The Future of Healthcare: Computerisation, Automation, and General Practice Services

Abstract

Background:
Despite widespread concern regarding new technology replacing jobs, or how technology will change the structure of jobs, we lack detailed real-world evidence about what can and cannot be automated. This lack of understanding can be confusing and dangerous for policymakers who often see healthcare-oriented roles as “low risk” for automation. However, within UK primary healthcare, automation is not regarded as a threat to staff, but as a necessary solution to tackle many issues it currently faces, such as staff shortages, increased demand and reduced budget to name a few.

Objective:
The purpose of this research is to understand the current state of automation, and the potential opportunities and challenges of further automation in the UK primary healthcare sector. Our first aim is to observe and collate a comprehensive understanding of what occupations and tasks occur in primary healthcare practices. No such similar dataset currently exists for analysing the work practices and tasks of NHS primary care staff. Our second aim is to use expert knowledge of the current state-of-the-art automation technologies as a guide to estimate what tasks and work practices could be susceptible to further automation opportunities going forward, and the potential effects on healthcare workflows.

Methods:
This project utilises a multi-method and mixed-method research design, comprising of two phases: a qualitative observational phase, and a quantitative data analysis phase; each phase addressing one of the two project aims. A critical part of the problem we propose to address is that a formal framework for measuring automation is somewhat lacking in the literature. The healthcare domain offers a further challenge in measuring automation because of a general lack of detailed, healthcare-specific, occupation and task observational data to form good insights about this notoriously misunderstood topic. Therefore, our first aim focuses on addressing this lack of data by collecting high quality, detailed task-specific data from UK primary healthcare practices. The second aim then proposes a formal framework for probabilistic inference of task and occupation-automation to gain valuable insights.

Results:
Our detailed fieldwork includes observing and documenting 16 unique occupations performing over 160 tasks across 7 primary care centres. Our initial findings are that tasks are often shared amongst staff and can include convoluted workflows which often vary between practices. The single most-used technology in primary healthcare is the desktop computer. Secondly, we have conducted a large-scale survey of 156 machine learning and robotics experts to assess what tasks are susceptible to automation given the current state-of-the-art technology available today. Further results and detailed analysis will be published towards the end of the project.

Conclusions:
To date, tasks observed in primary healthcare can be categorised into two types: structured or unstructured. We believe our analysis, will identify many of the structured tasks to be highly susceptible to automation using current technology, given that sufficient data can be collected. Further, we aim to identify work-
flows that, given the proper implementation of automation technologies, could save considerable staff resources.

**Keywords:** Qualitative Research; Supervised Machine Learning; Automation; Interdisciplinary Research; Task Performance And Analysis

**Introduction**

Computerisation and automation are rapidly changing employment practices across many sectors of the UK economy. The progress of advancements during the digital age has seen new technologies replacing and augmenting human labour at a diverse range of tasks, reshaping the experience of millions of customers and employees. For example, Amazon have recently launched a grocery store, Amazon Go, that uses computer vision to eliminate cashiers, of which over 3.5 million are employed in the United States [1,2]. In addition, Amazon and others also harness intelligent material moving robots to work alongside the 2.5 million freight and stock hand-labourers in warehouses and commercial buildings [3]. Looking ahead, continued progress in state-of-the-art machine learning and mobile robotics will cause further disruption to knowledge-based occupations that were previously thought of as less susceptible to automation. However, despite widespread concern regarding new technology replacing jobs, or how technology will change the structure of jobs, we lack detailed real-world evidence about what can, and cannot, be automated. This lack of understanding can be confusing and dangerous for policymakers who want to set effective policy to mitigate the consequences and foster potential benefits.

Dominant frameworks for measuring automation have previously focused on different “types” of occupations and the skills that are required to perform them[4–6]. These works conclude that occupations with so-called “routine” tasks are the most susceptible to automation; and specifically, that manual occupations are easier to automate than cognitive or knowledge-based occupations. Researchers at the University of Oxford analysed the U.S. Department of Labor Occupational database (O*NET) and found 47% of U.S. employment highly susceptible to automation over the next decade or two [4]. They propose a probabilistic machine learning approach using numerical occupation features that represent “bottlenecks” to computerisation and analysed over 700 occupations, producing an estimate of the probability of computerisation for each. Multiple follow-up studies applied these probabilities to other country’s employment data (assuming that an occupation’s risk of computerisation is comparable across countries). Such studies include a report from Deloitte [7] that identified 35% of current UK employment as being at high risk of becoming automated over the same time period; and a paper from the Bruegel Think Tank [8] estimates the share of jobs at high risk across Europe to range between 45 to more than 60%, with southern European workforces, e.g. Portugal and Romania, facing the highest exposure to potential automation.

