OXFORD MARTIN SCHOOL

OXFORD INTERNET INSTITUTE

UNIVERSITY OF OXFORD

THE FUTURE OF HEALTHCARE:

Computerisation, Automation and General Practice Services

Dr. Matthew Willis, Dr. Paul Duckworth, Logan Graham, Prof. Michael Osborne, Prof. Eric T. Meyer, Dr. Angela Coulter

June 2020

Sponsored by::

The Health Foundation

Contents

List of Figures

List of Tables

Executive Summary

NHS General Practice in England faces multiple challenges as evidenced from multiple national organisations, policy groups, think tanks, and universities. Some of these challenges include skill shortages, staff retention, increased workloads, and financial burdens. Various strategies have been proposed to ameliorate these challenges, from increasing financial investments and renewing infrastructure to reorganising and redesigning staff. There is certainly no single solution that will create a robust, sustainable, and high performing general practice of the future. The research program detailed in this report outlines the development and interpretation of an evidence base to inform the conversation about the use of technological support within primary care. Specifically, *to what extent can currently available automation technologies be applied in general practice? And, what kinds of work conducted by which type of staff are most amiable to automation?* By understanding the scope of automation and the location of where to best apply potential technologies, primary care organisations, policy makers, and industry can make an informed decision on work digitisation and workforce transformation.

We employ a machine learning methodology and expert survey estimates of the state-of-the-art in automation technologies, in order to model the extent to which tasks can be technically automated within primary care *today*. To improve the robustness of this analysis, our data describing tasks and work in primary care is based on ethnographic observations at six primary care health centres in England, involving over 350 hours spent in understanding the scope of work for each staff type commonly found in primary care. From this analysis we develop an "automatability score" to identify the specific healthcare tasks that have the potential to be mostly or completely replaced by automation techniques. We further describe which skills, knowledge and ability characteristics are likely to complement such work.

Our principal finding is that approximately 44% of all administrative and clerical tasks in general practice can be either mostly or completely automated. We are excited by this finding because it shows the potential for rethinking the future of primary care, and also the potential (beneficial) impact of appropriately implemented technologies on the workforce. We also find that no single occupation can be entirely automated, but that different staff types will be able to have more or less of their work automated. This is another exciting finding because it provides evidence to consider how occupations may be redesigned: from considering new ways of working to adding tasks that bring more value to the patient and the community. Lastly, we focus on a few select exemplar tasks that illustrate the need to carefully consider and think through the implications of automating work in general practice. For example, new ways of working with a machine may necessitate new ways of training and educating staff. The introduction of technology into organisations will always have intended and unintended consequences; we try to highlight and anticipate the former throughout our investigation.

Throughout the project we have found that the automation technologies that promise to make a positive improvement in the workforce are those which automate mundane tasks. Automating and removing some of these tasks could free up countless hours where more urgent, rewarding, or patient-centred tasks can take their place. Automation technologies that assist with tasks such as telephone answering, letter writing, document scanning, email monitoring, and information filtering show the most potential, and would represent a remarkable change to the way all staff work in primary care. However, we note that the work in general practice, much of which seems mundane and routine, is full of complexities and exceptions. We hope that

adopting a cautious and informed approach to task automation in general practice will allow for greater opportunities for mentorship, communication, multi-professional training, and increase opportunities to experiment with new ideas to improve patient care. At a minimum we believe reconsidering how the tasks we have highlighted are performed, and by whom, could contribute to an effective and efficient use of resources in primary care.

Introduction

1 Introduction

One of the most important ongoing conversations across society in the UK is the future of the NHS, and the role of automation and machine learning tools in that future. Automation is typically discussed as a threat in many areas of work, yet it presents an opportunity in the primary care sector of NHS in England. Certainly, automation may help address the multifactoral challenges currently faced by general practice. A recent report from the Health Foundation, The King's Fund, and the Nuffield Trust, shows the extent of the workforce shortage and its impact on patient care by stating that the shortfall of GPs will never be filled, and that other clinically trained staff will need to reconfigure how they work to reduce the workload on general practitioners (Beech et al., 2019). Previous work in this area has looked at the opinions of General Practitioners on automation (Blease et al., 2018), finding that GPs are skeptical about the capacity of technology to replace or perform most tasks in primary care better than humans, specifically: diagnosis, prognosis, personal treatment plans, referrals to other health professionals, and empathetic care. What is needed however, is empirical work to understand both the extent of automation in primary care using currently available technologies—i.e. no forecasting— and to show the "location" of automation, i.e. which parts of the staff work flow and task list would be most amenable to automation.

To this end, the multi-year research project detailed in this technical report provides much needed evidence for this important discussion on the role of automation and digitisation in the NHS in England. We believe that appropriate, evidence-based, and informed application of automation may provide breathing room for the overworked primary care sector and help control ever-expanding staff workloads.

1.1 Research Approach

A critical part of the problem we address in this project is that a formal framework for measuring task automation is essentially lacking in the literature. This is compounded in the healthcare domain by a general lack of detailed and healthcare-specific occupation and task data to inform good policy. We can only make pronouncements about workflows in primary care if we have granular data on what occupations in primary care do what kind of work, how they do it, and the various attributes and characteristics that are important to each task. We have focused on addressing this lack of data by collecting high-quality, qualitative, detailed task-specific data from multiple English primary healthcare practices. This presents us with several advantages. First, we can understand at a high level how general practice functions, and how tasks move through the practice, end-to-end. Second, we can understand at a low level the situated tasks and workflows of each staff type in primary care, and how they work to support the practice—and what that work means for the bigger picture. We use this detailed observational data to form the basis of a machine learning model to make probabilistic inferences about primary care tasks, in order to provide1 "ease-of-automation" scores at the task-based level.

The study was organised into the following key stages:

- **Literature Review**: Review of literature looking broadly at the current state of automation and its impact on work. The review then focuses on the unique opportunities and challenges of automation in primary care and shows what is currently known about different uses and applications of automation in healthcare.

- **Fieldwork**: Detailed qualitative observational work performed by Dr. Matthew Willis over a period of 12 months, involving visits to six NHS primary care practices in England. All practice staff occupational types were observed in order to understand each type of task. Focus groups were conducted at the end of fieldwork at each site to improve and validate the observational data captured.

- **Task specification:** Qualitative analysis to categorise the observed work into a formal concept of *tasks* within primary care.
- **Primary Care Survey**: Conducting an online survey of primary care staff aimed at further confirming and validating the tasks that were observed in primary care. This was done specifically to confirm the accuracy of the description of the task and identification of which staff member is responsible for each task.
- **Expert Automation Survey**: The largest online survey for automatability ever conducted, sent out to top academics and industry experts in machine learning, robotics and AI. Thousands of ratings were collected of how automatable specific tasks are today (not restricted to the healthcare domain).
- **Augmenting the Dataset**: Our primary care occupation-task data set, derived from the fieldwork stage, is augmented with numeric attributes available from a publicly available occupational survey produced for the US Department of Labor, the Occupational Network (O*NET) database.
- **Machine Learning Predictions**: The numeric O*NET attributes describing the skills, knowledge and abilities each task requires are used as input into a machine learning model, trained on the tasks where expert estimates were obtained (from the survey), in order to predict the automatability of healthcare tasks.
- **Machine Learning Insights**: Detailed analysis of the automatability of tasks performed in primary care, and the different occupations that perform what kinds of task. From this analysis, and our qualitative work, we provide insights into where policy recommendations may aid the application of automation technologies in primary care.
- **Critical Incidents**: Detailed critical incidents—examples drawn from factual encounters and observations—are supplied to add context and an appreciation of the complexity of the work in general practice.
- **Future of Primary Care**: The likely impact of task automation on the skills and task configuration of primary care workers is considered.

Each of these stages is expanded upon in this report, and further details can also be found in our published research protocol (Willis et al., 2019).

