Public Health England

Protecting and improving the nation's health

Vision for a congenital anomaly and rare disease registration service in England

About Public Health England

Public Health England exists to protect and improve the nation's health and wellbeing, and reduce health inequalities. It does this through advocacy, partnerships, world-class science, knowledge and intelligence, and the delivery of specialist public health services. PHE is an operationally autonomous executive agency of the Department of Health.

Public Health England
Wellington House
133-155 Waterloo Road
London SE1 8UG
Tel: 020 7654 8000
www.gov.uk/phe
Twitter: @PHE_uk
Facebook: www.facebook.com/PublicHealthEngland

Corporate member of
Plain English Campaign
Committed to clearer
communication
339

Contents

Introduction

Public Health England (PHE) has committed to the expansion of congenital anomaly registration in England from the current 49% of births[1] to cover the entire population by the end of 2015. Since 75% of rare diseases affect children[2] and the majority of these are either congenital or hereditary[3], the infrastructure and processes used to collect data on cases that are identified within the first few years of life can be readily expanded to include the remaining late-onset rare diseases.

This paper outlines the design of a single National Congenital Anomaly and Rare Disease Registration Service (NCARDRS) that will collect data on all congenital anomalies and rare diseases in England.

The vision

To create a single comprehensive national congenital anomaly and rare disease registration service that collects and quality-assures data on all rare diseases across the whole population in England as a resource to support individual patients, their families, research, service delivery, commissioning and public health.

Aims and objectives

The aim is to provide a comprehensive national registration service for all congenital anomalies and rare diseases diagnosed and treated in England as a source of data and information to:

- provide continuous epidemiological monitoring of the frequency, nature, cause and outcomes of congenital anomalies and rare diseases for the population of England
- provide a resource for clinicians to support and monitor their clinical practice including establishing systems to collect self-reported information from patients or their parents/carers

---

[1] www.binocar.org
[2] EURORDIS http://www.raredisease.org.uk/
[3] http://www.caris.wales.nhs.uk/rare-diseases-and-congenital-anomalies

- empower patients and their carers, through the provision of personalised information relevant to their disease
- support research into rare diseases and genomic medicine including basic science, diagnostics and therapeutics and public health
- monitor, evaluate and audit health and social care services, including the efficacy and outcomes of screening programmes
- inform the planning and commissioning of health and social care provision for pregnancies and infants affected by congenital anomalies and patients affected by rare diseases

Existing expertise

Much of the thinking for the design of the NCARDRS is based on our experience of creating and managing the National Cancer Registration Service (NCRS) in PHE. The new NCRS in the UK is now one of the largest, most timely and comprehensive cancer registration services anywhere in the world.

Although some uncommon types of cancer are also classified as rare diseases, we recognise that there are differences between cancer data collection and the requirements for a congenital anomay and rare disease registration service. Nevertheless, there are also similarities in the way that data is collected and collated. The ability to share infrastructure, facilities and teams, a robust information governance framework, and capitalise on existing relationships with large numbers of providers across the NHS, together with strong engagement with patients and carers, charities and clinicians potentially allows us to deliver a cost-effective solution. Table 1 compares the two services.

Table 1: Comparison of current NCRS functions and planned functions for the national congenital anomaly and rare disease registration service

Function	NCRS	National Congenital Anomaly and Rare Disease Registration Service
High case ascertainment, collection of numerous data items across the patient pathway	Data feeds from primary care through to palliative care. Lifelong follow up in place with records of significant clinical events from across a range of service augmented by patient information and data.	Data will include antenatal care information through to treatment and follow up.
Ability to handle disparate data feeds from multiple providers, both national and local	12 national feeds

500 local feeds

National registration system in place, ENCORE. | To scope in more detail but currently > 500 local feeds and numerous regional and national feeds. |
| Data liaison and links with local clinicians, laboratories and other NHS services | Federated expert registration teams across England. | Regional congenital anomaly registration teams currently in place in six regions. Additional teams required to ensure national coverage. |
| Quality assurance of the data | By registration staff | Dependent on disease; either by registration staff or clinical experts |
| Standardised single data collection system | ENCORE (English National Cancer Online Registration Environment)

