Predictive Modelling for Planned and Unplanned Care in the North East and Cumbria

Connected Health Cities – Durham Pathway Final Report

Contents

About the Funder

This project was delivered under the Connected Health Cities (CHC) initiative, commissioned by the Northern Health Science Alliance (NHSA) and funded by the Department of Health (DoH). This pathway grant was awarded by the CHC North East and North Cumbria to Durham University.

Connected Health Cities (CHC) unites local health data and advanced technology to improve health services for patients across the North of England.

By making better use of the information and technology that already exists in our health and social care system we are improving health and ensuring services are more joined-up. By working with patients, health practitioners and experts in digital health we are delivering research that is relevant, effective and having a real impact on public health.

This project was funded from October 2016 to December 2018 and received a no-cost extension to March 2019.

About the Project Team

This pathway project was delivered by a team from Durham University:

Professor Graham Towl – Principal Investigator (Dept of Psychology)
Dr Camila Caiado – Co-Investigator (Dept of Mathematical Sciences)
Dr Rachel Oughton – Research Associate (Dept of Mathematical Sciences)
Clare Collyer – Project Manager

The University worked alongside an external consultant, with extensive recent experience working in the NHS, Dr Ian Briggs.

With valuable input from Lisa Nattrass, Head of Information Governance, CDDFT

The key partners were County Durham and Darlington NHS Foundation Trust (CDDFT), South Tees Hospitals NHS Foundation Trust and Darlington Borough Council.

Other collaborating partners include AIMES, South Tyneside NHS Foundation Trust, Gateshead Health NHS Foundation Trust, Primary Healthcare Darlington, NECSU.

Enquiries about this project should be addressed to the Project Manager in the first instance: Clare.Collyer@Durham.ac.uk

1. Executive Summary

1.1 Key outcomes

a) The project delivered the necessary Information Governance Framework to enable data transfer from multiple NHS Trusts and GP practices, and adapted it for the new legislative environment. The documentation and process developed is already having impact elsewhere in the field and is currently being used by the team in another project. It is also being used by other researchers on another Health research project.
b) A planning tool was successful designed for Planned Care and developed in consultation and collaboration with General Practice (GP) managers. This was built, tested and developed using pseudonymised data and is now able to be successfully run using aggregated datasets. This simplifies the information governance procedures necessary and makes the app more easily accessible for practices.
c) The practice level app was developed to build a hub level model, combining datasets from multiple practices, which could inform service design across clusters of practices. This also works as a proof of concept for an area-wide app. The technical work has been completed to allow this to be run across the Primary Healthcare area, combining all of the practices, although this was not undertaken or tested.
d) This app is also being used by several practices to inform the next planning round. The team is still receiving requests to develop the app to help practices measure performance against key targets.
e) An unplanned care modelling app was developed to predict attendance at Accident and Emergency (A & E) departments. The A&E app was taken up by a Trust outside of the project.
f) The team procured an appropriate infrastructure to work securely with NHS data outside of the University's own systems. The team worked with the provider to develop this solution and was able to pass the learning from this process on to the team designing the regional Trusted Research Environment.
g) The work carried out and relationships built, under this project, has led to further funded work. This new funded work is directly building on the work started under this exemplar project.

1.2 Recommendations

a) Agreement and accessibility of generic/template documentation would reduce the need for extensive institutional drafting or review each time.
b) Future projects may also, benefit from the provision of regional expertise from an organisation like AHSN or similar. It would allow for concentration of expertise in health data research and reduce dependency on individual organisations. This would avoid duplication of effort regionally, mitigate impact of changes and thereby minimise delays and ultimately costs.
c) Future projects should ensure that there is funding allocated for NHS Trusts Information Governance Support, which proved invaluable during this project.

d) CHC showcasing events were very useful in sharing the work that was going on between project teams. Future projects within a programme may benefit from coming together earlier in the programme to raise awareness of crossover and shared problems – which would improve collaborative working, problem solving and lead to efficiencies. It could also potentially generate further collaborations and grant applications.

e) Potential future work to develop a regional picture would need an NHS/external organisation to bring together support and champion the benefits. Greta North Care Record (GNCR) developments could facilitate work of this type for future projects, providing a platform and establishing a trusted process.

f) Working with a consultant who has experience in the NHS organisations, and has the appropriate contacts and manages the links between the project team and collaborative partners from an external point of view has been key for this project. Future projects may well wish to consider replicating this model.

g) This project was funded to pay NHS Trusts for analyst time, which helped to keep the project on track. There was no funding to pay for datasets. Given the cost to NHS Trusts of collecting and maintaining this data, future projects may very well need to factor in data costs to the budget.

2. Introduction

The health sector is under increasing pressure, due to a range of changes in policy, funding constraints and changes in the demographics of the population. With the NHSE digital transformation agenda, there is increasing data available, but the NHS organisations are not always resourced or supported to carry out exploratory analysis and modelling of this. This project aimed to establish collaborative working between NHS Trusts and academics, to allow academic expertise to assist in understanding and modelling changes.

This project sought to produce practical modelling, planning and decision support tools that can be routinely used by appropriate health teams to forecast in both planned and unplanned care across health services in the North East. It sought to address the associated key infrastructure changes in Information Governance (IG), Information Technology (IT) and legal frameworks necessary for academic and health sector collaborations of this kind.

By building a collaborative approach, this project sought to bring academic, statistical expertise, to support the needs and requirements of healthcare professionals in potentially improving the outcomes and effectiveness in care delivery. Through collaboration with practice managers, NHS analysts and IG specialists, practical modelling, planning and decision-support tools have been built and made available to key NHS stakeholders across the North East of England.

3 Project Overview

This project successfully delivered the targeted outcomes

- The IT and IG infrastructures necessary to enable the secure data transfer and storage of datasets with several regional NHS Trusts and GP practices.
- Identifying the key data fields to develop effective working analytical predictive models for use by NHS Foundation Trusts to model unplanned care. In some cases apps for predicting A&E flow were created for individual Trusts as a proof of concept.
- The same exercise was conducted in a planned care setting. A GP focussed app based planning tool to support frontline practices, councils and planners to scenario plan service delivery as well as facilitating discussions on changing how services are designed to meet population changes and housebuilding development plans.

4. Project Aims

The main aims of this project were to:

1) set up the foundations necessary to facilitate data and information sharing between the local NHS trusts, primary care organisations, local authorities, and universities, and health economy planners

2) conduct modelling exercises to build decision support tools to predict flows and changes in planned and unplanned care and their associated costs and the potential changes needed in care delivery models.

5. Background exemplar work

This project was informed by exemplar work, which took place during 2015/16.

The exemplar work was developed with County Durham and Darlington NHS Foundation Trust (CDDFT), streamlining data sharing agreements, collaboration agreements, and other relevant tools for information sharing. The main analytic aim of the first exemplar project was to assess the predictability of daily A&E flow based on historical data for Darlington Memorial Hospital and, the University Hospital of North Durham. Statistical models were constructed for each hospital and their predictability assessed. While the analytical stage was successful, the core and most challenging part of the exemplar project was to identify suitable mechanisms for data and information sharing with the University and the Trusts.

This project sought to extend and further test the modelling approach to partner Trusts, as well as developing the infrastructure necessary to support and sustain collaborations going forward.