Literature Review

2 Literature Review

In this section we present a review of the vast literature concerning intelligent automation of work, with a specific focus on the UK healthcare domain. First, we provide some context for the state of automation and work, broadly. We then review the current state of NHS primary care, specifically the current pressures, problems, and challenges it is facing, in order to better understand where task automation might be applied and what kinds of problems technology might solve or create. Finally, we seek to understand the current state of automation technologies used in healthcare, worldwide.

To talk about automation and its potential impact on work is, in one sense, to talk about the history of technology. New technologies have disrupted, altered, and reconfigured how people work for as long as humans have used tools. There is a sense, however, that current automation technologies represent something different this time. They are reportedly supporting or replacing many kinds of *intelligent* work, with the potential to displace vast numbers of workers. Throughout this literature review, we examine the current state of knowledge concerning the following three questions:

1. What are the driving forces and current trends of automation in the workplace?
2. What, if anything, is unique about automation in the health care sector?
3. What are the opportunities and challenges for automation in general practice surgeries, and how might this effect the work of primary care staff?

2.1 Automation and Work

The potential for automating occupations has been the subject of public attention for centuries – for example, the Luddite riots of the early 1800s in response to new industrial machinery making it possible for mill owners to produce cloth with greater efficiency, and with far fewer workers. Technological changes during the early Industrial Revolution disproportionately hurt skilled workers, since cloth could be produced cheaply and in bulk using unskilled workers, women, and even children, in place of higher paid artisans (O'Rourke et al., 2008). Technology played a similarly disruptive part in the American "Agricultural Revolution" throughout the 19th century. This period saw the start of a huge decline in number of farm workers from 70 percent of the US labour force before the introduction of advanced agricultural technology, to only three percent today – who not only feed the population adequately, but also produces a large surplus for export (Rasmussen, 1982). Many are now asking whether the impact of current technology on jobs – representing a so-called "Fourth Industrial Revolution" (Maynard, 2015), or "Second Machine Age" (Brynjolfsson and McAfee, 2014) – will be any different. What will be the effect of recent advancements in digital technology on today's occupations, and how can we best prepare the labour force for the jobs of the future?

Many researchers have attempted to trace technological progress throughout history to understand its transformative effect on society as a whole (Mokyr, 1990; Morris, 2010). Work by (Morris, 2010) attempts to rank and compare individual human events and technology developments throughout history, concluding that the steam engine developments of James Watt and colleagues between 1765 and 1776 represented the most significant "bend in the curve of human history", leading to the first Industrial Revolution. The use of steam overcome the limitations of muscle power (human or animal) and generated huge amounts of useful energy on tap. Industrial mechanisation was followed by the widespread adoption of electricity, and thirdly, the use of electronics and information processing, with the latter revolution prompting *Time* magazine to declare the personal computer its "Machine of the Year" in 1982. The digital technologies that are ushering in a new "age of automation" are based on computer hardware, software, and networks. While the underlying technologies are decades old, their full effects on society are only now starting to be seen more clearly, as more and more tasks became computerised and automated. Such as financial services, accounting, law clerks, and other skills that require professional education. For the first time, ”intelligent” automation seems likely to substitute for intelligent human labour (Brynjolfsson and McAfee, 2014).

2.1.1 Automation in the Modern Day

Moore's Law has held up remarkably well for nearly four decades, leading (amongst other developments) to a recent shift in the types of work computers can do in substitution for human labour. This has caused a growing fear that automation will displace more jobs than can be reasonably absorbed. Indeed, we can see digital technology encroaching on traditional human labour jobs across nearly all sectors of the economy, in many cases performing tasks more efficiently and effectively. “What was once just a concern for manufacturing workers is now a concern for everyone whose work has any analytical or repetitive features” (Irving, 2017). This can be seen for example at Amazon, who recently opened a grocery store that uses machine learning and computer vision to eliminate the cashier's role, putting over 3.5 million jobs at risk in the US alone (Grewal et al., 2017; of Labor, 2017).

While there is a lot of public discussion and fear about “jobs being replaced by robots”, there is a large degree of uncertainty about what technology can and cannot automate. Certainly, many tasks once considered uniquely human have now been automated by a computer, although this depends on the level of difficulty of the task. In their 2004 book *The New Division of Labor*, Frank Levy and Richard Murnane placed information processing tasks on a spectrum, with arithmetic (rule-based operations) at one end and pattern recognition tasks (which cannot be easily reduced to rules or algorithms) at the other. (Levy and Murnane, 2005). Their example of a pattern recognition task that could *not* be computerized was driving a vehicle in traffic. And yet, only six years later, in 2010, Google announced that their autonomous car had been driving successfully, in traffic, on American roads and highways for millions of miles. We now see autonomous vehicles working alongside the 2.5 million freight and stock hand-labourers who move goods in US warehouses and commercial buildings (of Labor, 2017; Pooler, 2017). This uncertainty about what technology can (and will) automate has resulted in academics predicting everything from economic “misery” (Sachs and Kotlikoff, 2012) to the labour force simply adapting by shifting skill sets (Autor, 2015; Beaudry et al., 2016; Bessen, 2015); as well as something in-between (Susskind and Susskind, 2017).

Previous frameworks for measuring the likelihood of automation have focused only on the different “types” of occupations and the skills they require (Acemoglu and Restrepo, 2018; Autor and Handel, 2013). Researchers from MIT, Autor et al. (2003), suggest that work can be divided into a two-by-two matrix: cognitive versus manual, and routine versus non-routine. They show that “routine” occupations (e.g. manual labour) are the most susceptible to automation, and “manual” occupations are easier to automate than “cognitive” ones. However, as non-routine jobs increasingly come to be performed by sophisticated algorithms and machine-learning technologies, a commonly held belief is that digital technology will substitute for an ever-larger share of labour, and displace large numbers of jobs over the coming years. Building upon that work, (Frey and Osborne, 2017) presented a study into the automation of US occupations, focusing on bottle-necks to computerisation, providing estimates of the ”probability of automatability” of over 700 US occupations . They classified 47% of occupations currently performed in the US to be at “high risk” of being able to be performed by a computer within the next couple of decades. Several follow-up studies by other researchers have applied these probabilities to other countries' employment data. (Bowles, 2014), for example, estimate the share of “high risk” jobs across Europe to range between 45-60% , with southern European workforces (e.g. Portugal and Romania) facing the highest exposure to potential automation. In Finland, 35% of occupations have been estimated to be susceptible to automation (Pajarinen et al., 2014). In analysis of the UK, jobs in London were predicted to be at lower risk of automation than elsewhere in the country, with low-paid jobs at highest risk (Knowles-Cutler et al., 2014).

The conventional method reported in the literature for identifying and predicting automation is to hypothesize about the underlying dynamics driving automation, and extrapolate or speculate about future capabilities (Acemoglu and Restrepo, 2018; Autor and Handel, 2013; Bakhshi et al., 2017; Frey and Osborne, 2017; Goos et al., 2014). However, these approaches are somewhat subjective in relying on expert forecasts: given such experts may be unduly optimistic (or pessimistic) about future technological advancements. Our approach is to instead ask experts to provide estimates of *current* technology capabilities, rather than to speculate on the future. As part of this approach, we have collected thousands of automation ratings, at a granular task-level, from a survey of hundreds of automation and machine learning experts. We feel this is an accurate and reliable method that does not forecast but uses empirical evidence and expert knowledge of what is already automated given existing technology.

