Spatial and Spatio-temporal Epidemiology
Manuscript Draft

Manuscript Number: SSTE-D-13-00043

Title: Using Geographical Information Systems and Cartograms as a Health Service Quality Improvement Tool

Article Type: SI:GEOMED 2013

Keywords: Cartography; Visualisation; Health Services

Abstract: Background
In England, the National Health Service collects a dataset known as Hospital Episode Statistics, which holds patient information regarding their admission. With this, specific diseases can be spatially analysed to provide support for service implementation. A key challenge is to communicate these results to individuals, with variable levels of Geographical Information Systems (GIS) knowledge, in an effective manner to show trends in the data. This research describes the combination of a number of GIS methods and software tools to produce a novel technique of visualising disease admissions. The aim is to provide a tool that can support the ability of decision makers and service teams within health settings to develop services that are more efficient and better catered to the service population.
Materials and Methods
Initially, a standard choropleth of the study region (London) is used to visualise total admission values using the ESRI ArcGIS software. The mid-2010 population estimates of the London lower layer super output areas (LSOAs) are then used within the ScapeToad cartogram software tool with the aim of visualising geography at uniform population density. An interpolation surface, in this case ArcGIS' spline tool, allows the creation of a smooth surface over the LSOA centroids for admission values on both standard and cartogram geographies. The final product of this research is the Cartogram Interpolation Surface (Cart-IS).
Results
The method provides a series of outputs culminating in the Cart-IS, demonstrating the advantages and disadvantages of varying the geographic boundaries, population density and visualisation type. The cartogram effectively equalises the population density to remove visual bias from areas with a smaller population. Within the regular geography interpolation surface there are a number of artefacts (disproportionate high values), along with numerous other mid-value areas, which are side effects of the algorithm. These are removed or reduced within the Cart-IS the most extreme example being a decrease from approximately 29 to 3 admissions. The Cart-IS enables identification of key areas upon a uniform population density and a graded visual output of total admission values.
Conclusion
Overall, this methodology provides an insight into the combining of simple GIS tools to create a novel output, Cart-IS, in a health service context with the key aim of improving visualisation communication techniques which highlight the variation in small scale geographies across large regions to individuals with little technical understanding of the methods used.

***Highlights (for review)**

Highlights

- Small scale geographical variation is of great importance for service provision.
- Cartograms are useful when used effectively and put into context.
- Combined tools can achieve improved support in healthcare service improvement.

Corresponding Author:

Full Name: Mr Derryn A Lovett

Email: d.lovett@imperial.ac.uk

Telephone: +44(0)20 331 58130

Abstract

Background

In England, the National Health Service collects a dataset known as Hospital Episode Statistics, which holds patient information regarding their admission. With this, specific diseases can be spatially analysed to provide support for service implementation. A key challenge is to communicate these results to individuals, with variable levels of Geographical Information Systems (GIS) knowledge, in an effective manner to show trends in the data. This research describes the combination of a number of GIS methods and software tools to produce a novel technique of visualising disease admissions. The aim is to provide a tool that can support the ability of decision makers and service teams within health settings to develop services that are more efficient and better catered to the service population.

Materials and Methods

Initially, a standard choropleth of the study region (London) is used to visualise total admission values using the ESRI ArcGIS software. The mid-2010 population estimates of the London lower layer super output areas (LSOAs) are then used within the ScapeToad cartogram software tool with the aim of visualising geography at uniform population density. An interpolation surface, in this case ArcGIS' spline tool, allows the creation of a smooth surface over the LSOA centroids for admission values on both standard and cartogram geographies. The final product of this research is the Cartogram Interpolation Surface (Cart-IS).

Results

The method provides a series of outputs culminating in the Cart-IS, demonstrating the advantages and disadvantages of varying the geographic boundaries, population density and visualisation type. The cartogram effectively equalises the population density to remove visual bias from areas with a smaller population. Within the regular geography interpolation surface there are a number of artefacts (disproportionate high values), along with numerous other mid-value areas, which are side effects of the algorithm. These are removed or reduced within the Cart-IS the most extreme example being a decrease from approximately 29 to 3 admissions. The Cart-IS enables identification of key areas upon a uniform population density and a graded visual output of total admission values.

