Using routine geo-coded data to identify geographical heterogeneity to reduce disparities: case studies in UK.

Alan J Poots  
NIHR CLAHRC for NWL  
Imperial College London, Chelsea & Westminster Campus, London, UK, SW10 9NH  
a.poots@imperial.ac.uk

Stuart A Green  
NIHR CLAHRC for NWL  
Imperial College London, Chelsea & Westminster Campus, London, UK, SW10 9NH  
s.green@imperial.ac.uk

Ruth Barnes  
NIHR CLAHRC for NWL  
Department of Public Health, NHS Ealing, Perceval House, London, UK, W5 2HL,  
r.barnes@imperial.ac.uk

Derek Bell  
NIHR CLAHRC for NWL  
Imperial College London, Chelsea & Westminster Campus, London, UK, SW10 9NH  
d.bell@imperial.ac.uk

ABSTRACT

This paper outlines a structured argument for the use of routine health and demographic data to support the delivery of equitable services that are better aligned to the needs of the populations they serve. The paper describes case studies from a nationally funded research and quality improvement programme in London, UK as examples of targeting existing services, without top-down reconfiguration, using quality improvement methodology.

Three case studies are presented each demonstrating a differing use of geocoded routine data. The first demonstrates the use of a novel composite metric for the prospective targeting of service improvement; the second shows how routine geo-coded health data can be used to support the geographical location of services; the third demonstrates how routine data can be used to evaluate the impact of improvement initiatives on disparities in healthcare.

All methods provide a novel way of analyzing current service provision to ensure targeting of services where needed and contributing to the quality and cost challenges faced by healthcare providers and commissioners.

Categories and Subject Descriptors

H.2.8 [Information Systems]: Database Applications - Spatial databases and GIS.

General Terms

Management, Measurement, Performance, Design, Verification.

Keywords

Geographical heterogeneity, routine data, need, commissioning, health of the population, quality improvement, 3 variable display, GIS

1. INTRODUCTION

Quality improvement (QI) methodologies are becoming increasingly common in healthcare settings internationally. Most examples of the use of QI in healthcare come from the (acute) secondary sector where the QI tools are to specific and well-defined processes that are analogous to the industrial production line setting in which many of the tools originated.

In the UK, a preventative rather than responsive model of medicine, that ensures impacts spread beyond the individual service user, is known as “Public Health” or population health. This differs from the USA definition of public health, with the UK referring to the targeting of service to the wider population, analogous to the health of the population aspect of the “Triple Aim” of the Institute for Healthcare Improvement [1]. However, in the UK, many of the clinical services commissioned are not based on need but represent artefacts of historical commissioning or political boundaries. In the USA setting, this is typically driven by the “catchment” populations of the clients of insurance providers [5].

In order to achieve a commissioning regime that reduces disparities, clear identification of the population and need is of salient importance. Identifying the level of need within that defined population is essential so that services can be targeted effectively, and geographical information systems have been used to do this [5] [3] [7].

There is an interesting tension in the use of data in the UK: there is a large amount of routinely available data including demographics (e.g. ethnicity, age, gender, location socio-economic status), process data (e.g. procedures, Length of Stay, diagnoses), and outcome data (e.g. mortality, clinical response, readmissions). However, these data are often under-utilised despite the UK government’s commitment to improving access to data for research. This is likely due to sheer volume of data available and the difficulty in linking diverse databases. In addition, such data can be used for reconfiguration of services, and the evaluation of those services, with the same caveats.

Data collected routinely as part of service performance or quality assurance protocols can help to overcome this deluge of unlinked information. Routine data can act as proxy measures as indicators of need for service or in place of disease prevalence if actual data are not available. Furthermore, routine data are multi-advantageous: possessing a familiarity to decision-makers with a transparency of the metric, and having direct relevance to the system. Furthermore, routine data require no supplementary collection procedure, which reduces costs and expands the utility of extant data sets.

This paper seeks to show that by using routinely collected data, with which commissioners are familiar, geo-spatial analytical techniques allow the strategic determination of services; thereby aligning services to need and potentially reducing the disparity of healthcare provision. In addition, those using the data in this manner are responsible for that data, and as such, there is a natural feedback loop, where the quality of the data can be improved.
2. METHODS
Using a number of different geo-analyses the amenability of routine data to the provision of outputs for strategic planning of services is exemplified by several case studies, which provide the details of the individual methods used.

Analyses are conducted within the freeware QuantumGIS programme [12], using vector format data, and cross-tabulation of variables within the data tables.

