EXAMINING METHODS TO VISUALISE THE CANCER POPULATION USING CARTOGRAMS

WE ARE MACMILLAN. CANCER SUPPORT | Public Health England

20-year cancer prevalence in the UK for all cancers combined

S Jones[1], H McConnell[1], C White[2], D Huws[2], L Irvine[3], S Miller[4]
[1] Macmillan Cancer Support; [2] Public Health Wales, Welsh Cancer Intelligence and Surveillance Unit; [3] Public Health England, National Cancer Intelligence Network; [4] Public Health England, Knowledge and Intelligence Team (East)

Background

An estimated 2.5 million people are living with and beyond cancer in the UK, predicted to increase to 4 million by 2030[1]. The Macmillan-National Cancer Intelligence Network Cancer UK Prevalence Project aims to provide the most granular understanding of the cancer survivorship population in the UK, with outputs produced by sub-national geographies.

A key challenge is how to visualise and represent geographic variations in cancer prevalence appropriately in ways that are easily comprehensible.

Methods

We used the Public Health England National Cancer Data Repository to link UK cancer registrations to mortality records in order to identify people diagnosed with cancer between 1991 and 2010[2] and who were still alive on 31st December 2010. Counts are based on the first diagnosis of a specific cancer within the 20-year period.

Using geographical information systems (GIS) mapping software[3], the total 20-year cancer prevalence data (number and crude rate[4]) for all cancers combined (excluding non-melanoma skin cancer) was joined to the spatial layer of each local area in the UK[5], based on where people were resident at the time of diagnosis.

We visualised the data in different formats including: a standard choropleth (thematic) map; a non-contiguous cartogram[6]; a Demers-style cartogram[7] based on equal squares, and a population-equalised cartogram[8]. To allow for comparison, the same data classification (four classes based on the Jenks' Natural Breaks algorithm[9]) and colour theme has been applied across all maps, with darker green colours representing higher rates.

Results

There were 1.8 million people in the UK diagnosed with cancer between 1991 and 2010[2] and still alive in 2010, who were matched to a local area in the UK.

Legend for all maps
Crude rate per 100,000

- 1,250 – 2,490
- 2,500 – 3,000
- 3,010 – 3,500
- 3,510 – 5,040

All cancers excluding non-melanoma skin cancer. Data classified by natural breaks. Median = 2,998. Numbers have been rounded.

(i) Standard choropleth map (figure 1)

Based on the choropleth map, it could be perceived that cancer prevalence rates are high in the South West, South Coast and in Norfolk in England, as well as mid-Wales and parts of Scotland. Some of these large rural areas appear most prominent. Local variations, however, are not always easily distinguished in large conurbations made up of small local authorities, such as Greater Manchester and the West Midlands.