The project team also learned from their work on a Healthy New Towns (HNT) project in Bicester, where the GP practice app was initially developed and some of the issues around operational deployment were raised and addressed.

6. Aim 1: Data Sharing Foundations

Under this aim, the project sought to address three main challenges:

1) Information Governance (IG)

to verify that the correct information governance structure is in place in each institution and create a framework that can be reproduced in other institutions to ensure compliance with the Department of Health guidelines;

2) Information Technology Infrastructure

to identify the necessary IT infrastructure necessary for data sharing with one or more partners, and to ensure the solution chosen can be implemented by other Higher Education Institutions (HEIs);

3) Contractual Framework

to collate and produce a collaboration agreement pack for information and data sharing across multiple institutions that is compliant with the Information Sharing Gateway.

6.1 Information Governance

Key Principles in developing the Information Governance approach

Data Controllers: NHS trusts / primary care providers/ local authorities would remain data controllers of any data that is sent by them to Durham University.

Data Processors: The Durham University (DU) team are acting as Data Processors (DPs). The third party platform providers would be acting as subcontractor to the Data Processors.

Framework: The regional information sharing agreement framework would be the basis of the procedures used to demonstrate compliance with IG Standards. The project team would work with IG leads at NHS Trusts to identify and implement the most appropriate governance arrangements throughout the project.

Assurance: Due diligence processes undertaken by the DPs on the third party secure platform provider would be documented and incorporated into the contractual documentation, to give assurance to the Data Controller (DC).

Disclosure Control: No patient identifiable data was used during this work. In cases where pseudonymised datasets were used, the key for this was never passed to the project team. Additional datasets linked in to the health data were publically available datasets, which did not pose a risk of identifying patient records.

Legislation: Documentation would be updated as required by changes to legislation during the project (e.g. General Data Protection Regulations (GDPR), UK Data Protection Act 2018).

Outcomes: Documents developed would be shared across the CHC project and academic institutions to begin standardising data sharing approaches in the NENC area.

Overview

Preliminary work was carried out as proof of concept under Collaboration Agreements with Caldicott approval forms, using pseudonymised datasets. The progression of the work, through testing prototypes and refining models, as well as the changing legislative environment meant that the necessary documentation changed during the life of this project. Information Sharing Agreements were developed with the CDDFT IG lead to replace the Caldicott agreements.

With preliminary work, pseudonymised datasets were used in order to test and refine the modelling tools. Once the tools had been developed, it became possible to develop and run these with aggregated datasets, simplifying the approval process. The data storage processes and standards remained the same.

6.2 Information Technology Infrastructure

Key Principles in developing the information technology infrastructure

All analytical support platform infrastructure developed would be

a. ISO 27001 compliant
b. IG tool kit compliant
c. Capable of using secure data transfer protocols and established HCSN (former N3) connections

The testing and prototyping of the IT infrastructure would inform the specification of a future regional trusted research environment / trusted analytical platform

Overview

To ensure data was kept within an HCSN/N3 (or equivalent) environment, it was agreed that a secure infrastructure was key to the development of this work. In particular rather than pursuing a traditional University 'safe haven' approach, it was agreed to develop an analytical platform with an accredited NHS data management service provider. The use of an accredited NHS data management service would

- give assurance to the NHS Trusts on the security of the platform, and its continual compliance with national data standards
- enhance opportunities for secure cross-institutional working
- allow the team to explore a sustainable delivery model, informing regional developments.

The University procurement process was followed to find a GCloud approved data management service and a Trusted Research Platform specification was drafted which included accreditation, technical and service expectations for a solution. This can be found at Appendix A

AIMES were identified as a suitable supplier and a GCloud contract was agreed for their 'Infrastructure as a Service' product. This contract provided a key legal and governance link between the University and AIMES Ltd, which would become a key element of the due diligence requirements of the NHS Trusts and primary care providers.

Once the system was procured, the University project team worked with the supplier in order to commission the agreed technical specification and operationally test this solution. This has involved working with the supplier to

- find a solution to enable an NHS Trust to upload data directly to the TRE (e.g. via Secure File Transfer Protocol (SFTP) or equivalent), rather than by NHS email,
- develop the environment to incorporate the necessary analytical software,
- develop a standard operating procedure to give the required assurance to enable approved collaborators to potentially access the Durham University TRE for future joint analytical work.

6.3 Contractual Framework

Initially each trust operated a single collaboration agreement with the University. As the programme developed, a subcontract was subsequently signed between the University and multiple NHS Foundation Trusts to replace the earlier collaboration agreements. These contracts included payments to the NHS Trusts for their support with data extraction, testing and feedback on models as well as IG advice, recognising that the Trusts' data systems were funded to support care delivery and contract management, they are not resourced or paid to extract data for research.

Durham University also set up a contract with AIMES to deliver the IT infrastructure.

7. Aim 2: Build decision support tools to predict flows in urgent care and planned primary care and their associated costs

Work was focused in two distinct areas; one modelling approach was developed for Accident and Emergency (A&E) Admissions at key acute Hospitals in the region 'Unplanned care', and a second for Practice Managers and GPs within General Practice settings 'Planned care'.

7.1 Unplanned Care

A&E admissions data was received from several regional NHS Trusts. This was analysed by the statisticians to produce a meaningful predictive model of future A&E attendances. This model was developed in consultation with the NHS Trust analysts to ensure it addressed practical needs in planning and could be easily installed and used on site.

As a learning model, this was developed using historical data (usually 5+ years where available). Further developments are dependent on the completeness and quality of the data held within Trusts. To note, as a learning system approach, it was recognised the algorithms and models that underpinned the model would need to be recalibrated periodically to pick up any trends or changes in policy, work processes or population changes.

This was useful in detecting patterns and peaks in demand – which could be used to support or disprove anecdotal beliefs on service demand.

7.2 Planned Care

Darlington was a pathfinder site in a national NHSE initiative, 'Healthy New Towns'[1], which reviewed health provision and design in sites where extensive housing developments were planned. As part of the work around that programme, our team developed a planning support tool that could allow demand and activity changes in population workforce and house building programs to be modelled. This planning tool was developed in partnership with several GP practices and local authorities at 2 sites under the Healthy New Towns programme. It allows managers to model different scenarios and ways of working. As well as modelling the impact of large scale house building on GP practices, managers can also model the impact of GP retirements, changes to appointment lengths/ number of sessions and look at working on a hub basis to scenario plan group delivery options. It also allows practices to get visual representations of key targets. This has the potential to lead to efficiencies in delivery.

The model was built using a core dataset, with at least 12 months of aggregated appointment data. The data specification can be found at Appendix B.

Further detail on this app is given as a case study in Appendix C

[1] https://www.england.nhs.uk/ourwork/innovation/healthy-new-towns/

Potential future developments

With improvement recording of data, there is the potential to model different staffing models, understanding appointments by nature and use episodes of care to design the workforce to best support patient needs (e.g. the balance of GPs, nurses and Allied Health Professionals (AHPs).

We continue to work with key GP practices to make the app even more relevant to their needs. This includes incorporation of new key areas as they become more prominent (e.g. frailty).

