The Diagnostic Odyssey in rare diseases; a Task and Finish Group report for the Department of Health and Social Care

Authors: Eleanor Hay, Frances Elmslie, Peter Lanyon and Trevor Cole
April 2021

Keywords: rare diseases, diagnostic odyssey, outpatient diagnostic coding, Tuberous Sclerosis Complex, Bardet Biedl Syndrome, ANCA-associated Vasculitis, NCARDRS

INTRODUCTION

Background

Rare diseases can be defined as life-threatening or chronically debilitating conditions that have a frequency of less than 1 in 2000 in the population at risk. While individually rare, with greater recognition it is apparent that as many as 1 in 17 people will develop a rare disease over their lifetime, equating to approximately 3.5 million people in the UK alone. Over 75% of rare diseases affect children and currently 30% of patients die before reaching their 5th birthday [7]. Whilst it is estimated that 80% of rare diseases are genetic in origin, only a quarter of these have had their molecular basis defined. This has led to difficulty in recognising and effectively diagnosing these conditions [7].

The journey from initial presentation to the NHS through to receiving an accurate diagnosis is termed the diagnostic odyssey (DO). Whilst many individuals with rare diseases in the UK report excellent and timely patient journeys, a large European study, including a cohort of UK patients, looking at eight rare disease subsets, identified that 25% of patients had to wait between five and thirty years for a final diagnosis [8]. In these prolonged diagnostic odysseys, patients may have multiple NHS encounters with a variety of specialists and endure serial and often invasive investigations [4]. In a Rare Disease UK survey in 2016, it was found that on average from the time of first symptoms relating to the condition, to a formal diagnosis, individuals consulted five doctors and had three misdiagnoses [20]. This is frustrating and stressful for the patient and family and an ineffective and expensive use of resources. A delayed or misdiagnosis may also have serious implications for patient health and lead to missed opportunities for effective intervention.

The recently published 2021 Rare Diseases Framework has built on, and superseded, the 2013 UK strategy for Rare Diseases Report. The work and writing underlying this document was completed prior to the publication of the 2021 framework. In both the 2013 and 2021 documents however, a key recommendation is to reduce diagnostic delay and, in doing so, improve timely access to appropriate intervention[4]. The 2013 report set out several approaches to achieving this aim, including improved prenatal care and newborn screening programmes and carrier and cascade testing for relatives of individuals with known genetic conditions. Education of health care providers in recognising rare disease and understanding genomic testing was also noted, through massive open online courses (MOOC) and Health Education England funded programmes, ranging from stand-alone modules to full MSc courses in Genomic Medicine. Large sequencing projects, such as the Deciphering Developmental Disorders (DDD) and 100,000 Genomes project, have delivered

many new diagnoses to families who have endured long odysseys and has set in motion the development of an NHS clinical whole genome sequencing service, with live clinical testing for selected rare disease phenotypes available from November 2020. Reconfiguration of NHS England laboratories into seven Genomic Laboratory Hubs has additionally met the aim to ensure equal access to high quality genomic testing [4].

The 2021 Framework notes that, despite making significant progress, there remain areas for further development [5]. The framework outlines a future vision, shared by all four UK nations, to address health inequalities, improve the quality and availability of care, and improve the lives of people living with rare diseases. The four key priorities highlighted include helping patients to get a final diagnosis faster, increasing awareness of rare diseases among healthcare professionals, better coordination of care and improving access to specialist care, treatments and drugs. In the related document, 'Genome UK: the future of healthcare', incorporating genomics into routine healthcare for the provision of diagnosis and personalised medicine for rare disease, is also highlighted as one of the three pillars supporting the strategy to create the most advanced genomic healthcare system in the world [6]. The framework however, also raises greater awareness that a significant proportion of rare diseases, at least 20%, do not have a genomic aetiology but face many of the same issues.

This focus on strategy implementation related to genomic advances reflects the fact that up to 80% of rare diseases are genetic in origin. However, implementation also needs to be inclusive of the needs of the remaining 20% of people living with rare non-genetic diseases, as further emphasized in the 2021 UK Rare Diseases Framework [5]. An exemplar for this non-genetic group are the rare autoimmune rheumatic diseases, conditions which predominantly have onset in later life and share multisystem clinical features which occur when the body's own immune system attacks healthy tissues, akin to "friendly fire". The unmet needs of people living with these conditions, in particular related to diagnostic delay, coordination of care and lack of knowledge, are highlighted in a 2018 report by the Rare Autoimmune Rheumatic Disease Alliance (RAIRDA) [1]. A key finding was that almost half (46%) of the 2300 respondents reported waiting more than three years from first symptom to receiving their correct diagnosis, with fewer than 1 in 10 being diagnosed in less than three months. Of concern, the report did not find evidence of any recent improvement in this delay amongst those diagnosed in the last three years.

To inform implementation of the new Framework and to understand whether the UK strategy has been successful in leading to improvements in the DO, there is a need to develop an effective and efficient methodology to measure changes in the DO over time. This is particularly challenging in the case of rare diseases as, by definition, they affect only a small fraction of the population. For example, the small number of cases of some individual conditions may not allow for sufficient statistical power to detect improvements in the DO following any policy implementation. Additionally, collation of data across multiple NHS Trusts presents a logistical challenge.

While several definitions of the DO have been proposed, for the purposes of this report we have considered this to:

- begin at the time point when an individual first seeks NHS care with features (clinical symptoms, examination findings or test results) that are later classified as likely to have been a presenting feature of their rare disease

- end at the time point of definitive diagnosis, be that clinical, biochemical or genetic

We also considered that the DO metric could be further characterised by two inter-dependant components

- the duration of the odyssey e.g. time from beginning to end
- the circuitousness of the odyssey e.g. the number of consultations or different specialties involved in that person's healthcare, from beginning to end

The 2015 Policy Innovation Research Unit (PIRU) report sought to address whether an accurate, robust and financially plausible method could be developed that could effectively measure the duration of rare disease diagnostic odysseys and therefore evaluate the impact of interventions and policies [3]. Whilst a number of potential measures were considered, including the use of clinical databases, rare disease registries and hospital administrative records, there was no unifying suggestion as to how best to approach this analysis.

In 2015/2016 the UK Genetic Testing Network (UKGTN) Rare Disease Service Improvement Working Group evaluated practical solutions to this analysis [4]. A small pilot of rare disease patients highlighted the challenges of obtaining and evaluating data when multiple specialities, databases and different hospital trusts were involved.

In-house and RD SI working group discussions together with literature reviews illustrated how existing routinely collected healthcare databases, such as the Clinical Practice Research Datalink (CPRD), could potentially be used to obtain data on the duration of the DO.