From these studies, healthcare-oriented roles are often estimated to be at low risk of automation. This is, in part, due to many healthcare tasks requiring a high level of skills that align with the bottlenecks to computerisation identified in [4], e.g. assisting and caring for others, manual dexterity, social perception, originality, negotiation and persuasion. A secondary reason for these low risk estimates is a general lack of empirical data describing work practices, work flows and the skills required to perform many healthcare roles, in what is a largely interrupt-driven environment – something we have observed first hand in the fieldwork conducted.
Automated technologies are often speculated to target or displace vulnerable, low-skilled workers. However, healthcare is one of the few economic sectors where automation is seen as a huge opportunity [9,10]. Specifically, the UK NHS primary care system currently faces numerous building pressures including: staff shortages, increasing workloads, increased demand, reduced budget, skill shortages, and decreased patient consultation time [10–12]. It is generally accepted that automation may address some of these pressures. However, there is also a potential threat that through increased automation of tasks, the roles performed by healthcare staff will need to be reconfigured which may ultimately affect the patients’ relationship with their general practitioner.

This project turns the lens to look at automation in NHS primary care, analysing in more detail than ever before, the potential of automation on primary healthcare professional’s tasks and workflows. A key aspect of our approach is that we start with tasks, rather than occupations, to understand what technically can be automated and how an occupation’s work might be impacted as a result. By collecting granular task-level data, we believe we capture a more accurate effect of automation since it is tasks, rather than entire occupations, that are automated by new technologies. We also believe this provides the most valuable real-world policy insights, with recommendations over entire workflows potentially saving considerable resources.

Health sector automation will undoubtedly involve some routine tasks like scheduling or some laboratory tasks, it is also likely to involve technologies that are uniquely developed, and as yet in their infancy. Current applications of automation in healthcare is a rich and well covered topic with decades of history. The literature includes examples such as electronic medical records; personal health records; remote test ordering and repeat prescriptions; check-in and booking systems; patient access to appointment systems; telehealth and telemedicine systems; physician order entry; much of the pharmacists work [13]; automating the collection of data from patients in the waiting room [14]; reducing provider to provider communication [15]. Additionally, countless different software systems each have an element of automation, e.g. computerized physician order entry (CPOE) is a decade old technology that has helped to automate workflows such as requesting lab work, checking for allergies, and electronic prescribing [16]. The key question is what is meant by “automation”. Clearly, a sophisticated CPOE system can automate some, if not most, of a clinician’s work increasing efficiency. But it is not completely automated, and a human clinician is still required to keep notes, converses with colleagues, works directly with the patient, reference materials, and likely works in other systems besides CPOE. This example gets at the core of our study: there are automated and partially automated systems to be found throughout primary care, yet work remains in understanding their effects on occupations in order to provide further policy and workflow recommendations.

Prior work also exists in the area of mobile robotics in healthcare. Areas such as robot assisted logistics, telepresence and companion robots, education and communication robots, motivational persuasive robots, ageing society robots, and home assistance robots all introduce automation healthcare [17]. Although many robotic approaches are under development for the healthcare domain, our study does not focus on recommending what, if any, specific type of technology will automate a task, we merely seek to understand to what extent healthcare tasks are technically possible to automate using currently available technology, and interpret and disseminate the effects of this automation in the healthcare domain.

This 2-year study is organized by the two aims of this protocol: First, to observe and collate a comprehensive understanding of what occupations and tasks occur in primary healthcare practices. Second, to use expert knowledge of the current state-of-the-art automation technologies as a guide to estimate what tasks and work practices could be susceptible to further automation opportunities going forward.
Methodology

Our approach constitutes a multi-method and mixed-method research design: Aim 1 can be regarded as multi-methodological in that we use multiple techniques for gathering data within each paradigm; jointly, the two aims together can be seen as a mixed-method, as both quantitative and qualitative approaches to data collection and analysis, are used throughout [18]. A critical part of the problem we intend to address is that a formal framework for measuring automation is somewhat lacking in the literature. This is compounded in the healthcare domain by a general lack of detailed, healthcare-specific, occupation and task data to form good policy insights. Therefore, our first aim focuses on addressing this lack of data by collecting qualitative, high-quality, detailed task-specific data from multiple UK primary healthcare practices. The second aim then proposes a formal, quantitative framework for probabilistic inference of task and occupation-level automation.