Conclusion

Overall, this methodology provides an insight into the combining of simple GIS tools to create a novel output, Cart-IS, in a health service context with the key aim of improving visualisation communication techniques which highlight the variation in small scale geographies across large regions to individuals with little technical understanding of the methods used.

Keywords: Cartography; Visualisation; Health Services

Introduction

The use of small-area data to describe variations in healthcare was developed in the 1970s by John Wennberg (Wennberg & Gittelsohn 1973); over recent years the increased use and capabilities of technology, combined with improved access to data , has allowed the development of atlases that visualise geocoded health data . The production of atlases in the United Kingdom (UK) (NHS Right Care 2010), building on similar American atlases produced by the Dartmouth Institute(Dartmouth Medical School 1998), have raised questions about why there are variations in clinical outcomes at regional and local levels. Subsequently significant strides have been made to reduce unwarranted variation in clinical outcome through improving the quality and consistency of care processes e.g. Chronic Obstructive Pulmonary Disease (COPD) Care Bundle (Hopkinson et al. 2012). Whilst atlases have been instrumental in highlighting the disparities in health outcomes nationally there still remains some concern about how these data are translated into strategies to improve outcomes (Joyce 2009). With the potential for more and more data to be available for such analysis, including patient reported outcome measures (PROMs) and patient reported experience measures (PREMs), the question arises at how technologies can be used to "intelligently" interpret the data to improve the quality of services that will translate into improved outcomes and patient experience. A key challenge is how to present service activity data, such as Emergency Department (ED) admissions, in such a way as to be helpful to commissioners, clinicians, healthcare managers and policy makers to support initiatives to improve the quality of services (e.g.(Welch & Allen 2006; Noble et al. 2012; Green et al. 2012; Future Hospital Commission 2013) .

Geographical Information Systems (GIS) have a long history of use within health care, particularly public health (Wennberg & Gittelsohn 1973; Twigg 1990; Bullen et al. 1996; Higgs & Gould 2001; Nacul et al. 2011), to visualise the epidemiology of disease, delivery of health services and allocation of resources. Whilst GIS can often be perceived as simple visualisation of data through mapping, this belies the complexity of data and the interaction with the geographical or spatial plane used for display. Communicating this complexity effectively and accurately provides opportunities for the "viewer" to interpret the data and identify patterns. When visualising datasets, the resolution of display is critical to the information portrayed, with larger geographic aggregations masking heterogeneity at finer scales, and potentially causing incorrect interpretation of the data. However, operating at finer resolutions can result in a loss of data in areas of relatively small geographies and, when using choropleths, importance can mistakenly be assigned to areas of larger extent, rather than areas of higher value (Pickle & Carr 2010). In this paper, we explore the use of a combination of methods as a possible solution to this visualisation problem, using health service utilisation data derived from Hospital Episode Statistics (HES) data as an example.

Study Setting

The National Health Service (NHS) in the UK collates a large amount of data generated from each interaction a patient makes with the health service. Data are reported to the Health and Social Care Information Centre, which processes 125 million individual episodes of care each year, in the form of HES (Health & Social Care Information Centre 2013). The most common activity data collected are outpatient department attendances, elective admissions to hospital and ED attendances and subsequent admissions. Although anonymised, to protect the privacy of patients, the data include demographic and clinical information that can be used by researchers to analyse temporal and geographical trends, as well as more complex analyses. Numerous published studies have demonstrated the utility of HES data, including the identification of population and healthcare factors associated with increased COPD admissions (Calderón-Larrañaga et al. 2011), increases in emergency admissions related to sickle cell disease (Aljuburi et al. 2012) and identifying links between the smoking ban and a reduction in asthma-related emergency admissions (Millett et al. 2013). The continuing development of new technologies and novel platforms for data visualisation have led to the development of explicit strategies in the UK to harness data for research through greater access to patient data across research and industry (Department for Business Innovation and Skills 2012). Furthermore, UK policies have been developed that explicitly focus on improving the delivery and effectiveness of clinical services through the use of routine data and metrics (Department of Health 2008).

Material and Methods

Software

ESRI's ArcGIS 10.1 for Desktop Advanced (ESRI 2012) and the open source software ScapeToad (Choros Laboratory 2008) are used to combine simple GIS methods and tools, manipulating the data to provide a novel methodological output.