Choice of exemplar conditions

Examples of conditions documented in the PIRU systematic review frequently were single system disorders, or disorders which would likely present to a single or small number of specialities [3]. Some of the disorders would be detected in national screening programmes (inborn errors of metabolism including congenital hypothyroidism or homocystinuria) or had simple diagnostic tests (Cystic Fibrosis, Duchenne muscular dystrophy or Fabry's disease) and these were not thought to encompass the true challenges of the diagnostic odyssey in rare disorders [3]. We therefore considered that it was important to include multisystem diseases in our pilot, in order to assess variable clinical and temporal presentations to multiple medical specialities at different ages.

An initial pilot utilising a manual review of paper records for two rare disorder cases (CHARGE syndrome and Bardet Biedl Syndrome) ascertained that single institution hospital records were too incomplete to provide meaningful data, as patients with rare disorders are often seen in multiple specialities that are frequently not co-located. In addition, a requirement from the pilot was that only those patients with an assured diagnosis be included to ensure that the diagnostic odyssey had come to a "correct end point". If this has not been the case a rapid but incorrect diagnosis would mask the true length of the DO.

In this report three rare conditions were chosen as exemplars for the DO faced by individuals and their families. In two (Tuberous Sclerosis Complex and Bardet Biedl syndrome) the diagnosis could be confirmed with genetic testing. In the third, a non-genetic disorder (ANCA-associated Vasculitis)

clinical and biological markers could provide assurance of the correct diagnosis. The DO incurred by patients with these three rare genetic and non-genetic conditions are felt to be representative of many rare diseases and highlight the challenges of securing a timely diagnosis.

Description of conditions selected

Tuberous Sclerosis Complex

Tuberous sclerosis Complex (TSC) is an extremely variable multi-system genetic disorder that affects around 8000 people in the UK and has an incidence of 1 in 6000-10,000. It is characterised by abnormalities of the skin, brain, eye, kidney and heart and is often associated with seizures and a range of cognitive and neurodevelopmental disorders [10]. Complications such as treatment resistant epilepsy, renal angiomyolipoma and an overgrowth of smooth muscle in the lungs (LAM) can significantly reduce life expectancy.

Advances in genetic sequencing technologies over the last decade have improved the molecular diagnostic yield such that, in 75-90% of patients, a pathogenic (disease-causing) gene change in *TSC1* or *TSC2* is identified. These gene changes may occur for the first time in an individual or be inherited in an autosomal dominant fashion from a variably affected parent, who may not be aware that they have the condition. In the 10-25% without a recognised gene change, the 2012 International Tuberous Sclerosis Complex Diagnostic Criteria (11 major and 6 minor features) provide guidance on the clinical diagnosis of TSC [10].

In the past, presentation of TSC was often in childhood following the onset of a seizure disorder. As there is greater recognition of key features of the condition, presentations have now broadened to include antenatal referral for cardiac rhabdomyomas, children and adults with hypopigmented patches, young women with renal angiomyolipomata and family members of a relative with a known *TSC1/2* mutation. Unfortunately, more mildly affected individuals may go decades without a diagnosis due to the insidious onset of many non-specific symptoms that are not recognised in aggregate [2].

Pathogenic variants can now be used independently of clinical signs to diagnose individuals with TSC, facilitating earlier diagnosis in children in whom few clinical signs may be evident. This is important, as earlier access to surveillance and treatment has the potential to improve outcomes of individuals with TSC and reduce the rate of complications. In the TS2000 study, the impact of earlier diagnosis and intervention of TSC-induced infantile spasms was evident in an improvement in neurodevelopmental outcomes [10]. A confirmed diagnosis of TSC can also entitle the individual to management with oral or topical mTOR inhibitors, such as Everolimus, that target the specific pathway disrupted in TSC [10]. Additionally, a genetic diagnosis can enable other members of the family to be tested and permit prenatal or preimplantation genetic diagnosis.

Bardet Biedl Syndrome

Bardet Biedl syndrome (BBS) is another multi-system genetic condition with a prevalence in non-consanguineous populations of between 1 in 100,000 and 1 in 160,000. The features of BBS are variable, though most develop truncal obesity and a specific eye phenotype that leads to night blindness in childhood, progressing to loss of visual acuity. 75% of adults become legally blind, often from adolescence. Additional features include postaxial polydactyly, developmental delay,

intellectual disability, hypogonadotrophic hypogonadism in males, genitourinary abnormalities in females and structural or functional impairment of the kidneys [9].

Diagnosis can be made on the basis of clinical features, but more recently the genetic basis of BBS has been further elucidated. To date, 19 genes for BBS have been documented, although around 40% of diagnoses are accounted for by *BBS1* and *BBS10* alone and 20% have no pathogenic variant detected [9], suggesting that further genes may be responsible. Inheritance is autosomal recessive and parents are usually carriers of the gene change and therefore have a 25% (1 in 4) chance of recurrence in future pregnancies.

As many of the features of BBS are non-specific (developmental delay and obesity) or do not manifest until later in childhood (visual and renal issues), diagnosis is often delayed. Whilst there is currently no management to prevent or reverse the eye changes, early diagnosis can improve visual and psychological outcomes by preparing children for visual loss through Braille/visual support. Prompt access to educational support and surveillance and intervention for endocrine and renal manifestations is also beneficial [9] . Understanding the cause of children's weight gain and targeting this early can additionally reduce or prevent the associated co-morbidities, such as hypertension and diabetes mellitus. Finally, recognising the genetic cause in an individual permits accurate counselling of the family on recurrence and provides the option for prenatal testing or pre-implantation genetic diagnosis, if desired.

ANCA-associated Vasculitis

The rare autoimmune rheumatic diseases can be grouped into two categories, *Autoimmune Connective Tissue Diseases* (e.g. Systemic Lupus Erythematosus, Systemic Sclerosis, Myositis, and Sjogren's Syndrome) and *Systemic Vasculitis* (e.g. ANCA-associated Vasculitis, Giant Cell Arteritis, Takayasu's Arteritis, Behçets Syndrome). ANCA-associated vasculitis (AAV) was selected as an ideal exemplar for these conditions. AAV is a late-onset (median age 55 years) non-genetic condition in which vascular inflammation leads to irreversible organ damage to kidneys, lungs, gut, peripheral nerves, hearing and vision. The estimated incidence is 23 per million person-years [14] and more than half of the 1500 new cases each year are diagnosed during an episode of unscheduled emergency in-patient care [15]. The estimated prevalence is 25/100,000 persons, with 13,000 people living with AAV in the UK [14].

AAV comprises three sub-types that share overlapping clinical features; Granulomatosis with Polyangiitis (GPA, previously called Wegener's Granulomatosis), Microscopic Polyangiitis (MPA) and Eosinophilic Granulomatosis with Polyangiitis (EGPA, previously called Churg Strauss Syndrome), with each type having their own ICD code that maps one-one with a single orphacode.