Figure 1: Study design

AIM 1
We first address the general lack of detailed healthcare data that is situated in the NHS health system. Our aim is to observe work practices and collect data about how each occupation performs tasks in NHS primary care and general practice services. This detailed and rich data collection is guided by interviews, document collection, photographs, detailed field notes, and shadowing each occupation in general practice health centres across England. This data is then qualitatively analysed and organised into a dataset that supports the second aim of the project: analysing the data using machine learning techniques to infer probabilities of automation.

Field Work
To understand the work practices of all primary care staff, from partner general practitioners to receptionists, we employ an ethnographic method to: observe situated practices, ask questions, gather documents, write detailed field notes and catalogue each occupation with as much clarity as possible. Time spent on site at each health centre ranges from three days to over a week. Prior to starting the fieldwork, the field researcher will work with practice staff to build a schedule where time can be made available with each occupational type. Since this project is interested in the tasks each occupation performs, we do not need to observe every member of staff, but a representative subset if there are multiple employees of each occupation type. In developing this schedule, time will also be made available for the field researcher to attend GP meetings, chronic disease clinics, and other special events that showcase other occupational tasks of primary care. To date we have recruited seven practices with task data collected on every
occupation type in each practice; adequate to reach saturation of observed unique occupation and task descriptions.

The field researcher will focus on four streams of data collection when in the field. Including:

1. Observation of day-to-day work and tasks performed by staff members. This includes asking detailed questions and/or behavioural queries to understand specific skills required to accomplish tasks; or the description of specific computer use and software configuration; or the specific order in which filing must be performed. That is, any details about identified routine tasks;
2. Collection of documents such as training manuals, job description documents, policy and protocol manuals and other organisational documents that describe work tasks and how the practice is to be run. This can extend to photographs of documents or information scattered throughout the practice, in order to understand how work and tasks are documented and distributed;
3. Photographs of work spaces and of manual physical tasks in the general practice;
4. Audio recorded discussions of work processes or tasks taking place, and any specific required skills necessary to perform day to day tasks.

**Focus Groups**

At the end of the field researcher’s observations and initial data collection the field researcher will conduct a focus group with all primary care staff at the facility, given staff availability. The focus group, while providing additional data, also serves as a validation technique. During the focus group the field researcher shows representations and descriptions of tasks that were observed during fieldwork. They will enquire if the tasks are being represented accurately, and if not, what additional information is needed for an accurate portrayal of that occupation’s work. In addition to task validation, the focus group allows for discussion of healthcare professionals perceived benefits, opportunities, and challenges to automation of work in primary care. The field researcher also presents several different scenarios that involve the automation of different types of work in the health centre. These scenarios are intended to generate discussion between participants in the focus group about potential changes in work due to automation.

**Task validation and workload measurement survey**

A second validation technique utilised in this study is through the distribution of a survey to support the accuracy of task descriptions for each occupation; and to provide rating of how automating a task would impact workload. The survey uses tasks gathered through the fieldwork, focus groups, and interviews described in the first aim. Survey respondents are shown the set of tasks that we understand are commonly performed by their occupation. For a subset of tasks, they are asked “If it were possible to fully automate the above task, i.e. entirely performed by a computer or a collection of technologies, how would it influence your daily workload?” We provide the following categories:

1. I do not perform this task.
2. There would be no change in my workload.
3. There would be little change in my day to day workload, and would not save much time.
4. Automating this task would provide me time to work on other tasks in my workload.
5. Automating this task would eliminate a core aspect of the work identified in my job description.

**AIM 2**

The second aim of the project is to develop a formal, quantitative framework for inferring the automation potential of tasks and occupations observed in aim 1. First, we augment the collected healthcare task dataset with existing high-dimensional data about skills, knowledge, and abilities required to perform each task (120 numeric variables). Second, we conduct the largest and most comprehensive survey of machine
learning, robotics and Artificial Intelligence (AI) experts to elicit expert estimates regarding the current state of automation of real-world tasks, not specifically restricted to the healthcare domain. Third, the estimates are used to train a probabilistic machine learning model to identify patterns connecting task automation potential to the occupation and task characteristic variables. The three steps: augmenting task dataset, automation expert survey, and developing the machine learning model, are discussed in detail below.

