

Is it Possible to Implement a Rare Disease Case-Finding Tool in Primary Care? A UK-Based Pilot Study

Orlando Buendia (✉ orlando@mendelian.co)

Mendelian <https://orcid.org/0000-0002-0217-6646>

Sneha Shankar

Mendelian

Hadley Mahon

Mendelian

Connor Toal

Mendelian

Lara Menzies

Mendelian

Pradeep Ravichandran

Mendelian

Jane Roper

The maples health centre

Jag Taghar

Parsonage surgery

Rudy Benfredj

Mendelian

Will Evans

Mendelian

Research Article

Keywords: Rare disease, primary care, electronic health records, database analysis

Posted Date: August 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-638180/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Introduction:

This study implemented MendelScan, a primary care rare disease case-finding tool, into a UK National Health Service (NHS) population. Rare disease diagnosis is challenging due to disease complexity and low physician awareness. The 2021 UK Rare Diseases Framework highlights a global need for faster diagnosis to improve clinical outcomes as a key priority.

Methods & Results:

A UK primary care locality with 68,705 patients was examined. MendelScan encodes diagnostic/screening criteria for multiple rare diseases, mapping clinical terms to appropriate SNOMED CT codes (UK primary care standardised clinical terminology) to create digital algorithms. These algorithms were applied to a pseudo-anonymised structured data extract of the electronic health records (EHR) in this locality to "flag" at-risk patients who may require further evaluation. All flagged patients then underwent internal clinical review (a doctor reviewing each EHR flagged by the algorithm, removing all cases with a clear diagnosis that explains the clinical features that led to the patient being flagged); for those that passed this review, a report was returned to their GP.

55 of 76 disease criteria flagged at least one patient. 227 (0.33%) of the total 68,705 of EHR were flagged; 18 EHR were already diagnosed with the disease (The highlighted EHR has a diagnostic code for the same RD it was screened for. e.g Behcet's disease algorithm identifying an EHR with a SNOMED CT code Behcet's disease). 75/227 (33%) EHR passed our internal review. Thirty-six reports were returned to the GP. Feedback was available for 28/36 of the reports sent. GP categorised nine reports as "Reasonable possible diagnosis" (advance for investigation), six reports as "diagnosis has already been excluded", ten reports as "patient has a clear alternative aetiology", and three reports as "Other" (patient left study locality, unable to re identify accurately). All the 9 cases considered as "reasonable possible diagnosis" had a further actionable evaluation.

Conclusions:

This pilot demonstrates that implementing such a tool is feasible at a population level. The case-finding tool identified credible cases which were subsequently referred for further investigation. Future work includes performance-based validation studies of diagnostic algorithms and the scalability of the tool.

1 Background:

Rare diseases (RD) are individually rare but collectively common [1], with an estimated of 6,000-8000 RD and affecting 3.5–5.9% or 263–446 million persons globally [2]. RD are heterogeneous in aetiology, frequently chronic and debilitating [3]. There is no universal definition of a rare disease, with most legislative frameworks using point prevalence. In the UK and the European Union (EU) a rare disorder is defined as affecting fewer than 1 in 2000 persons [4,5]. 71.9% are considered genetic diseases and 69.9% have a pediatric-onset [2]. Of all RD, 149 diseases (4.2%) have a prevalence in the range of 1-5 per 10,000, these account for 77.3 - 80.7% of the total population of patients affected. Collectively RD are a significant burden to healthcare systems and society, the economic burden of 379 RD with a combined prevalence of 15.5 million in the US, was estimated to be \$966 billion in 2019. [6].

Rare disease diagnosis is challenging meaning that patients frequently remain without a correct diagnosis for extended periods. This hunt for a diagnosis has its own term: the diagnostic odyssey [7]. During this diagnostic odyssey, patients typically experience numerous primary care visits, specialist clinic reviews, investigations, invasive interventions, misdiagnoses and inappropriate treatments [8]. A cohort of patients with RD in the UK and US reported a diagnostic delay of an average 5.6 and 7.6 years respectively, with patients typically visiting eight physicians (four primary care and four specialists) and receiving two to three misdiagnoses [9]. Similarly, an EU survey reported that 40% of patients with RD were initially incorrectly diagnosed, and a quarter experienced a diagnostic odyssey of more than 5 years [10].

The reason for this diagnostic delay is multifactorial with issues related to the heterogeneity of many rare diseases, clinicians, and healthcare systems. No individual clinician can be expected to know all RD, and the adage "When you hear hoofbeats, do not expect to see Zebras" [11], describes a well held approach to considering the differential diagnosis of a clinical problem, but is not helpful for diagnosing patients with RD. Enabling clinicians, especially those in primary care, to identify unusual patterns and revisit diagnoses is crucial to reducing the diagnostic odyssey for patients with RD [12].

Early diagnosis is central to achieving better patient outcomes [13]. It enables an improved assessment of prognosis, optimization of care, access to therapies, linkage to patient organizations, easier access to social and educational support, as well as more accurate disease information [12]. It also brings clarity and understanding to the challenging, puzzling and costly diagnostic odyssey for patients and their families [14]. Furthermore, an accurate diagnosis enables the patient to contribute to the broader understanding of their disease, through for example patient registries, engagement with research and therapy development [15].

In the UK, addressing this diagnostic odyssey is the first of four key priorities in the The UK Rare Diseases Framework, published in January 2021. The priority 'Helping rare disease patients get a final diagnosis faster' is underpinned by five themes and proposes using data and digital technologies as a solution to enable a more timely diagnosis [16].

2 Methods:

Mendelian is a UK-based health data analytics company focused on shortening the diagnostic odyssey of rare and hard-to-diagnose diseases. Mendelian has developed a digital case-finding tool, "MendelScan", that can analyse structured clinical vocabulary, such as SNOMED CT codes [17] from primary care electronic health records (EHR) and highlight patterns of data that correspond to an increased likelihood of the patient being affected by certain RD. This enables the identification of those at risk and assists their clinician in accessing the correct diagnostic pathway. The "MendelScan" system is summarised in Figure 1.