Data

Chronic Obstructive Pulmonary Disease and bronchiectasis (hereafter COPD) is a lung disease that over time narrows the airways, leading to shortness of breath. The ICD-10 (international classification of diseases, tenth version) codes J40-J44, J47 and J96.X were used to specify emergency admissions of COPD (Health & Social Care Information Centre 2013) in all England during the 2010/11 financial year.

Geography and Demography

Several spatial resolutions are used to visualise this data: Strategic Health Authorities (SHAs); Primary Care Trusts (PCTs) and Lower Layer Super Output Areas (LSOAs) to allow comparison of detail and change as the focus draws on finer scales. In this paper, to provide sufficient but not over exhaustive analysis, the use of LSOAs and PCTs has been restricted to London SHA. Whilst PCTs have now been replaced by Clinical Commissioning Groups (CCGs), any conclusions drawn can still be considered relevant due to the focus on a finer scale of geography and the comparison between the two. In 2010 there were 10 SHAs, 151 PCTs and 32,482 LSOAs in England, of which 1 SHA, 32 PCTs and 4765 LSOAs were in London. Population estimates at LSOA level for mid-2010, also created by the Office for National Statistics (ONS), are based on 2001 census population data (Office for National Statistics 2010).

Analytical Approach

Data obtained from HES was provided at LSOA level, retaining anonymity, allowing it to be directly linked to the ONS LSOA boundary dataset (Office for National Statistics 2001). These boundaries then aggregated to PCT and SHA geographies (UK Data Service Census Support 2011) using LSOA centroids, a necessity as the 2004 PCT and SHA boundaries are not coterminous with each other or 2001 LSOAs.

Cartograms created using ScapeToad based on the mid-2010 population estimates of LSOAs in London SHA, were used to visualise these data accounting for differences in population density, using the same classification as the normal geography to allow direct comparison. This software tool uses the Gastner/Newman method (Gastner & Newman 2004) to perform an iterative diffusion algorithm using the chosen value attribute upon a square grid placed over the original geography. Once completed, interpolation surfaces were created, one each for the normal and cartogram geographies, using the LSOA centroids and both included all LSOAs within 10km of London SHA, to prevent possible edge effects

(Gatrell et al. 1996). The outputs of which are clipped back to the London SHA extent before being classified with a colour gradient to clearly depict the variation within the dataset. The culmination is a novel display methodology, the Cartogram Interpolation Surface (Cart-IS).

Preference Testing

The product of these methods was then used as the basis for a series of informal interviews conducted at the GEOMED 2013 conference to uncover the opinion of professionals already working within this field of expertise.

Results

Figure 1. Choropleths of total COPD emergency admissions from 2010/11 at (a) SHA level for England, (b) PCT level in London and (c) LSOA level in London.

Figures 1A to C display the decreasing scales at which data have been aggregated, moving through SHA and PCT to LSOA.

Figure 2. Choropleth cartogram of total COPD emergency admissions from 2010/11 at LSOA level in London.

Critically, however, LSOA geographies fluctuate in size relative to population density. The effect of this on visual output puts aesthetic importance on areas of lower population density where geographies are larger, though not necessarily of more significance. Figure 2 presents a possible solution to this issue using the deformation properties of a cartogram, which deforms the geographies based on the mid-2010 population estimates. Clearly visible is this apparent swelling of areas in central London, where population sizes are larger i.e. more dense, while the area of LSOAs on the outskirts decreases, leading to an overall equalising of LSOA area size variation. While this does, to best effect, solve the issue of visual priority, the quintile classification still limits the user’s visual understanding of data variation.

Figure 3. Interpolation Surfaces of total COPD emergency admissions from 2010/11 at LSOA level in London for (a) normal geography and (b) cartogram geography (Cart-IS).

An alternative method to visualise the variation is in the use of an interpolation surface, which estimates surface values at unsampled points based on known surface values of surrounding points. In this example the ArcGIS spline tool is utilised where the classification is based on the minimum and maximum of the dataset. Initially produced upon normal geography, figure 3(a) presents a reasonable portrayal of the data.