Recent successes in automation have come about in large part thanks to technologies known as "narrow Artificial Intelligence", commonly achieved using machine learning technologies, and more specifically, supervised learning. This narrower scope is seen as more feasible than the much more ambitious goal of an "Artificial General Intelligence" (AGI) [1]. Artificial Intelligence (AI) is referred to as the ability for a machine to reproduce human behaviour, regardless of the technology used to achieve it. Machine learning, a subset of AI, can be considered as a family of algorithms where machines learn to recognise patterns from provided data, for example identify objects in an image or translate from one language to another (Irving, 2017). Andrew Ng cites the recent transformational nature on industries such as search, advertising, e-commerce, finance, logistics, media, and more from simply being able to transform one input into a new output (Ng, 2016). However, he warns against the current hype surrounding machine learning; that it will profoundly impact employment and allow people to work 4 hour weeks. Deep Learning is a further subset of machine learning: a family of supervised machine-learning algorithms where successive layers of data in a neural network are combined to learn increasingly abstract concepts, such as the difference between a cat and dog or identifying suspicious behaviour (Ng, 2016).

Supervised ML has allowed machines to replicate many tasks once believed to require a human, such as identifying hand-written digits in images, recommending favorable movies, and even driving a car in traffic. These tasks would be extremely difficult to automate without machine learning. That is, historically, to automate a particular task (such as recognising a particular person in a photo), a researcher would be required to identify and understand every detail of the problem and write an algorithm to deal with every possible input example – defining every possible different angle, distance, age, clothing, hair style, lighting, etc. of the person in the photo. Defining every possible input in an algorithm can thus represent an intractable problem: taking more time and money to write the code, than would be saved by just asking a person to look at the photos instead. However, by using machine learning, researchers can compile a set of data examples (eg photos of people) and labels (identifying which is the correct person), and allow an algorithm to identify patterns within that data, i.e. to "learn" from the data. This allows massive scaling up of effort: for example, using facial recognition software to scan footage of a football crowd for known violent fans, rather than relying only on police spotters. It is reasonable to postulate that, using these techniques, any task where a suitable training dataset can be captured can be automated by a computer (Frey and Osborne, 2017). Since collecting data (e.g. about who does well in job interviews) is often much easier and quicker than fully understanding and

[1] However, in a recent survey of AI experts the time frame of a high-level artificial intelligence being developed by 2040-2050 was given as a one in two chance, and a nine in ten chance by 2075 (Müller and Bostrom, 2014)

writing an algorithm that can capture every detail about a human task (ie how an interviewer comes to their decision in every conceivable case), the rate at which we can automate a task using machine learning is often orders of magnitude faster.

Agrawal et al. (2018) have drawn an analogy between current machine-learning technologies and "Prediction Machines". They observe that most human tasks can be re-formulated and characterised to include a prediction element and a judgment element, and given enough data, a machine-learning algorithm can be trained to automate the predictive element of the task. For example, a Canadian legal firm, Blue J. Legal, is using prediction to help accountants and tax lawyers predict how courts are likely to rule on a given set of facts and client circumstances, years into the future. Narrative Science, an artificial intelligence technology company, produce corporate earnings previews for Forbes.com that are considered indistinguishable from human-authored articles (Irving, 2017). As the cost of prediction falls, and as machines become increasingly good at it, many white-collar "intelligent" jobs – such as business journalism – are being re-defined in this way, leaving human labour to work specifically on those parts of a task considered to require human judgment – or to continue to work in areas where data is either so scare or difficult to collect, that machine-learning systems can't be trained effectively.

2.1.2 Automation of the Physical Domain

Computers are encroaching on an increasing number of jobs that require advanced pattern recognition and complex communication skills previously considered uniquely human traits. For example, traditional financial investors, once considered as at low risk of automation, are being replaced by smart-balancing and individualized "robo-advisors", putting over 900,000 US financial analysts and sales managers at risk of losing their jobs. Similarly, a legal startup, Casetext, enables lawyers to upload briefs and have "AI" do the case research work of hundreds of paralegals (Irving, 2017). Although these examples exhibit intelligent-like behaviours, they are restricted to the software domain of information processing. Humans still currently hold the high ground in the physical (hardware) domain (Brynjolfsson and McAfee, 2012). Machine learning algorithms, and the data required to train them, have advanced far faster than the ability to build physical machines (e.g. robots) with human-like capabilities, such as fine motor skills, or machines that can deal with unpredictable and dynamical real-world environments. Indeed, roboticists have found it terrifically difficult to substitute a physical machine for even a low-skilled manual worker. For example iRobot's Roomba can't physically do all the tasks a cleaner does; it simply vacuums the floor. Human-level perception, dexterity and flexibility still present huge advantages over machines, providing crucial bottlenecks to the capabilities of current state-of-the-art autonomous machines. Moravec's paradox (named for Hans Moravec) describes the discovery by artificial intelligence and robotics researchers that, contrary to traditional assumptions, "high-level reasoning requires very little computation, but low-level sensorimotor skills, require enormous computational reseources" (Brynjolfsson and McAfee, 2014). This, in part, explains the current lack of substitution of domesticated, autonomous robots for physical human labour. That said, companies are attempting to bridge this gap.

2.1.3 Restructuring Occupations

We have shown examples where machines are substituting for more and more intelligent work, but are currently under-performing at specific physical or sensory abilities. It is therefore of utmost importance to policymakers and business leaders to understand what is (and what is not) at risk of being performed by a computer. As has been noted, "The root of our problems is not that we're in a Great Recession, or a Great Stagnation, but rather that we are in the early throes of a Great Restructuring. So it's urgent that we understand these phenomena, discuss their implications, and come up with strategies that allow human workers to race ahead with machines instead of racing against them." (Brynjolfsson and McAfee, 2012).

The greatest benefits to companies arise when they use digital technologies to reorganize decision-making authority, incentives systems, information flows, hiring systems, and other aspects of their management and organizational processes in order to leverage ever-advancing digital technology, augmented by uniquely human skills (Bresnahan et al., 2002). Identifying aspects of an organization that can be re-formulated as prediction (rather than judgment), means those tasks can be substituted for machine-capital. This widely accepted accepted understanding has been referred to as a "race with machines" strategy (Brynjolfsson and McAfee, 2012), as involving "prediction machines" (Agrawal et al., 2018), or referred to as the decomposition of the professional services (Susskind and Susskind, 2017).

Under this scenario, work is reorganised to eliminate a lot of repetitive routine tasks, regardless of whether they are cognitive or manual, leaving behind residual tasks that require relatively more judgment, skills, and training. The demand for machine-complementary skills, such as creative thinking and judgment, is therefore likely to rise – resulting in an increasing demand for more educated workers, and reduced demand for less-skilled workers (Autor, 2015). However, as argued by (Susskind and Susskind, 2017), the logical long-term future of this scenario – given more abundantly available knowledge via ever-increasing ways technology connects us – is that almost *all* professional service jobs will be dismantled. People will no longer be required to hold practical expertise in specific areas, since this knowledge (of the performance of financial markets, or of legal codes, etc.) will be readily available to all via digital technologies. Studies have shown that companies with significant IT investments have typically made the biggest organisational changes, usually with a lag of five to seven years before seeing the full performance benefits (Brynjolfsson and Hitt, 2003; Brynjolfsson et al., 2002). The lags reflect the time it takes for managers and workers to understand new ways to use the technology. There are many examples of this in the literature, from simply "paving cowpaths" (non-formalised ways of work and interaction) across Boston, to positioning the first electric-powered motors in the centre of a factory which previously relied on centralised-steam power (Brynjolfsson and McAfee, 2014).

It is widely accepted that tasks rather than entire occupations are increasingly being automated, leaving those residual tasks that are not yet possible or appropriate to automate (Bresnahan et al., 2002; Brynjolfsson and McAfee, 2012, 2014; Duckworth et al., 2019; Susskind and Susskind, 2017). The current scholarship therefore suggests that occupations are better understood as evolving combinations of detailed tasks, skills, and/or environments – rather than as rigid and immutable (Acemoglu and Restrepo, 2018; Arntz et al., 2016; Duckworth et al., 2019; Manyika et al., 2017). With more granular job data available than ever before, from sources such as the US National Center for O*NET Development (National Center for O*NET Development, 2018), thousands of occupations can be broken down into hundreds of constituent components, relating to the skills, knowledge and abilities required to perform them. Similarly, the specific tasks performed by each occupations can be analysed with respect to these components.

