Scottish Atlas of Variation

Methodology
Contents

Introduction ........................................................................................................................................... 3
Standardisation ..................................................................................................................................... 3
Classification - significance of difference from Scotland ................................................................. 4
Classification - percentage rates ......................................................................................................... 5
Funnel plot .......................................................................................................................................... 5
Deprivation analysis ............................................................................................................................. 7
Glossary .............................................................................................................................................. 8
Contact ............................................................................................................................................. 10
Introduction

This document has been created to explain the statistical methodologies that have been used in the Scottish Atlas of Variation and to assist users with interpretation of the visualisations. There is also a User Guide available to help users navigate through the Scottish Atlas of Variation dashboard. The User Guide can be found, alongside the Metadata document and data files, at: [www.isdscotland.org/products-and-services/scottish-atlas-of-variation/supporting-documentation/](http://www.isdscotland.org/products-and-services/scottish-atlas-of-variation/supporting-documentation/)

Standardisation

Differences in population structure between regions can strongly affect the numbers of events (such as number of procedures or diagnoses). For example, an area with an older population will have higher rates of hip replacements. The aim of standardisation is to provide a summary ‘adjusted’ rate to take into account underlying differences in the structure (e.g. age and sex) of a study population relative to a ‘reference’ or standard population.

The crude rate is the number of events (e.g. hip replacement procedures) per head of population and is calculated for a certain time period (e.g. year). A standardised rate is calculated by adjusting the crude rate to take into account the structure of the population. There are two types of standardisation – direct and indirect. They both rely on reference to a single standard population. A directly standardised rate can be compared with the standardised rate for another population of interest (e.g. the comparison of two Local Authority areas).

The rates of procedures for specific age-sex groups in each region are calculated, and then applied to the European Standard Population 2013 and summed across all groups to give the sex and age-adjusted count for the area. The sex and age-adjusted number of events is divided by the total standard population for the whole age range included in the indicator, and multiplied by 100,000 to give the sex and age-standardised rate for the area.

For indicators that specify a certain age cohort, such as cataract procedures for people aged 65 years and over, a truncated standardisation is calculated by only taking into account only the 65+ population.
Classification - significance of difference from Scotland

On the dashboard tabs where standardisation has been used, the colour classification indicates how significantly different the standardised rates for the Local Authorities or Health Boards are from the Scotland value, and in which direction. Looking only at the absolute value of the indicator would sometimes be misleading; for example, in a Health Board with a small population, a large difference from the Scotland value could occur by random variation alone. By basing the classification on significance, we can avoid giving undue importance to such large differences.

Significance is evaluated based on whether the Scotland standardised rate falls within the confidence interval of the regional value, at two confidence levels: 95% and 99.8%. The confidence interval for directly standardised rates (DSR) is calculated using the ‘Dobson’ method recommended by APHO\(^1\). This uses the variance of the DSR, obtained using a Poisson assumption to the binomial variance of the individual age-specific rates. It also uses Poisson confidence limits for the total number of procedures, calculated using Byar’s approximation.

Where the population of a region is more than 10% of the Scottish population, a correction is applied to the confidence interval to adjust for the substantial non-independence between its rate and the Scotland rate\(^2\).

---


\(^2\) Lillian J. Hayes and Geoffrey Berry (2006), *Comparing the part with the whole: should overlap be ignored in public health measures?* Journal of Public Health 28(3) 278–282. [https://doi.org/10.1093/pubmed/fdl038](https://doi.org/10.1093/pubmed/fdl038)
Classification - percentage rates

On the tabs where a crude % value has been used, the colour classification indicates the range of % rates from lowest to highest for the selected procedure and year. The categories are shaded from dark purple (lowest % rate achieved) to light blue (highest % rate achieved).

The % rates are classified by ranking the rates for all locations (Local Authorities or Health Boards) in order from lowest to highest and the range of values is divided into 5 equal categories. For example, if the minimum rate is 0% and the maximum rate is 100%, each colour category will be 20%. The locations will be coloured according to the % banding they fall into. Similarly, if the minimum rate was 25% and the maximum rate was 75%, the range would be 50% and each colour category would represent 10%.

The size of the categories will dynamically update when a different procedure is selected as the minimum and maximum % rate will change. This methodology demonstrates the different ranges of variation, but it is important to note that the colour classification between maps should not be compared.

The actual % rates can be found in the tooltips when a mouse is hovered over an area on the map.

Funnel plot

Funnel plots are intended to distinguish random from non-random variation in distributions in the same way that statistical process control (SPC) charts do for time-series. A funnel plot is a scatter plot that displays a measure of something (e.g. hip replacement rates) associated with categorical items (e.g. Local Authorities) on the Y axis and sample sizes along the X axis. There is a data point for each item and lines to mark the boundaries or random variation. The boundary lines start out far apart from each other on the left where the number of events are small and converge to the right where the number of events is
large - this gives the boundary lines the shape of a funnel. This allows us to visualise variation in the indicator values as a function of measurement precision, since rates are “measured” more accurately in areas with larger populations.