The combination of both the cartogram and spline produces a population-based geography with a stretched classification surface, figure 3(b), Cart-IS. As with the spline for normal geography a guard area has been included to avoid edge effects, though was not utilised in the creation of the cartogram. However in comparison to the normal geography spline, the Cart-IS has removed some of the higher values, previously seeming of higher priority though were, as mentioned, an artefact of the tool. Continuing to be represented as in figure 1(c) is the very low, almost uniform values, that occur in PCTs to the South East of the city centre whose significance is more noticeable in comparison to neighbouring PCTs that contain a relatively sporadic distribution.

During presentation of these techniques to the 66 delegates of the GEOMED 2013 conference, a series of informal interviews were undertaken to ask delegates which of the four London area LSOA maps (Figures 1c, 2, 3a and b) presented here they believed most effectively communicated the data (Ritchie & Lewis 2003) (see Appendix 1). What was most prominent from discussions had with many of the delegates was the benefit in keeping normal geography and simple methods for communication purposes. However, half of the 14 who answered, when asked to choose one map, agreed that the Cart-IS method was the best option, but only if explanation and story of how and why the methods had been used was included.

Discussion

Data at SHA level shows some geographical heterogeneity but whilst it is useful for broad brush benchmarking, it is ineffectual in identifying specific areas that could be targeted for improvement and masks regions of high value surrounded by low values, for instance North Devon. The standard unit for large area analysis has been the PCT level, utilised for regional benchmarking of clinical outcomes, as demonstrated by the various atlases. The analysis at the LSOA level provides greater detail of the geographical heterogeneity within each PCT, as shown for London in Figure 1C and highlights that trends at this level can be more clearly represented at smaller geographies with the data that is already available.

The use of a cartogram transformation to reshape the geographic regions so that the chosen attribute, population density, has uniform density depends on the density of the framework upon which this deformation occurs. A dense framework re-distribution can leave the visualisation unrecognisable, while a less dense framework will not ensure that the variable is uniform (Gastner & Newman 2004).

Following on from the LSOA, PCT and SHA regions, the use of an interpolation surface allows the visualisation of the variability in the dataset. However, the nature of the tool is to use the nearest n number of points to create a smooth surface across the region of study using, in this case, the LSOA centroids. What this means is that areas of lower population density have larger space over which to smooth values, a side effect of which is anomalous values that are artefacts of the tool and do not represent the underlying data. Additionally groupings or stand-alone high values in areas of high population density are underrepresented. Edge effects in the context of this spatial analysis tool are overcome through the inclusion of values that lie outside of the study region (Gatrell et al. 1996). By including these additional values in a spatial analysis task such as a spline, you often remove some, if not all, of the anomalous values that are a side effect of the tool.

The preference testing that was undertaken in support of this research in the most part reinforces the methods and conclusions presented. However, due to the expert focus of the sample group there is space for expansion and development of this feedback that would provide greater value.

Future

Developing this work further it is envisaged that similar analyses would be undertaken at both regional and national scales for the purpose of influencing strategic planning and targeting of population level interventions. Combining these results with same method analyses on different diagnosis codes could provide an overview of the state of similar or related symptoms. Other than varying region and topic, though these could be used for this purpose, further analysis of the patterns seen using datasets such as the UK indices of deprivation would be advantageous for intended users to understand the trends demonstrated.

While the debate around the use of cartograms and their value as an evaluation tool remains, their continued use to distort visualisations based on population or other values, e.g. indices of deprivation, would allow continued discussion and teaching to inform the larger audience of both their benefits and problems.

Conclusion

The combined methods demonstrate an effective tool for the visualisation of variation in small scale geographies across larger areas to individuals with little technical understanding of the methodology used. The difference between normal and cartogram geographies warrants the additional explanation and teaching, be that through lectures, tutorials or user guides, which would allow individuals to assess and effectively implement changes based on the data. The Cart-IS method, which provides a mapping product that forms a visual narrative for decision makers is thus presented for consideration and suitability testing.

Acknowledgements:

Disclaimer:

This article presents independent research commissioned by the National Institute for Health Research (NIHR) under the Collaborations for Leadership in Applied Health Research and Care (CLAHRC) programme for North West London. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

References

Aljuburi, G. et al., 2012. Trends in hospital admissions for sickle cell disease in England, 2001/02-2009/10. *Journal of public health (Oxford, England)*, 34(4), pp.570–6.