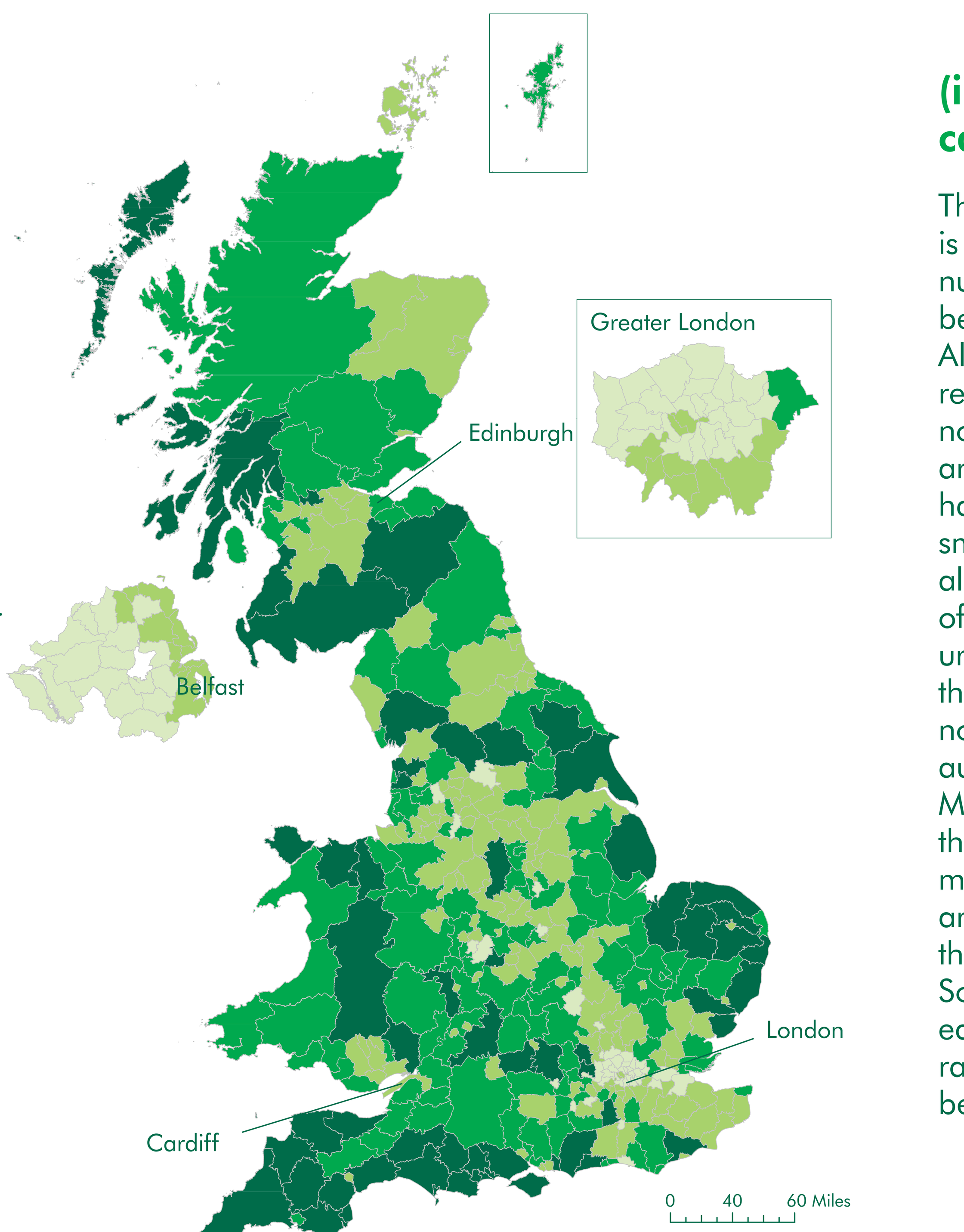

Figure 1: Choropleth map representing people living with and beyond cancer in the UK in 2010 by local authority

(ii) Non-contiguous cartogram (figure 2)

The size of each local authority is scaled proportionally to the number of people living with and beyond cancer in each area. Although the shapes of local areas remain the same, the topology is not maintained. Some large rural areas, particularly the Scottish Isles, have shrunk in accordance to the small population. This method also increases the size of some of the more densely populated urban areas – in some cases this is to such an extent that they now overlap (e.g. London local authorities and those in the West Midlands conurbation). In spite of this, the spatial data is presented more fairly in that the large rural areas no longer visually dominate the maps. The South West and South Coast of England remain easily identified as having high rates of people living with and beyond cancer in 2010.