Ensures standard operating procedures and as a national system overcomes problems with regional/local boundaries. | Planned |
| Feedback to clinical teams | Supports the Cancer Outcomes and Services Dataset (COSD), National Cancer Audit and data QA. | Planned |

Patient and public engagement	Through engagement with voluntary sector and the patient information portal pilot. Patient Information material.	Through engagement with voluntary sector and the patient portal. Patient information material.
Information governance (IG) and data security	Compliance with IG guidelines for handling patient identifiable data. Data collected under section 251 of NHS Act 2006.	Current CARs collect data through section 251. Would seek approval to expand coverage.
Data access	Aggregated data made sufficiently de-identified to be publicly available. PHE Office of Data Release manages all requests for data release to ensure IG requirements are met. Secure Safe Haven in place as a secure physical space with controlled access to data.	Aggregated data made sufficiently de-identified to be publicly available. PHE Office of Data Release manages all requests for data release to ensure IG requirements are met. Secure Safe Haven in place as a secure physical space with controlled access to data.

Although related, we make a distinction between a disease register and a disease registration service.

Registers

Population-level registers have traditionally been used in public health as a source of data to support epidemiology and surveillance. Data collection has focused on a small, but high-quality longitudinal dataset for one or more diseases while the continuous follow-up and accretion of information on registered cases has been limited. In general users request snapshots of data from the register to perform analyses on a cohort; time series analysis would usually involve using sequential snapshots. The restricted breadth of the dataset and the relative tardiness of the data collection create registers that although useful to identify cases, mean that it was almost always necessary to collect additional extra data items to answer more detailed specific questions or pursue high-resolution studies.

Registration service

Unlike a register, the registration service aims to collect data along the whole care pathway in near-real time, as a resource that can be used to support many uses. The dataset is much more extensive with an emphasis placed on data linkage from a wide range of sources, where the quality and the provenance of items collected is known. So for example, recognising that a tumour may be correctly staged at diagnosis in a variety of ways, the registration service will record several different data items for one tumour recording whether this stage is based on imaging, pathology, clinical assessment etc. Similarly, no one date will be defined as the date of diagnosis rather the dates of all relevant episodes on a referral pathway will be recorded. The interpretation of the collected data then forms an integral part of the future analysis.

The creation of the single NCRS in England has shown that it is possible to create a population register that is timely, with more comprehensive data collection that can meet the needs of nearly all users. Figure 1 gives an overview of the practical aspects of cancer data collection.

Figure 1: National Cancer Registration Service
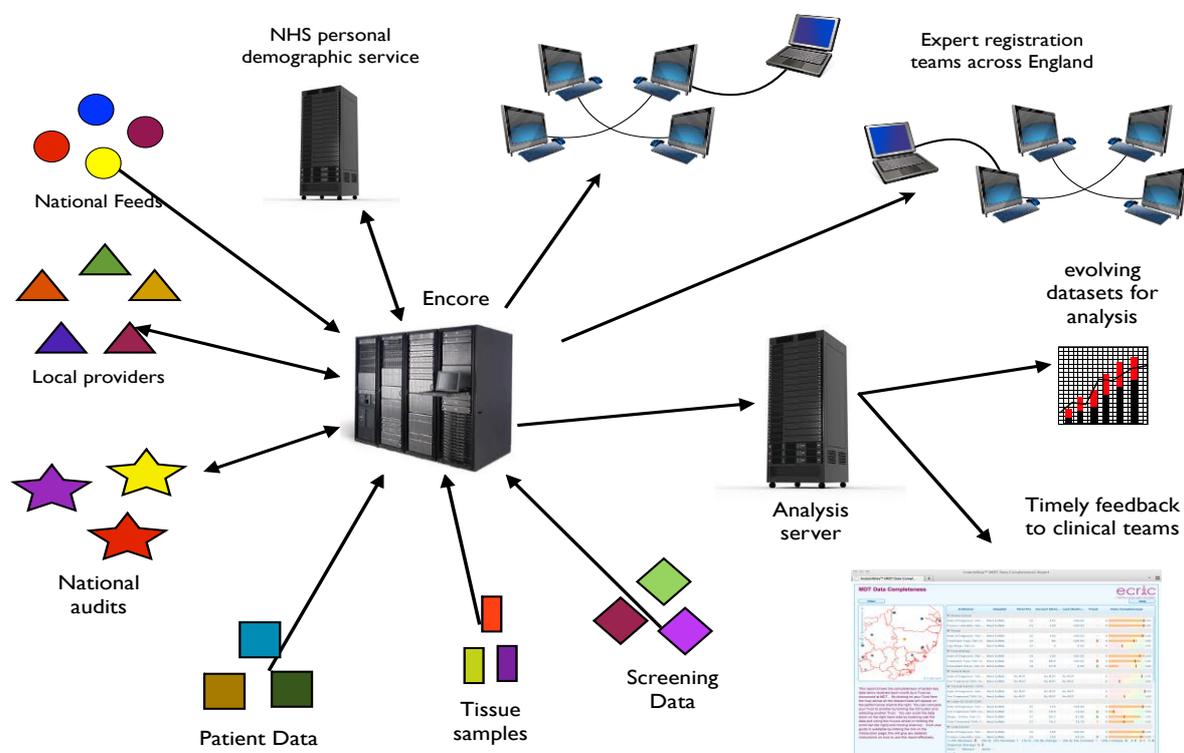
A National Congenital Anomaly and Rare Disease Registration Service