Bullen, N., Moon, G. & Jones, K., 1996. Defining localities for health planning: a GIS approach. *Social Science & Medicine*, 42(6), pp.801–16. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8778994.

Calderón-Larrañaga, A. et al., 2011. Association of population and primary healthcare factors with hospital admission rates for chronic obstructive pulmonary disease in England: national cross-sectional study. *Thorax*, 66(3), pp.191–6.

Choros Laboratory, 2008. ScapeToad. Available at: http://scapetoad.choros.ch/.

Dartmouth Medical School, 1998. *Dartmouth Atlas of Health Care*, Dartmouth Medical School.

Department for Business Innovation and Skills, 2012. *Strategy for UK Life Sciences*, London.

Department of Health, 2008. *High Quality Care for All: NHS Next Stage Review Final Report*, London.

ESRI, 2012. ArcGIS 10.1 for Desktop Advanced. Available at: http://www.esri.com/software/arcgis/arcgis-for-desktop.

Future Hospital Commission, 2013. *Future Hospital Commission: Caring for medical patients*, Available at: http://www.rcplondon.ac.uk/sites/default/files/future-hospital-commission-report.pdf.

Gastner, M.T. & Newman, M.E.J., 2004. Diffusion-based method for producing density-equalizing maps Previous Methods for Constructing Cartograms. *PNAS*, 101(20), pp.7499–7504.

Gatrell, A.C. et al., 1996. Spatial Point Pattern Analysis and Its Application in Geographical Epidemiology. *Transactions of the Institute of British Geographers*, 21(1), p.256. Available at: http://www.jstor.org/stable/622936?origin=crossref.

Green, S. a et al., 2012. Mapping mental health service access: achieving equity through quality improvement. *Journal of public health (Oxford, England)*, pp.1–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22961469 [Accessed May 24, 2013].

Health & Social Care Information Centre, 2013. Hospital Episode Statistics. *Health & Social Care Information Centre*. Available at: http://www.hscic.gov.uk/hes.

Higgs, G. & Gould, M., 2001. Is there a role for GIS in the “new NHS”? *Health & Place*, 7(3), pp.247–59. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11439259.

Hopkinson, N.S. et al., 2012. Designing and implementing a COPD discharge care bundle. *Thorax*, 67(1), pp.90–2. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3240777&tool=pmcentrez&rendertype=abstract [Accessed October 17, 2013].

Joyce, K., 2009. “To me it’s just another tool to help understand the evidence”: public health decision-makers’ perceptions of the value of geographical information systems (GIS). *Health & Place*, 15(3), pp.831–840. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19268622 [Accessed March 5, 2013].

Millett, C. et al., 2013. Hospital admissions for childhood asthma after smoke-free legislation in England. *Pediatrics*, 131(2), pp.e495–501.

Nacul, L. et al., 2011. COPD in England: a comparison of expected, model-based prevalence and observed prevalence from general practice data. *Journal of Public Health*, 33(1), pp.108–16. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20522452 [Accessed April 16, 2013].

NHS Right Care, 2010. The NHS Atlas of Variation in Healthcare.

Noble, D. et al., 2012. Feasibility study of geospatial mapping of chronic disease risk to inform public health commissioning. *BMJ open*, 2(1), p.e000711. Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3282296&tool=pmcentrez&rendertype=abstract [Accessed October 18, 2013].

Office for National Statistics, 2001. 2001 Boundaries. Available at: http://www.ons.gov.uk/ons/guide-method/geography/products/census/spatial/2001/index.html.

Office for National Statistics, 2010. Lower Super Output Area Population Estimates - Mid-2010. Available at: http://www.ons.gov.uk/ons/rel/sape/soa-mid-year-pop-est-engl-wales-exp/mid-2010-release/rft---mid-2010-lsoa-population-estimates.zip.

Pickle, L.W. & Carr, D.B., 2010. Visualizing health data with micromaps. *Spatial and spatio-temporal epidemiology*, 1(2-3), pp.143–50. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22749470 [Accessed October 15, 2013].

Ritchie, J. & Lewis, J., 2003. *Qualitative Research Practice: A Guide for Social Science Students and Researchers*, SAGE Publications Ltd.