Diagnosis is associated with the detection of autoantibodies directed against proteins within the cytoplasm of neutrophil white blood cells (Antineutrophil cytoplasmic antibodies, ANCA), and it is likely that these antibodies are involved in pathogenesis. Diagnosis is established by ANCA detection, in the context of appropriate clinical features, supported by tissue biopsy. Initial treatment to induce remission is high dose steroids and either the chemotherapeutic drug cyclophosphamide (AAV is the most common indication for the use of non-cancer chemotherapy) or rituximab, a high-cost biologic commissioned by NHS England for this indication. Optimal care also involves coordination across a

range of disciplines, including rheumatology, nephrology, respiratory, dermatology, ENT and ophthalmology.

AAV has the highest mortality of any rare autoimmune disease, with one in seven people dying within a year after diagnosis [13]. Reducing diagnostic delay and improving outcomes have been identified as top priorities by people living with vasculitis. L ike many rare diseases, it can be difficult to diagnose, mimicking other conditions, including infection and malignancy. The initial presentation can be nonspecific and include sinusitis, fatigue, joint pain, weight loss, fever, cough and skin rashes, or renal failure, which is often asymptomatic. Delays in diagnosis are associated with greater risk of irreversible organ damage, which adversely affects clinical outcomes and also increases costs to the NHS, including renal replacement therapy, which is required by 20% of people who survive the first year after diagnosis.

Recent research undertaken in primary care records has found evidence of increased health-seeking behaviour in the five years before diagnosis amongst people subsequently diagnosed with GPA subtype compared with those without this diagnosis [21] This study also emphasised the difficulties of developing risk prediction models in primary care and highlighted opportunities for earlier diagnosis in secondary care, which dovetails with the work in this pilot. This accords with the 2018 RAIRDA report; of the respondents with any type of vasculitis, more than 40% reported a delay of at least one year from onset of symptoms to diagnosis. These complementary studies support the existence of a diagnostic odyssey of ill health for a long time without a diagnosis, often with this only being achieved when there is further clinical deterioration leading to unscheduled emergency admission [15].

Whilst progress is being made, to address the remaining diagnostic delay for rare diseases it will be important to improve education for healthcare professionals, increase population awareness of the condition and develop condition specific pathways to support diagnosis, including genetic testing.

Aims

The Diagnostic Odyssey Task and Finish Group aims to propose standard methods for estimating the diagnostic odyssey for individuals with rare diseases. This methodology would then facilitate appropriate measurements for the effectiveness of the new UK Rare Diseases Framework and other interventions.

At outset, we recognised that the ideal methodology for assessing the diagnostic odyssey should include the ability to investigate pathways to diagnosis across the whole healthcare system e.g. from primary into secondary care, where the majority of rare diseases will be diagnosed.

In addition, any proposed methodology should have a robust, reliable and reproducible process for:

- ascertaining cases, with a high level of diagnostic certainty
- estimating their likely date of diagnosis
- evaluating healthcare activity prior to diagnosis across multiple specialties and multiple NHS locations

Because of the low prevalence of these conditions, we considered that the methodology should also have the future ability to be scaled up and automated, to address the DO for that condition across

the whole of population of the United Kingdom e.g. utilise routinely collected or nationally mandated NHS healthcare dataset.

For this pilot study, we considered the available sources of routinely collected healthcare data.

1. Primary care

The Clinical Practice Research Datalink (CPRD) is a longitudinal general practice database which collects de-identified patient data from a network of GP practices across the UK, including 11 million currently registered patients. It includes information on demographics, diagnoses (coded using Read codes and more recently SNOMED-CT), referrals, medications, and tests and is considered to be representative of the UK population. Since the PIRU report, research using CPRD data has been published on the health-seeking behaviour of patients prior to their diagnosis for two rare autoimmune diseases [21] [17]. These studies identified increased health care activity up to five years prior to diagnosis, compared to those without this diagnosis, providing further support for the importance of reducing the diagnostic odyssey. However, they both highlighted the problems with trying to develop risk prediction models for rare diseases in primary care [17]. These include there being no specific symptoms, either alone or in combination, with sufficient discrimination to support development of a prediction model, which is particularly the case when initial symptoms are also shared with common non-rare conditions. Secondly, the rarer the disease, the lower the positive predictive value (i.e. the lower the probability) that an individual flagged at risk either has or will develop the disease.

2. Secondary care

Secondary care activity related to the DO can be ascertained by using Hospital Episode Statistics (HES) data, which include episodes of in-patient care and attendances at Accident and Emergency (A+E) and outpatients. It is known from previous research in ANCA-associated vasculitis that there is a higher frequency of outpatient attendances in the year prior to diagnosis in people who are subsequently diagnosed with this condition, compared to a control group who do not have this diagnosis. However, although this increased secondary care activity prior to diagnosis is likely to be a component of the diagnostic odyssey, it has not been possible to validate or confirm this assumption within the available routinely collected healthcare data. This is for two main reasons. Firstly, 95% of NHS outpatient attendances are not ascribed an ICD code, meaning that the clinical reason (e.g. symptoms, physical signs and test results) for these consultations is not available in the routinely available data. Secondly, although there is linkage between CPRD and HES, because of patient confidentiality it is not possible to view outpatient clinic letters within the CPRD record, to confirm whether a particular consultation will have been for clinical features that were likely to be related to the subsequent rare disease diagnosis [13]. Whilst all in-patient activity (including day cases) is coded in HES, it is not appropriate to assume that the date at which a particular diagnosis code occurs in HES is the correct diagnosis date, unless there is a very high prior certainty that diagnosis or initial treatment would usually involve an inpatient admission or day case treatment close to the time of diagnosis, or the date of diagnosis can be obtained from another (or linked) source. This is unlikely to be the case for many rare genetic conditions, where the majority of initial consultations and diagnostic evaluations will take place on an outpatient basis.

We therefore chose to study the DO in secondary care, starting with a pilot study utilising small cohorts of patients with both a secure rare disease diagnosis and a clinician-confirmed date of diagnosis. We subsequently aimed to study the DO in primary care for these subjects by seeking the legal permissions and funding to enable linkage with their primary care data via The General Practice Extraction Service (GPES). Although initial discussions took place with NHS Digital about accessing this data, time and financial constraints prevented this work from proceeding within the remit of this working group.

METHODOLOGY

Case ascertainment

Three clinicians working in NHS Trusts providing clinical services for patients with TS, BBS and AAV identified an unselected cohort of patients under their care who had a date of diagnosis available in the medical record. It was agreed to study the DO of up to 20 individual patients in each of these cohorts. This comprised:

- 14 patients from two randomly selected national Bardet Biedl clinics
- 6 randomly selected patients with TSC from one specialist centre
- 20 patients with ANCA-associated vasculitis from Nottingham University Hospitals NHS Trust, who had previously been identified during a Trust audit.