We believe there is a good opportunity to combine these data sources in order to evaluate the automation of tasks using these high dimensional skills, knowledge and abilities components made available by O*NET. Two recent reports use such a task-first approach, however, they are unfortunately limited with respect to their methodological clarity and reproducibility. The report by McKinsey (Manyika et al., 2017), uses an unclear approach to model the opinions of automation potential of an unknown number of industry-based experts who are unfamiliar with and work in the frontier of automation technology. Another, by the OECD (Arntz et al., 2016), derives task-level estimates from occupation-level estimates, and uses worker (not task) characteristics in their inference steps. These works, released by private institutions, are not required to be transparent or peer-reviewed. This makes them, at best, difficult to reproduce, and at worse, untrustworthy estimates of future automation of the workplace.

2.2 Making the Case for Automation in Primary Care: Unique Challenges and Opportunities

Up to this point, our review has outlined some of the core concepts in automation and how technologies have assisted, replaced, or augmented human labour. Now, we turn our attention to the healthcare profession, which has typically been thought of as one of the most challenging areas in which to apply automation technologies because of the need for social intelligence, empathy, conversation, creativity, and improvisation. However, we posit that automation applied to healthcare is an opportunity and a welcome tool. In contrast to the previous section on automation and work, in which automation was presented as a technology that typically replaces workers, this is unlikely to occur in the healthcare industry as it requires many human traits that are considered bottlenecks to computerization (Frey and Osborne, 2017).

That said, the unique needs of primary care produce a promising field in which technology can be advantageously implemented. General practitioners (GPs) represent the largest population of British doctors and are also at the front line of numerous issues in healthcare. As the BMJ has stated in a recent headline: ”if general practice fails, the whole NHS fails” (Roland and Everington, 2016). The NHS in England is under remarkable pressure to improve services, cut costs, and address low staff morale. These challenges must be met despite the shadow cast by the longest financial crunch in the history of the NHS, concurrent with a dramatic increase in service demand and use, lagging performance in mental health services (TUC, 2018), staff shortages, an ageing population, and increasing wait times (Hopson, 2016).

The General Practice Forward View (2016) focuses on the current terrain of primary care by scanning the “horizon” of general practices and addressing the most relevant issues and challenges to primary care. It is clear that the future of primary care must address the challenges of workload through a renewed infrastructure, which requires growth in the workforce, increased financial investment, and a redesign of care delivery models to provide more support to practices. Each of these areas of focus in the General Practice Forward View, as well as the Five Year Forward Plan, stem from the aforementioned challenges (Baird et al., 2016; Hopson, 2016; Martin et al., 2016) in primary care. Although the challenges are numerous, the informed implementation of automation technologies can be used to alleviate workload stressors and allow time to invest in redesigned models of care, among other benefits. The next section on administrative work discusses some specific problems in primary care that may be addressed by automation.

2.2.1 Administrative and Bureaucratic Work in Primary Care

In 2014/15, a report was commissioned by NHS England from the Primary Care Foundation and NHS Alliance to investigate bureaucracy and workload problems in primary care. The report identified several burdensome areas for practices, including the 27% of practice time spent on activities that concern GPs “getting paid”, which refers to tasks such as filling out reports and entering data into various information systems that report to other agencies and institutions. Other areas include processing information from hospitals (26%), keeping up to date with changes, reporting information, and supporting patients dealing with the NHS. Additionally, about 27% of appointments were considered avoidable, such as directing patients where else to go in the NHS, answering questions about benefit appeals, or informing patients of no action needed on test results and other continuity of care appointments (Clay and Stern, 2015).

The 2014/15 Primary Care Foundation report reveals an increase in GP workload, as well as a shift in practice priorities and time spent when compared to the 2006/07 British Medical Association (BMA), NHS Employers, and Department of Health survey of General Practice Workload. In the 06/07 survey, GPs indicated they spent about 72% of their time on "essential services". There is a strong shift in the 2014/15 report which concluded that over 50% of practice time was spent on bureaucracy. To address some of these workload and time concerns, the GP Forward View report calls for streamlining certain administrative tasks and the automation of common tasks and appointment software (NHS, 2016).

In 2016 The BMA also designed a survey of GPs in England, conducted by ICM Research. Of the over 5,000 GP respondents, 84% described the workload as excessive and preventive of quality and safe delivery of care. Only one GP in ten described the workload as manageable, and permissible of quality and safe care. The result of this strain is that 38% of partners consider closing their patient lists to better manage workload (Byrne et al., 2016).

A similar survey undertaken in 2016 of GPs from the Policy Research Unit in Commissioning and the Healthcare System (PRUComm) to measure stressors and job satisfaction, among other parameters. Of the 14 stressors evaluated, the most challenging are workload and meeting requirements of external bodies. Finding a locum and emergency call interruptions were reported as the least stressful. However, every single stressor was at its highest point (on a 5-point scale) since the beginning of the National GP Worklife Survey in 1998. Job satisfaction declined from 4.5 points in 2012 to 4.1 points (on a seven point scale) in 2015, with the most significant areas for dissatisfaction being hours worked and remuneration (Gibson et al., 2016).

Another study from Hobbs et al. (2016) focused only on clinical work resulted in similar findings. A retrospective analysis of 100 million consultations revealed that the clinical workload of primary care had increased overall by 16% between 2007 and 2014. The conclusion of this study was that the current delivery of Primary Care is reaching a saturation point. Most notably, the authors call attention to the fact that they only studied the clinical workload, and that other administrative and professional activities have likely also increased (Hobbs et al., 2016). This claim is indeed supported by the previous surveys. These results all confirm that primary care workloads have increased in recent years, with significant impacts on primary care staff. This phenomenon reflects the general trend in increased administrative and bureaucratic work in public sector occupations as they grow over time (Hitchcock and Sundorph, 2017).

Robotic process automation (RPA) is an application of machine learning developed to replace this kind of repetitive administrative and clerical workflow. RPA can be thought of as a type of software robot that reproduces repetitive tasks using the interface and other tools to control applications on a desktop computer (James Manyika et al., 2017). Repetitive workflows are amenable to RPA, especially in the areas of finance and book keeping, information technology management, and human resources. RPA has successfully been utilised to replace a lost ATM card, process claims and payments, monitor system error messages and fix simple problems, and process invoices and respond to routine requests from customers and suppliers (Davenport and Kirby, 2016). Although we discuss more examples of automation specific to healthcare in a later section, the administrative and clerical work in primary care is similar to office work outside of healthcare in which RPA has already been proven a successful tool.

We consider this context of the state of the workforce and the rise on administrative and reporting work across all staff to be important for the analysis of what is automatable in primary care. Attention should be given to the workloads of each individual occupation, understanding the administrative and clerical work that occurs throughout primary care and how this work factors into maintaining quality healthcare. In the following sections we explore the research around automation and its application to healthcare, specifically NHS primary care when relevant.

2.3 The Platform of Automation: Information Communication Technologies, Computerisation, and Data Availability

One of the first reviews of the impact and use of automation and computerization in healthcare was conducted by Leontief and Duchin (1986). In their study, reported in their book *The Future Impact of Automation on Workers*, extensive economic data and computer-based modelling were used to explore how automation and technology were expected to change the areas of manufacturing, employment, office work, education, and healthcare. The chapter on technological change in healthcare describes an automation environment that may seem trivial in 2018, but which was revolutionary at the time. Computers were used extensively in hospitals for bookkeeping, billing, and inventory control and patient records. Computers were most widely adopted following the introduction of Medicare and Medicaid in the United States in 1966, which doubled the paperwork per patient. At the time of the study, 20-30% of hospital costs were related to handling of this kind of information–an administrative burden that could be reduced by the introduction of more computers. Furthermore, the authors showed that hospital laundries and kitchens were made more efficient through automated equipment, computer controls, and electronic management systems. While they noted that health professionals were reluctant to identify specific cost savings from the application of computers to the delivery of care, further case studies indicated benefits in reduction of congestion and quality of care. Certainly, computers have improved speed and accuracy of test equipment and clinical labs (Leontief and Duchin, 1986).