In brief, we outline each of the key components of the service.

Information governance and data security

The NCARDRS will be managed by the National Disease Registration Division in PHE. To ensure that we comply with the information governance requirements for handling patient-identifiable data only specialist, trained registration staff will work on the data and they will be based in the secure physical infrastructure inside PHE.

Keeping information safe is a very high priority for PHE. There are very strict procedures in place to prevent unauthorised access to personal information. These include physically and electronically protected data storage equipment, strict security for disease registration buildings, comprehensive staff training on policies and procedures for the protection of information, and compliance with all relevant laws and policies including the requirements of the Data Protection Act 1998 regarding the receipt, storage and transfer of personal data.

Data sources

Our aim is to achieve high levels of ascertainment and completeness of information by collecting information from multiple sources. Data relating to the same case will be derived from multiple independent sources as this maximises the details available. It also facilitates the continued data collection over the patient's entire life.

To help build this network of data suppliers we will work with a wide range of groups to identify the sources that might be relevant. Within PHE we already have access to data collected by the existing regional congenital anomaly registries, non-cancer screening programmes, the cancer registration service and neonatal infectious disease surveillance service.

Externally, we will work with NHS England and others to secure data feeds from existing registers of rare diseases and testing laboratories, and other relevant data from the Health and Social Care Information Centre, academic and 3rd sector organisations. Through the Department of Health and Office

for National Statistics we will investigate methods to identify non-viable births. Figure 2 illustrates the breadth of data sources available.

Central national processing system

The numerous existing congenital anomaly and rare disease registries all have disparate IT systems – and there is no overarching consistent data dictionary or processing system that can be used. We will therefore develop a single national data collection system, drawing on the lessons learnt from the design of ENCORE, the system that supports cancer registration data collection and quality assurance.

There are estimated to be around 8,000 different rare diseases worldwide. One of the most significant challenges is how to develop a data collection system that can hold data items on so many different diseases and yet allow similarities between cases to be identified and investigated. We recognise that database technologies that allow us to maintain a highly flexible data structure will be used so that specialist data items for individual diseases can be readily incorporated.

Data definitions inevitably vary between diseases but we will establish a minimum mandatory set of patient demographic indicators for all cases, to ensure information is consistent and adheres to the standard data collection requirements adopted across Europe. The European Orphanet coding system and nomenclature will be used for rare diseases and, until the International Classification of Diseases (ICD) 11 is released, ICD 10 will be used for congenital anomalies.

Figure 2: Main Data Sources for the NRDRS

Key: PINK local data/ GREEN regional data/ YELLOW national data/ORANGE research studies. Please note this is not an exhaustive list.