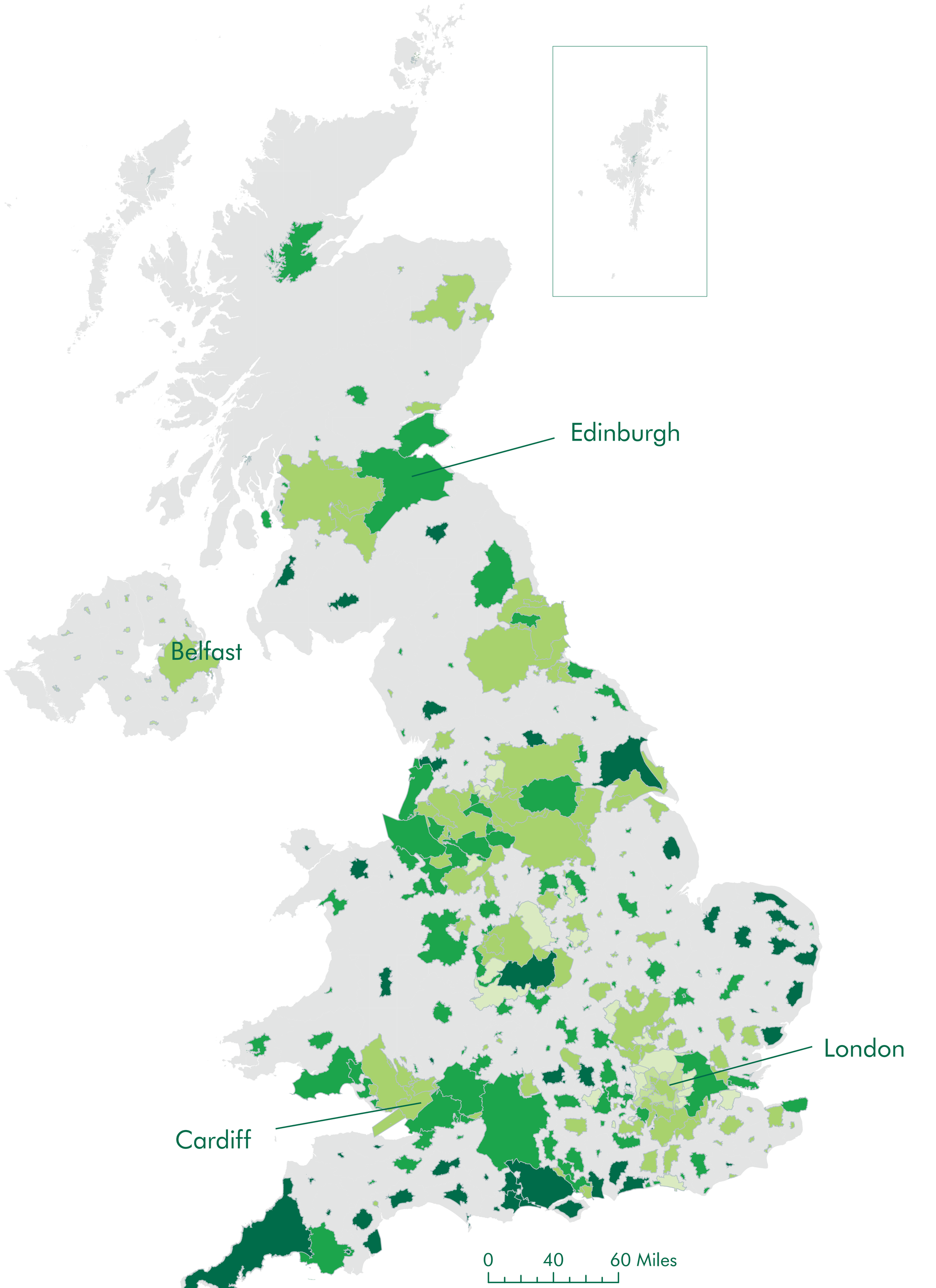

Figure 2: Non-contiguous cartogram representing people living with and beyond cancer in the UK in 2010 by local authority

(iii) Equal square cartogram (figure 3)

All local areas have been replaced by a uniform square grid. Although the shapes of local areas are lost, there are no over-lapping features. True topology has not been maintained, but there is still in place some schematic topological relationship.

This presents the data from a non-biased starting point enabling geographical variations to be easily and quickly identified. There is a high density of local authorities in the South West of England clearly visible with high rates of people living with and beyond cancer in 2010.