2.1 Data sources
Boundary line data for Lower Layer Super Output Areas (LSOAs) [10] across England are used as analytical units. LSOAs are a geographical hierarchy designed to improve the reporting of small area statistics in England and Wales [10]. LSOAs were generated automatically to be as consistent in population size as possible from 2001 census data, this means they are of variable area [10]. Most of the UK government’s statistics are released freely at this geography, including the “Index of Multiple Deprivation” (IMD) – a composite metric of 38 separate indicators of socio-economic status across 7 domains for all LSOAs in England [2]. Given the well-established association between socio-economic status and health, this IMD metric can be used as a proxy measure for need for some diseases, with the lemma: higher IMD is indicative of higher need in that area. The population age structure is also available at LSOA level [9].

The Improving Access to Psychological Therapies (IAPT) service in England began with pilot work in 2006, and since then has been rolling out across all catchment populations. Each individual service maintains its own data and reports data of a standardized format to the central team [6] [11]. This paper uses data extracted from two services based within Northwest London.

Every hospital trust England also has its own data warehouse in which admissions data, clinical data, and outcome data are stored. These data are aggregated centrally, with less lag, as the Secondary Uses Service (SUS) data [8]. SUS data can be used to examine many disease conditions; in this paper, data pertaining to Chronic Obstructive Pulmonary Disease (COPD) are used as an exemplar of the techniques.

3. CASE STUDIES
3.1 Prospective analyses for targeting action
The IAPT service in the borough of Brent, London, is running a QI initiative to increase referrals, particularly focusing on the older population (over 65s year olds), who are under-represented in their current service uptake. Thus, there are three variables that available for display: a proxy of need, the IMD score; the adult population over 65 years old; and the location of previous referrals to the service. This is a classic three variable visualization problem, and an effective manner to overcome this is via the creation of composite and bespoke metrics.

In this instance, because of the particular interest in the needs of older people, a bespoke metric combining IMD score and Agedness of population can be constructed easily. This metric was scaled to the IMD scores found within London as a whole; thus, it is a relative measure calculated by dividing the IMD for an LSOA by the maximum IMD score found in London, and multiplying this by the proportion of the population who are over 65. All London was used because the service plans to roll the learning out to other boroughs that will also have to target activity appropriately and the calculation is no more arduous for the whole patch than for the individual borough. The maximum value of this derived metric is obtained only if the IMD score is the maximum and the population is entirely over 65 in that LSOA. The metric is thus a scaled IMD weighted by population age, and is a continuous representation of the classification described in the retrospective analysis. In practice, it is the comparison between values is relevant. For display purposes, any number of categories of this continuous metric can be selected.

Anonymised residential postcode and date of referral for the period March 2010 to May 2012 were extracted from the service database.

Overlaying the IMD-Agedness score metric with a mask based on any referral allows the investigation of whether previous referrals have originated in high need areas, and can highlight gaps in which there is an estimated high need and yet no referrals (see Figure 1A and 1B). These continuous overlays are simple, and provide greater flexibility for display.

Figure 1A. Referrals to IAPT Brent service shown in green, overlaying a composite metric of estimated need where darker red represents a higher level; and Figure 1B. A mask showing the estimated need in areas where referrals have occurred.
3.2 Prospective analyses for locating services
The IAPT data so far discussed have been aggregated, partly to preserve anonymity and to allow for a more effective display of large datasets. However, disaggregated data can also be effectively used in decision-making.

Admissions to Hillingdon Hospital, London for acute exacerbations of COPD during a 6-month period (November 2011 to April 2012) were extracted from SUS data, including residential postcodes. Data were cross-referenced to referrals to the Pulmonary Rehabilitation (PR) service. Anonymised admissions were mapped by residential postcode to provide a geographical distribution of patients that were referred to PR and those that were not (see Figure 2). These point data were overlaid on to mapped IMD scores, estimating need, within the borough of Hillingdon.

These data can be used for the prospective allocation of service locations, and to determine whether non-referrals are more, or less, common in higher IMD areas.

Figure 2. Admissions for Acute Exacerbations of COPD in Hillingdon, London. Displayed by subsequent referral to Pulmonary Rehabilitation service (red not referred), this can be used to site new clinics. Higher estimated need is shown in darker blue.

3.3 Retrospective analyses for evaluation
The IAPT service in the borough of Westminster, London, ran a QI initiative to increase referrals; this provided an opportunity to evaluate whether access exhibits disparity through assessment of routine service data in conjunction with a proxy measure for need [4].

The analysis found that all of the LSOAs in Westminster that were categorised as high IMD/low referral converted to high IMD/high or medium referral [4]. This demonstrated that the QI initiative had delivered better access to some of the areas with higher estimated need in Westminster; and that there are opportunities for further improvement [4]. Commissioners can use this information to evaluate the effectiveness of the service and to build appropriate requirements into new contracts. For instance, it was found that two LSOAs have low estimated need but high referrals in the post-funding phase of the data; investigating what is happening in these LSOA will be of interest.