The pilot study took place between January 2019 and October 2020. The primary objective was to assess the feasibility of applying "MendelScan" with seventy-six rare disease algorithms (see Appendix 1), at a small scale in a primary care environment in the Lower Lea Valley (LLV) primary care GP Federation.

The process for delivering "MendelScan" into the selected primary care federation involved establishing agreements, deploying the algorithms into a pseudonymised data set, manually reviewing the EHR identified by the algorithm, delivering the reports to GP and collecting their feedback. Figure 2 summarises the implementation process.

2.1 Primary care EHR access

2.1.1 Ethics and information governance

To enable data access and confidence in this study an independent ethical analysis of this approach was commissioned [18]. Building on the outcome and recommendations of this report, and in compliance with information governance legislation, a data sharing agreement (DSA) was agreed between stakeholders (Mendelian, Medeanalytics, East & North Hertfordshire NHS trust & Lea valley primary care network) The DSA is a contract that stipulates the rules regarding usage and handling of data. Finally, a Data Protection Impact Assessment (DPIA) was drafted. identifying and minimising the data protection risks of the project [19].

2.1.2 Data Transfer

Data transfer involved Medanalytics creating a data set of patients' EHR, removing personal identifiers. Information removed included: Names, addresses and individuals who had opted out of sharing through the national data opt-out. For the remainder

EHR a pseudonym, with a unique numeric identifier was created. This pseudonymised dataset of records was sent to Mendelian for analysis.

2.2 Algorithm deployment

Not all of the 7,000 - 9,000 rare diseases are appropriate for “MendelScan”. Mendelian developed an approach to stratify, which rare diseases are more likely to be suitable for primary care records analysis in the following steps:

1. Analysing the suitability of a RD using a scoring system based on features of the disease, the benefit of early diagnosis and the likelihood that relevant clinical characteristics would be captured in the primary care EHR.

- We developed a scoring system to identify rare diseases with a suitable profile for primary care electronic health records analysis. Scoring variables and metrics used are given in Table 1.

2. Performing a systematic literature review, searching for peer-reviewed screening or diagnostic criteria for the selected RD.

- We developed a scoring system for identified criterias to assess its suitability for primary care electronic health records analysis. Scoring variables and metrics used are given in Table 2.

3. Digitising the selected criteria into a numeric algorithm using structured data SNOMED CT codes, based on a combined scoring across several individual data points of information from the EHR. Data points are given in Table 3. We did not interrogate data held in unstructured formats (Free text) such as letters or consultation notes.

The “MendelScan” case-finding tool checked the seventy-six disease algorithms against the pseudonymised EHR data extracts flagging patients who met the algorithms’ criteria of being at risk of one of the RDs. Patients’ structured EHR flagged at risk by “MendelScan” for a rare disease, were then reviewed by a clinician and returned to the GP if a plausible alternative explanation for the clinical features could not be found.

2.3 Internal review of identified cases’ EHR

We performed an anonymous, two-round manual review process for each EHR identified by any of the seventy-six algorithms deployed. In round one, a medical doctor reviewed each EHR and assigned to each case one of three outcomes:

- Rule-in: The medical doctor considers that there is enough clinical evidence to suspect the highlighted RD for this case.
- Rule-out: The medical doctor considers there are other diagnoses recorded in the EHR that explains the highlighted features.
- Already diagnosed: The highlighted EHR has a diagnostic code for the same RD it was screened for.

In round two, rule-in cases were further reviewed by a GP, geneticist or an expert in a particular rare disease and further assigned a rule-in or rule-out outcome. For each rule-in case, a patient report was generated and sent to their GP practice. The review process is summarised in Figure 3.

2.4 Returning reports to GP

A report for each of the ‘rule-in’ patients was returned to their GP by email. The report included the unique patient identifier, to enable matching to the patient’s full EHR, an explanation of the condition, the reasons why this patient was flagged, and suggested next steps.

2.5 GP feedback on reports

Feedback from the GP was requested at two stages. The first, 'patient report feedback', was requested as soon as the GP completed evaluating the patient's report and EHR. This consisted of an online questionnaire, accessed through a link on each patient report (Appendix 2).

The first question asked the GP the main outcome of the report. See Table 4.

The second, patient outcome feedback was requested 3 months later requesting the result of those advanced for further evaluation. Figure 4.

3 Results:

Delivering "MendelScan" into a primary care locality involved a process starting with setting up the agreements through to receiving feedback from the reports sent to GPs. The main results are summarised in Figure 5.

3.1 Primary care EHR access

3.1.1 Ethics and information governance

The ethics report helped to clarify challenges and potential risks in data management for stakeholder to minimise. The information governance process (Data Sharing Agreement and the Data Protection Impact Assessment) took almost a year to complete. There were no standard documents to enable data transfer for commercial service providers in the NHS. With no standardised documents nor any previous experience of such processes in either Mendelian nor the LLV, this took longer than anticipated to agree the legal and information governance documentation.

3.1.2 Data transfer

The process of data transfer led to about a one-third reduction in the number of EHR. This was either due no clinical events being recorded in a patient's EHR, death of the patient, deregistration or issues related to the quality of the data (incomplete or an unsound record). After the data cleaning process, the de-anonymised data set included 68,705 patients' EHR.

3.2 Algorithm deployment

We analysed 259 RD, selected due to the following assessment through the disease and algorithm assessment process. 76 RD met the criteria to be deployed. All 76 algorithms were deployed successfully into the de-anonymised data set. All the algorithms identified at least one EHR. 227 (0.33% of the total population) EHR were identified in total; 18 EHR had an existing diagnostic code for the flagged RD. In the population of 152 EHR were found with a diagnostic code for any of the 76 rare diseases investigated.

3.3 Internal Review of identified EHR

75 of the 227 EHR (33%) passed our internal review process.

3.4 Send reports to GP

36 of the 75 (48%) patient reports were returned to the GP practices. The return of reports was batched to manage GP workload (a maximum of 5 reports per week per GP practice). A GP Federation Research Nurse, with access to the full EHR across multiple GP practices, did an initial review of reports for the majority of the practices in the study.