Clearly, the work of Leontief and Duchin showed that computers played an important role at a time when healthcare was undergoing profound change, from technological developments to organisation and delivery of care. Computers were just being adopted and applied to many administrative and non-clinical tasks. The authors noted that new computer-based tests and technologies for diagnosis were controversial (and expensive) at the time. And the question is still relevant: how does technology impact the quality of care and medical outcomes? In 2018, (Kruse and Beane, 2018) conducted a systematic review to determine if the projected impacts of information technology on healthcare hold true. From their review of the literature, they found that 81% of the published research showed positive results that improve outcomes.

Another study in the theme of ICT benefits, which projects future potential benefits, comes from (Weiner et al., 2013). The authors performed a comprehensive literature review on how health information technology and e-health applications might impact the future of demand for physician services. They estimated that if health information technology were fully implemented in at least 30% of community physician offices, demand on service would be reduced by about 4-9% .

While ICTs and computerization have already enabled dramatic improvement in healthcare, the data these systems generate and store continue to fuel change and growth in healthcare. Clinicians gather data from patients in the form of vitals and medical tests, doctors code letters and patient summaries using codes and taxonomies that produce additional medical data, and administrative and support staff produce data on appointments, clinics, and patient demographics. These are just some of the many ways in which the accumulated information can become data. These data are then used to inform and advance both quality of care and information systems. Automation, decision support systems, and all the other technologies discussed in this report, are made possible and functional through leveraging data (Obermeyer and Emanuel, 2016).

Indeed, healthcare has become a very data-rich industry, from pharmaceutical information and health tracking data to the ubiquity of electronic health records and medical devices that gather biomedical data from the patient. Most information is made available in machine-readable digital formats. However, interoperability of systems remains the primary challenge to leveraging these data. Reviewing the opportunities and challenges of data use in primary care, (de Lusignan and van Weel, 2006) share five opportunities and five challenges for use of this immense data resource in primary care.

Opportunities:

1. Growing volumes of routinely recorded data.
2. Improving data quality.
3. Technological progress enabling large datasets to be processed.
4. The potential to link clinical data in family practice with other data including genetic databases.
5. An established body of know-how within the international health informatics community.

Challenges:

1. Research methods for working with large primary care datasets are limited.
2. How to infer meaning from data.
3. Pace of change in medicine and technology.
4. Integrating systems where there is often no reliable unique identifier and between health (person-based records) and social care (care-based records, e.g. child protection).
5. Achieving appropriate levels of information security, confidentiality, and privacy.

The authors' work shows the potential in the volume of the data available, which is current and continuously updated, accurately representative of the patient or their interactions with the healthcare system, and inclusive of many different variables. However, all this potential in the data is lost without the ability to properly analyse the data, draw insights, create knowledge, and share this information in a way that is responsible and secure. This is particularly true of data privacy and all the associated work that come along with it: governance, privacy, security, and sharing challenges. These all create additional costs by way of specialised staff and additional organisational policy. These steps are necessary to approach full automation, representing challenges not only in primary care, but across the healthcare sector.

Current data collection and management systems are designed within a specific context to support clinical activity for reporting, liability, and documentation purposes. However, these data have proved useful for secondary data analysis, including many of the machine-learning and AI techniques discussed throughout this review. These secondary uses of data have created challenges for machine learning. (Johnson et al., 2016) focused on critical care data to point out three challenges that data present to machine learning applications: the compartmentalization of data, which makes it difficult to share, acquire, and integrate different data silos; the corruption of data, including missing or incorrect data, which compromises the results of any quantitative framework built with it; and the complexity of data. While the storage of complex data may present technical challenges for a data management system, the ability to utilise these data will enable multiple pathways for modelling, estimation, prediction, and other machine-learning insights.

Since the 1960s, many tasks and responsibilities across healthcare have been automated by the digitisation and computerisation of healthcare. These technological systems continue to expand and mature and in turn social organizations continue to adapt to the capabilities and features of emerging technologies. As shown in our previous discussion of automation and work, automation technologies are maturing to a point where more cognitive-based tasks (that is, those involving working with knowledge and information) can be automated to a degree. Raising an important point to investigate the interaction of general practice staff with technologies that are automating decisions and other cognitive work.

2.4 Automating Decisions, Automation Bias, and Unintended Consequences

The work of decision making and knowledge management can be automated through the application of ICTs – including Clinical Decision Support Systems (CDSS). A CDSS is an information system designed to support clinicians through the process of making decisions on a diagnosis or recommending a treatment or change in therapy. The type of system can range in complexity, thus there is a range in what is automated from the clinician's decision-making tasks. A simple CDSS only checks the input from the clinician to confirm the input is valid and within range of the specified field (eg numerical entry for heart rate and blood pressure in a standard format), producing an error or notification if the data is invalid. Automated CDSS are developed for specific clinical specialisations and involve the use of data models, standardised medical knowledge ontologies, and other medical and clinical knowledge, along with data inputs from electronic medical records to support diagnoses based on the systems guidelines. Complex CDSS use a series of computational, data mining, and statistical methods to support complex reasoning on classifications of disease, or predictions of a disease or patient concern (O'Sullivan et al., 2014).

While CDSS benefit the clinical decision-making process, there are several complications. They are easily susceptible to automation bias, which ocurrs when clinicians rely too heavily on decision systems to the detriment of their own reasoning (Goddard et al., 2011). A related concern about the use of automation systems is that over-reliance on automation can create an inability to detect system errors. When automated systems make more decisions autonomously, this can remove users from the loop, reducing awareness of other related processes, workers' task status, and system states (Robin Felder, Majd Alwan, 2008).

The current interest in machine learning approaches to clinical decision making and support have increased awareness about how new machine-learning powered technologies will impact the workflow and decision making of medical professionals. Cabitza, Rasoini, and Gensini (2017) offer four potential unintended consequences of machine learning in medicine. The first consequence is the reduction of skill in users, which is known from previous generations of automation. Over-reliance on the machine could atrophy the decisions and skills of clinicians. Second, data used to inform decisions and insights may be from a different context, which has its own biases and requires context that may be difficult to describe to a machine. Removal of context from the data may confound the interpretation of diagnostic, therapeutic, and prognostic outputs. Third, the authors note the intrinsic uncertainty in medicine, which will influence the interpretation of observed phenomena by different medical professionals. This uncertainly can be a challenge for machine-learning algorithms, and the challenge of uncertainty in medicine is often underrated. Fourth, these tools and algorithms can be a black box to users, creating a tension between accuracy and interpretation. It is important that decision support systems should provide physicians with visualisations, descriptions, or other explanations about how decisions are being made by the system. However, even as better explanatory tools become commonplace, the authors suggest that physicians build skills in the critical appraisal and use of machine-learning based tools for medicine.

When considering the ways in which technology can learn and make autonomous decisions apart from humans, Challen, Denny, Pitt, Gompels, Edwards, and Tsaneva-Atanasova (2019) provide a framework for clinical automation issues in medicine that impact quality and safety. While these are clinically oriented concerns of the technology in use, we can learn from these issues, as any automated technology based on similar machine learning approaches would inherit similar concerns for general practice. Though explored in depth in (Challen et al., 2019), we provide a brief summary here. They are:

1. **Short term considerations**

- The mismatch between data used to train the system, and also any bias in the data, leads to *distributional shift* in which the system is making a prediction that is not applicable to the current situation.
- In the content of clinical use, the *insensitivity to impact* is when the system does not take into account false positive or false negative predictions.
- *Black box decision making* is when the decisions and predictions made by an autonomous system cannot be audited or interpreted, and only the final output is made visible.
- *Unsafe failure mode* is when the system is making predictions and decisions from insufficient information.