8. Barriers to delivery

The project highlighted some of the practicalities in time sensitive work between NHS Trusts and universities (Appendix D). A key learning point from this project has been the delivery impact that occurs from **administrative process delays**. Delays in sign off of documentation were not due to contentious issues, but rather to institutional processes. Given the number of institutional and legislative changes taking place, the delay in formalising agreements has at times necessitated further changes. In the period between agreement and signature, new protocols and systems are introduced, and existing standards and frameworks become obsolete.

Due to the exemplar nature of this work, the process of contract and ISA sign off was more challenging, and iterative. It highlighted some resource pressures as well as areas for improvements. Given these challenges, and in order to take better advantage of time sensitive opportunities, a more responsive, streamlined approach is needed.

Recommendation: agreement and accessibility of generic/template documentation for work of this nature (as far as practical) would reduce the need for extensive institutional drafting or review each time.

An additional barrier was the **changing legislative environment** (General Data Protection Regulation (GDPR) and Data Protection Act (2018) implementation). As well as leading to an extensive review of the governance documentation, these changes also delayed the project due to the uncertainty they generated in the IG community. With unclear definitions, the legal basis for processing data was debated with NHS Trust staff in the context of this work, as well as regionally. A consent based model was not considered practical for this work but as the community gained familiarity and understanding of GDPR, and with ICO guidance at a CHC event, it became clear that there were other appropriate legal bases for processing.

Furthermore, this new legislation added significant challenges to NHS Trusts in the conduct of their usual work and this exacerbated delays to the project as staff were understandably focused on institutional core business. While the specific funding of NHS IG expertise for this project certainly aided our ability to deliver this, the impact was still notable.

Attendance at CHC events in Manchester showed that this had been a challenge to project teams across the programme. While it is accepted that the external environment cannot be controlled, and that this work is exemplar by nature, lessons have been learned nonetheless.

Recommendation: Future projects would however benefit from the provision of regional expertise from an organisation like AHSN or similar. It would allow for concentration of expertise in health data research and reduce dependency on individual organisations. This would avoid duplication of effort regionally, mitigate impact of changes and thereby minimise delay.

Recommendation: Future projects should ensure that there is funding allocated for NHS Trusts Information Governance Support, which proved invaluable during this project.

Recommendation: CHC showcasing events were very useful in sharing the work that was going on between project teams. Future projects within a programme may benefit from coming together earlier in the programme to raise awareness of crossover and shared problems – which would improve collaborative working, problem solving and lead to efficiencies. It could also potentially generate further collaborations and grant applications.

Linking datasets – it had been hoped to link the regional A&E datasets to other NHS Trusts to develop a regional model to understand patient flow across the region, recognising the fact that demand for unplanned care at specific sites can be impacted by other factors, such as regional service demand and ambulance diverts. Looking at hospitals or Trusts in isolation did not allow us to model this patient flow and it would be a potential development opportunity to look at this regional picture.

Moving from a distinct Trust-based model to a multi-institutional regional one changed the basis of data processing – it was perceived to be more research-focused. While such a model could potentially improve service delivery and efficiency for these Trusts, given that the NHS is not one institution, there was not the demand for this development and, without internal champions, this area of approval and connectivity was not pursued.

The project showed that creating a model which pulled together data from multiple institutions was technically possible with the Healthy New Towns work in Darlington. Having received and tested practice level models, a 'hub' level model was developed and installed allowing managers to get an understanding of service planning at a cluster level. The technical functionality is set up to allow this to be extended to

This success shows that it is technically possible to create a regional model, although the governance around linking A&E datasets was beyond the feasibility of this project.

Recommendation: Potential future work to develop a regional picture would need an NHS/external organisation to bring together support and champion the benefits. GNCR developments could facilitate work of this type for future projects, providing a platform and establishing a trusted process.

Changes within the NHS A number of changes took place during this project which impacted on the timescales and requirements. These included a change of infrastructure (N3 to HSCN) and accreditation processes (suspension of the IG Toolkit and subsequent introduction of the NHS Data Security and Protection Toolkit) as well as organisational changes with potential mergers of NHS Trusts and NHS digital developments.

Due to the dynamic nature of NHS organisations, potential collaborations were affected by organisational and staffing restructures. In practice, the loss of key staff reduced the

engagement with the project and organisation changes on a wider scale reduced the visibility of the project. Without champions within the institutions, and with so much taking place, partnerships faltered and timescales slipped in some instances.

Recommendation: Working with a consultant who knows the NHS organisations, has the appropriate contacts and manages the links between the project team and collaborative partners from an external point of view has been key for this project. Future projects may well wish to consider replicating this model.

Recommendation: This project was funded to pay NHS Trusts for analyst time, which helped to keep the project on track. There was no funding to pay for datasets. Given the cost to NHS Trusts of collecting and maintaining this data, future projects may need to factor in data costs to the budget.

Variation in datasets and operating systems The prototype models were designed using some core datasets. While these varied in content and quality, the team has been able to identify the minimum datasets required for functionality. An example is given at Appendix B. Discussions with GP practices about potential functionality improvements have highlighted analysis which practices would find useful but which are not possible given that the data is not currently captured at source. Understanding the analysis that is possible with this data may help to improve data recording at practice level.

9. Limitations

The proposed pathway looking at Dementia/Frail elderly data and inclusion of social care/local authority was not feasible due to access to the necessary data. This project focused on Planned care as an alternative pathway.

The proposed regional level work was not feasible in the timeframe. Although technical solutions to regional pseudonymisation were identified, challenges remain in securing data controller agreement to release identifiable data for work of this nature which sites outside of direct patient care.

While the project gained additional partners through the course of the work, there was also the loss of engagement of one Trust. We believe that this is due to organisational and personnel changes and pressures rather than project-related matters as this happened early in the programme of work in the context of changes at the Trust involved.

10. Sustainability and Scalability

As the project progressed, expressions of interest were received from other NHS Trusts, GP practices and groups, as well as local authorities. In some cases this lead to the signing of agreements and the production of analytical models. This highlights that the outcomes of this project are successful and that there is diverse interest across the region in these. As such, the team is keen to ensure their sustainability.

An unsupported version, free to NHS organisations, which practices can plugin their own data to will be made available. A potential platform for this would be the Great North Care Record platform.

While this model designed to be as user friendly as possible, it is understood that some practices may seek a supported or bespoke solution. It has become clear that the free-to-NHS nature of this app is attractive to practices who may not have the resources or will to pay for a supported model. The development team is not resourced to offer this support, nor is it part of the core business of an academic institution. It is also apparent that there is commercial interest, in both the product and algorithms behind it, as well as in selling this as a supported tool.

In light of this, we have explored a number of options. In order to protect the resource as a free to NHS tool, trademarking applications are underway.

Discussions have been held with a number of commercial providers who could potentially host this and offer user support. This may be at a cost to practices, or at limited scale via alternative grant applications and funding. No agreements are in place to date and the feasibility of this option has not been fully evaluated.

Furthermore, this project has led to further work, and the award of additional grant funding to continue to collaborate with one of the Trusts on modelling within the healthcare setting, which shows that the fundamental bases of collaborative work have been successfully established.

11. Conclusion

This project has included a consideration of governance and what good governance looks like in relation to such a patient data related project. Some of the key learning in terms of good governance is detailed in the Executive Summary (p4.-p5).