39 (52%) of reports were not delivered due to clinical pressures associated with the COVID-19 pandemic and restructuring of the primary care organisation during the study leading to further disruption and then needing to review contracting and data sharing agreements. Feedback reported that: the batching of reports was welcomed both to manage the workload and enable the review work to be planned efficiently; the reports were re-identified within the primary care data system successfully; and found to be clear and laid out with the reasons for flagging that patient and supporting evidence for this easy to interpret, enabling a quick and targeted approach to challenge/confirm the conclusions.

3.5 GP feedback on reports

Reports were returned in batches of 5, welcomed by GPs both to manage the workload and also to enable the review work to be planned easily. The initial Patient report feedback was available for 28 of the 36 reports sent to the GP. The 8 outstanding patients were lost to follow up as the due to the delays in the pilot related to the COVID-19 pandemic, and the reconfiguring of the primary care organisation leading to practices dropping out of the study locality. (Table 5).

GPs re-identified 27/28 (96%) the reports within the primary care data system, so the practice knew the names and details of each patient. Feedback from clinicians commented that the reports were well laid out and the reason for the patient flag clear, making the review in primary care straightforward.. The process of reviewing a report in primary care was between ten and thirty minutes, with disease, patient complexity and familiarity of the patient to the clinician factors influencing time to review. The follow up feedback (patient outcome feedback) is available for six of the nine patients. The three patients for whom data is not available, were lost to follow up as the practices they were registered moved from the study locality in the primary care reorganisation and no longer responded to feedback requests. Time for feedback varied from one month to eight months depending on the engagement of the primary care practice and the nature of the confirmatory test performed. For example, alpha-1-antitrypsin serum concentration levels are readily available in primary care, the 22q11 deletion syndrome took much longer as the referral from primary care to clinical genetics was further delayed due to the pandemic before testing was performed. Summary outcomes are detailed in Table 6.

4 Discussion

While this pilot was limited in scope, it demonstrated the potential of this approach in three critical areas:

1. It validated that a rare disease case-finding program could be designed within appropriate guidelines for data privacy and protection.
2. It proved the feasibility of using phenotypes documented within EHR as the basis for case finding algorithms.
3. It revealed that rare disease case finding is possible without significant disruption to the GP workflow or local specialist referral volumes.

There are two additional elements critical to the overall potential of EHR case finding that were not demonstrated in this pilot. The first is the clinical validity of the algorithms in the form of a statistically significant measure of accuracy such as positive predictive value, sensitivity, and specificity. This pilot was not powered to demonstrate this and indeed given the incidence of rare diseases, a very large sample size, possibly unachievable in a pilot or study population, would be required to determine the accuracy at conventional statistical significance of an individual disease algorithm. However, significant potential exists in future work to demonstrate the clinical validity of MendelScan's multi-disease portfolio of algorithms as a broad rare disease tool, and more the analytical validity of individual disease algorithms in large research databases. This work is ongoing.

The second element not demonstrated is the scalability of the approach. To demonstrate scalability in the future, we will need to show the ability to:

- Gain direct access to dynamic EHR in an appropriate, pseudonymised form.

- Demonstrate value to patients and efficiency for GP. Further work to explore the acceptability and value of “MendelScan” to clinicians and patients is ongoing.
- Maintain quality with reduced reliance on manual review by internal clinicians.

4.1 Challenges and barriers for digital health deployment in rare diseases.

4.1.1 Data access considerations

The current uptake of digital health solutions has historically been hindered due to a variety of reasons. One of the considerations often cited relates to information governance such as data privacy, IT systems security and data confidentiality [20]. These are partly due to technical inability to hold the data safely or due to the dynamic aspect of the legal framework.

In the current environment, it remains difficult to share and transfer patient health information, among healthcare professionals, from different organisations. This prevents a seamless care environment and breaks the promise of a potential continuity of care across the spectrum of the caregivers and predisposes the system to delays in diagnosis.

A specific lesson learned from this pilot is to establish a pathway to accelerate engagement with the final signatory stakeholders for information governance. The initial contracting process and production of legal documentation took almost a year, as there were no standard documents for data transfer for commercial service provision in the UK NHS. It is vital to establish clear criteria for final confirmation from the ultimate signatories, as commercial data access remains a relatively novel concept in the NHS and requires significant consideration.

4.1.2 Primary care EHR

To enable pseudo-anonymisation of the EHR data only structured/coded data was extracted. This structured data can be more accurate with numerical values such as weight, height, blood pressure and laboratory values readily comparable from one patient to the next. Diagnoses and clinical features are typically captured by physicians in UK primary care. Those that have been encoded and therefore appearing in a structured data extract, are typically those that are recorded with a greater degree of confidence, free text entries typically recording information when there is greater uncertainty. Consequently, one can consider the quality of this coded data as a strength, however, this means that the majority of information in the EHR was not examined by “MendelScan”. Free text diagnoses, clinical features, and the content of secondary care correspondence, often rich in information, was largely not examined. For example, secondary care correspondence often rich in pertinent information may only generate two pieces of structured data: the presence of the letter and possibly a single diagnostic code. The motivators and context for coding in primary care also needs to be considered, including the demonstration that certain tasks have been performed for reimbursement (e.g., Quality and Outcomes Framework) and to have pertinent information quickly to hand for future consultations.

4.1.3 Physician-related barriers for deployment

The implementation of any digital health tool into routine clinical practice faces challenges. However, “MendelScan” had some specific challenges related to working in the RD space and the timing of this pilot. Firstly, a lack of awareness that RD diagnosis is a primary care issue. The combined prevalence of RD, affecting 3.5-5.9% of the population [2], is not widely known and most GPs have seen a relatively small number of RD patients.

Secondly, the sheer number of rare diseases can be daunting and lead the generalist to think that awareness of their features is impossible [12]. However, many features and presentations are shared, and empowering GPs to consider rare diseases as part of their diagnostic work up is an important component to impact rare disease diagnostic rates. “MendelScan” is one such

solution presenting targeted information directly relevant to a specific patient to their clinician, equipping them with the information to take things forward.

