MEDINFO 2010
C. Safran et al. (Eds.)
IOS Press, 2010

doi:10.3233/978-1-60750-588-4-496

Obesity Atlas and Methodbox: Towards an Open Framework for Sharing Public Health Intelligence Workflows

Sarah Thew, Paul Jarvis, John Ainsworth and Iain Buchan

Northwest Institute for Bio-Health Informatics, School of Community Based Medicine,
University of Manchester, United Kingdom

Abstract

The large growth in data sources relevant to public health has not been matched by a growth in human resource for producing intelligence to support decisions or generate new insights. There is a need to bring scarce public health expertise into closer alignment with data and data processing methods to support timely public health analysis. The difficulties of developing and sharing this expertise in large organisations such as the UK's National Health Service have long been recognised. We report findings in this area across two projects Obesity Atlas and Methodbox, which are developing and sharing best practice between Public Health Analysts in England, and we address the relevant generic knowledge management problems in the Public Health community.

***Keywords**:*

Public health intelligence, Workflow, e-Lab, Knowledge management

Introduction

Most public health analysis in England takes place in Primary Care Trusts (PCTs), which are responsible for commissioning health services for populations of around 300,000 people. Each PCT employs analysts who collect and analyse local and national health-related data to support local decision-making. Analysts from the different PCTs belong to informal public health intelligence networks which meet regularly to exchange ideas and share expertise. They also communicate by frequent ad hoc telephone and email conversations.

Despite regular communications within and between PCTs' analysts, we identified substantial difficulties in finding information for particular analyses in a timely manner. For example, analysts reported issues such as the difficulty of locating an email sent a year ago which discussed the calculation of an annual statistic, or uncertainty around deciding how and when to apply different versions of a method. All of these difficulties can be characterised as knowledge management problems. The term 'knowledge management' refers to a broad group of activities around the creation, sharing and adoption of expertise and understanding. Knowledge management looks beyond the sharing of information (e.g. the definition of a statistical method) and considers the experience and tacit knowledge required to make use of the information, such as how and when to apply the method, or who is the relevant local expert. Such knowledge, often tacit to individuals or teams, can be difficult to transfer to new staff and is easily lost to an organisation when people move to new jobs. This folk knowledge (making connections, shared learning through problem solving) is often highly contextual and held only in people's minds or reflected in e-mails or informal communication channels such as forums or instant messaging – ideal for supporting the processes of knowledge exchange and problem solving, but poor long term repositories, which cannot be searched and cannot generally provide reliable, accessible context data.

Bate and Robert's review of knowledge management practices in the UK National Health Service (NHS) [1] compared the relative maturity of private sector companies with what they describe as a 'naïveté' around knowledge management within the NHS. In particular, they observed a policy of 'top down' imposition of networks intended to support the sharing of best practice and expertise which was considered unlikely to be successful or self sustaining, and contrast this with a recent change of approach within the private sector to support more organic, grass-roots knowledge sharing activities and naturally-occurring networks

Previous reviews of knowledge management practices in the NHS identified successful approaches as well as pitfalls. The CHAIN (Contact, Help, Advice and Information Networks) project [2] succeeded in fostering the development of a community, and helping users find people with particular areas of expertise. This study highlighted the importance of the social dimensions of knowledge management – whilst there was a technical element (the development of a messaging system), the success of the project depended on vigorous facilitation by an email list administrator, who created new connections between pre-existing groups.

Studies of knowledge management have tended to focus on technologies for storing knowledge [1]; there have been few in-depth case-studies of the human factors and the wider processes of managing knowledge. Storey [3] observes that often, there can be too much focus on the technology, a sense that if only the right technology is in place, people will be able to make use of it, but this has repeatedly been shown not to be the case – successful knowledge management projects build on existing networks. Consequently, we report the engineering of two pieces of knowledge management software as extensions of a user community, involving users from inception.