2. **Medium term considerations**

- When a system is seen as being always correct, or its predictions are not questioned, this is known as *Automation complacency*.
- A form of bias in the data, when a system is informed by historical data and does not adapt to changes in data or practice is classified as *Reinforcement of outmoded practice*.
- Systems that reinforce the outcomes they were designed to impact are *Self-fulfilling predictions*.

3. **Long term considerations**

- When a system performing a narrow function does not take into account a broader context they can have *Negative side effects*.
- Reinforcement-based learning systems use a reward based function to nudge the system into the preferred end goals. A continuously learning system may find an unexpected way to achieve the reward without meeting the goals, which is referred to as *Reward hacking*.
- *Unsafe explorations* are when learning systems are able to explore and test boundaries in ways that lead to unsafe outcomes.

- *Unscalable oversight* refers to systems that require so much human oversight that it becomes prohibitive in either time or expense to monitor a system so closely.

Of particular interest to this project are the short term challenges of black box decision making and the medium term challenges of automation complacency and self-fulfilling systems. While we primarily focus on administrative and clerical tasks which have less life threatening outcomes for patients, there is potential for negative consequences from administrative systems to impact patient quality of care. Furthermore, in terms of longer term risks there exists the potential for automated systems to waste enormous resource in relation to data loss or making medical information so unreliable as to have catastrophic consequences for the healthcare system as a whole. The above classifications of automation issues from Challen et al. (2019) are only the negative consequences which we can foresee. There will undoubtedly be unforeseen consequences and outcomes that lead to quality and safety issues if systems are incorrectly designed and used. These issues may only arise after the system has been in use for years, and newer features and newer data become available to the system. This scenario, however, shows the challenges of having complex social systems collaborate with increasingly complex technology where the agency of the automated system can be opaque.

It is clearly important to be mindful of the potential pitfalls of automation, particularly given how apparently beneficial automation of repetitive tasks can be. It is important to prevent automation bias in primary care. Given the current stresses faced by GPs, any negative unintended consequences will put them at a further disadvantage, and have a lasting impact on the rest of NHS. We must be aware of the full range of what tasks can be automated, as well as how automation of certain tasks may influence the overall work of primary care occupations. Just because something can be automated, should it be? And if it is, how does it shape the knowledge and skill of an occupation?

2.5 The Scope of Automation in NHS and Primary Care in England

Several reports have considered the potential for automation in the NHS, and attempted to identify the areas with the greatest potential for improvement by automation and artificial intelligence technologies. In this section, we present the current discussion on the potential of automation and AI in the NHS.

A report in 2018 from the policy group *Reform* report on AI in the NHS addresses the previously identified problem of administrative workload in NHS, and that AI and similar technologies promise to reduce burden of excessive administrative work. Figures from the British Medical Association put the administrative burden at a 15% increase in work, and others cite a 70%. Estimates by the Royal College of Nursing range from 17-19% of nurse time spend on non-essential paperwork (Harwich and Laycock, 2018). Similar surveys report a similar range on non-clinical and non-essential tasks from 15 percent to over 50 percent (Byrne et al., 2016; Clay and Stern, 2015). Although each of these sources uses different methods for identifying administrative work, time spent on this work, influence on workload, and overall bureaucracy, they each arrive at similar conclusions. In general, the administrative workload of GPs and other primary care staff members has been on the rise, and all of this additional work is superfluous to their core function of caring for the patient. However, aside from identifying administrative work as a useful area to automate, no specific administrative tasks or methods of automation of administrative task are mentioned in these reports.

Though not specific to primary care, the public service think tank Reform identifies several potential areas where the NHS may benefit from applications of AI to become more efficient and produce better outcomes. Aside from administrative work, other areas of healthcare that might benefit from AI applications include:

1. Health promotion using wearables and other health trackers.
2. Prevention through the identification of patterns of co-morbidities at the population level.
3. Supporting cognitive processing tasks such as processing a large corpus of medical publications and patient data to keep the clinician updated on new treatments.
4. Improved analytics and diagnostics of radiology and medical imaging.
5. Improved treatment from therapy robots and AI software applications to support mental health therapies.
6. Using AI to help patients manage chronic diseases such as diabetes, using real-time data analysis to tailor treatment to the individual.

Automation and AI technologies in these areas would clearly benefit the clinician and patient. Each of these areas entail the use of computers and technologies to assist the clinician's ability to diagnose and treat patients. However, we emphasize that they do not address the underlying challenges of primary care we identify at the start of this report. Indeed, if improperly supported these technologies may exacerbate some of the occupational stressors already evident in primary care.

In 2018, the Institute for Public Policy Research published a report on "Better Health and Care for All" (Darzi et al., 2018). The document lays out a 10-point plan for reform and investment in the NHS. The second point is to create a digital-first healthcare system, calling for full automation of all tasks that can possibly be automated. The report lists applications for automation such as bedside robots to assist with meals, mobilization assistance, social engagement, and rehabilitation. Other tools like virtual reality would be used for rehabilitation and therapy. Machine learning is mentioned as a way that will improve accuracy and productivity of image analysis including x-rays, CT scans, and MRIs. Similar technologies will support an overall better form of diagnosis and treatment of patients. The report advocates a technologically deterministic approach: to redesign care pathways around new technologies. However, while the report mentions several clinical workflows that can be further computerized, it contains no information on administrative work. The existing literature on automation bias would suggest that a measured response and cautionary approach to automation in healthcare would be wise.

2.6 Current Applications of Automation in Healthcare

Now that we have sketched the vision of automation and AI as specific to the NHS, we highlight some of the current uses of these technologies in the automation of many different healthcare processes more broadly. Although not always specific to work in primary care, these examples show what is possible and what may influence primary care in the future.

There are many AI, computational, and machine-learning approaches that are applied to different aspects of healthcare, from diagnosis and treatment to prevention and prognosis. A recent review of medical literature by (Jiang et al., 2017) set out to understand what machine learning and computational approaches are used in medicine, and to which areas of medicine these techniques are applied. They found the following to be the most commonly used techniques: support vector machines, neural networks, logistic regression, discriminant analysis, random forest, linear regression, naïve Bayes, nearest neighbour, decision trees, hidden Markov, deep belief networks, deep neural networks, convolutional neural networks, and recurrent neural networks. These techniques are applied to problems and analyses in the following areas: neoplasms; the nervous system, cardiovascular, urogenital, digestive, respiratory, and endocrine systems; skin; pregnancy; and nutrition. Currently most of these applications apply exclusively to specialists (i.e. not to GPS). However, as these techniques and the software to perform these analyses become more democratised and mature, tools powered by these technologies may be made increasingly available to general practitioners to support primary care clinics.

Although the computational and mathematical techniques outlined by Jiang et al. (2017) are primarily for the diagnosis, prognosis, and treatment of patients using medical data, other computational techniques exist that aim to enhance human-computer interaction and augment the capabilities of computer systems and (hardware and software) robots. These include natural-language processing, machine vision, and speech recognition and generation (Harwich and Laycock, 2018). A combination of machine learning techniques in tandem with natural-language processing and other tools enhances human interaction with computers, promising the greatest impact in the NHS, especially in primary care.

As an example, Hardy, Holford, Hall, and Bracken (2004) developed an algorithm for identifying the beginning and ending records of pregnancies in the automated medical records of the General Practice Research Database (GPRD). They found computational assistance from an algorithm to be useful for scouring pregnancy identification databases for time codes and outcomes. This work would have been unrealistically large and time-consuming for humans to complete without technological assistance, or by only using paper-based records.

Another way for current automation technologies to assist is in the prognosis of end-of-life care, as physicians typically overestimate patient survival time with a serious illness (gra). Data from electronic health records and a deep learning method were used to evaluate patients and present clinicians with those that would benefit most from palliative care. Currently being piloted, the system would remove lengthy chart reviews and physician referrals from the process (Avati et al., 2017).

