

# Using routinely collected data to inform service design and delivery: trials, tribulations and successes.

Sarahjane Jones<sup>1</sup>, Eloise Phillips<sup>1</sup>, Jessica Runacres<sup>1</sup>

<sup>1</sup>Centre for Social Care, Health and Related Research, Birmingham City University

## Abstract

Increases in connectivity and digital technologies in healthcare provide a growing pool of data, data that, with the right analytics, could provide new insights into disease and care services that were previously unavailable to us. We present our experiences of accessing and examining routinely collected data, the trials, tribulations and successes.

## Introduction

In 2013, in an effort to embrace the information age and technology revolution, the Secretary of State for Health challenged the NHS to go paperless by 2018 with expectations that cost efficiencies and improved services would follow.<sup>1</sup> The delivery of this vision has been the responsibility of the NHS Commissioning Board,<sup>2</sup> who laid out plans to support better data and informed commissioning. The growing volume of data and increased integration has the potential to deliver previously inaccessible insights that could support improvements in quality and safety.

## Methods

### **Project 1: service specific data**

In this example, we sought to access data relating to the case management service across the West Midlands region. Case management is a community-delivered service designed to reduce the number of unnecessary hospital admissions amongst patients with repeat admissions. The purpose of this study was to assess and understand the distribution of A&E attendances and hospital admissions of this patient population.

### **Project 2: software specific data**

In this example, we sought to access data from a specific software that was used by multiple organisations for the delivery of a 'social prescribing' type service. Social prescribing is a concept that aims to provide holistic care to patients presenting in healthcare whose needs are greater than the traditional clinical treatment options. The purpose of this study was to describe the user population and characterise the care options available.

### **Project 3: organisation specific data**

In this example, in collaboration with University Hospitals Coventry and Warwickshire NHS Trust, London South Bank University and Wolfram Research, we sought to establish the relationship between nurse staffing and patient outcomes across a single organisation from multiple database sources.

## Results

Across the three projects, we found a range of issues from identifying the appropriate data sources, to poor quality data input, to poorly organised and managed data.

### **Project 1: service specific data**

Case management is a community-delivered service. For some patients, the service in the community is delivered by the same organisation that provides their acute services. Some of these Trusts also had integrated data systems, whereby the nature of the service provision in the community was known by the database in A&E. We characterised these as simple single Trust providers. Some patients received care in the community by the same organisation that provided their acute services, but data systems were not integrated. We characterised these as complex single Trust providers. For some patients, their community services were provided by a different to organisation to those delivering local acute services and therefore, because of data protection laws, it was unlikely the acute provider would have intelligence on the case managed patients. We characterised these as double Trust access. The classification of organisations was essential in enabling the team to establish which organisations could reasonably provide the data required. Data were obtained from two simple single Trust providers. Analysis of one Trust has indicated that the case managed population is typically over 70 years old and most often arrive at A&E via emergency services. A large proportion of those arriving by ambulance are admitted to hospital and this is statistically significant.<sup>3</sup>

### **Project 2: software specific data**

All data recorded by an intermediary social prescribing service were extracted. Of 32 variables, three were incomplete (not populated with data) more than 90% of the time and 16 were incomplete more than 75% of the time. Only three variables were fully complete and these were automatically populated. No variable that required manual data input was completed 100% of the time including essential variables such as patient date of birth and postcode. The data variables were designed with the expectation that they would be relevant to all patients in the service, including a large range of demographic data including employment and disability status and some baseline data patient wellbeing.

### **Project 3: organisation specific data**

In this project, we aspired to use multiple data sources within a single organisation to establish the presence of relationships, if any, between nurse staffing and patient outcomes. This single organisation had more than 400 databases, from hospital-wide systems such as Datix, to team specific spreadsheets, with some repetition in data collection. When combining databases, data co-ordination was time intensive; by this we mean, the time required to align data relating to different areas to of the hospital so that area-specific correlations could be analysed. This study found more than 40 relationships between nurse staffing and outcomes as well as provided a proof of concept for the use of routinely collected administrative data to inform safe staffing levels.<sup>4</sup>

## Discussion

These projects demonstrate that accessing and analysing routinely collected data can deliver new knowledge on current service provision. However, there are difficulties in

obtaining and handling this data that need to be considered when planning studies. Understanding the number of organisations and the way they interact with other organisations with regards to a defined patient population is important in selecting organisations to approach to access data and further thought needs to be given to how this selection bias affects the outcomes of the study.

When using databases that have required manual inputting, thought needs to be given to the quality of the data and the fullness of the content provided. Variables that are incomplete more than 90% of the time can indicate one of two things; the variable is not necessary/redundant, or the data inputter doesn't value the variable and therefore doesn't spend time completing the field. Understanding which is essential if data collection is to be seen as a valuable exercise and produce robust, useable data.

Data is becoming easier to collect, store and manage. Too much data, structured and organisation in isolation of other systems can create difficulties in data integration at a later date. Data redundancy, although it has some value such as safeguarding the loss of data in one location, can be inefficient, particularly in the case of time intensive data collection, as well as introduce inconsistencies where data is poorly controlled.

## Conclusion

Using routinely collected data to support service design and delivery can provide invaluable insight into how the service is performing and where improvements can be made. However, it is not without its difficulties. Consideration needs to be given to data protection laws and information governance, and how these can impact access to data regarding patient populations that traverse multiple organisations. Consideration is also required for the data inputter experience and the value they perceive of the field as this can have serious consequences for data quality. Finally, when data is being collected, thought needs to be given to the structure of various databases in the event of integration. The problems with using routinely collected data presented here are by no means exclusive and we will present a more comprehensive narrative of the trials, tribulations and successes in our presentation.

## References

1. Department of Health. Jeremy Hunt challenges NHS to go paperless. 2013. Available at: <http://webarchive.nationalarchives.gov.uk/20150402110949/http://digitalchallenge.dh.gov.uk/2013/01/16/paperless/>
2. NHS Commissioning Board. Everyone counts: planning for patients 2013/14. 2013. Available at: <http://webarchive.nationalarchives.gov.uk/20150402110949/http://england.nhs.uk/everyonecounts/>
3. Phillips E, Jones S, Runaces J, Lintern M, and Radford M. Understanding the distribution of A&E attendances and hospital admissions for the case managed population: a single case cross sectional study. Applied Nursing Research. 2016 (online first)
4. Leary A, Cook R, Jones S, Smith J, Gough M, Maxwell E, Punshon G and Radford M. Mining routinely collected acute data to reveal non-linear relationships between nurse staffing levels and outcomes. BMJ 2016 (accepted, awaiting publication)