This project also shone a light upon some key enablers of research project effectiveness. For example, the inclusion of a consultant to the project with both knowledge and experience of the regional NHS organisational and political landscape proved critical in enabling the delivery of the project outcomes. The importance of being able to draw upon the support and good counsel of the regional IG expertise proved invaluable too. These two observations perhaps underscore the critical importance of the relationships of trust which the project clearly achieved mutual benefits from.

The key outcomes from the project are listed in the Executive Summary (1.1).

Appendices

Appendix A: TRE Specification

Durham University

Connected Health Cities Health Analytics Trusted Research Platform

Introduction and Requirements

Durham University require a secure and trusted environment in which to process and analyse data relating to attendance in Urgent and Emergency Care in hospitals, walk-in centres and GP practices. The project will produce statistical models for use in producing forecasts.

The data handled initially will be anonymised, however as work progresses the intention will be to move to pseudonymised and potentially patient identifiable data.

The key challenges in providing a safe haven for the building and testing of the models is in providing an environment that has the appropriate controls in place to protect the data and assure all partners that their data will be properly looked after.

Approach

The North East and North Cumbria Connected Health Cities programme is looking to establish a regional health data sharing and analytics platform, however given the likely the timescales for the delivery of this, Durham will establish their own shared platform limited to the purposes of the Durham project. This may involve interaction with other partner universities. This will provide a proof of concept and provide an opportunity for learning for the wider regional platform.

Durham University are working as an exemplar within the Connected Health Cities programme and as such would expect to share the design patterns and implementation approaches for information sharing.

Key Features Required

Ability to access analytics platform through a secure connection which is not located on the N3 (or other secure) network. This can be from a location that is physically secure and has undergone a physical security risk assessment.

Ability to receive data from partners over the N3 network with the data being held within the N3 environment. Only pseudonymised or anonymised information will be held or used within the analytics environment.

The solution must be capable of partners on N3 securely accessing analytical data as results. This may not be a day 1 requirement.

Support for the following software:

- R, RStudio, Shiny and R Packages
- Microsoft Office (Excel and PowerPoint)
- Microsoft SQL Server
- QlikView
- Python

As part of the process Durham University will be specifying the interoperability standards that will be used however at this point the following communication mechanisms will be required:

- SFTP
- HTTPS

Solution Architecture

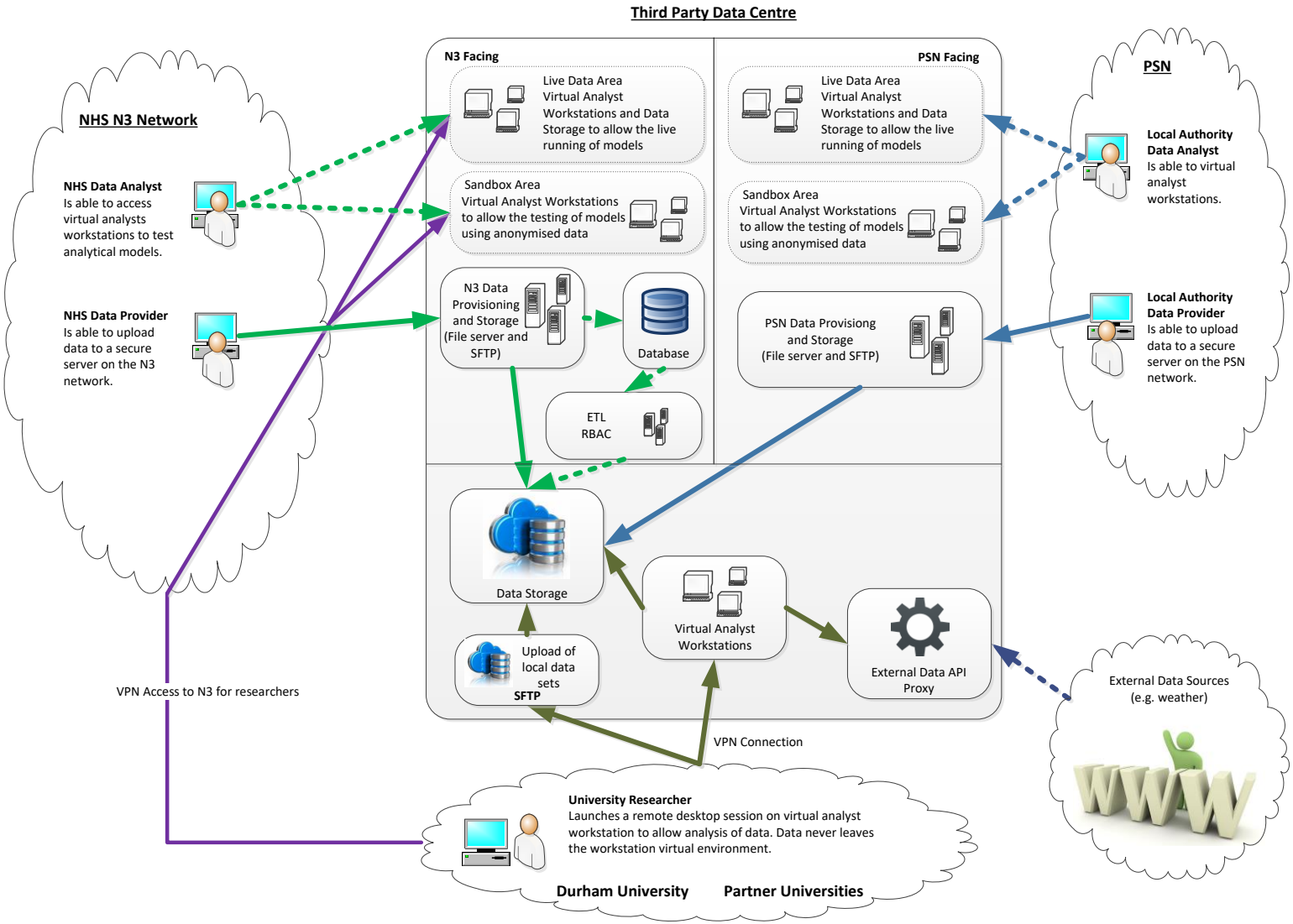
The solution architecture is split into a potential final solution which encompasses the full range of functionality that may be required. It is anticipated that additional funding would need to be secured prior to this full solution being put in place, however it is useful to have as target operating model to ensure that the initial solution and any subsequent amendments are compatible with the desired end point. It is however likely that this end point will be reviewed and amended over time.

As the research broadens, more information will be required from different sources. This includes secure local authority datasets that may be sourced from PSN based systems or alike rather than N3. The solution should be adaptable to deal with this sort of changed requirement. It is likely that the information sharing will also need to encompass linking with commissioning support units.