Similar techniques are being applied to predict patients at risk for chronic conditions such as heart disease and diabetes. After comparing logistic regression, support vector machines, and boosting, Wu, Roy, and Stewart (2010) were able to predict heart failure six months before clinical diagnosis. Machine learning has also enjoyed success when applied to situations that involve the analysis of medical images, typically assisting in artefact identification, classification, and other support tools to assist radiologists in the reading and interpretation of medical images. Another important effect of applying machine learning to medical imagery is that it will help radiologists process more images – an important consideration, given the greater frequency of images produced by imaging technologies, means increasing time spent reading and interpreting these images by radiologists (Wang and Summers, 2012). Another approach to saving physicians time, and supporting diagnostic utility, is in the use of deep convolutional neural networks to classify erythrocytes (red blood cells). Typically, this classification work is labour intensive and suffers from observer bias, slide distribution errors, sampling issues, recording errors, and requires highly trained personnel. The neural network employed by Durant, Olson, Schulz, and Torres (2017) was correct in classification 90.60% of the time, with the potential for increased accuracy given larger data sets. Finally, Han, Park, Lim, Kim, Na, Park, and Chang (2018) also use a deep convolutional neural network to diagnose onychomycosis, resulting in a highly accurate and significant score that was higher than that of 42 dermatologists doing the same assessments manually.

Challen, Denny, Pitt, Gompels, Edwards, and Tsaneva-Atanasova (2019) put current and emerging automation technologies for clinical application on a spectrum of low to high automation. The typical CDSS systems previously described are on the low end of automation, as they are rule based, and the system has no autonomy over how it learns and adapts. In contrast, clinical technologies such as diagnostic support, tailored radiotherapy, automated triage, referral prioritisation, and multifactorial risk prediction employ a supervised machine learning approach that is data driven, based on prior information the system has from the electronic medical record and other data sources. The authors put these clinical applications in the medium category of automation. On the high automation end of the spectrum are personalised drug protocols, autonomous ventilators, and AI insulin pumps. These systems are machine learning applications that are reinforced by data the system can monitor and use to drive it toward a goal (or reward) that is defined for the system. These highly automated clinical systems enjoy a greater range of autonomy than rule based systems, and are in the longer end of technological development than CDSS systems or risk prediction systems, which are already in wide use.

From these multiple examples we can conclude that machine learning and related technologies will continue to offer valuable ways to assist physicians throughout the process of their work from prognosis to treatment. Although these algorithms and techniques are expensive to develop they can easily scale.

2.7 Summary: Considering the Context of Automation and Information Technology in the NHS

The use of automation technology requires both caution and deep consideration of the context and environment these technologies will be used in. A 2016 report from The Nuffield Trust summarises the relationship between information technology and the NHS: “The British health system has been a laggard when it comes to the uptake of health information technologies, being slower to embrace digital opportunities than other sectors of the economy. After the failure of the over-ambitious attempt to introduce a single, centrally mandated electronic care record, the NHS adopted a much more cautious approach ” (p. 18). (Coulter and Mearns, 2016).

Furthermore, 2017 findings from a Select Committee inquiry into the long-term sustainability of the NHS and adult social care elaborate on the social and organisational challenges of technology, implementation, and use in the NHS (Milner et al., 2017). The Committee finds that ”there is a worrying absence of a credible strategy to encourage the uptake of innovation and technology at scale across the NHS. It is not clear who is ultimately responsible for driving innovation and ensuring consistency in the assessment and the adoption of new technological approaches. The provision of appropriate training and development of strong leaders to support this agenda within the NHS will be critical to its success” (p. 67). The Committee’s report also provides empirical evidence of past IT failures, specifically: ”The failure of the care.data project illustrates the inevitable consequences of failing to grapple with important issues relating to personal privacy. NHS Digital and all those responsible for data sharing in the NHS should seek to engage the public effectively in advance of any future large-scale sharing of personal data. Public engagement on data sharing needs to become a priority at a local level for staff in hospitals and the community, and not be left to remote national bodies.” (p. 69). Moving forward, any implementation of information technology or automation technology should consider the history of IT in the NHS as well as the social needs of implementation, organisational support of the technology, and social acceptance including education and continued support.

Fieldwork details

Observation and Data Collection

3 Fieldwork details: Observation and Data Collection

What does the average primary care health centre look like? There are 7,454 General Practice (GP) services in England, as of March 2017. Finding traits to identify and represent all these practices is a significant undertaking: a task we did not attempt in this study. Instead we opted for maximum diversity, purposefully recruiting practices in different geographic regions with different catchment sizes, including both independent small practices and large multi-practice partnerships.

To understand the detailed work in the complex and fast paced terrain of primary care we employed fieldwork to gather empirical data. Dr. Matthew Willis undertook all ethnographic fieldwork. He observed six primary care practices across England, spending on average one week at each practice. The study practices were located in Oxfordshire, Yorkshire, Berkshire, Surrey, and the West Midlands. Their list sizes ranged from 5,000 to 24,000 patients, with an average of around 11,500 patients per practice. Details are presented in Table 1.

The focus of the fieldwork was to understand the work practices and tasks performed by different staff groups at each primary care practice. Namely the typical occupational types that are found in most general practices, such as receptionists, seceretaries, general practitioners of different traning levels, practice manegers, and administrators. This did not require every individual member of staff in the practice, but rather a representative sample of each occupation at each practice. For example, if there were 12 receptionists working in a practice, the researcher observed and interviewed only a sample of them, depending on availability and consent, until the work of a receptionist was clearly understood and could be accurately represented. In total this resulted in 65 sessions of observation ranging in duration from 45 to 80 minutes.

Interviews were semi-structured, and questioning took place while the participant was working on their daily routine or showing the field researcher a specific task. The field researcher was careful to suspend questioning if critical events occurred that required immediate focus and attention of the participant.

The field researcher was also present for a number of clinical appointments to observe the clinician, when patients consented to the presence of the researcher. During clinical appointments the researcher sat in the corner of the room watching the general practitioner or other clinicians, taking notes on how they worked and what tools they used during patient appointments. Queries about work practices were noted during the clinical encounter and asked after the patient had left the room.

Documents describing occupational responsibilities or work tasks were collected, along with photos or samples of other relevant documents, tools, or software that provided further clarification. The field researcher kept detailed field notes of observations and interviews. A spreadsheet was created, delineating each occupation and the tasks they perform. This spreadsheet was then used at subsequent field observations and shown to staff to help confirm each task's representativeness, and the accuracy of that occupation.

At the end of fieldwork at each site the researcher conducted a focus group that was open to all practice staff. The purpose of the focus groups was to verify and check the accuracy of observations and to seek staff views on the use of automation and artificial intelligence at their practice. By the end of the fieldwork, the researcher had a clear sense that data saturation had been reached. No new tasks emerged in subsequent field visits, confirming that the core scope of work and responsibilities of each occupation had been accurately recorded.

Individual tasks were coded and further categorised to note other relevant pieces of information such as if any tools were used to complete the task, the type of software required, and other features of the task, including conceptual inductive categories. These inductive categories emerged from the data through comparing and clustering of similar concepts and ideas. This process continued through iteration until all data were clustered and represented into groups of categories.

Table 1: Details of the six fieldwork Primary Care practices.