Local registration teams

Despite the wealth of data available from other specific sources that can be used to populate a national register, the ability to investigate and pursue local hard-to-find data is essential for comprehensive registration. Experience indicates that data quality is enhanced where regional teams develop a good working relationship with their hospital trusts, health professionals and clinicians. We will therefore use specialist data liaison teams based in the individual local registration offices across England. These teams will work closely with local providers, clinical teams and others in their catchment area to help get data from disparate local systems and then follow up any data anomalies to ensure that the information we hold is accurate. They will work alongside, and be managed by, colleagues working in cancer registration so will be able to draw on the expertise and resources available to these larger teams.

Quality assurance (QA)

Because there are relatively few cancer types, and a dataset that is largely similar for all cancer types, it is relatively easy for a cancer registration office to have all the resources it needs to register all cases. Unlike the registration teams in cancer, with relatively few different cancer types, the registration teams that collect data on rare diseases cannot be expected to have sufficient expertise to understand the subtleties of the data associated with every rare disease. We therefore propose wherever possible to use the distributed expertise across the clinical and academic community to provide a QA function for specific diseases or groups of disease.

The registration system will be built to allow individual QA experts to log into the system securely so that they can review the details of cases in a form that does not reveal the individual patient identity. Inconsistent or key data items that have been missed can be flagged to the registration teams so that they can be investigated further.

Patient and public engagement

In relation to congenital anomalies and rare diseases it must be recognised that the term 'patient' refers to a broad range of people, including pregnant women and their unborn child, babies, children, young people (and therefore their parents/carers) and adults.

The NCARDRS has the potential to significantly improve healthcare provision for patients and their families through better planning, coordination and implementation of services. Patient and public engagement is key to the development of this work, and should include the engagement and involvement of each specific group of patients. In addition to involving patients in the communication and development of the service, there is also a larger role that we can perform for patients, their families and carers.

Patient organisations, where one exists for that condition, are well placed to provide information and support to patients. Working with patients and patient organisations can provide us with richer data regarding the natural history of illnesses. In turn, through a high-quality registration service, patients could be signposted to relevant voluntary organisations and registration itself will improve data quality and understanding, potentially creating a catalyst for the development of new patient organisations where they are lacking. This can all help support patients, families and carers to understand and manage conditions, and can help inform decision-making.

Patients could also enhance registration and improve the quality of the data through the ability to self-register. These groups, particularly those dealing with rare diseases, often hold the most information about their condition.

The service could build additional functions, such as a web-based patient portal, to provide a platform for patients, families and carers; a way to securely share their own information, find additional information and access peer to peer support. This concept is currently being piloted within cancer registration for brain tumour patients: www.myregistry.nhs.uk is a partnership project involving NCRS, and the charities CRUK and brainstrust.

Sharing data with others

Rare disease research has historically been highly fragmented by data type, by research institution and by disease. Individual efforts often have little interoperability and it is almost impossible to link different data sets. The NCARDRS will provide this interoperability and therefore enable researchers to have a better overview of the disease they are studying. Secure processes will be established so that researchers and academics can access the data, this includes provision of a safe haven. Importantly, there are also opportunities for this work to support and align with Genomics England's 100,000 Genome Project and wider translational research.

Although PHE only has the remit to develop a registration service to cover England we acknowledge the potential value of a UK wide registry and the importance of integration with systems in Scotland, Wales and Northern Ireland. Initial discussions have taken place with colleagues in the devolved nations and as the work develops we will work in colloration to maximise interoperability between systems and establish appropriate data-sharing agreements

European and international registration

There are significant and important developments in Europe on rare disease registration. We will work to ensure a two-way interface with other international repositories, for example EUROCAT, RDConnect and the proposed European Platform for rare diseases. New and emerging European standards will be incorporated into the development of our national service contributing to the commitments outlined in the UK Strategy for Rare Diseases published in November 2013.[4]

---

[4] https://www.gov.uk/government/publications/rare-diseases-strategy

Challenges

Finding cases

The aim of a population-level register is to collect information on all cases that meet the inclusion criteria for the register, to ensure high case ascertainment. With rare diseases, apart from the uncommonness of the condition itself, there is no single or small number of inclusion criteria that create a case definition; there are also no predictable diagnostic pathways, centres or clinics in the NHS that can be used to ensure that all cases are found.