The end point solution is defined pictorially below

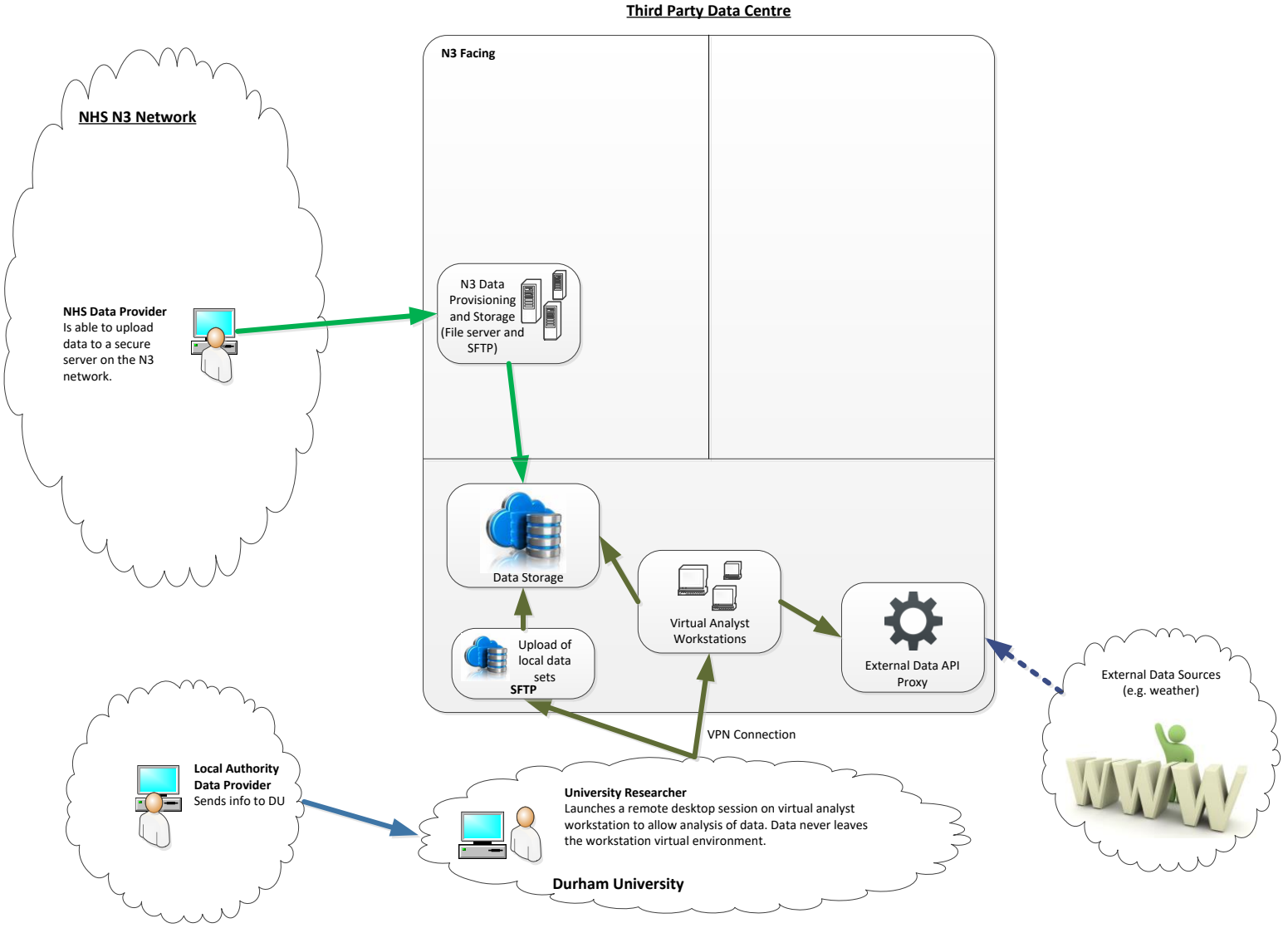
Appendix A– TRE specification

Target Logical Operating Architecture



Appendix A– TRE specification

Initial Target Operating Architecture



Supplier Requirements

The supplier is required to be:

- ISO9001 Accredited
- ISO27001 Accredited
- NHS Digital Information Governance Toolkit Level 2 or higher compliant.
- NHS N3 Accredited and an N3 aggregator

The data centre used to host the service must be within England and must be accredited at with a recognised data centre accreditation body. Data must be stored in England at all times.

The supplier should be experienced in the provision of trusted research platforms to the health research sector, having domain knowledge of the requirements for securing research data and providing appropriate governance.

Detailed Requirements

Service Level Agreements

The service shall be available 24x7 with an uptime expectancy of the 97% per month.

Availability of service and Disaster Recovery Estimated RTO and RPO

It is anticipated that the maximum RTO be 2 hours.

The RPO should be to a previous day's backup.

Pricing model and capacity for expansion

The pricing model should allow for expansion – initially there will be only a small number of organisations using the service, this however could expand quite rapidly.

Data Protection

Data Segregation (within our environment and between our environment and other customers) shall be delivered using logical segregation.

Authentication of users (on N3 and gaining remote access to analytics platform).

Exit strategy.

A clear exit strategy will be required to allow the University to obtain any non-personal information from the servers.

Environment Specification

The following specifications should be used to provide the anticipated cost of the service. At this stage this is the best estimate of the likely usage of the service. It is anticipated that the service will be procured as service credits which can then be used against any service within the scope of the Trusted Research Environment. This is to facilitate agility in the use of the service.

Analytics Zone Management Services

Item	Detail	QTY	Months
Domain Controller	ACVS.B2 2CPU 4GB RAM 120 HDD (premium) SVR 2016 – Weekly Back up	1	24
RRAS Server	ACVS.B2 2CPU 4GB RAM 120 HDD (premium) SVR 2016 – Weekly Back up	1	24
File Server	ACVS.B2 2CPU 4GB RAM 120 HDD (premium) SVR 2016 – Weekly Back up	1	24
Shared File Service	500GB Additional Disk Premium	1	24

Analytics Environment

Item	Detail	QTY	Months
VDI D plus 500GB	8CPU 32GB RAM 120 HDD (premium) SVR 2016 – Additional 500GB (Premium)	2	24
MS Office	OfficeProPlus ALNG LicSAPk MVL SAL	2	24
VDI B plus 1TB	4CPU 16GB RAM 120 HDD (premium) SVR 2016 – Additional 1TB (Premium)	2	14
MS Office	OfficeProPlus ALNG LicSAPk MVL SAL	2	14

Data Provisioning Management Zone

Item	Detail	QTY	Months
Domain Controller	2CPU 4GB RAM 120 HDD (premium) SVR 2016	1	12
SFTP Server	2CPU 4GB RAM 120 HDD (premium) SVR 2016 – 500GB Additional Disk	1	24
SQL for ETL	4CPU 8GB RAM 120 HDD (premium) SVR 2016 – 500GB Additional disk	1	12

Data Provisioning Environment

Item	Detail	QTY	Months
Operational Management	VDI – DATA – ETL – MGMT – Set up - AD	1	24
Bronze support	STD SLA – 30 Hours per year engineer support Mon-Fri 9-5	1	24

Procurement Approach

It is intended that the procurement of the service will be done using the UK Government Digital Marketplace. The service will be purchased based on an upfront purchase of cloud credits which will then be used over a 24 month period.