Data sharing with NHS Digital

Following confirmation between NHS Digital and a Caldicott Guardian that there was a legal basis for data sharing, the unique NHS identifiers and date of diagnosis for each patient was securely transmitted to NHS Digital. For all patients, the date of diagnosis was obtained from the medical records available to the relevant managing clinician. This date may have been before confirmation of the diagnosis by molecular genetic or biological markers as these tests may not have been available at the time of the original diagnosis or may have been instigated at a later date as confirmation of a preceding clinical diagnosis.

NHS Digital subsequently provided each clinician with an analysis based on HES data for each patient, for three years prior to the diagnosis date and for the two years following diagnosis. This included details of hospital attendances (outpatient, Accident and Emergency, inpatient and daycase), treatment function code (specialty) and hospital trust prior to the date of diagnosis, and the International Classification of Diseases (ICD-10) diagnostic codes associated with any inpatient or day case episodes. Many rare disorders (including BBS), lack a unique ICD-10 diagnostic code, hence a selection of several disorder specific terms that could identify the diagnoses in digital records were provided, alongside the likely treating specialties for each disorder (table 1).

Within this data set, the number and pattern of prior attendances in NHS secondary care services was assessed by the relevant clinician, informed by their knowledge of the individual cases and supported by information already collected in their medical records, to establish whether these attendances were likely to have been related to the DO.

Table 1: Case specific clinical features, specialties and Hospital Trusts provided to NHS Digital to facilitate data collection		
TUBEROUS SCLEROSIS COMPLEX (ICD-10 code Q85.1)		
Clinical features	**Speciality**	**Hospital Trust**
1.Cutaneous Neurocutanous markers (angiofibroma, hypomelanotic macules, confetti lesions, shagreen patch, ungula fibromas, café-au-lait macules) **2.Nervous system** Infantile spasms Seizure disorder MRI imaging anomalies (Cortical dysplasia, subependymal nodules, subependymal giant cell astrocytoma) **3.Renal** Renal angiomyolipoma Renal cysts **4. General/community paediatrics** Developmental delay Neurodevelopmental disorders **5.Eye** Retinal nodular hamartomas Retinal achromic patch **6.Chest** Lymphangioleiomyomatosis **7.Cardiovascular** Cardiac rhabdomyoma (7)	Dermatology General Paediatrics Neurology Nephrology Community Paediatrics Ophthalmology Respiratory Fetal Medicine Unit Paediatric cardiology Clinical Genetics	Detailed clinical information from single TSC specialist clinic
BARDET-BIEDL SYNDROME (No specific ICD-10)		
Clinical features	**Speciality**	**Hospital Trust**
1.Endocrine Obesity Diabetes Hypogonadism Ammenorrhea Delayed Puberty Hypoplastic Genitalia Small Testes **2.Eyes** Visual Impairment Retinal disease Nystagmus, Glaucoma Cataract **3.Renal** Kidney Failure Dysplasic Cystic Kidney Disease **4.Cardiovascular** Congenital heart Disease Cardiomyopathy **5.Gastrointestinal** Hirschprungs Bilary Atresia **6.Plastics** Digital anomalies – Polydactyly, Syndactyly **7.General/community paediatrics** Developmental Delay (9)	Endocrinology (adult or paediatrics) Ophthalmology (adult or paediatrics) Orthoptics Nephrology (adult or paediatrics) Dietetics Diabetes clinics Paediatric surgery / paediatric urology Hand or plastic surgery (adult or paediatrics) General Paediatrics Clinical Genetics	Birmingham Women and Children's Hospital NHS Foundation Trust Smaller contribution from other Trust within West Midlands Deanery

ANCA-ASSOCIATED VASCULITIS (ICD codes M301, M313, M317)		
Clinical features	**Speciality**	**Hospital Trust**
1. General Myalgia, Arthralgia or arthritis Fever Weight Loss **2. Cutaneous** Infarct Purpura Ulcer, Gangrene, Extensive tissue necrosis Other skin vasculitis **3. Mucous membranes/eyes** Mouth ulcers/granulomata Genital ulcers Adnexal inflammation Significant proptosis Red eye (Epi)scleritis/conjunctivitis Blepharitis/Keratitis/Uveitis Blurred vision Sudden visual loss Retinal vasculitis/vessel thrombosis Retinal exudates/haemorrhages **4. ENT** Bloody nasal discharge/nasal crusts/ulcers and/or granulomata Paranasal sinus involvement Subglottic stenosis Conductive/sensorineural hearing loss **5. Chest** Wheeze Nodules or cavities Pleural effusion/pleurisy Infiltrate Endobronchial involvement Massive haemoptysis/alveolar haemorrhage Respiratory failure **6. Cardiovascular** Loss of pulses Valvular heart disease Pericarditis Ischaemic cardiac pain Cardiomyopathy Congestive cardiac failure **7. Abdominal** Peritonism Bloody diarrhoea Ischaemic abdominal pain **8. Renal** Hypertension Proteinuria/ Haematuria Rise in creatinine **9. Nervous system** Headache Meningitis Organic confusion Seizures (not hypertensive) Stroke Cranial nerve palsy Sensory peripheral neuropathy Motor mononeuritis multiplex	General Medicine Rheumatology Dermatology Ophthalmology Ear, Nose and Throat (ENT Respiratory Medicine Cardiology Gastroenterology Nephrology Neurology Clinical Haematology Symptom list adapted from Clinical features in the Birmingham Vasculitis Activity Score (BVAS) [11]	Nottingham University Hospitals NHS Trust

RESULTS AND IMPLICATIONS

Tuberous Sclerosis Complex

Tuberous sclerosis complex was selected as an exemplar because it can present to a range of different specialists at different ages and there is evidence from the medical literature and the patient support groups that diagnosis can be delayed. Initially, eleven patients from two specialist Tuberous Sclerosis clinics were selected at random. Insufficient information was available for the patients from one centre, so a detailed analysis was conducted for six patients only, of which three were adults and three children. Contrary to initial expectations, there was little evidence of a diagnostic odyssey in five of the six patients from a detailed analysis of local hospital records, but the diagnosis was not reliably recorded in HES data despite the diagnosis being well known.

Patient 1 was diagnosed with TSC after presenting with seizures at the age of 10 months. Investigations at this time led to the diagnosis of TSC within a few weeks of presentation. However, the diagnosis did not appear in HES until the patient was 14 years old.

The diagnosis of TSC was suspected in *Patient 2* following the detection of multiple cardiac rhabdomyomas. Postnatal investigations confirmed the diagnosis and was recorded in HES at the point of discharge from hospital soon after birth.