Obesity Atlas

Like many western countries, for more than half a century, the UK has recorded the heights and weights of children in order to monitor their growth [4]. More recently these data have been used to produce public health reports for obesity surveillance and local planning [5]. The Obesity Atlas project aimed to develop software to automate the cleaning and analysis of these data.

We identified a number of knowledge management requirements. First and foremost was the need to support a disparate group of Public Health Analysts, from different organisations and backgrounds, to converge on a single workflow for data cleaning & analysis and interpretation of the resulting child obesity profiles for their localities. The analysts work independently in separate geographic areas and are used to determining their preferred analysis approach in isolation. Whilst there were aspects of the workflow on which it was easy to reach a consensus (e.g. which summary statistics to use), other steps were more ambiguous (e.g. handling missing and possibly erroneous data – some analysts may choose to exclude records with any missing data, other analysts may decide to keep those records depending on exactly which fields are missing, or alternatively may choose to go back to the data supplier for clarification). Therefore agreeing a shared workflow for such decisions requires the analysts to negotiate and share their tacit knowledge. In other circumstances, rather than working with the analysts to agree a common workflow, the system has been built to provide some flexibility, for example providing obesity profiles based on different definitions of child obesity.

We used a variety of requirements-gathering techniques to facilitate this negotiation process, including group discussions and meetings, one-to-one formal meetings, informal telephone conversations, and workshops allowing users to interact with prototypes. The use of prototypes was particularly successful, users were able to respond to and interact with simple prototype versions of the software. This interaction encouraged and inspired participants to discuss additional functionality that was incorporated into the system.

The analysts identified two workflow types, data pre-processing and data querying. The data pre-processing workflow included data categorisation, cleaning, standardisation and integration. Once uploaded and pre-processed, the dataset is ready to be queried. There were four functional types of data query: statistical analysis; thematic mapping, charting, and locality profiling. The latter is a context-specific superset of the former three.

The most sophisticated query was the locality profiling in the context of creating a consistent PCT-level profile of child obesity. A profile is a document resulting from a series of data queries. The profiles summarised child obesity statistics for the PCT in a standard format that could be compared with other PCTs. It also compared obesity prevalence by socio-demographic groups and geographical areas within the PCT. Thematic maps showed the spatial distribution of obesity; and charts and accompanying statistical modelling results reflected relations (or not) of obesity with deprivation, ethnicity, or other groups specified by the user. The profile is produced automatically following the upload of a dataset, saving each analyst several days of work, and provided to the user as a downloadable document of approximately 30 A4 pages. Again, the content and format of the document was designed in consultation with the PCT analysts. The document was designed such that components can be extracted and reused in official public health reports. An example component (a thematic map) showing the relative levels of child obesity within Bury PCT sub-regions is shown in Figure 1.

Obesity Atlas V1.0 has been released and is being used by public health decision makers and data analysts across the 10 Greater Manchester PCTs; it is informing their decisions on local public health policy to monitor and reduce childhood obesity[6, 7]. Early feedback is very positive and ongoing improvements and support will be delivered by the Northwest e-Health programme[8], which is set up to sustain health informatics innovations as reliable services for the NHS.

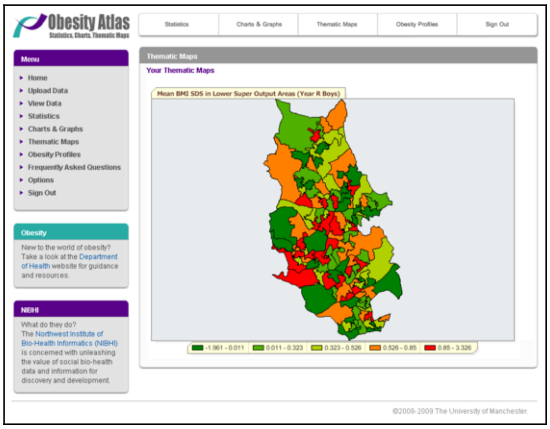

Figure 1- Thematic mapping in Obesity Atlas

The Obesity Atlas project is an example of combining the expertise and knowledge of many public health analysts to converge on an agreed workflow and produce a piece of software that dramatically reduces the burden of analysis on the PCT staff. From a knowledge management viewpoint, the process has enabled a group of experts to work together, sharing their tacit expertise to agree on a set of best practices for this analysis process, which have then been automated. This expertise has now been captured, holding it within the organisation, so that if any of the analysts involved in the project were to leave, the relevant organisational memory would persist.