Appendix B: Example Data Specification

File 1: (ONS/NHS England or Practice) List size input/local demographics
Filename: darl_demo.csv
Variables: AgeGroup, Practice, Gender, Total
AgeGroup: 0 to 9, 10 to 19, ..., 70 to 79, 80 plus
Practice: Practice names – A,B,C
Gender: Male, Female
Total: Counts

File 2: Appointment Rates (Practice)
Filename: darl_monthly.csv
Variables: Year, Gender, AgeGroup, MonthName, AppsForOne, AppsForOne_nosat, AppsTotal, AppsTotal_nosat, Practice
Year: Year number, at least one full year of data for calibration needed
Gender: Male, Female
AgeGroup: 0 to 9, 10 to 19, ..., 70 to 79, 80 plus
MonthName: Shortened Month name (e.g. Jan, Feb, Mar)
AppsForOne: Average number of appointments per one person (positive float, often range between 0 and 2)
AppsForOne_nosat: Same as AppsForOne excluding Saturdays (positive float, often range between 0 and 2)
AppsTotal: Total number of appointments for given subset. Count data.
AppsTotal_nosat: Same as AppsTotal excluding Saturdays. Count data.
Practice: Practice names – A,B,C

File 3: (Practice)
Filename: apps_for_100.csv (Practice)
Variables: Gender, AgeGroup, RotaType, AppsFor100, AppsFor100_nosat, UsingYear, Practice
Gender: Female, Male
AgeGroup: 0 to 9, 10 to 19, ..., 70 to 79, 80 plus
RotaType: Staff member by role (Consultant, Gp, HCA, Nurse, Nurse Practitioner, etc)
AppsFor100: Average number of appointments per 100 people (0.1 to 483)
AppsFor100_nosat: Same as AppsForOne excluding Saturdays (0.1 to 483)
UsingYear: All
Practice: Practice A, B, C

File 4:
Filename: darlington_proj_ons.csv (ONS)
Variables: AgeGroup, Gender, AREA, Popn, Year
AgeGroup: 0 to 9, 10 to 19, ..., 70 to 79, 80 plus, All
Gender: Female, Male
Area: Darlington
Popn: 0.7 to 2.24
Year: 2016, 2017...., 2039

File 5: (Practice)
Filename: ltc_app2cond.csv
Variables: _, Gender, AgeGroup, Condition 1, Condition 2, RotaType, AppsOne_only, AppsOne_mult, Site, Group1, Group2
_: row number (1-10000)
Gender: Female, Male
AgeGroup: 0 to 9, 10 to 19, ..., 70 to 79, 80 plus
Condition1: Patient first LTC (Asthma, COPD etc)
Condition2: Second LTC (as above)
RotaType: Staff member by role (Consultant, Gp, HCA, Nurse, Nurse Practitioner, etc)
AppsOne_only: Average number of appointments per one person (positive float, range between 0 and 34)
AppsOne_mult
Site: Practice A, B
Group1: LTC group for condition1 (Cardiovascular, Mental Health, Respiratory, None)
Group2: LTC group for condition2 (Cardiovascular, Mental Health, Respiratory, None)

File 6: (Practice)
Filename: ltc_totals2cond.csv
Variables: Gender, AgeGroup, Condition1, Condition2, NCond_only, CondOnly_for100, NCond_mult, CondMult_for 100, Site, Group 1, Group 2
Gender: Female, Male
AgeGroup: 0 to 9, 10 to 19, ..., 70 to 79, 80 plus
Condition1: Patient first LTC (Asthma, COPD etc)
Condition2: Second LTC (as above)
NCond_only:
CondOnly_for100:
NCond_mult:
CondMult_for100:
Site: Practice A, C
Group1: LTC group for condition1 (Cardiovascular, Mental Health, Respiratory, None)
Group2: LTC group for condition2 (Cardiovascular, Mental Health, Respiratory, None)

File 7: (ONS, estimate, developer)
Filename: new_households.csv
AgeGroup: 0 to 9, 10 to 19, ..., 70 to 79, 80 plus, All
Practice: Practice A, B, C
perc.household: percentage

Appendix C: Planned Care App

Data informed General Practice patient services planning

The app

The Connected Health Cities (CHC) team at Durham University has developed an app that enables General Practitioner (GP) managers to analyse patterns in their data. In particular, it allows managers and clinical teams to view predictions of future GP practice activity based on current data and on interactively specified housing growth and staff planning scenarios. Users can also choose to group practices together to form 'hubs', to investigate the ramifications of such strategies.

The app has been created using Shiny (Chang, Cheng, Allaire, Xie, & McPherson, 2018) a package for building dashboards in the R language (R Core Team, 2013). By using Shiny, we have been able to develop a user-friendly interface using familiar objects such as sliders, checkboxes and drop-down menus. This software is being developed in conjunction with GPs and Practice Managers (PMs) from Bicester, Darlington and Bishop Auckland, each of whom have been provided with a working app from the Durham University research team. The app allows users to better understand current levels of activity and to investigate how different scenarios are likely to affect their practice. For example, one of the practice managers involved indicated surprise at the number of GP and nurse appointments allocated to women aged in their thirties – more than to any other age group – and that this had prompted her to think further about whether or not any clinics could be set up to cater to this demand more specifically and thus potentially more efficiently and effectively.

Population forecast

The app requires several datasets. The office for national statistics produces regional population forecasts for the next 20 years, split by gender and in age bands of 10 years. These are combined with user-specified housing development scenarios (see Figure 1) and projections of the demographic build-up of new residents to produce a detailed practice list projection. In the example in Figure 2, it is clear that young families are expected to account for a significant proportion of the new householders, with those aged 0-9 years and 30-39 years showing the greatest increases. The model includes a 'new patient factor', reflecting the extra workload a patient incurs in their first year with a practice.

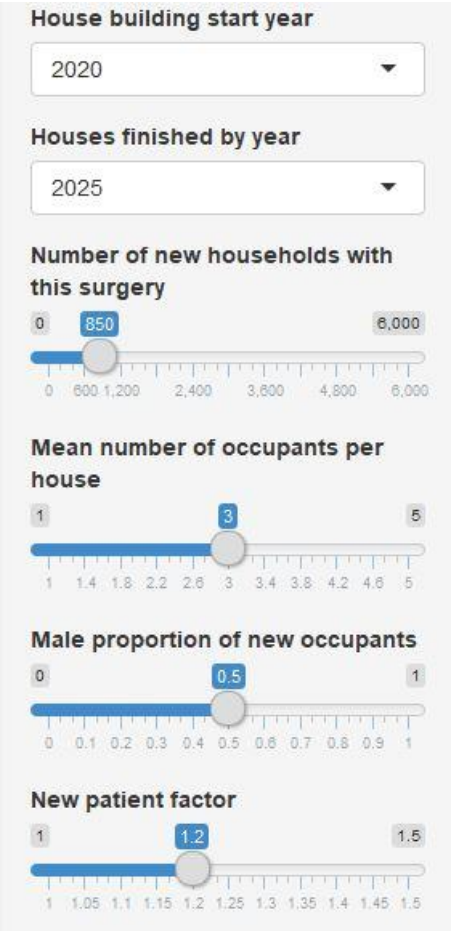


Figure 1: The app inputs by which a user specifies a housing scenario.