4. DISCUSSION
The comparison of quality measures across populations is not new, yet the combination of population level metrics of estimated need with quality improvement methodologies is novel, and can be used to help enhance service planning and delivery. Furthermore, the use of fine resolution geographies allows specific targeting whilst still providing manageable analysis and presentation. The work presented focuses on access, but follow up work on uptake of service and sub-group analysis is being conducted, examining even when appropriate levels of access (referral) are made, there are issues with service utilisation.

The use of routine data demonstrates that there is a capability to provide decision support materials without access to full data sets, and indeed even when given partial data sets due to confidentiality restrictions, analyses can be undertaken, and meaningful results derived at useful resolutions. In addition, the establishment of a baseline coupled with routine data, allows an “evaluate-as-you-go” method, moving away from the implement-evaluate model, which is ineffective for the appropriate handling of temporal variation, and possibly problematic for complex interventions, e.g. [13].

There are some constraints with the use of such data. The foremost being that the use of proxies is problematic: not all conditions will be amenable to proxy by IMD, the linear model of higher IMD being equal to greater need is simplistic, no index will explain all the variation in healthcare need [7], and whilst there are other proxies available, all must be used with similar caution. There are ways to ameliorate this shortcoming, e.g. by combination with additional data types. For example, certain types of illness are more likely to occur in certain groups, e.g. Sickle Cell Disease is more common in Black and Minority Ethnicity (BME) populations. An additional problem is that of aggregation, and whilst the LSOAs provide a fine resolution at which to work, they still mask differences within those areas [9]. In London, this is of particular relevance as within many highly affluent districts there are areas of poverty that would be hidden by such treatment.

Whilst there are cases in which linkages of data-sets results in incomplete, inaccurate, or contradictory data; the data sets presented in these analyses are fortunate in possession of quality assurance mechanisms in their collection, and all include geo-spatial references as part of the data set. For data sets in which this in not the case, the researcher must choose between data loss and data-estimation (through interpolation or prediction from other variables). Relatedly, the possession of geo-spatial references should not automatically lead a researcher to present data as a map; the end use needs consideration as there is potential for inadequate or misleading presentation and interpretation; and other techniques may prove more appropriate.
Another consideration is that for display purposes a smaller number of categories must be used, to reduce visual confusion: this again homogenises the very heterogeneities that were set out to be examined. As with many applications of data visualisation analyses, care must be given to the display of the results, considering the number of displayed classes of information to prevent visual clutter.

Furthermore, there may be entirely satisfactory explanations for low numbers of referrals where they are expected them to be higher and reacting to the broad data picture without an understanding of the context can be as problematic as reacting to the context without understanding the broader picture. In some cases, there might be an alternative provision by the third sector (charitable) organisations, or through additional support services available to particular cultures (In Westminster the Chinese Cultural Centre provides such an option), and this would not constitute a problem. In other cases, there can be cultural taboos that prevent access to a service and it may necessary to find different ways of increasing referrals in such cases, this is likely in Brent. The need for partnership working across disciplines and professions (e.g. academia, community groups, government, and commissioners) for the full interpretation of such analyses is highlighted by these case studies.

Despite these limitations, the use of the routine data can help commissioners and decision makers to have a more complete understanding of health needs and service provision by marrying data to context, in a systems thinking approach. Because the methods shown here use data that are already available to a service, and a freeware analysis platform, it shows how benefits from that data can be enhanced to the service at low cost. One use that has not been demonstrated here is the mapping of economic data pertaining to the service but there is no reason why this cannot be performed. In addition, the reconfiguration of the Public Health function in England, risks a potential loss of a skilled workforce that would have traditionally undertaken such analyses. The methods outlined here can support the current analytical regime, and are a viable partial solution should such a skill loss occur.

5. CONCLUSIONS

This paper shows that routinely available data coupled with geo-spatial analysis can be utilised in several ways to analyse service provision; thereby outlining a method to align service provision based on health needs of the population, thereby reducing the disparity of healthcare provision and potentially addressing wider health disparity. This is independent of the healthcare setting as it provides decision makers with an easy to interpret format of familiar data (that which is used to inform them about service performance), and so both commissioners in the UK and Insurance providers and Accountable Care Organisations, in the USA can identify high need areas and act according to their remits. The Triple Aim of the IHI is analogous to the current role of Public Health departments in the UK, and no doubt, parallels and comparisons will be made between these models of care.

The use of routine data for strategy should not come at the expense of clinical judgments, however, and the exact tactics to be employed once the heterogeneities of each care regime are known will rightly depend on context and case. This tension is well known in Public Health departments in the UK, and one that is yet to be resolved.

6. ACKNOWLEDGMENTS

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 authors and not necessarily those of the NHS, the NIHR or the Department of Health.

The authors thank the project teams of CLAHRC for NWL at IAPT Brent, Hillingdon Pulmonary Rehabilitation Collaborative, and IAPT Westminster; and the two anonymous reviewers.

7. REFERENCES