Thirdly, the restructuring of primary care organisations during the pilot led to further disruption with contracting and data sharing agreements and responsibilities needing to change. Health care services are not static organisations and such changes need to be considered. Despite this, some of the restructuring changes such as the introductions of integrated care systems, an organisation challenged to improve public health across populations of 1-2 million, may be well placed for encouraging such rare disease case finding. Fifthly, the COVID-19 pandemic significantly affected the implementation of this pilot, protracting its roll-out, leading to additional challenges related to contracting and practices dropping out due to additional clinical pressures. Despite the negative impact of the COVID-19 pandemic on this pilot's rollout, we are optimistic that the wider user of routinely collected health data for research and quality improvement as well as the adoption of digital health technologies and the blending of these into routine healthcare pathways that has occurred during the pandemic will help facilitate the use of technologies such as "MendelScan" in the future.

Finally, one cautionary lesson learned is that it is likely that the primary care federation data chosen for this pilot is better structured and the data stakeholders and system architecture more ready for the Mendelian approach than other primary care regions. More work needs to be done in prioritizing which primary care healthcare systems have the capabilities to adopt "MendelScan" as a case-finding tool for RD.

5 Conclusions:

This pilot demonstrates that implementing a novel digital case-finding tool, "MendelScan", for RD in UK primary care is feasible, with minimal impact on workload or system resources. "MendelScan" identified credible cases, which subsequently were investigated in primary care or referred for further investigation aiming for earlier diagnosis. This study also highlighted challenges in implementing such a tool, including the restructuring of NHS organisations and shifting priorities due to outside pressures such as the COVID-19 pandemic.

Further research is ongoing, in the form of retrospective and prospective studies focusing on v evaluating MendelScan's analytical and clinical validity. Additional studies to evaluate the cost-benefit of early diagnosis and the impact of "MendelScan" on clinical practice and the acceptability and perception of this novel approach among specialists, GP and patients are in process or planned.

A digital health approach, such as "MendelScan", could be an invaluable tool to address the rare disease diagnostic odyssey, flagging those potential zebras amongst the horses. In addition, through the use of routinely collected EHR data, it can be scaled quickly and adopted broadly helping to ensure equality of access to a correct diagnosis.

Declarations

- Ethics approval and consent to participate: "Not applicable", this study was performed as part of a quality improvement exercise as part of routine care and therefore ethics approval was not deemed necessary.

- Consent for publication: "Not applicable".

- Availability of data and materials: "Data sharing not applicable to this article as no datasets were generated or analysed during the current study".

- Funding: All funds were covered by Mendelian Ltd

- Competing interests: "The authors declare that they have no competing interests".

- Authors' contributions: OB created the algorithms and drafted the manuscript. SS drafted the manuscript. HM edited the manuscript. CT edited the manuscript. PR edited the manuscript, LM edited the manuscript, RB managed the data sharing

agreement. WE participated in its design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

- Acknowledgements: The success and final outcome of this implementation pilot required the collaboration of Lower Lea Valley primary care network, Eastern Academic Health Science Network, East and north Hertfordshire clinical commissioning group CCG and Innovate UK. J This implementation pilot wouldn't be possible without the help and contribution of all Mendelian staff. We want to express gratitude to the experts for some particular rare disease who gave us their support through the algorithm development such as Dr Derralynn Hughes, Dr Robin Lachmann and Dr Lakshminarayan R Ranganath.

6 References:

1. Somanadhan S, Nicholson E, Dorris E, Brinkley A, Kennan A, Treacy E, Atif A, Ennis S, McGrath V, Mitchell D, O'Sullivan G, Power J, Lawlor A, Harkin P, Lynch SA, Watt P, Daly A, Donnelly S, Kroll T. Rare Disease Research Partnership (RAinDRoP): a collaborative approach to identify research priorities for rare diseases in Ireland. *HRB Open Res.* 2020 Nov 11;3:13. doi: 10.12688/hrbopenres.13017.2. PMID: 33299965; PMCID: PMC7702160.
2. Nguengang Wakap S, Lambert DM, Olry A, Rodwell C, Gueydan C, Lanneau V, Murphy D, Le Cam Y, Rath A. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *Eur J Hum Genet.* 2020 Feb;28(2):165-173. doi: 10.1038/s41431-019-0508-0. Epub 2019 Sep 16. PMID: 31527858; PMCID: PMC6974615.
3. Richter T, Nestler-Parr S, Babela R, Khan ZM, Tesoro T, Molsen E, Hughes DA; International Society for Pharmacoeconomics and Outcomes Research Rare Disease Special Interest Group. Rare Disease Terminology and Definitions-A Systematic Global Review: Report of the ISPOR Rare Disease Special Interest Group. *Value Health.* 2015 Sep;18(6):906-14. doi: 10.1016/j.jval.2015.05.008. Epub 2015 Aug 18. PMID: 26409619.
4. Moliner AM, Waligora J. The European Union Policy in the Field of Rare Diseases. *Adv Exp Med Biol.* 2017;1031:561-587. doi: 10.1007/978-3-319-67144-4_30. PMID: 29214592.
5. The UK Rare Diseases Framework [Internet]. GOV.UK. 2021 [cited 2021 May 18]. Available from: <https://www.gov.uk/government/publications/uk-rare-diseases-framework/the-uk-rare-diseases-framework>
6. [Internet]. EverylifeFoundation.org. 2021 [cited 2021 May 26]. Available from: https://everylifefoundation.org/wp-content/uploads/2021/02/The_National_Economic_Burden_of_Rare_Disease_Study_Summary_Report_February_2021.pdf
7. Kuiper, GA., Meijer, O.L.M., Langereis, E.J. *et al.* Failure to shorten the diagnostic delay in two ultra-orphan diseases (mucopolysaccharidosis types I and III): potential causes and implications. *Orphanet J Rare Dis* 13, 2 (2018). <https://doi.org/10.1186/s13023-017-0733-y>
8. Zurynski, Y., Deverell, M., Dalkeith, T. *et al.* Australian children living with rare diseases: experiences of diagnosis and perceived consequences of diagnostic delays. *Orphanet J Rare Dis* 12, 68 (2017). <https://doi.org/10.1186/s13023-017-0622-4>
9. [Internet]. Globalgenes.org. 2021 [cited 2021 Aug 2]. Available from: <https://globalgenes.org/wp-content/uploads/2013/04/ShireReport-1.pdf>
10. [Internet]. Eurordis.org. 2021 [cited 2021 May 25]. Available from: https://www.eurordis.org/IMG/pdf/Fact_Sheet_Eurordiscare2.pdf
11. Sotos John. Dr. Zebra. The Celebrated Web Site. The clinical medicine classic Zebra Cards. <http://www.sotos.com/> (accessed 21 May 2021). [Google Scholar] [Ref list]
12. Evans WR, Rafi I. Rare diseases in general practice: recognising the zebras among the horses. *Br J Gen Pract.* 2016 Nov;66(652):550-551. doi: 10.3399/bjgp16X687625. PMID: 27789486; PMCID: PMC5072891.
13. Laraway S, Breen C, Mercer J, Jones S, Wraith JE. Does early use of enzyme replacement therapy alter the natural history of mucopolysaccharidosis I? Experience in three siblings. *Mol Genet Metab.* 2013 Jul;109(3):315-6. doi: 10.1016/j.ymgme.2013.04.023. Epub 2013 May 10. PMID: 23721889.
14. Kohlschütter A, van den Bussche H. Frühzeitige Diagnose einer seltenen Krankheit bei Kindern durch bessere Kommunikation zwischen Eltern, niedergelassenen Ärzten und spezialisierten Zentren [Early diagnosis of a rare disease in