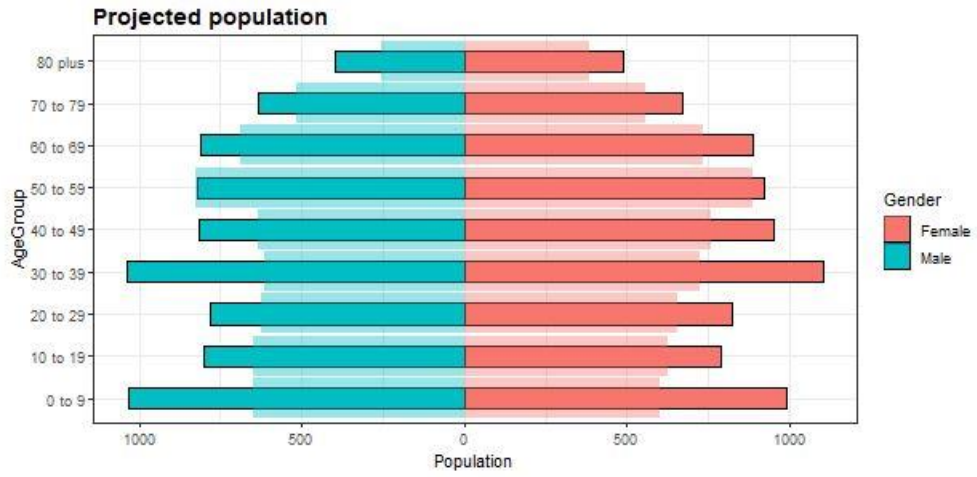


Figure 2: A population projection plot for a given year. The projected population is shown by the narrower, darker, outlined bars. For comparison, the current population is shown by feinter, wider bars.

The GP practices were especially interested in anticipating changes to the numbers of patients with long term conditions (LTCs). Using their practice QOF long term conditions registers we were able to add this functionality, using the prevalence of each LTC within each demographic group with the population forecast to produce forecasts specific to each LTC or group of LTCs (the general groups chosen were Cardiovascular, Respiratory and

Mental Health). Figure 3 shows the projected population of people with a Cardiovascular disease (this includes LTCs such as Stroke, hypertension and chronic heart disease).

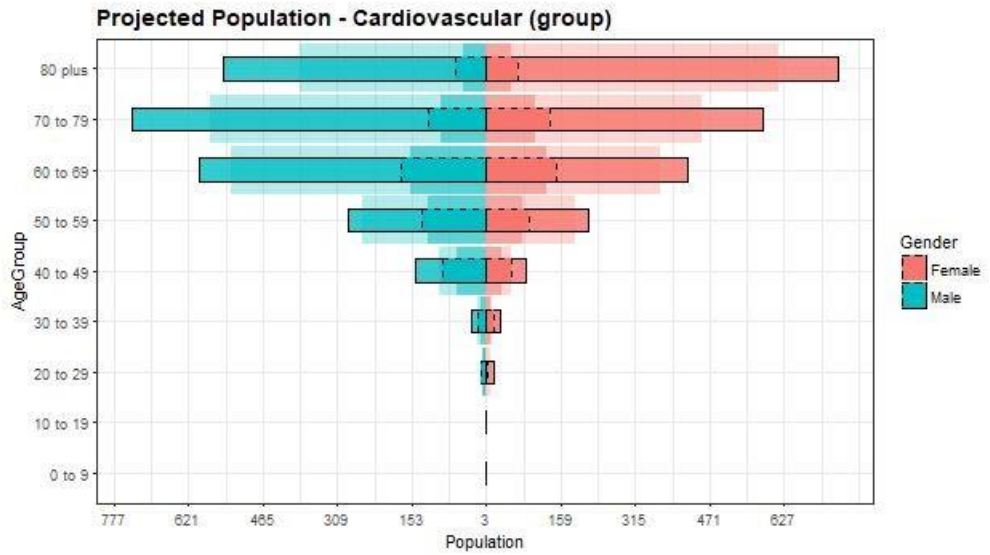


Figure 3: An LTC population projection for a given year for patients with cardiovascular disease. The projected population is shown by the narrower, darker, outlined bars. For comparison, the current population is shown by feinter, wider bars. The dashed lines show those patients with just a cardiovascular condition and the rest of each bar represents those patients with at least one other LTC.

The app allows the user to choose a second LTC, and will then display results for patients with both conditions. Understanding the prevalence of combinations of conditions is of particular interest for pharmacists, who would be concerned about medication clashes for patients with multiple LTCs.

Appointment data

As well as population data, for each practice we use appointment data covering at least a full year. This allows us to investigate rates of appointments within each demographic or LTC group, and to produce forecasts of practice activity. Where a hub scenario is created, the app aggregates the data from the practices involved.

By combining the demographic data (ie. Everyone registered with each practice, whether or not they have had any appointments) with this appointment data we can show the rates of appointments for each type of practitioner, as in Figure 4.

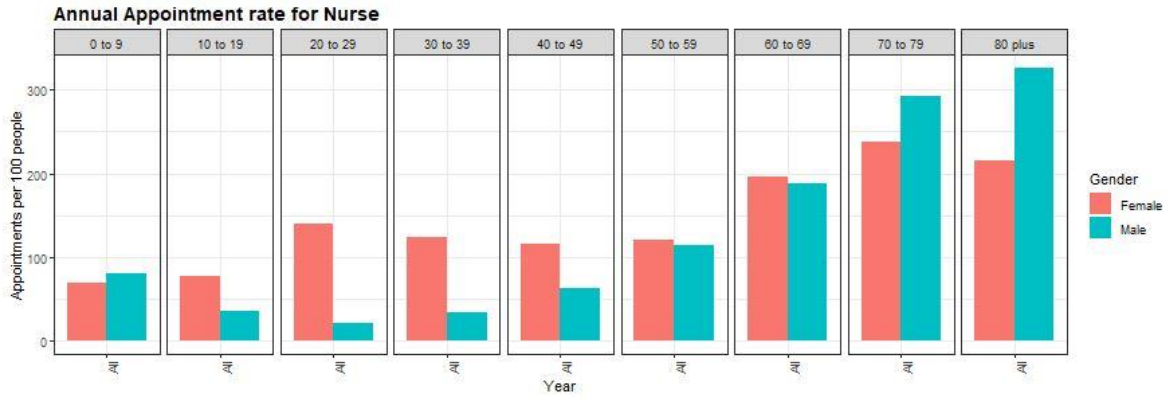


Figure 4: Nurse appointments per 100 people of each demographic group.

The forecasted population can then be used with these appointment rates to forecast the annual number of appointments with each practitioner type, as in Figure 5.

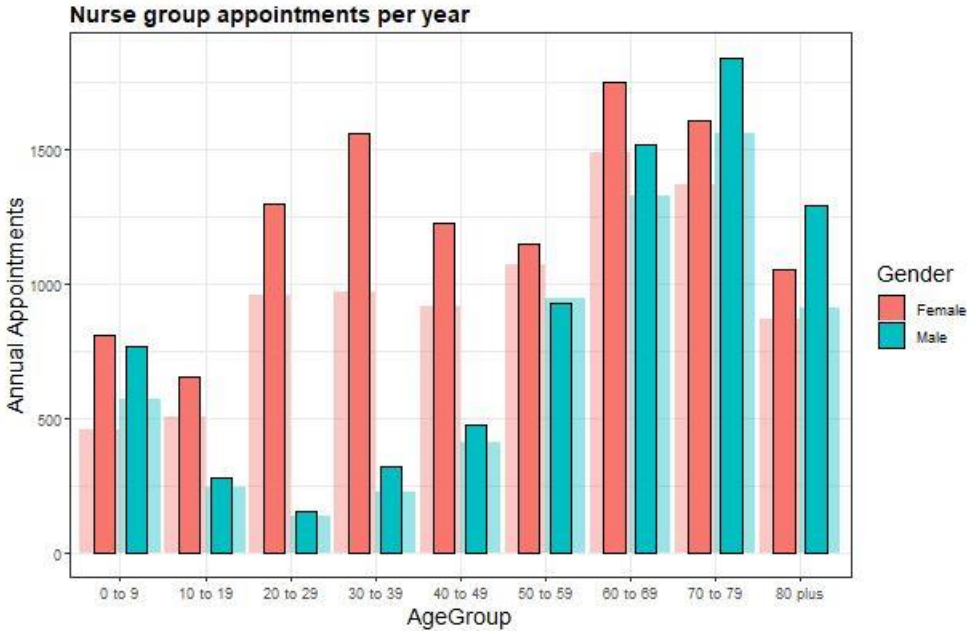


Figure 5: Projected nurse appointments for a given year. The projected population is shown by the narrower, darker, outlined bars. For comparison, the current population is shown by feinter, wider bars