Patient 3 did experience a delay in diagnosis. This patient presented with seizures in adolescence. Over a period of ten years, she had a number of evaluations, but it was only after an admission and detailed re-evaluation of her problems that the diagnosis of TSC was made. Data from HES suggests a 4 year delay in diagnosis, but the medical records indicate a longer path to diagnosis.

Patient 4 was also found to have cardiac rhabdomyomas at birth, once again leading to the suspicion of TSC which was confirmed by brain MRI. The diagnosis did not appear in HES until 8 years later.

Patient 5 was diagnosed in childhood. He was the oldest of the group and his early medical records were not available. However, the first mention of his diagnosis appeared in HES data 20 years after his first presentation to the specialist TSC clinic.

Patient 6 presented with infantile spasms at 3 months. Investigations confirmed the diagnosis of TSC and this was recorded in HES two months later.

Summary of TSC results

1. In this small cohort the majority (5 of 6) of patients did not experience a diagnostic odyssey, with time from initial presentation with related symptoms to clinical diagnosis being less than six months.
2. Even when the clinical diagnosis was clear and marked in the notes, in three cases there was a considerable lag in the diagnosis being reflected in HES data.

Implications of TSC results

Although these data are very limited, they indicate that in patients presenting with the cardinal features of TSC (e.g. cardiac rhabdomyomas or infantile spasms) the diagnosis is not significantly

delayed. However, the patient presenting with simple seizures did experience a delay possibly related to the quality of brain imaging at the time of first presentation. It was apparent that the diagnosis is unreliably recorded in HES despite multiple contacts with health professionals. One approach to improving digital recording of diagnosis in TSC would be to mandate that all rare disease (using TSC as a representative disorder) be on a national register, such as the National Congenital Anomaly and Rare Disease Registration Service (NCARDRS).

Bardet Biedl Syndrome

A total of 14 patients with a formal diagnosis of BBS made between 2003 and 2015 were randomly selected from those attending the specialist BBS clinic at Birmingham Women and Children's NHS Foundation Trust. HES data were tabulated by NHS digital and reviewed alongside the clinical notes by the treating consultant.

Initially, data were reviewed in relation to birth cohort (table 2), comparing the age that patients received a clinical diagnosis when born in 1991-2000 to 2001-2010.

Table 2 BBS: Date of clinical diagnosis in relation to year of birth

In Birth Cohort	Number of patients	Ages at clinical diagnosis in years	Mean Age at diagnosis in years
1991-2000	5	6-18 years	12
2001-2010	9	0-5 years	2

This illustrated a clear relationship of a much longer times to clinical diagnosis in the earlier birth cohort, whereas for children born since 2001 there has been a more rapid progression to a clinical diagnosis, indicating a shorter DO.

Data were then reviewed in relation to the year in which a diagnosis was established (clinical or molecular) and organised into two different cohorts A (diagnosed 2003-2008) and B (diagnosed 2010-2015) (table 3).

Table 3 BBS: Details of date of clinical/molecular diagnosis and HES related activity up to 2 years pre and post diagnosis

Diagnostic cohort	**A (2003-2008)**							**B (2010-2015)**						
Patient	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Age clinical dx (years)	6	5	3	1.8	2	1.9	4	0.5	0.7	12	13	18	14	0.25
Time from clinical to molecular dx (months)	120	76	88	72	65	21	10	32	6	24	7	25	12	7
HES activity 2 years pre clinical diagnosis	3	14	24	12	5	8	11	0	0	11	2	4	6	0
HES activity 1 years pre clinical diagnosis	4	17	16	25	3	10	21	10	28	16	7	5	20	10
HES activity 1 years post clinical diagnosis	3	9	3	10	5	5	25	9	41*	3	2	5	7	18
HES activity 2 years post clinical diagnosis	5	18	5	12	12	3	11	16	69*	13	7	6	6	22

Values with * represented as 30+ in the line graph to aid visualisation of all patient data

Patients 1-7, Cohort A, were diagnosed during the time period 2003-2008 and Patients 8-14, the Cohort B, from 2010 to 2015. The age at clinical diagnosis varied from infancy to 18 years. Importently, the mean time to receive genetic confirmation of a BBS diagnosis in the early cohort was 65 months, considerably greater than that of the later cohort, with a mean of 16 months (figure 1). There are likely to be several reasons for this, not least that genetic testing has become more available and affordable and that the number of BBS genes included on the relevant panels has increased. In addition, it is likely that increased awareness of the genetic testing on offer and confidence in the correct clinical diagnosis have played a role.

There is, however, room for improvement with evidence of multiple outpatient appointments in the 2 years prior to diagnosis. Reviewing HES data in relation to the date of clinical diagnosis (figure 2), revealed a general trend for hospital interactions to increase from 2 years to 1 year prior to diagnosis and then fall following diagnosis. This is not, however, universal and in five cases the number of HES indicators remained the same or increased at 1 year following diagnosis. At 2 years post diagnosis, there also appeared to be a subsequent rise in HES related activity, possibly due to additional assessments being appropriately arranged with related specialities. For example, one outlier (Patient 9) had a significant rise in HES activity due to the requirement for extensive renal monitoring as a result of the identification of impaired renal function following the diagnosis of BBS.

Reviewing out-patient consultations in the 5 year period prior to diagnosis, 70% (365 of 524) of events were in key specialties (as highlighted in Table 1), indicating that these are representative of secondary care interactions with children with BBS.