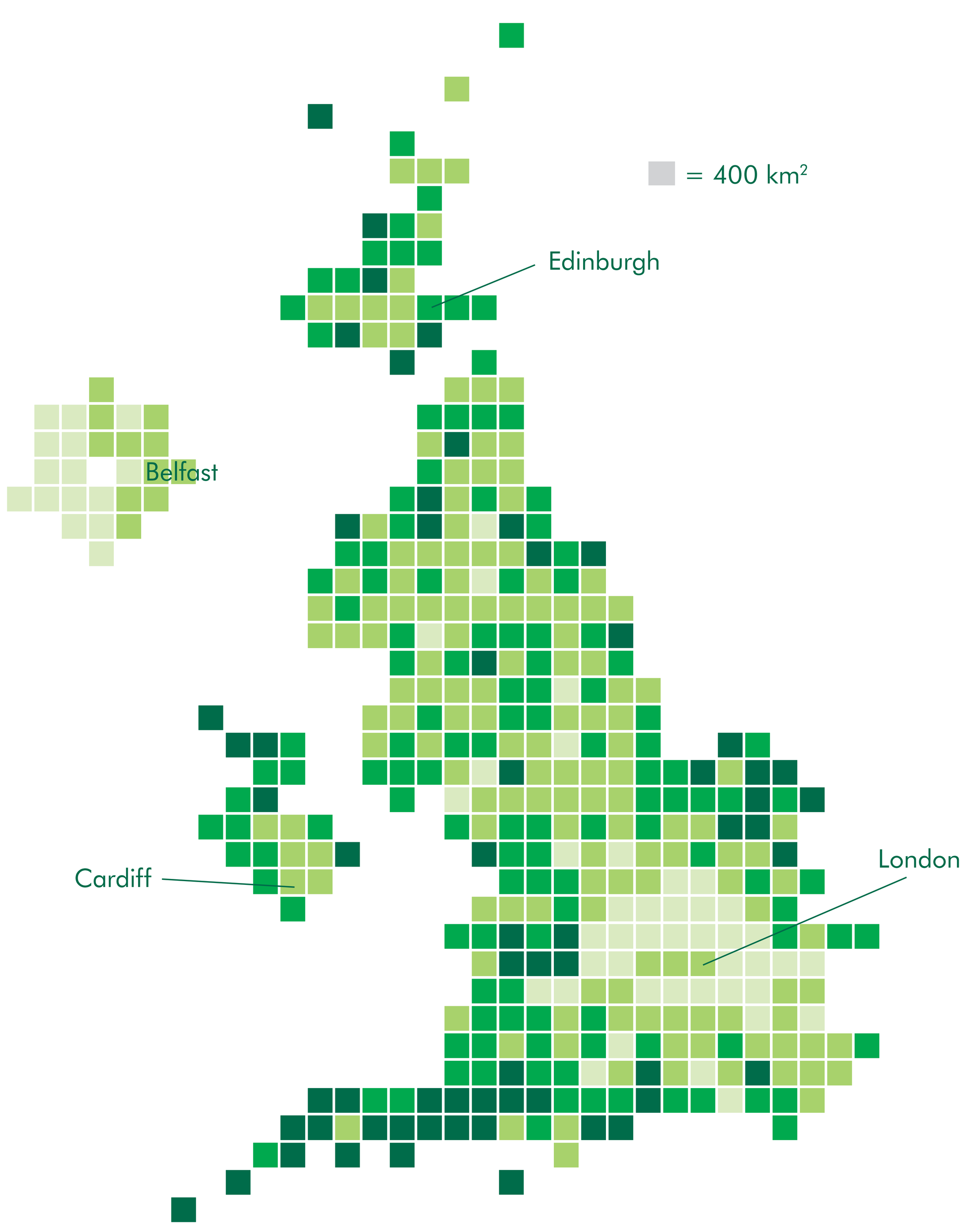

Figure 3: Equal-square cartogram representing people living with and beyond cancer in the UK in 2010 by local authority

(iv) Population-equalised cartogram (figure 4)

Local authorities are re-sized according to a population attribute – in this case the number of people living with and beyond cancer. Despite the distortion, topological relationships and borders have been maintained. The addition of a thematic classification has resulted in a bivariate map, in which both absolute numbers and rates are now visualised.