We found that although practice managers had a reasonable instincts for the appointment rates as shown in Figure 4, they were often surprised by these projections. In particular, the number of appointments for females aged 20-50 was often unexpectedly high. This is perhaps surprising since these, rather than the rates of appointments, are the 'raw data' being seen in the practices.

Another area of interest is the rate and number of appointments for the LTC patients. Figure 6 shows an example of the projected number of appointments for patients with dementia. The appointment rate used here is the number of appointments per year for each person with dementia in each age group and gender, rather than the general population rates like those shown in Figure 4.

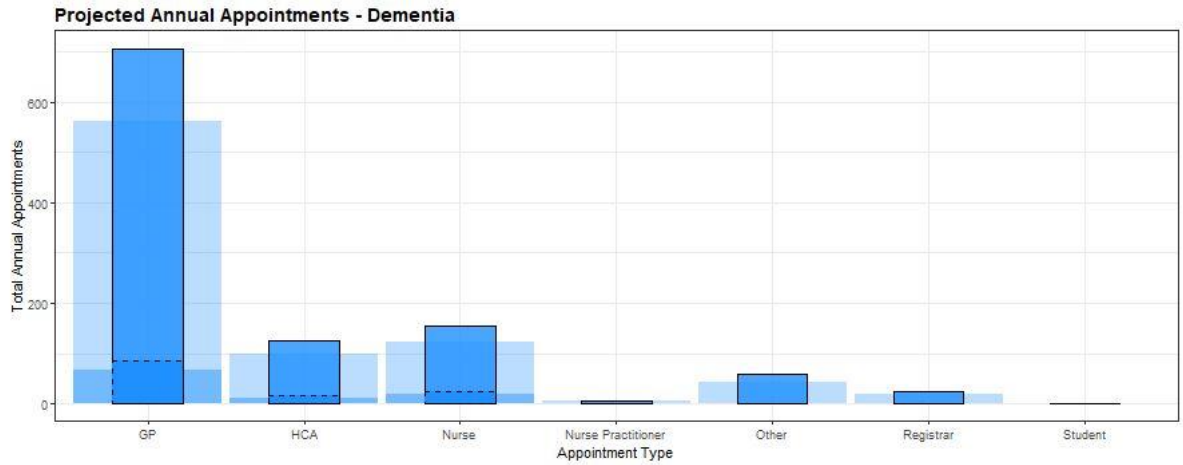


Figure 6: A projection of annual appointments for those patients with dementia, for a given year. The projected population is shown by the narrower, darker, outlined bars. For comparison, the current population is shown by feinter, wider bars. The dashed lines show those patients with dementia, and the rest of each bar represents those patients with at least one other LTC.

In this model, it is assumed that the prevalence of each condition will remain the same within each age and gender group, and that the rate of appointments will remain the same within each condition, age and gender group.

It is worth noting that in some cases the numbers of people with conditions, especially in some of the younger age categories, is very small, and these rates are therefore not as robust as the general population rates. In cases where a patient has multiple LTCs, we are also unable to tell from the appointment data which condition was the cause of the appointment, if any.

Staffing patterns

An area of concern for the practice managers we spoke to was understanding the impact of the projected appointments on their staffing, and the potential cost of any locum cover. In response to this, we introduced a set of inputs for the user to be able to describe their current staffing patterns, in order to see the impact of population changes. Figure 7 shows the inputs used to specify the GP working patterns, and how much a locum GP would cost.

Figure 7: A set of inputs to describe the staffing patterns for GPs.

The app has similar sets of inputs for healthcare assistants (HCAs) and nurses, and produces a table showing the current workload in terms of full time equivalent (FTE) staff, the projected workload for the chosen year, and the cost of covering the difference with locum staff.

Further work / developments

At present we are able to analyse the appointments by the type of practitioner, but a more useful facility would be to be able to see the reason for the appointment, or what took place, eg. a chest infection, or a wound dressing. The reason for this is that this information is not often recorded by the practices with whom we have worked, and so our recommendation is that surgeries begin recording the reason for each appointment. With this facility, it would be simpler to plan specialist clinics, eg. for wound dressings, or to see where care is being provided by inappropriate staff.

Appendix D: Learning

Learning from the past

Critical success factors

- **Collaboration** with H&SC organisations on **priority** needs
- Engagement with **information governance** specialists at the start of Research and innovation
- Understanding **skills** and **roles** in each organisation
- **Project management** approach to collaborative work

Critical failure factors

- Lack of understanding of **information** availability and **resources**
- Not adequately **engaging** key **stakeholders**: public, clinicians, service managers
- Competitive not **collaborative** bids for **grants**
- Data **structure** and transfer complex, inconsistent **tags**

CONNECTED HEALTH CITIES

Supporting the Great North Care Record

Managing the present

Critical success factors

- Develop common **systems, language** and **collaboration** frameworks for academics and analysts
- Develop a stakeholder **trusted analytical platform with secure information sharing infrastructure**
- Trial and **share** experience with **new** data **analytics** approaches
- A mechanism to ensure **H&SC share in the economic opportunity their population data can bring to improve regional services**
- Developing **better academic 'back office and management processes** to facilitate collaboration with H&SC

Critical failure factors

- Promoting local organisational focussed approaches that are not **interoperable**, or open standard based
- Not **recognising the cultural differences and performance targets between Academic and H&SC organisations**
- Losing the citizens trust with a significant breach in **information sharing** protocols
- Not establishing effective regional **leadership** and **governance** arrangements
- Bidding for funding opportunities as **universities** not as a **health economy** for funding

CONNECTED HEALTH CITIES

Supporting the Great North Care Record

Creating the future

- **Joint development of** new predictive and **analytical approaches focused on prioritised H&SC problems and opportunities with maximum benefit to regional population**
- **Collaborative** development of **big data** based research grants across **universities** and **healthcare and social care**
- **New approaches to data sharing, managing legal compliance and IG management of health and social care data from specialist** organisations
- Developing a **shared trusted analytical platform** for **e-collaboration** between **universities**, and **health** and **social care** organisations
- Working with health and social care organisations to **assure** and **standardise** information **sharing, systems and** approaches
- **Universities partners in a GNCR infrastructure** that includes key back office co-ordination – legal, IG, comms, duly authorised researcher and analyst role developments
- A mechanism to ensure **H&SC share in the economic opportunity their population data can bring to improve regional services**

CONNECTED HEALTH CITIES

Supporting the Great North Care Record