Twigg, L., 1990. Health based geographical information systems: their potential examined in the light of existing data sources. *Social science & medicine (1982)*, 30(1), pp.143–55. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2305277.

UK Data Service Census Support, 2011. English Health Boundaries. Available at: http://edina.ac.uk/census/.

Welch, S.J. & Allen, T.L., 2006. Data-driven quality improvement in the Emergency Department at a level one trauma and tertiary care hospital. *The Journal of emergency medicine*, 30(3), pp.269–76. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16677976 [Accessed October 18, 2013].

Wennberg, J. & Gittelsohn, A., 1973. Small Area Variations in Health Care Delivery. , pp.1102–1108.

Appendix 1 – Informal Interview Results

Interview Question: Which map do you feel most effectively communicates the data?

Figure	1(c)	2	3(a)	3(b)
Title	Total COPD emergency admissions from 2010/11 LSOA level in London	Cartogram of total COPD emergency admissions from 2010/11 at LSOA level in London	Interpolation surface of total COPD emergency admissions from 2010/11 at LSOA level in London	Interpolation surface of total COPD emergency admissions from 2010/11 at LSOA level in London upon cartogram geography (Cart-IS)
Number of Votes	5	1	1	7

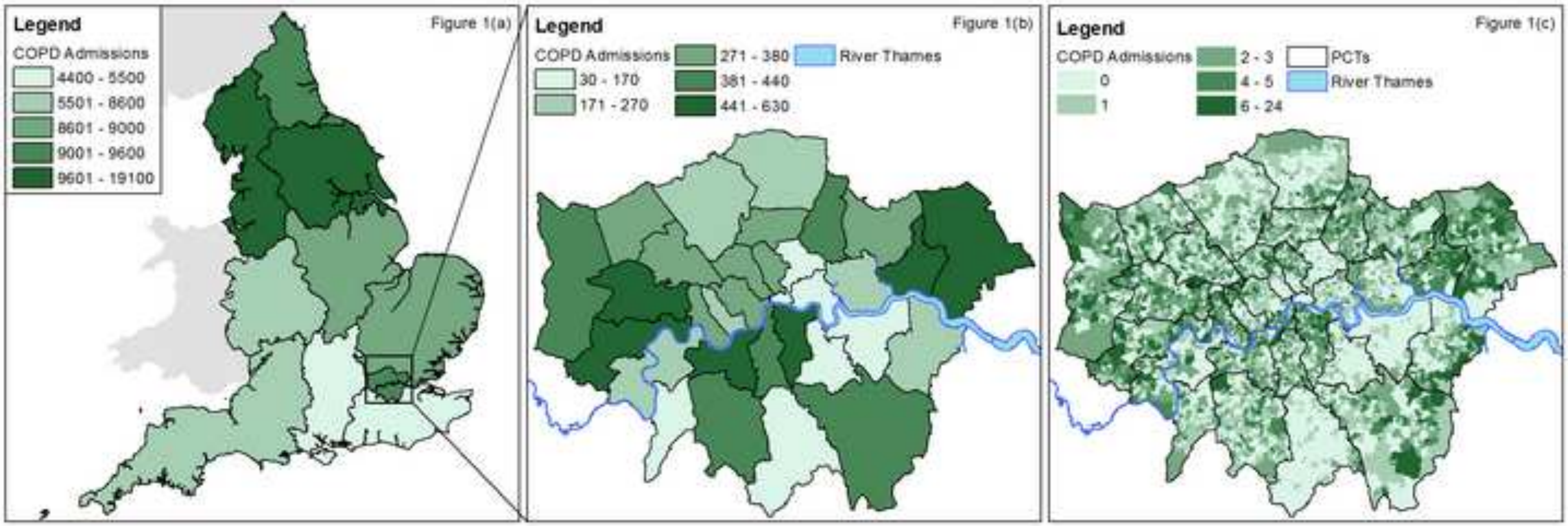
General Feedback Points:

1 – Figure 1(c) preferred as stand-alone item, however once Cart-IS (Figure 3(b)) explained and the need for the story described, it was understood why the methods had been applied and the majority of the time preference swayed towards Cart-IS.