Typically a funnel plot includes two sets of boundary lines; one set for 95% confidence intervals (calculated as standard error x 1.96) which are known as warning limits, and one set for 99.8% confidence intervals (calculated as standard error x 3.09), known as control limits. These are limits between which 95% or 99.8% of the observed rates are expected to fall, if the true underlying rate in all areas is equal to that of Scotland. This allows us to see how well the variation appears to be controlled overall by the factors accounted for in the standardisation (age and sex).

The limits are calculated on the assumption that the observed number of events is Poisson-distributed based on the expected number, with the Poisson confidence limits calculated using Byar’s approximation. The use of warning and control limits corresponds conceptually to the significance of difference from Scotland, but the results do not correspond exactly. The limits can, however, potentially highlight areas where substantial deviation from the national average may warrant further examination.

---

4 For a directly standardised rate, the limits should be regarded as approximate.
Deprivation analysis

The Scottish Index of Multiple Deprivation (SIMD) is the Scottish Government's official tool for identifying areas in Scotland concentrations of deprivation by incorporating several different aspects of deprivation (multiple-deprivations) and combining them into a single index.

Concentrations of deprivation are identified in SIMD at Data Zone level and can be analysed using this small geographical unit. The use of data for such small areas helps to identify 'pockets' (or concentrations) of deprivation that may be missed in analyses based on larger areas such as council wards or local authorities.

By identifying small areas where there are concentrations of multiple deprivation, the SIMD can be used to target policies and resources at the places with greatest need. **The SIMD identifies deprived areas, not deprived individuals.**

**Within-Scotland population-weighted quintiles** are used in this analysis and are calculated by ranking all Data Zones in Scotland from most to least deprived and then splitting this into 5 deprivation quintiles with approximately 20% of the Scotland population in each quintile.

Some parts of the country will not be represented across the deprivation spectrum. In **SIMD 2016**, Orkney and Shetland NHS Boards had no Data Zones in the most deprived quintile (quintile 1) when calculated on the within-Scotland basis. Also in SIMD 2016, Western Isles NHS Board fell within quintiles 2, 3 and 4 when calculated on the within-Scotland basis; there were no Data Zones in quintiles 1 and 5.

More information available at: [https://www2.gov.scot/Topics/Statistics/SIMD](https://www2.gov.scot/Topics/Statistics/SIMD)

Analysis for deprivation in the Atlas maps shows age-sex standardised rates by SIMD quintile.
## Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health Board of Residence</td>
<td>NHS Board area in Scotland in which the patient usually resides (at the time of admission to hospital).</td>
</tr>
<tr>
<td>Local Authority</td>
<td>The local government area in which the patient usually resides (at the time of admission to hospital).</td>
</tr>
<tr>
<td>Fold-variation</td>
<td>The ‘fold-variation’ value identifies the difference between the highest and lowest standardised rates for the diagnosis or procedure and geography level selected.</td>
</tr>
<tr>
<td>Crude rate</td>
<td>The crude rate in this analysis is defined as the total events (such as number of procedures) divided by the mid-year total population of the selected geography, and multiplied by 100,000. This provides a rate per 100,000 population.</td>
</tr>
<tr>
<td>Standardisation</td>
<td>A standardised rate is calculated by adjusting the crude rate to take into account underlying differences in the structure (such as age and sex) of a study population (the population of interest, e.g. Health Board or Local Authority) relative to a ‘reference’ or standard population (European standard population).</td>
</tr>
<tr>
<td>Standard error</td>
<td>The standard error is the variability of a set of means. It is used to construct confidence intervals around estimates of population means and in significance testing.</td>
</tr>
<tr>
<td>Confidence intervals</td>
<td>The confidence interval indicates a range of values which is likely to include the ‘true’ value for the population (for example, Health Board of Residence or Local Authority). Wider confidence intervals indicate that there is a lack of certainty due to a smaller sample size (or population). A narrow confidence interval means that the value is likely to be more accurate due to a larger sample size. The 95% confidence interval means that the range of values has a 95 in 100 chance of including the ‘true’ value.</td>
</tr>
<tr>
<td>Poisson</td>
<td>A discrete frequency distribution which gives the probability of a number of independent events occurring in a fixed time with equal mean and variance.</td>
</tr>
<tr>
<td>Byar’s approximation</td>
<td>A computationally simple approximation that gives very accurate approximations to the exact Poisson probabilities even for small counts.</td>
</tr>
</tbody>
</table>
Within-Scotland population-weighted quintiles are used in this analysis and are calculated by ranking all Data Zones in Scotland from most to least deprived and then splitting this into 5 deprivation quintiles with approximately 20% of the Scotland population in each quintile.

[https://www2.gov.scot/Topics/Statistics/SIMD](https://www2.gov.scot/Topics/Statistics/SIMD)
Contact

Please contact us if you have any questions, comments or suggestions regarding the Scottish Atlas of Variation by sending an email to:

nss.scottishatlasofvariation@nhs.net

Alternatively, please get in touch with the following contacts:

Emma McNair, Information Consultant
Phone: 0131 275 6668
Email: emma.mcnaire@nhs.net

Graham McGowan, Principal Information Analyst
Phone: 0141 282 2233
Email: graham.mcgowan@nhs.net

Catherine Perkins, Senior Information Analyst
Phone: 0131 275 7710
Email: catherine.perkins1@nhs.net

Siân Guy, Senior Information Analyst
Phone: 0141 282 2313
Email: sian.guy@nhs.net