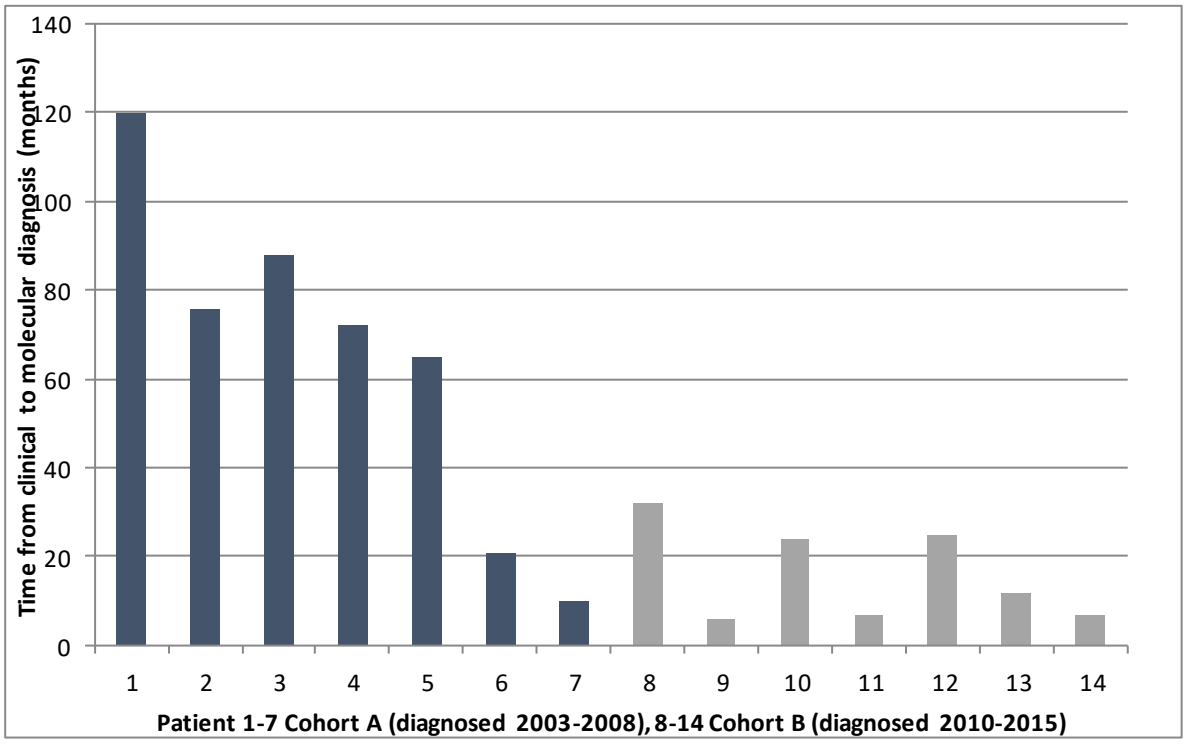

Figure 1: Bar chart demonstrating time taken from clinical diagnosis of BBS to confirmation with genetic testing in the two cohorts, A 2003-2008 (navy) and B 2010-2015 (grey)

In addition those children with more recent clinical diagnoses, including those in the 1991-2000 birth cohort, had a shortened period to molecular confirmation. This suggests some improvement, with a shortening of diagnostic odyssey for BBS and more rapid instigation of confirmatory genetic testing.

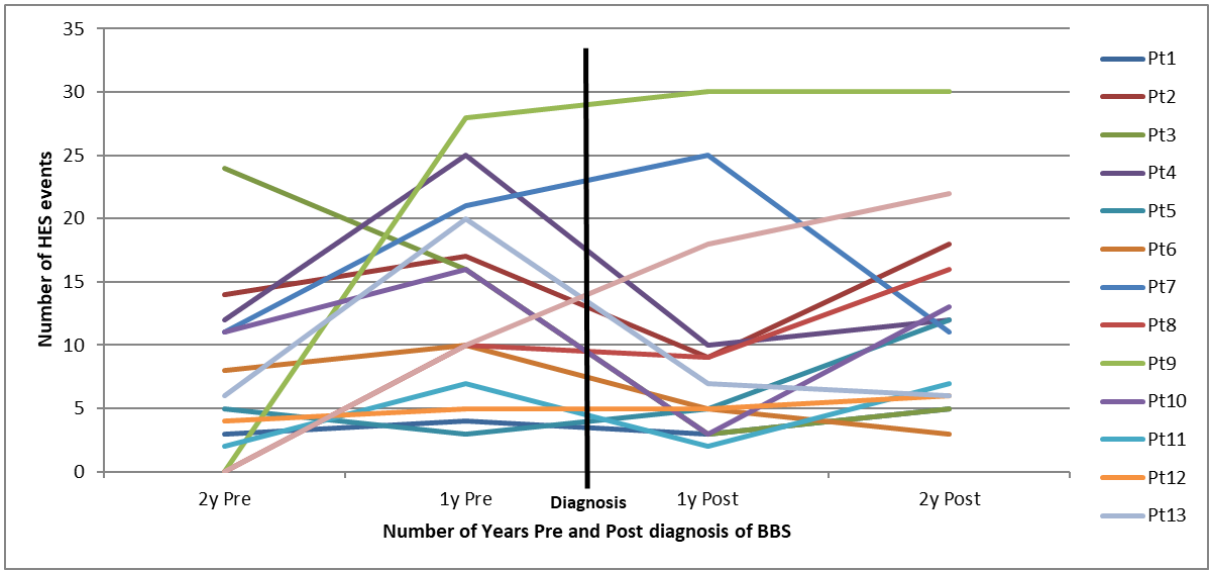

Figure 2: Line plot demonstrating HES activity up to 2 years pre- and post- clinical diagnosis for all 14 patients in the pilot study

Summary of BBS results

1. There is reduction in time to receive a clinical diagnosis from a mean of 12 years in the early birth cohort (1991-2000) to 2 years in the later birth cohort (2001-2010).
2. The time taken to confirm a clinical diagnosis molecularly has reduced by a mean of 47 months between the two diagnostic cohorts A (2003-2008) and B (2010-2015).
3. There is a general trend for HES activity to reduce from 1 year prior to diagnosis to 1 year following diagnosis.
4. An additional rise in HES activity is evident in most patients 2 years following diagnosis
5. Where outpatient records have HES data recording symptoms, all align to those submitted to NHS digital.
6. 70% of HES data activity five years prior to diagnosis is with specialities related to the management of symptoms of BBS.

Implications of BBS results

Encouragingly, this small cohort suggests that there has been a reduction in the time taken to receive a genetic diagnosis in recent years; however, it was not possible to extrapolate this directly from the HES data. This is in part due to most disease activity being in an outpatient setting where HES data is less routinely recorded but also as a result of BBS not having a specific ICD code – instead Q87.89 (Other specified congenital malformation syndrome, not elsewhere classified) is recorded. In the future, it is possible that more specific coding practices for rare disease (such as snomed or orphaned) are adopted. Whilst there is a suggestion that HES activity increases at around the point of diagnosis and falls thereafter, this is not universal and the HES trend is therefore unlikely to be useful as an accurate surrogate marker for BBS diagnosis alone, unless coding practices change. As there is currently no specific ICD term for BBS, the use of alternatives such as orphanet (OPRHA:110

– Bardet-Biedl syndrome) should be actively encouraged, and this should be in both the outpatient and inpatient setting.

One option to improve data extrapolation to mark the start and end of diagnostic odyssey might be to devise an algorithm that includes symptom /reason for attendance, speciality seen and number of interactions. These parameters could be syndrome specific but larger studies would be required to investigate whether applying such algorithms to HES data within NCARDRS can be utilised to monitor change in the duration of DO over time and the see whether this reflects the trend in reduced time to achieving genetic diagnosis that we have visualised here.

It would also be important to evaluate the temporal position of molecular diagnosis on the timeline of the diagnostic odyssey. A reduction in the number and time period of the interactions prior to the clinical diagnosis and a reduction in the gap between the clinical and molecular diagnosis would indicate improving clinical evaluation and utilisation of molecular tests in the DO, especially where specific BBS gene panel are employed. However, with increasing utilisation of agnostic exomic or genomic sequencing earlier in a patient's pathway it may become apparent that the greatest change to the DO could be achieved by wider utilisation of genomic testing. To assess whether this is indeed the case, evaluation of the DO baseline, with a workable algorithym, prior to widespread agnostic whole genome sequencing would provide an important reference point for future comparison.