2 – Majority of those interviewed acknowledged preference of spline over chloropleth.

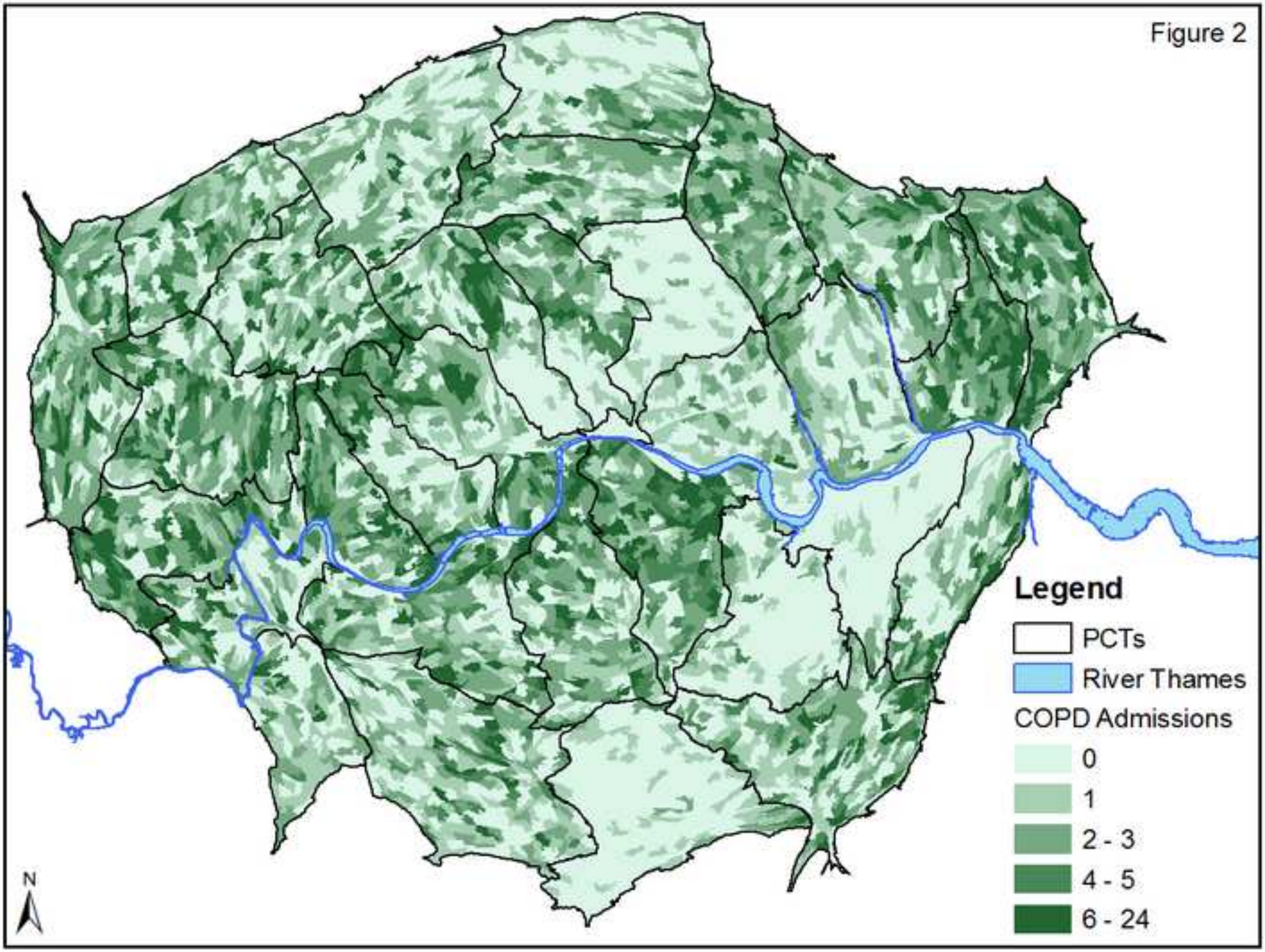
3 – It was noted that a common feeling that individuals do not have time, or are unwilling, to understand, learn or read the story and want a single output.

Figure
Click here to download high resolution image



Figure

Click here to download high resolution image



Figure

Click here to download high resolution image