Augmenting Task Dataset
The detailed observational and qualitative healthcare-specific data captured in Aim 1, is transformed into a matrix of occupational roles and the tasks performed, with each row representing a unique occupation-task pair. We then augment these identified occupation-task pairs with numeric attributes from a publicly available occupational survey produced for the US Department of Labor called the Occupational Network (O*NET) 2016 database. O*NET provides key features of an occupation as a standardised and measurable set of variables, and also provides open-ended descriptions of specific tasks each occupation performs; its strengths and weaknesses are reviewed in [19]. The database contains information on more than 1,000 US occupations using a modified form of the Standard Occupation Classification (SOC) system; comprising of over 2,000 detailed work activities and nearly 20,000 individual occupation-specific tasks arranged into a hierarchical structure. A simplified hierarchy of the O*NET taxonomy is presented in Fig 2.

The O*NET occupational variables include:
- 35 skill attributes, such as “coordination”, “critical thinking” and “time management”;
- 33 knowledge attributes, such as “mathematics”, “clerical” and “sales and marketing”;
- 52 abilities, such as “depth perception” and “speech recognition”.

The variables described as “bottlenecks to computerisation” used in previous literature [4] are a subset of the 120 O*NET variables used in our study.

The numerical attributes are designed to provide an accurate representation of an exemplar employee within each O*NET occupation, where each occupation is also represented by the collection of tasks they are required to perform. We assume that an occupation’s skills, knowledge, and abilities informs those needed to perform the occupation’s list of tasks. Next, we aggregate the occupation variables into work activity variables by taking a weighted average of an activity’s tasks, normalising over the combined weight of the task’s relative importance to its occupation and to its work activity.

We manually match the observed healthcare tasks to their corresponding “work activities” within the O*NET hierarchy, allowing for a one-to-many weighted mapping. This allows for the observed healthcare tasks to be augmented with high quality O*NET variables regarding the level of skills, knowledge and abilities required to perform such tasks. The result of these manipulations is that each observed healthcare occupation-task is represented by a vector of 120 numeric attributes, which are vertically stacked to become a training data matrix for our proposed machine learning model.
The second step is the elicitation of expert knowledge of state-of-the-art automation technologies. In order to obtain estimates about how automatable our healthcare specific tasks are, we surveyed machine learning, robotic, and AI experts at the forefront of research and commercially available technology. The survey is designed so each participant is presented with five O*NET occupations, (with occupations chosen to be representative of the feature space, with an emphasis on high employment and hence familiar occupations). Survey participants are asked to rate how automatable the five "most important" tasks are (task importance is relative to occupation, as defined in O*NET). We frame the survey by asking the following question: "Do you believe that technology exists today that could automate these tasks?", participants rate each task on the following scale:

0. Unsure
1. Not Automatable Today
2. Mostly Not Automatable Today (Human Does Most of It)
3. Could be Mostly Automated Today (Human Still Needed)
4. Completely Automatable Today

Our demographic is specifically technology experts, as opposed to healthcare experts, because we believe that annotating tasks requires little-to-no subject matter knowledge. If respondents felt any doubt in their ability to assess the automatability of a task, they can select the "Unsure" option. Survey responses and demographics can be seen in Fig 3.

We aggregate the categorical responses for each task into a continuous numeric variable representing the automation potential of the work activity level (corresponding to the tasks' parent in the O*NET hierarchy – see Fig 2.). This can be computed by taking a weighted average of the responder scores valued 1 to 4, and...
twice-normalising the task-importance score. First, for each task across its corresponding occupation, and second, across the work activity. A score of 4 represents a fully automatable work activity, and 1 represents an activity that cannot be automated using currently available technology.

We believe this expert survey provides a robust ground truth estimate as to what extent activities are automatable using currently available technology. One important note is that the survey results provide a measure of what can be automated using technology, with no prediction of future technological advancements, i.e. not what necessarily will be automated given technology uptake or societal pressures.

**Machine Learning Model**

Finally, we use a machine learning framework to learn a functional mapping between the skills/knowledge/ability feature vectors of a work activity and the ground truth automation scores elicited from our expert survey. Gaussian processes [20] are a particular modelling tool that have a natural advantage in this scenario and also offer advantages to policymakers, such as providing formal estimate of uncertainty contained in the model. The algorithm uses the trends and patterns it has learned from labelled data to provide a smoothly varying, probabilistic assessment of automatability as a function of the input variables. For the Gaussian process, this function is non-linear, meaning that it flexibly adapts to the patterns inherent in the training data.