Based on the population-equalised cartogram, it can be observed that whilst there are high numbers of people living with and beyond cancer in areas such as London and Birmingham (signified by their respective increased size), the prevalence rate in these areas is relatively low (as indicated by the lighter shade of green). Conversely, some of the large rural areas in Wales and Scotland are still identified as having a relatively high rate, but in absolute terms the number is low, as represented by their shrunken size. The South West and Norfolk areas in England remain prominent reflecting both a relatively high rate and a high number of people living with and beyond cancer.

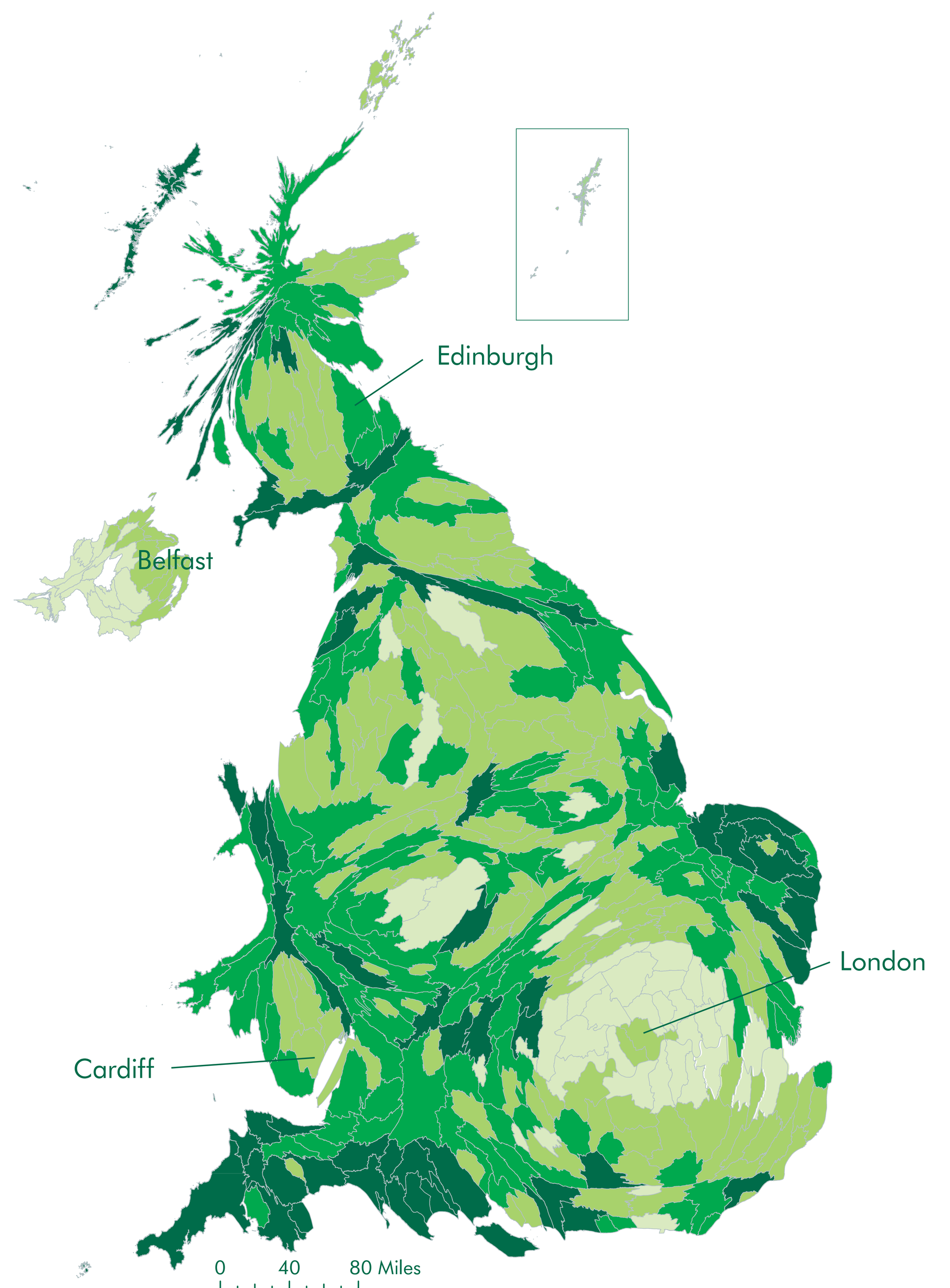

Figure 4: Population-equalised density cartogram representing people living with and beyond cancer in the UK in 2010 by local authority

Conclusions

Whilst choropleth maps represent data in ways that are familiar to an audience, and show the true geography and topology, there can sometimes be a risk of data misinterpretation and visual bias.

Cartograms can help to remove visual bias and present data in a fairer and more comprehensible way by allowing for quick visual comparisons, although true topology is sacrificed.

Population-equalised cartograms (figure 4), and resulting bivariate maps, can not only highlight the relative spatial variations in cancer prevalence but also show where the greatest demand for health services may lie by emphasizing the absolute number (by increasing the size of an area). Presenting data in an alternative way such as this, provided it is explained appropriately for the audience, can allow users to to quickly understand and plan for better service delivery.

This work could be extended to include prevalence data based on age-standardised rates which would make areas even more comparable. In the maps presented here, the South West region in England, for example, has high numbers and rates of people living with and beyond cancer but this is likely to be largely due to the high proportion of the population being aged 65 and over.

Acknowledgements