We will therefore need to work with colleagues across PHE, the NHS and HSCIC to intercept a diverse range of data feeds and work very closely with existing experts to collate details on all cases.

Collecting relevant data items

One of the most significant challenges for a generic rare disease registry is to gather the richness of specific data items on each individual disease. Many diseases will have data items that are only relevant to that disease or a few others, while there will also be other characteristics that are poorly defined and difficult to include in any formalised data schema. Not only is the data very variable but it is impossible to define all possible data items for every disease in advance, not least because new diseases are constantly being described and key data requirements for individual diseases change.

One approach would be only to collect the minimum data that is required to include the case on the register and then to provide a link to extended external datasets held by others. While this minimises the work for the registration teams and reduces the data collected centrally, it creates significant external dependencies on other data collections, prevents the analysis of shared features of individual cases and detracts from the value of a single national rare disease register.

Our preferred approach is to use a relatively new industry-standard method of data storage called a graph (or no-SQL) database that has become popular to address the Big Data questions associated with social media and other very large complex datasets. A graph database allows us to hold

extensive unstructured data associated with each individual and their disease without the need to pre-define a schema. Tools are then available to analyse and perform complex queries on this disparate data. Each case would be defined as a node with the core patient demographics and then the remainder of the data on that case would be stored alongside the node.

The role of the registration staff in PHE will be to find cases from disparate sources, reconcile and link duplicates, to code and classify the case, and to quality assure the data associated with each case to check that it is in the correct format and consistent.

Information governance and consent

Accuracy is particularly important when collating data on rare and uncommon conditions, because every case provides important insights; small errors can significantly skew analyses. The quality of data across the NHS is highly variable, and based on the experience from the NCRS linking data from disparate sources is difficult because identifiers are not always correct and the NHS number is not used consistently, for a significant proportion of congenital anomaly cases the NHS number is not applicable. Even with the most reliable data a number of cases require manual intervention.

In order to ensure that the coverage is as complete as possible, we propose to continue to use the existing permissions under Section 251 of NHS Act 2006 to collect patient-identifiable data into the register. Where these permissions need further clarification we will work with the Confidentiality Advisory Group to ensure we have consistent and transparent inclusion criteria and ultimately we propose to use Orphanet and the future release of ICD 11 as a reference list for inclusion.

The absolute right for any patients to opt out if they wish will remain integral to our service. We will work with patients and patient organisations to develop patient information so that patients are informed about the nature, purpose and use of the data held in the registry. Clear communication outlining the benefits and considerations of being part of the registration service is vital. If a patient does not want us to hold information about them they can opt out at any time and this will not affect the care they receive in any way.

Next steps

This paper outlines our vision for a National Congenital Anomaly and Rare Disease Registration Service, illustrated visually in Figure 3. It provides an overview of the key components of the service and how we propose to overcome some of the potential challenges.

Our aim is to ensure the infrastructure is in place for national coverage by the end of 2015. Table 2 shows the high-level Gantt chart of the project, outlining the key areas of work and timescales.

The business case for resource to implement congenital anomaly registration across England was approved at the end of June. Recurrent funding is now in place to ensure continuation of data collection in regions where regional teams already exist, and establishment of new regional teams to expand and achieve national coverage. Funding has also been approved for the development of a single, standardised national data collection system.