Methodbox

Our team is now working on a project which takes a wider view of analysts' working practices. Rather than focussing on one single workflow, the Methodbox project [9] looks at supporting PCT analysts and academic researchers in sharing analysis expertise and methods – applied in the first instance to large health surveys. Methodbox draws on the experience of managing knowledge in e-Science settings, particularly on sharing workflows and using social networking technologies to share expertise [10]. The initial motivation for the project was the recognition that understanding the obesity epidemic at the population level requires expertise from disparate disciplines including medicine, biology and the social sciences, as well as the involvement of academic and NHS researchers – experts who are interested in learning from other disciplines but who find accessing another field challenging for numerous reasons including language, methods, tacit assumptions and background to the work.

The first stage of our project was to interview NHS analysts and academic researchers who either currently use or have expressed an interest in using Health Survey for England (HSE) [11] in order to understand current working practices and barriers to use of the survey. HSE is the largest scale annual health survey in the UK and is intended for both health policy and scientific purposes. Its datasets, although freely available, are complex and perceived to be under-used. We held two workshop sessions at regular meetings of North West England Public Health Analysts, followed by longer interviews with three volunteers. We also interviewed five academic users of HSE, three epidemiologists and two social scientists.

Many of their concerns relate to problems accessing expertise and knowledge about the specifics of how to work with HSE. The most frequently cited reasons for not using existing data were methodological concerns over how to go about using the data, for example understanding the circumstances in which to apply weightings to a survey – tacit expertise that is not easily codified into clear-cut rules. Interviewees also reported being overwhelmed at the size of the dataset, difficulties working with the documentation and a lack of worked examples. Similarly, whilst the expert academic users we interviewed had developed coping strategies for many of these problems, novice academic users also felt the lack of examples and found the size and complexity of the survey and metadata problematic. They were equally unsure of how to access expertise.

These concerns have informed the design of Methodbox. Methodbox provides tools to support the individual in working with HSE data, including improved metadata searching and the ability to download selected slices of a survey rather than entire datasets, but also provides support for collaboration and networking in order to address many of the knowledge management concerns expressed by the analysts. Methodbox allows users to share their expertise with other analysts, whether that be a comment about a variable, a script demonstrating how to harmonise educational variables across several years or a publication they have used, but it does this in a rich context specific way, creating links between the data, for example:

- An analyst's comment about a variable is associated with that variable, so that any search returning the variable will also display the comment to other users. A second user can react to the comment, adding their experiences or suggestions.
- When uploading a harmonisation script the user can link the script to all the survey variables used within that script, providing a rich example for other users. Another analyst who refines the script can upload a new version, linking back to the original script and adding comments to develop the discussion.
- A user can upload details of a publication, perhaps providing evidence in support of a method they have chosen to use, or uploading their own publication, and associating it with a script and a set of variables, used in generating the published results.

Thus Methodbox allows users to interact with each other, in an environment that supports richly contextual information sharing, and allows the searching and indexing of content.

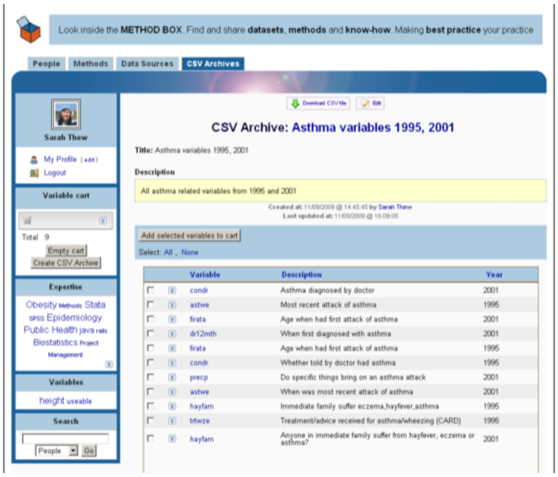

Figure 2- Methodbox Prototype showing a selection of asthma related variables collated by a user from Health Survey for England 1995 and 2001