This work is part of the Macmillan Cancer Support and Public Health England's National Cancer Intelligence Network Partnership Work Plan. Data are sourced and presented in collaboration with the Welsh Cancer Intelligence and Surveillance Unit, Health Intelligence Division, Public Health Wales, the Scottish Cancer Registry and the Northern Ireland Cancer Registry which is funded by the Public Health Agency for Northern Ireland.

For more information please contact Samuel Jones, evidence@macmillan.org.uk

June 2015

N. Ireland Cancer Registry | NHS National Services Scotland | GIG Cymru NHS Wales Iechyd Cyhoeddus Cymru Public Health Wales | Uned Deallusrwydd a Gwyliadwriaeth Canser Cymru Welsh Cancer Intelligence and Surveillance Unit

References and notes:

1. Maddams J, Utley M, Møller H. Projections of cancer prevalence in the United Kingdom, 2010-2040. Br J Cancer 2012; 107: 1195-1202. (Projections scenario 1). Macmillan analysis based on extrapolation of 2010 and 2020 projections that the number of people living with cancer will hit an estimated 2.5 million in 2015
2. Northern Ireland data is based on cancer registrations collected for people diagnosed within cancer between 1993 and 2010
3. Spatial analysis and cartographic output was produced using a build of ESRI ArcMap 10.0 called MicromarketerG3 provided by Experian. Additional output was configured using Indiemapper, a free service from Axis Maps and available at http://indiemapper.com/.
4. Rates referred to in this poster are crude rates per 100,000 population
5. Local Authority for England and Wales; Local Council Area for Scotland and Local Government District for Northern Ireland. Boundaries correct for 2010 in line with the prevalence dataset.
6. Olson, Judy M. 1976. Non contiguous Area Cartograms. Professional Geographer 28:371–80
7. National Center for Geographic Information and Analysis. Cartogram Types, available at http://www.ncgia.ucsb.edu/projects/Cartogram_Central/types.html [accessed April 2015].
8. Gastner M, Newman M. Diffusion-based method for producing density-equalizing maps. Proceedings of the National Academy of Sciences of the USA 2004; 101: 7499-7504
9. de Smith, Goodchild, Longley. Univariate classification schemes in Geospatial Analysis—A Comprehensive Guide, 3rd edition; 2006–2009.