- children through better communication between parents, physicians and academic centers]. *Z Evid Fortbild Qual Gesundheitswes.* 2019 May;141-142:18-23. German. doi: 10.1016/j.zefq.2019.02.008. Epub 2019 Mar 14. PMID: 30878243.
15. Ng DM, Hooper AJ, Bellgard MI, Burnett JR. The role of patient registries for rare genetic lipid disorders. *Curr Opin Lipidol.* 2018 Apr;29(2):156-162. doi: 10.1097/MOL.0000000000000485. PMID: 29351106.
16. Assets.publishing.service.gov.uk. 2021 [cited 2021 May 26]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/950651/the-UK-rare-diseases-framework.pdf
17. SNOMED CT - NHS Digital [Internet]. NHS Digital. 2021 [cited 2021 Jun 3]. Available from: <https://digital.nhs.uk/services/terminology-and-classifications/snomed-ct>
18. Mendelian.co. 2021 [cited 2021 Jun 17]. Available from: <https://www.mendelian.co/pdf/ethics-report-ph-mendelian.pdf>
19. Data protection impact assessments [Internet]. Ico.org.uk. 2021 [cited 2021 Jun 3]. Available from: <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/data-protection-impact-assessments/>
20. Williams MG, Stott R, Bromwich N, *et al*/Determinants of and barriers to adoption of digital therapeutics for mental health at scale in the NHS *BMJ Innovations* 2020;**6**:92-98.

Tables

Table 1 - Variables for RD suitability of primary care records analysis.

Disease scoring	
Variable	Metric
Is the disease multisystemic? (\geq than three organ systems involved)	Yes / No
Is there an absence of significant mortality ($>30\%$) before five years of age?	Yes / No
Does the disease have a high specialized service pathway in the NHS?	Yes / No
Does the disease have signs and symptoms that are progressive and potentially missed?	Yes / No
How many clinical features are likely to be encoded in the primary care EHR?	# of clinical features
What is the prevalence per 100,000.	prevalence

Table 2 - Variables for criteria suitability of primary care records analysis.

Criteria scoring	
Variable	Metric
Has a criterion been identified in the r literature review?	Yes / No
Does the criterion have findings that will be captured in structured data in the primary care EHR?	Yes / No
Does the criterion use a classification/scoring system?	Yes / No
Do the specific features/findings at the criterion are explained in the peer-reviewed article?	Yes / No
Are there further studies that validate the performance of the criterion?	Yes / No

Table 3 - Electronic health record (EHR) code types used.

EHR code type	Description
Vitals	Physiological values such as blood pressure, weight, height, and BMI
Demographics	Patients demographic information such as age, sex and ethnicity
Problem list	Patients list of active medical issues
Past medical history	Diagnosis of chronic conditions diagnosed previously
Diagnosis	Diagnosis and diagnostic codes
Referrals	Referral ordered and admissions
Medications	Medications ordered and currently taken
Lab results	Numerical results of any laboratory tests

Table 4 - Patient report feedback.

Please indicate which options below best describe this report. Tick box multiple answer questions.
Reasonable possible diagnosis (Advance for further evaluation)
Diagnosis has already been excluded (Disease highlighted in the report was already studied)
Patient has clear alternative diagnosis that explains the clinical features flagged
Patient has left GP practice
Unable to accurately identify patient

Table 5 - Patient report feedback results

	Reasonable possible diagnosis (advanced for investigation)	Diagnosis has already been excluded	GP believes the patient has a clear alternative aetiology	Other (Patient no longer at the practice or unable to correctly identify patient record)
Number of EHR	9	6	10	3

Table 6 - Patient outcome feedback results

Patient EHR	Suggested disease	Action taken by GP	Outcome
1	Classic Homocystinuria	Discuss with patient	Not available
2	Fabry Disease	Discuss with patient	Not available
3	Alpha-1 antitrypsin deficiency	Alpha-1-antitrypsin level	Normal
4	Alpha-1 antitrypsin deficiency	Alpha-1-antitrypsin level	Normal
5	Loeys-Dietz syndrome	Referral to cardiologist	Seen by Cardiologist, pending cardiologist letter
6	Loeys-Dietz syndrome	Referral to cardiologist	Seen by Cardiologist, pending cardiologist letter
7	Beckwith-Wiedemann syndrome	Discuss with patient	Not available
8	Alpha-1 antitrypsin deficiency	Alpha-1-antitrypsin level	Normal
9	DiGeorge Syndrome - 22q11 deletion syndrome	Referral to clinical geneticist	- Seen by Clinical geneticist - Microarray negative