A prototype system has been developed which is being tested with academic public health researchers and NHS analysts. Based on the experiences of other relevant knowledge management projects we are focusing on building existing communities [12] by working alongside NHS Action Learning groups, PhD supervisors & students, and topic-specific research groups. We believe the system will encourage wider sharing and improved communication between different communities. Whilst a social scientist and an epidemiologist, both interested in public health issues around obesity, may superficially appear to have much common ground, the methods, language and tacit assumptions underlying their analyses can be quite different. Methodbox aims to bridge the division of knowledge that results from this specialisation by providing rich, worked examples of analyses from different disciplines.

e-Labs and Work/Research Objects

The Obesity Atlas and Methodbox activities feed into the general e-Labs initiative, focused on healthcare by NorthWest e-Health [8]. An e-Lab is a set of integrated components that, used together, form a distributed and collaborative space, enabling in-silico investigations. An e-Lab brings together data, methods and people sharing a currency of Work Objects (digital resources encapsulating the inputs, processes and outputs of an exploration or problem-solving activity). In the e-Science context a Work Object (WO) is known as a Research Object. A WO is an entity that an e-Lab: creates, stores, accesses and manages; exchanges with other e-Labs; publishes to external sites, deposits in external resources; and displays through work- benches. The motivation behind WOs (and the associated services that produce and consume them) is to improve the curation, accessibility and repeatability of research and business intelligence processes. A WO might be:

- A template equity audit containing a collection of workflows with instructions, examples with default input data, a tutorial and links to external learning resources on the methodology;
- A reproducible research article with the workflows and data required to reproduce the results described in the article.
- A reproducible and easily repurposed annual public health report containing: the report and supporting files such as slides, spreadsheets and graphics; commentary around the report such as copies of emails and press releases/articles; pointers to data sources and extracts used in analyses; data extraction, cleaning, derivation and statistical analysis workflows behind the analyses; bibliography; project management resources.

From a knowledge management perspective, WOs address the concerns that: knowledge is heavily context dependent; and knowledge-transfer relies on common ground.

Research/experimental/business intelligence knowledge depends on context, making it's codification and exchange difficult [12]. Some redundancy of knowledge between two parties is necessary for knowledge transfer to take place [13], without any overlap it is almost impossible for anything but the most basic information to be transferred. WOs can provide the rich context and common ground necessary to take, for example, a method beyond a flat set of instructions, to provide worked, interactive, re-playable examples – the learner can interact with the research object, changing parameters, supplying new raw data, and learning by doing.

A case study by Quintas [12] observed that rigid formal reporting and recording of progress, without a social dimension is not an effective way to transfer knowledge. There is a requirement for joint activities - "working together and gaining shared experiential knowledge", WOs provide an alternative way to support collaborative working and, longer term can provide rich examples for teaching.

Conclusion

Obesity Atlas demonstrated the need for combining related public health intelligence tasks into workflows that result in near-complete reports. Methodbox demonstrated the need for a generic framework for finding, sharing and reusing public health intelligence workflows – with an initial focus on the under use of large-scale, general health surveys. The UK NHS is using the NorthWest e-Health programme to harness e-Science activities such as these into a generic framework, e-Lab, for combining data, methods and people more effectively – to produce more health intelligence from the same number of analysts.

WOs and e-Labs, however, are not familiar concepts to most public health analysts. Adoption will require technical and cultural change; change to working practices, as well as assumptions about what it is preferable to share or keep confidential. Obesity Atlas and Methodbox provide steps towards public health e-Labs. Obesity Atlas provides an example of a single packaged reusable workflow, and Methodbox encourages the sharing of methods and expertise.