ANCA-associated vasculitis

Twenty cases were selected from amongst a cohort of patients previously identified during a Trust audit of AAV care [16]. These cases included each of the three AAV subtypes and also cases which their treating clinician considered anecdotally to have had either a short or long DO. All available prior HES activity was tabulated by NHS digital for each patient. This prior HES activity was then reviewed by their treating clinician, to ascertain from their existing knowledge of the case presentation or review of medical records, whether each episode of prior secondary care activity was likely to have constituted a component of the DO. Prior to this review, secondary care activity was considered likely to be relevant to the DO if it had occurred in a specialty whose treatment function code matched to an organ system where symptoms would be likely to arise in people with AAV. This had been pre-specified based according to the clinical features that are included assessment of patients with the Birmingham Vasculitis Activity Score (BVAS) [17].

Of the 20 patients, three had no activity recorded in HES prior to the date of diagnosis. Of the 17 with HES activity, 12 had outpatient, Accident and Emergency or in-patient activity in HES in the two years preceding diagnosis. These were often spread across multiple NHS Trusts, and out of a total of 88 attendances, 79 were in the predefined relevant specialties. On review, it was considered that in 11 out of 12 patients this activity was most likely to have been associated with clinical features relevant to the subsequent diagnosis. However, this statement should not be viewed as implying that a rare disease diagnosis could or should have necessarily been made at an earlier time point in any specific individual case. Importantly, none of these previous admissions had an associated ICD code of any AAV condition prior to the stated diagnosis date, which provides additional validity of the methodology of case selection.

Summary of ANCA-associated vasculitis results

1. Approximately one third of the pilot group had no secondary care (HES) activity in the 2 years prior to diagnosis. Whilst this might be an indicator of the fact that their illness had only a very short period of symptom onset, it is more likely that their DO existed solely in primary care. This highlights the limitations of only utilising HES data, and the importance of exploring linkage with primary care data.
2. Amongst the patients with HES activity in the 2 years prior to diagnosis, the majority (90%) was in relevant specialties
3. In all except one patient, this activity was for clinical features related to the subsequent diagnosis.

Whilst the pilot has only studied one rare rheumatic disease, it is likely that these results will be applicable to other rare autoimmune diseases which are known to have similar diagnostic delays, as outlined in the RAIRDA report.

Implications of ANCA-associated vasculitis results

We have demonstrated that it is possible to study the length and tortuosity of the AAV DO in secondary care using the routinely collected NHS data (HES). Whilst this will only be relevant to the cohort of patients who do have prior secondary care activity, this pilot indicates that this is likely to be a significant proportion (40-50%) of all cases. Our findings also indicate that it is likely that prior HES activity in relevant pre-identified specialties could potentially be assumed to be related to the DO. These results pave the way for larger, repeatable studies utilising this methodology, to assess changes in the DO with time, and following specific interventions.

RECOMMENDATIONS

The following recommendations promote strategies that aim to support the implementation of the 2021 UK Rare Diseases Framework and pillar 1 of the Genome UK proposals. In particular, they highlight key points regarding methodologies for future evaluation of the DO in secondary care.

1. The most appropriate way to study the DO for specific rare diseases would be to utilise the routinely collected healthcare data for the whole of the UK. This national approach would have the power to identify an appropriate number of rare diseases cases, with no selection bias, from the whole population.

2. This methodology will require the ability to accurately identify cases, and their diagnosis date, within the routinely collected NHS data.

3. For some rare diseases, cases can be identified where there is a specific ICD coded in-patient or daycase admission, or a specific OPCS (Classification of Interventions and Procedures) coded procedure or treatment, or genetic test, where the timing of any of these is known to closely match the diagnosis date for that specific condition. For example, most cases of ANCA-associated vasculitis are diagnosed during an acute hospital admission and/or will have an infusion of a cytotoxic or biologic drug as a day case shortly after the diagnosis,

enabling case identification. This is an automated and rapid process but relies on accurate and timely coding of the patient event.

4. Case identification at national level using the above methodology, and future analysis, could potentially be undertaken within the National Congenital Anomalies and Rare Diseases Registration Service (NCARDRS), enabled by legal permissions (CAG 10-02(d)/2015). This feasibility is supported by the progress of an existing NCARDRS project (RECORDER - Registration of Complex Rare Diseases - Exemplars in Rheumatology) [18] which has validated methodologies for identification and registration of rare autoimmune diseases and their diagnosis date [19].

5. Evaluation of the DO in secondary care could be undertaken with access to HES data within NCARDRS.

6. The availability of a primary care data through linkage to GPES or CPRD would also enable evaluation of the DO in primary care and should be explored further.

7. For rare diseases where the majority of cases are diagnosed during out-patient care, which is not routinely coded, alternative approaches will be required to enable identification of cases and diagnosis date. Potential solutions include embedding the reporting of rare disease cases to NCARDRS into routine NHS clinical care. This could be facilitated by NHS England's commissioning levers, for example by including this requirement within specialised commissioning Service Specifications.

8. These methodologies would enable an assessment of variation in the DO amongst different cohorts of patients with the same disease. For example, cohorts could be assembled and compared according to:
   - Year of diagnosis, to assess changes in the diagnostic process over time
   - Location at diagnosis, to assess whether there is equity regardless of geography
   - Place of NHS care, to assess performance and variation between NHS providers
   - Timing of genomic testing within the DO

9. They could also be utilised to assess the impact of COVID-19 on people with rare diseases, by measuring rates of new diagnoses and their diagnostic pathways. A similar approach was utilised to quantify the risk of death among people with rare autoimmune rheumatic diseases during the 2020 COVID-19 pandemic compared with the general population [12].

10. Changes in the diagnostic odyssey could also be evaluated after specific interventions at either national or regional/local level. For example, this might include specific education and awareness raising amongst specific key specialties identified in the DO. This could also include quality improvement initiatives to reduce the time from laboratory tests (e.g., testing for genetic, metabolic or autoimmune abnormalities) to diagnosis and definitive treatment. These improvements would be facilitated by the development of NICE Quality Standards for rare diseases. The introduction of the National Test Directory by NHS England

should see improve nationwide access to standardised genetic testing, although access to additional technologies is required to detect low level mosaicism, which is likely to contribute significantly to the 15% of individuals with a clinical but not molecular diagnosis of TSC.

11. Recommendations about the development and use of algorithms using the routinely collected healthcare data to improve speed of diagnosis and reduce the diagnostic odyssey (a commitment within the UK Strategy) were beyond the scope of this working group. Previous studies have highlighted the difficulties of developing risk prediction models for rare diseases in primary care [17].