Figures

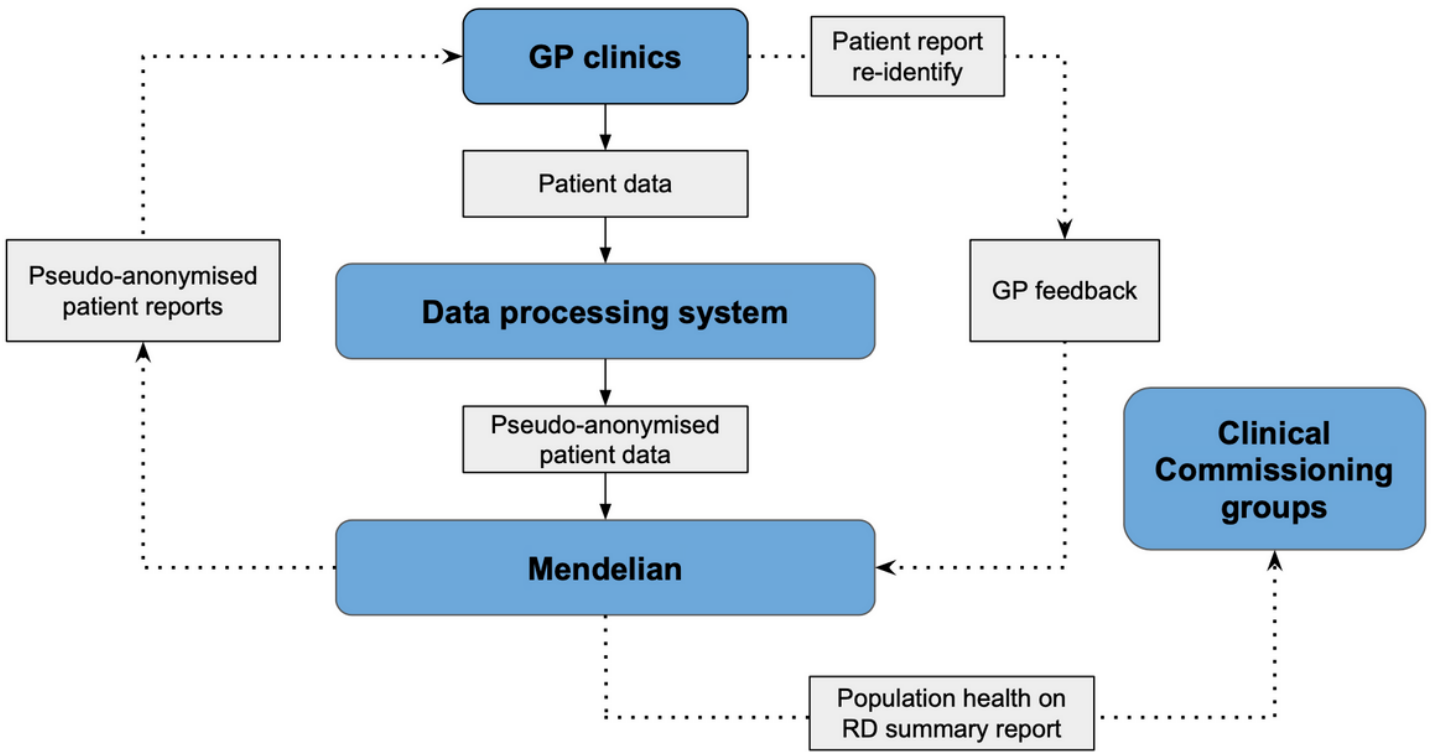


Figure 1

“MendelScan” data flow and systems integration