The next steps include developing knowledge management practices alongside the e-Lab software by embedding e-Lab development in the continuing professional development of public health practice and science. For example, action learning groups in healthcare organisations can produce and use WOs as learning materials as well as operational tools. Continuing professional development credits could be earned by users transforming WOs in ways that demonstrate key competences. The provenance and audit services of e-Labs are easily harnessed for this purpose. These services are essential for the information governance required by healthcare organisations. As WOs contain rich contextual metadata, an e-Lab is a context-specific alternative to (or extension of) a conventional file store – in this respect organisational memory is increased by keeping work as WOs rather than separate files.

e-Lab could also be used to engineer the reward environment for working across organisations and disciplines. For example, a Social Scientist could demonstrate the impact of his/her method for measuring material deprivation by showing the consumption of a WO they produced as a template for this methodology. Given that such environments are increasingly global, e-Lab could be a step towards international interoperability of public health intelligence. Given also that future healthcare is likely to be more model-rich [14], clinical decision-support and public health intelligence may need to interoperate in a similar way.

Acknowledgements

The Obesity Atlas project was jointly funded by the ten Greater Manchester Primary Care Trusts.

The Methodbox project was funded by the *Economic and Social Research Council* under the *National Centre for e-Social Science* programme.

References

[1] Bate SP, "Knowledge management and communities of practice in the private sector: lessons for modernising the National Health Service in England and Wales," Public Administration, vol. 80, p. 643, 2002.

[2] Russell J, Greenhalgh T, Boynton P, and Rigby M, "Soft networks for bridging the gap between research

and practice: illuminative evaluation of CHAIN," BMJ, vol. 328, p. 1174, May 15 2004.

[3] Storey J and Barnett E, "Knowledge Management Initiatives: Learning from Failure," in Managing Knowledge, Second Edition ed S. Little and T. Ray, Eds.: The Open University, 2005.

[4] de Onis M, Wijnhoven TM, and Onyango AW, "Worldwide practices in child growth monitoring," J Pediatr, vol. 144, pp. 461-5, Apr 2004.

[5] Patterson L, Jarvis P, Verma A, Harrison R, and Buchan I, "Measuring children and monitoring obesity: surveys of English Primary Care Trusts 2004-06," J Public Health (Oxf), vol. 28, pp. 330-6, Dec 2006.

[6] Jarvis P, "The Development of a Public Health Intelligence Tool for Child Obesity in Greater Manchester" School of Community Based Medicine, University of Manchester Doctor of Philosophy, 2009

[7] Obesity Atlas,https://www.obesityatlas.org.uk/, Accessed: 15th October 2009

[8] North West e-Health, http://www.nweh.org.uk/Default.aspx, Accessed: 13th October 2009

[9] Methodbox,http://www.methodbox.org, Accessed: 14th October 2009

[10] De Roure D, Goble C, Bhagat J, Cruickshank D, Goderis A, Michaelides D, and Newman D, "myExperiment: Defining the Social Virtual Research Environment," in eScience, 2008. eScience '08. IEEE Fourth International Conference on, 2008, pp. 182-189.

[11] UKDA, Health Survey for England, http://www.esds.ac.uk/government/hse/, Accessed: 6th May 2009

[12] Quintas P, "Managing Knowledge and Innovation Across Boundaries," in Managing Knowledge - An Essential Reader, 2nd Edition ed, S. Little and T. Ray, Eds. London: The Open University, 2005.

[13] Nonaka I and Takeguchi H, The Knowledge Creating Company. Oxford: Oxford University Press, 1995.

[14] Buchan I, Winn J, and Bishop C, "A Unified Modelling Approach to Data Intensive Healthcare," in The Fourth Paradigm: Data Intensive Scientific Discovery: Microsoft Press, 2009.

Address for Correspondence

Sarah Thew
Sarah.Thew@Manchester.ac.uk
Northwest Institute for Bio-Health Informatics,
School of Community Based Medicine,
University of Manchester, United Kingdom