Figure 3: The National Congenital Anomaly and Rare Disease Registration Service
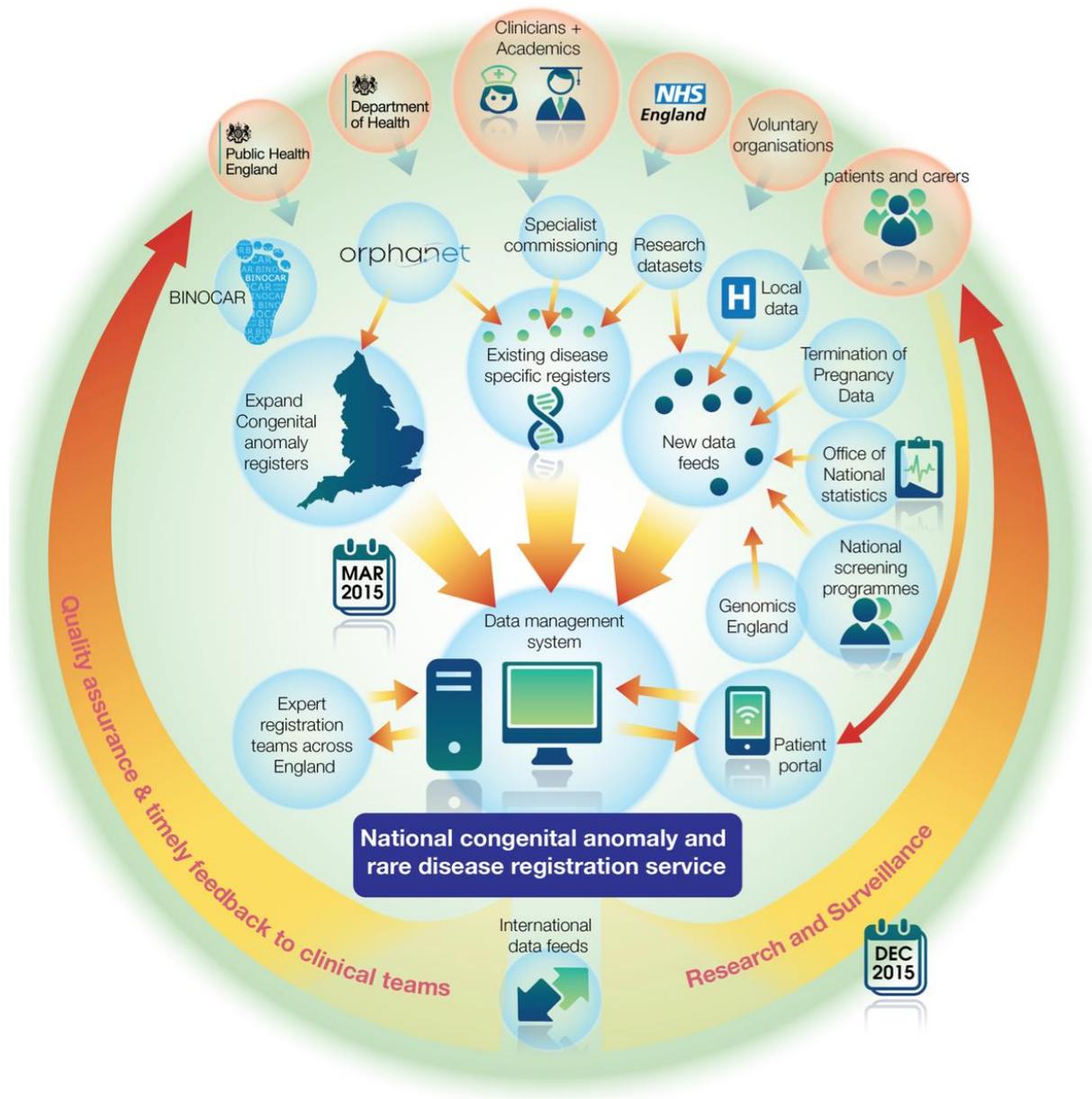
Table 2: Project Gantt Chart

	Q1 2014/15	Q2 2014/15	Q3 2014/15	Q4 2014/15	Q1 2015/16	Q2 2015/16	Q3 2015/16	Q4 2015/16
Agree project plan, confirm timescales, identify funding and appoint project manager.	▮	▮						
Secure necessary information governance permissions and establish data transfer processes		▮	▮	▮				
Scope current congenital anomaly register systems and processes		▮	▮					
Scope other individual rare disease registers (funding permitting)			▮	▮				
Strengthen regional registration teams; moving existing CAR teams to PHE		▮	▮	▮	▮			
Establish new regional teams in the 5 regions which currently have no team in place		▮	▮	▮	▮			
Design and build national data management system		▮	▮	▮	▮	▮		
Develop systems to enable quality assurance and feedback to clinical teams and patients					▮	▮	▮	▮
Raise awareness and engage with patients, clinicians, voluntary organisations & other stakeholders	▮	▮	▮	▮	▮	▮	▮	▮
Build patient portal						▮	▮	▮