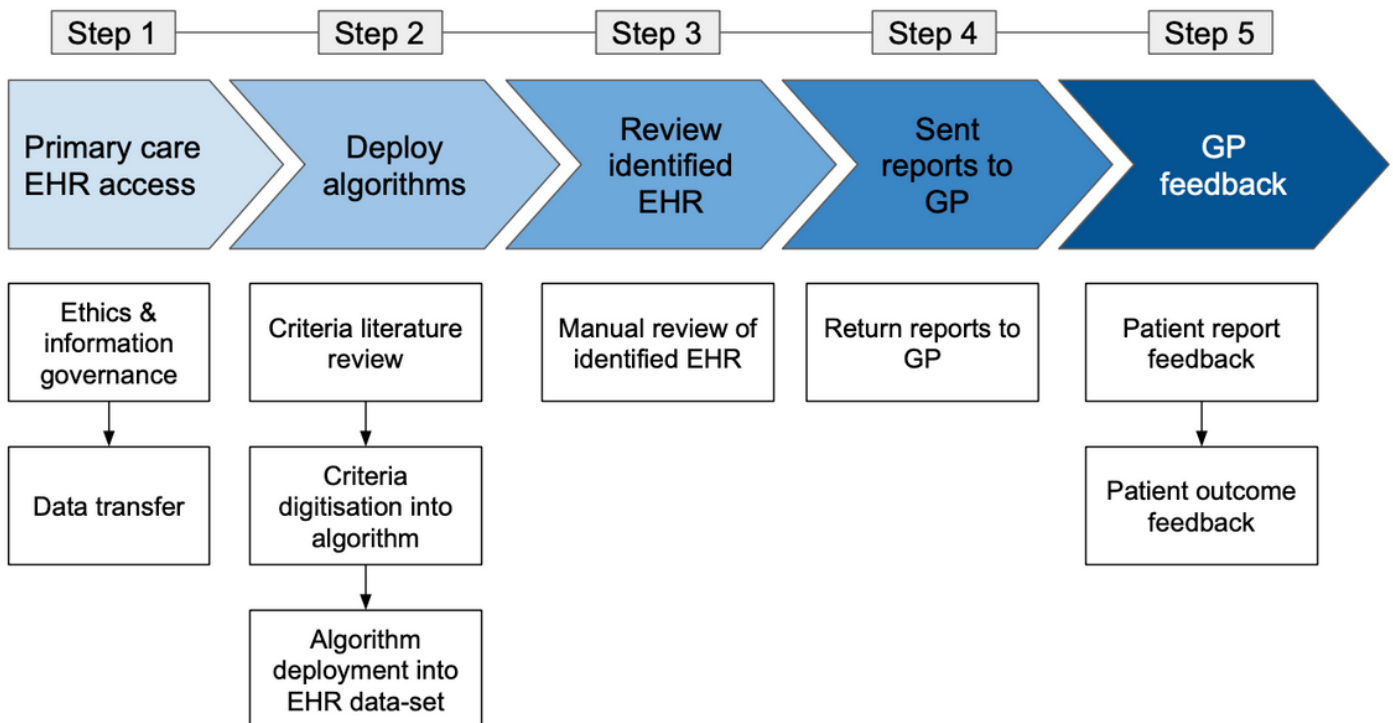


Figure 2

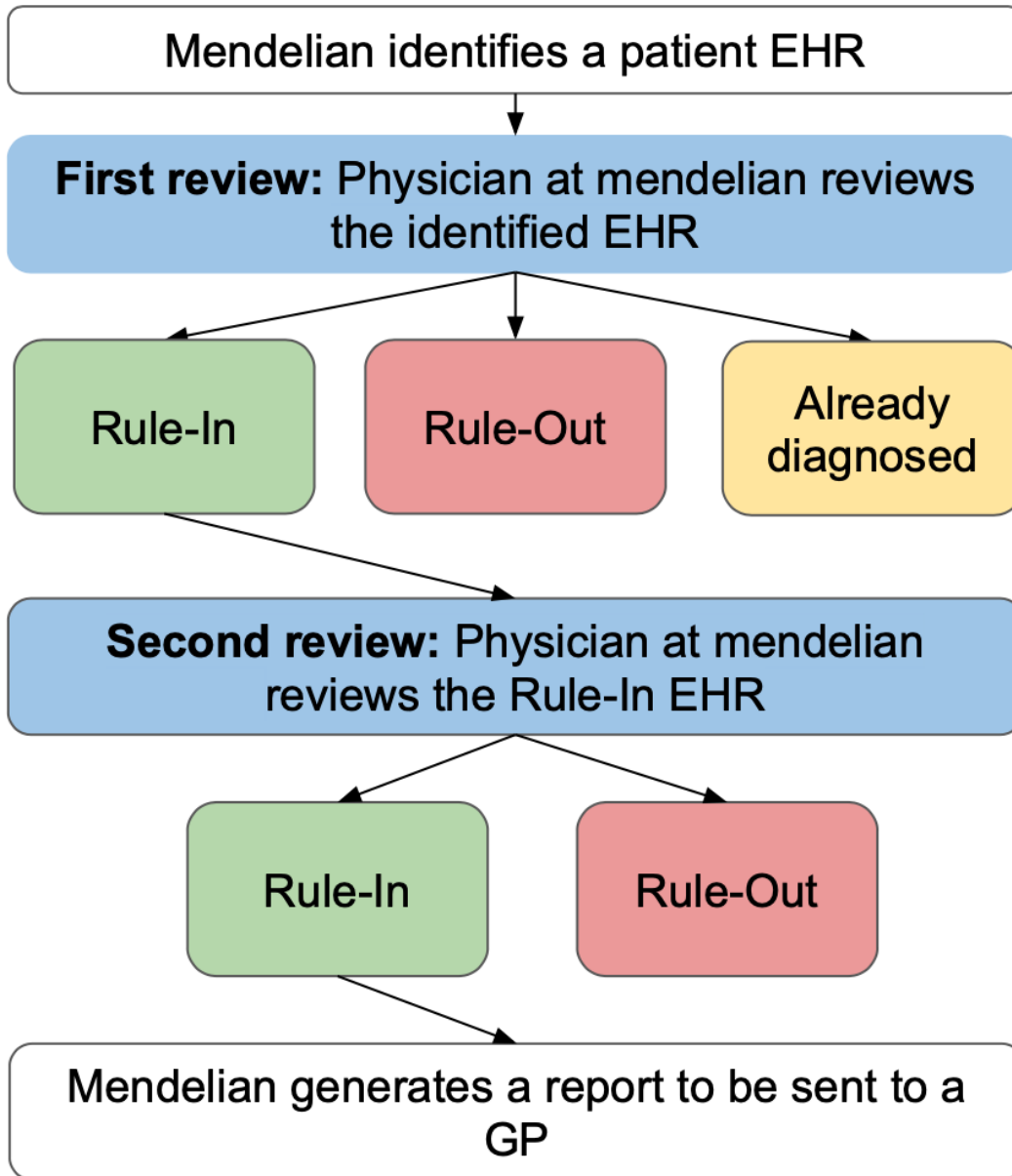


Figure 3

Internal review process methods.