12. Overall, our results demonstrate the feasibility of studying the diagnostic odyssey for rare diseases in secondary care using Hospital Epsiode Statistics (HES) data. Further evaluation of this at national population level is now justified, ideally in combination with access to national primary care data. The methodology can be applied to conditions which do not have genomics data, and is therefore equally applicable to non-genetic rare diseases. This will support the vision of the UK Rare Diseases Framework for all rare disease patients across the UK to get a final diagnosis faster.

ACKNOWLEDGMENTS

We would like to thank all those who have contributed from the Department of Health and Social Care, including, Monika Preuss, Sylvia Pratt, Amy Chadwick and Emily Whamond, from NHS England, including Jane Deller and Sarah Watson and from NHS Digital, including Bettina Mavrommatis, Laura Markendale, Tia Cheang, Netta Hollings and Paul McIntosh.

REFERENCES

1 Rare Autoimmune Rheumatic Disease Alliance, 'Reduce, Improve, Empower. Addressing the Shared Needs of Rare Autoimmune Rheumatic Diseases', (2018).
2 Tuberous Sclerosis Association, 'What Is Tsc?' <http://www.tuberous-sclerosis.org/about-tsc.html>2019].
3 Nick Black, Fred Martineau, and Tommaso Manacorda, 'Diagnostic Odyssey for Rare Diseases: Exploration of Potential Indicators. Policy Innovation Research Unit (Piru)', (2013-2015).
4 Department of Health and Social Care, *UK Strategy for Rare Diseases* (2013).
5 Department of Health and Social Care, 'The Uk Rare Diseases Framework'2021) <https://www.gov.uk/government/publications/uk-rare-diseases-framework>.
6 Department of Health and Social Care, Department for Business, Energy & Industrial Strategy, Office for Life Sciences, and Lord Bethell of Romford, 'Genome Uk: The Future of Healthcare '2020) <https://www.gov.uk/government/publications/genome-uk-the-future-of-healthcare>.
7 Genomics England, '100,000 Genome Project' <https://www.genomicsengland.co.uk/about-genomics-england/the-100000-genomes-project/> [Accessed October 2018].

8 EURORDISCARE, 'The Voice of 12,000 Patients: Experiences and Expectations of Rare Disease Patients on Diagnosis and Care in Europe Survey of the Delay in Diagnosis for Eight Rare Diseases in Europe', (2006).
9 E. Forsythe, and P. L. Beales, 'Bardet-Biedl Syndrome', *Eur J Hum Genet,* 21 (2013), 8-13.
10 D. A. Krueger, H. Northrup, and Group International Tuberous Sclerosis Complex Consensus, 'Tuberous Sclerosis Complex Surveillance and Management: Recommendations of the 2012 International Tuberous Sclerosis Complex Consensus Conference', *Pediatr Neurol,* 49 (2013), 255-65.
11 C. Mukhtyar, R. Lee, D. Brown, D. Carruthers, B. Dasgupta, S. Dubey, O. Flossmann, C. Hall, J. Hollywood, D. Jayne, R. Jones, P. Lanyon, A. Muir, D. Scott, L. Young, and R. A. Luqmani, 'Modification and Validation of the Birmingham Vasculitis Activity Score (Version 3)', *Ann Rheum Dis,* 68 (2009), 1827-32.
12 E. Peach, M. Rutter, P. Lanyon, M. J. Grainge, R. Hubbard, J. Aston, M. Bythell, S. Stevens, and F. Pearce, 'Risk of Death among People with Rare Autoimmune Diseases Compared with the General Population in England During the 2020 Covid-19 Pandemic', *Rheumatology (Oxford),* 60 (2021), 1902-09.
13 F. A. Pearce, M. J. Grainge, P. C. Lanyon, R. A. Watts, and R. B. Hubbard, 'The Incidence, Prevalence and Mortality of Granulomatosis with Polyangiitis in the Uk Clinical Practice Research Datalink', *Rheumatology (Oxford),* 56 (2017), 589-96.
14 F. A. Pearce, R. B. Hubbard, M. J. Grainge, R. A. Watts, A. Abhishek, and P. C. Lanyon, 'Can Granulomatosis with Polyangiitis Be Diagnosed Earlier in Primary Care? A Case-Control Study', *QJM,* 111 (2018), 39-45.
15 F. A. Pearce, P. C. Lanyon, M. J. Grainge, R. Shaunak, A. Mahr, R. B. Hubbard, and R. A. Watts, 'Incidence of Anca-Associated Vasculitis in a Uk Mixed Ethnicity Population', *Rheumatology (Oxford),* 55 (2016), 1656-63.
16 F. A. Pearce, C. McGrath, R. Sandhu, J. Packham, R. A. Watts, B. Rhodes, R. Al-Jayyousi, L. Harper, K. Obrenovic, and P. Lanyon, 'Outcomes and Compliance with Standards of Care in Anti-Neutrophil Cytoplasmic Antibody-Associated Vasculitis-Insights from a Large Multiregion Audit', *Rheumatol Adv Pract,* 2 (2018), rky025.
17 F. Pearce, P. C. Lanyon, and R. A. Watts, 'Can Prediction Models in Primary Care Enable Earlier Diagnosis of Rare Rheumatic Diseases?', *Rheumatology (Oxford),* 57 (2018), 2065-66.
18 Fiona A. Pearce, Peter C Lanyon, and Mary Bythell, 'Using Health Data Science and Large-Scale Data to Improve the Lives of People with Rare Autoimmune Rheumatic Diseases and Other Rare Diseases 2021 https://www.hdruk.org/projects/using-health-data-science-and-large-scale-data-to-improve-the-lives-of-people-with-rare-autoimmune-rheumatic-diseases-and-other-rare-diseases/
19 FA. Pearce, M. Rutter, B. Griffiths, C. Mukhtyar, R. Al-Jayoussi, R. Watts, J. Aston, M. Bythell,, S. Stevens, P.C. Lanyon. 'Validation of Methods to Enable National Registration for Rare Autoimmune Rheumatic Diseases', *Rheumatology (Oxford), 59 (2020), 036 Issue Supplement_2, April 2020,*
20 Rare Disease UK Rare Disease Policy Board, 'The Rare Reality – an Insight into the Patient and Family Experience of Rare Disease ', (2016).
21 F. Rees, M. Doherty, P. Lanyon, G. Davenport, R. D. Riley, W. Zhang, and M. J. Grainge, 'Early Clinical Features in Systemic Lupus Erythematosus: Can They Be Used to Achieve Earlier Diagnosis? A Risk Prediction Model', *Arthritis Care Res (Hoboken),* 69 (2017), 833-41.