We train the Gaussian process model on 314 work activities present in O*NET, for which we had expert annotations. In brief, we optimise the Radial Basis Function kernel hyperparameters by minimising the negative marginal log likelihood as described in [9]. Once trained, the model allows us to estimate the automatability of “unlabelled” work activities, i.e. for those activities where expert labels were prohibitively difficult to obtain.

**Preliminary Results**

We have recruited seven general practice medical centres to date. We have begun looking at each occupation and their work practices in primary care. The following are some preliminary findings from the project.

We have identified 16 unique occupational roles in primary care. These 16 occupations conduct all of the work that occurs in primary care and have currently catalogued over 160 unique tasks. In general, each occupation performs between 10-20 tasks regularly. For practice staff (non-clinical occupations) there are around three to eight tasks that require the collaboration of another person to complete. For example, signing off prescriptions or letters, reviewing documents, gathering signatures from multiple people, or entering a portion of data into an electronic system.

Aside from face-to-face meetings and phone calls, we observed that nearly every other task relies in some way on a desktop computer. This heavy use of desktop computers is an important indicator for future automation, since it is likely that software-based automation will be large driver of further automation, through emerging technologies such as robotic process automation. The electronic medical record (EMR), for example, is what most staff in primary care spend the majority of their time interacting with, however, we have observed that different practices use the software very differently from one another.

We have classified the 160 observed tasks within primary care as either highly-structured or unstructured tasks. “Structured” defined as being performed in the exact same way multiple times, using the same set of steps or workflow. “Unstructured” tasks involve an element of improvisation, creativity or negotiation.
Completing an unstructured task is often straightforward and may not have an easily defined end goal, although many can be categorized as “responding to a patient request”.

Going forward, we anticipate how a task could be restructured; what technologies the task might require; how much time an occupation spends performing tasks; how multiple staff collaborate on tasks. All of which highlight important factors used to address the second aim of our research.

Inductive qualitative content analysis has been performed as part of Aim 1. The results of this analysis have produced categories we continue to build on and plan to use in future analysis. These categories will help us to identify work that lends itself to being automated or presents a technological challenge to being automated. Specifically, we are interested in potential correlations between the identified categories of primary care work and their association with probability of task automation. This will help identify categories of work, or entire workflows that are closely correlated with high or low probabilities of task automation, in order to propose future automated workflow design within the healthcare domain.

Discussion
We propose this two phase study to assess the impacts of automation on UK NHS primary healthcare. As such, we address two issues: one is an inherent lack of detailed data on the work practices of primary healthcare staff; the second, is the development of a formal representation for estimating automation and its impact on tasks and occupations.

As external researchers, we have been granted a comprehensive level of access to primary healthcare sites across the UK and use this access to build a detailed dataset of occupations and the tasks they perform. To the best of our knowledge, no such similar dataset currently exists for analysing the work practices and tasks of NHS primary care staff.

This study aims to advance earlier research from the University of Oxford [4], on automation and its effects on employment. Specifically, we address the understanding that entire occupations are unlikely to be automated in their entirety, rather their composite tasks, by taking a task-level approach to modelling automation [21–23], and identify where efficient workflows could arise. We are also keen to inform policy decisions and best practices in primary care concerning the design and configuration of occupational workloads and tasks in highly automated environments. For example, the reduction of back office paperwork may eliminate much of the workload of receptionists and prescription clerks. From our analysis we intend to provide insights as to other, less-automatable work that occupations may engage with given a reconfigured role.

We acknowledge the challenge, and commonly held belief, that typical healthcare related occupations are associated with a low risk of automation. Through our detailed task level data collection and expert elicitation, we plan to gain better insights into which tasks are technically automatable. In turn, our machine learning model will learn which tasks do require a human to manually perform them, and we highlight specific attributes that drive higher or lower automatability estimates using sensitivity analysis.

From initial fieldwork, we have found that many forms of automation already exist in healthcare. We have witnessed these forms of automation increase the productivity of human employees, however they do not remove tasks for them entirely. Indeed, some of these forms of automation have actually created more work or allowed humans to process tasks more efficiently, but with more administrative work being produced that needs to be processed. We anticipate that our analysis will inform the design and re-
configuration of work processes in primary healthcare, and lead to recommendations of new automated processes.

Acknowledgements

With thanks to: The Health Foundation (award #7559), the Oxford Martin Programme on Technology and Employment, NESTA, MIRI, the Rhodes Trust.

References