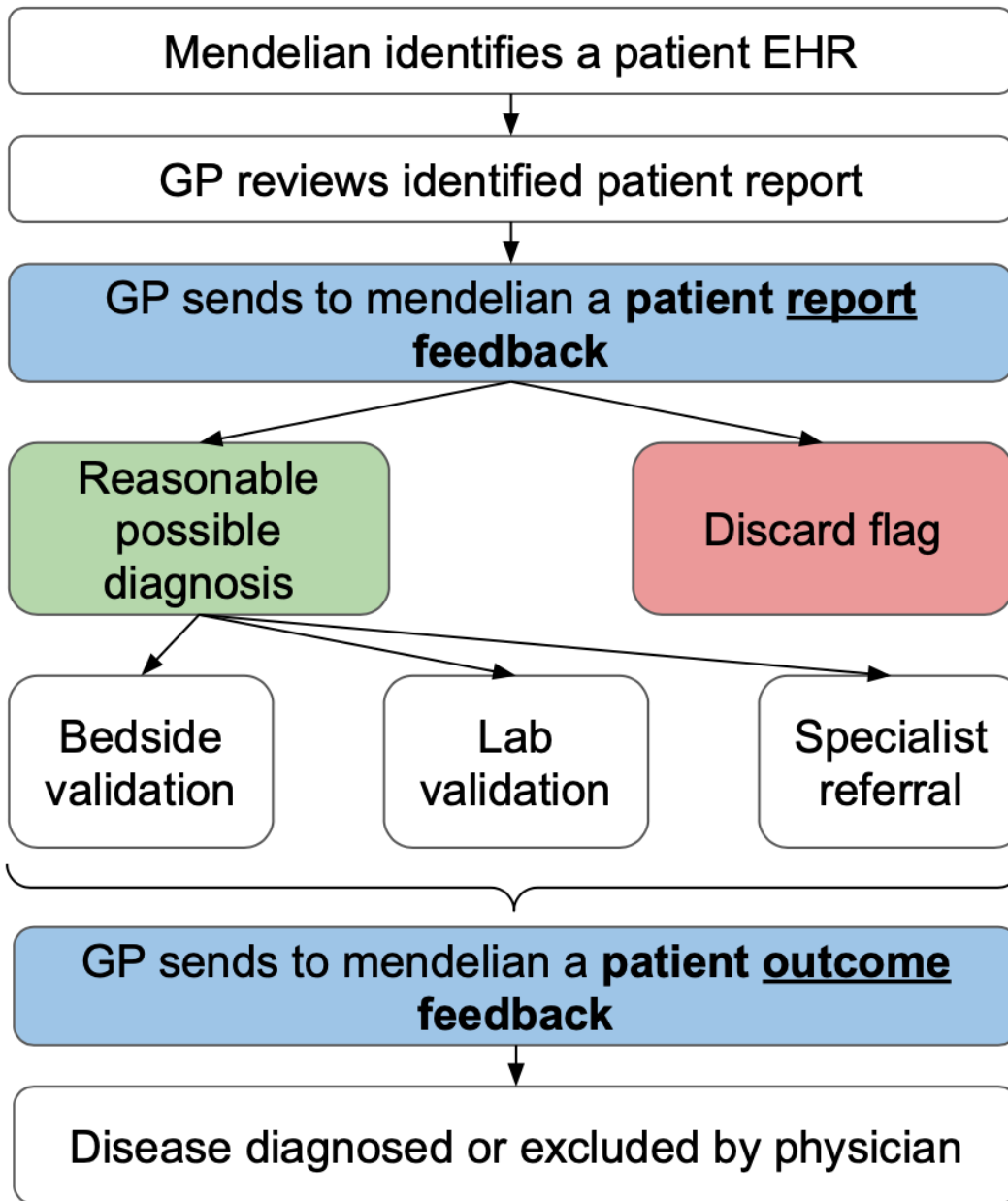


Figure 4

Feedback flowchart.

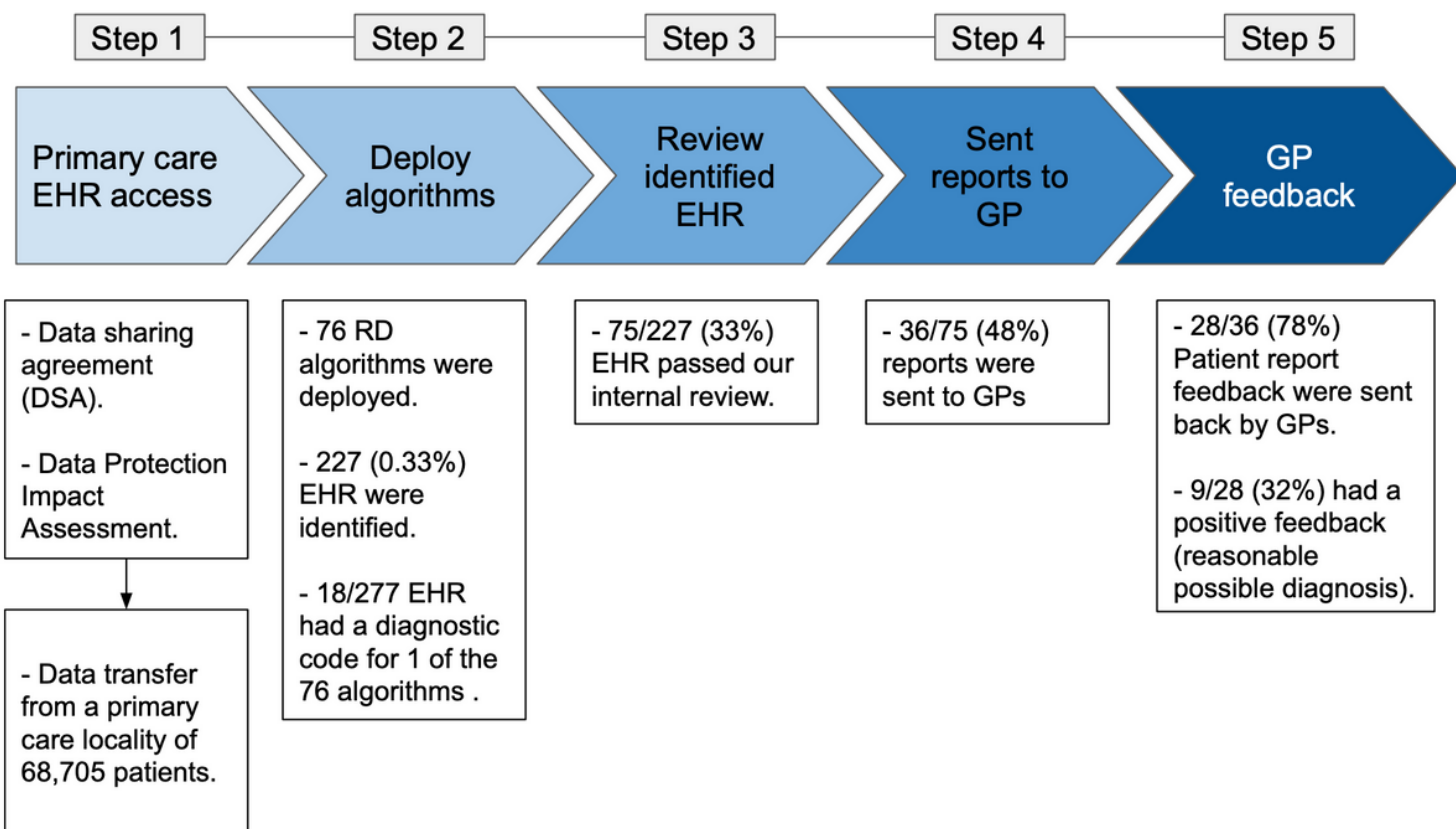


Figure 5

Results overview of RD "MendelScan" implementation steps.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Appendix.docx](#)