Sources: General Medical Practitioners 2014, Welsh Government; Mid-Year Population Estimates, ONS

Figure 2: Regional Variations in GP and ONS Population Estimates in Wales, 2013



Sources: General Medical Practitioners, 2014 – Registered patients and number of GP practitioners by practice 2013, Welsh Government; Mid-Year Population Estimates, ONS 2013.

[1] See http://jpubhealth.oxfordjournals.org/content/31/4/582.full.pdf+html
[2] http://gov.wales/statistics-and-research/general-medical-practitioners/?lang=en
[3] http://fampra.oxfordjournals.org/content/22/5/529.full.pdf+html

Previous Research using the WDS

The WDS forms an important research resource and has been used in a variety of ways to obtain demographic and residential information about the population being studied. Research carried out using the Secure Anonymised Information Linkage (SAIL) Databank has utilised RALFs from the WDS to examine whether children who move house frequently have poorer health and educational outcomes[4], whilst other studies have utilised the WDS to assess the effects of the built environment on the health of those resident in the area[5 6].

The WDS has also been used as part of the Suicide Information Database-Cymru (SID-Cymru), which aimed to develop a protocol for a population-based, routinely collected data linkage study to explore the risks and patterns of a patient's healthcare contact prior to suicide in order to identify opportunities for intervention[7]. Other studies have used the WDS to obtain population denominators to examine health factors in specific areas of Wales [8 9 10].

As the WDS dataset enables a person's deprivation score to be determined based on their lower layer super output area (LSOA) of residence, all projects carried out through ADRC-W have the ability to add a deprivation measure.

Examples of previous research that have utilised deprivation scores to assess outcomes have included analysing socioeconomic patterns of harmful alcohol consumption[11] and the influence of social deprivation and air pollutants on asthma[12]. The WDS has also been utilised in a series of Welsh Government data linkage demonstration projects[13] to evaluate interventions targeted at some of the most deprived communities in Wales, such as the Flying Start programme and the Home Energy Efficiency Scheme.

The Administrative Data Research Centre Wales (ADRC-W) is one of four UK centres that along with the Administrative Data Service make up the Administrative Data Research Network (ADRN). The ADRN provides a safe, secure and transparent data linkage service for accredited, approved research using de-identified UK administrative data.

ADRC-W does not hold datasets. It works closely with government departments to make them available to researchers but this is negotiated on a case by case basis.

This report may be cited as: ADRC-W (2016) Welsh Demographic Survey. Administrative Data Research Centre Wales, Data Brief No. 6.

This report, along with other titles in this series is downloadable free from ADRN at https://adrn.ac.uk/about/research-centre-wales/resources/ or by emailing ADRCWales@Swansea.ac.uk.

[4] http://www.saildatabank.com/media/18995/final_r3.pdf
[5] http://www.sciencedirect.com/science/article/pii/S1353829211001687
[6] http://jpubhealth.oxfordjournals.org/content/31/4/582.full.pdf+html
[7] http://bmjopen.bmj.com/content/4/11/e006780.full?rss=1
[8] /media/19225/20110510_copdinabmu_rgtp_v1.docx
[9] /media/19271/20111124_copdinabmu_fluvacc_tplmrg_v1afinal_withoutadmissions.pdf
[10] /media/19317/20120725_copdinabmu_fluvaccpartiib_lmrgtp_v1afinal.pdf
[11] http://bmjopen.bmj.com/content/3/4/e002337
[12] http://erj.ersjournals.com/content/40/3/785
[13] http://gov.wales/statistics-and-research/data-linking-demonstration-projects/?lang=en