Region	**Patient Registry**	**Characteristics**
Surrey	11k	Financially strained, interested in new ideas to apply to the practice, recently adopted the EMIS health system. Are currently training staff to get full benefits from digitisation, focused on QOF for financial reimbursement, serve a low socio-economic area.
Oxfordshire	8k	Diverse patient population, 50% non-white, mostly Asian-Caribbean, almost even split on gender demographics, many tasks shared across different administrative occupations. Frequent cross-training between occupations (dual roles) leading to increased responsibilities. Issues with efficiency and optimal workflows.
Oxfordshire	5k	Older population, average patient here is more likely to be economically disadvantaged. Small staff size so many tasks are shared and overlap with all administrative workloads. Staff must cover a wide range of tasks and are not as specialised. GPs write their own letters, different distribution of tasks because of size of clinic.
Yorkshire	12k	Ageing and elderly population, high rates of diabetes and smoking related diseases shape the clinics they offer. Frequent cross training between occupations and upskilling. Recently joined super-practice.
Berkshire	24k	Aging and elderly population, large population of commuters in the area, interest in private insurance, highly research active, large staff size, very focused tasks for each occupation and very little sharing of similar tasks, quick to adopt new technologies, technically competent staff, many tasks outsourced to super-practice main site.
West Midlands	9k	Many general clerical/admin tasks outsourced to super-practice main site. Observation took place during a CQC inspection, heavy document and policy reviews conducted by the practice manager and lead GP, work is structured by the EMIS workflow, staff largely stick to what their job description/task set entails with little overlap or sharing of tasks.

Tasks in Primary Care

Classification of Work Tasks

4 Tasks in Primary Care:
Classification of Work Tasks

Most of the work we observed in primary care relies on multiple people: A specific task may be initiated by one staff member and completed by another. This relies on a high amount of communication and cooperation between primary care staff. Staff that understand the entire workflow or lifespan of tasks can better appreciate how their work fits into the bigger picture of primary care.

Observing and gathering data on tasks in primary care presents at least three major challenges: First, like primary care itself, tasks are greatly varied. Variance in tasks can occur in the order in which parts of the task are performed, how long they take, the occupational role of the person performing the task, the "importance" of the task or how time-critical it is, and how many individuals become involved in completing it. This variance can occur in the same task being performed by different staff members within the same practice, and the same task often varies from one practice to the next.

There are further complexities that emerge due to the task being performed in a healthcare context, which in itself contains multiple complex contexts such as the patients' needs, backgrounds, and particular problems (Greenhalgh et al., 2007). Additionally, what is considered to be a general administrative task can easily become a clinical task requiring specialist medical knowledge from a clinician. This is what makes work in healthcare different and exceptional, when compared to other fields with similar task descriptions.

Second, perhaps one of the most challenging aspects to our research is that the same occupation does not always perform the same set of tasks across practices. The field researcher found that some tasks that receptionists routinely perform in one practice were performed by administrators and secretaries at another practice. The general practices we observed in this study differed considerably in their organisational forms, from single-site practices working independently to larger groups or "super-practices", who share services across multiple sites. The issue of "task fit", i.e. allocating tasks to the correct person at the practice, involves both matching the task to the most appropriate pay grade and making sure clinicians are shielded from unnecessary administrative work. However, as we shall see later in this report, GPs and other clinicians are involved in a considerable amount of administrative work, i.e. paperwork. The driving force behind the assigning of tasks to occupations in this study is to accurately represent the core identity and "scope of work" of each observed occupation. In primary care, this translates into the same task being assigned to multiple occupations when it requires a degree of collaborative working.

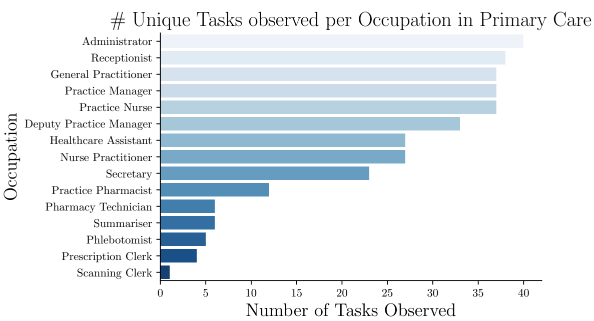
The third issue is the fundamental difference between "clinical" and "administrative" tasks. There is an abundance of administrative and clerical tasks that occur in primary care. While the field researcher observed, catalogued, and categorized nearly every clinical interaction and treatment that takes place in the domain of primary care, there are far fewer purely clinical oriented tasks than their administrative counterparts. Also, for reasons that will become apparent in Section 5, the O*NET database that is used for our analysis, provides more terminology and language to describe administrative tasks and office work than it does nuanced clinical work.

This raises issues regarding how to define a "task", what level of granularity tasks should be regarded at, and the overall "size" of tasks. Descriptions of such tasks could be too broad, such as "provide healthcare", or "operate a primary care practice", or too specific. Details such as where a staff member clicks on a software interface, or how many times they scroll in a document, and what elements of an interface are used to accomplish the task are rather unnecessary details for our purposes.

Formulating an appropriate level of task granularity and description requires calibration between generality and specificity. To this end we have relied on the US Department of Labor O*NET database and its descriptions of "Detailed Work Activities" (DWAs) to help gauge granularity. The following are examples of DWAs: Answer telephones to direct calls or provide information; Process healthcare paperwork; Enter information into databases or software programs; Prepare outgoing mail; Proofread documents, records, or other files to ensure accuracy; Read materials to determine needed actions; Code data or other information; Diagnose medical conditions; Document client health progress; and Interact with patients to build rapport or provide emotional support. Each of these DWAs are examples that were manually connected to our observed tasks. The "observed tasks" are those tasks that emerged from our empirical data and were verified through interviews, focus groups, and a survey. Example observed tasks include: Register new patients to the practice; Scan letters and other documents; Create a query/report in clinical electronic medical record; Check for errors in paperwork; Talk to other staff about incentive schemes; Cleaning workspace and objects after patient examination or treatment.

In some cases, the DWA and observed task mirror each other perfectly. In other instances, multiple DWAs would make up one observed task, requiring the addition of greater specificity in the O*NET classification system. For example, the single observed task: "Writing notes about a prescription renewal and assigning that task to a GP", may match the DWAs of: "using an operating system, typing documents, relaying information to others, communicating with others, and assigning tasks or work". This process naturally leads to an analytical technique of constantly comparing, measuring, and weighing tasks to create an accurate depiction of the work that was observed in primary care. During our fieldwork we observed 137 unique tasks, and 333 unique task-occupation pairs. Each unique task was considered as one of two categories: clinical or administrative. We describe each of these categories in more detail next.

With the complexities of the work classification in mind, we present Figure 1 below: It displays the number of unique tasks observed for each occupational group. It is important to note that this figure does not inform us about the duration or frequency of the tasks observed, nor of an occupation's workload, effort, or any part-time / full-time differences. For example, a scanning clerk may work full-time or part-time on a single task (scanning documents) that fills up the entirety of their working hours, whereas an administrator may perform a great many different unique tasks, but some of these may take only a few minutes to complete, or may only be done infrequently alongside more core tasks that make up a dominant portion of the administrator's workload.



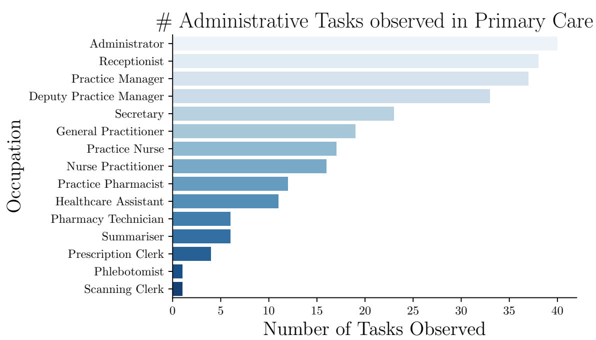


Figure 1: (left:) The number of identified unique tasks performed by each occupation in our empirical data. (right:) Restricted to Administrative tasks only.

4.1 Task Validation

Three problems, or complexities, of the classification of tasks have been previously identified. They concern variance within surgeries for who does what task among all staff; variance within occupations not always performing the same tasks across primary care; and the ability of clinical and administrative tasks to transform between the two classifications, meaning a task can start as an administrative based task and then become a clinical task requiring medical knowledge. However, despite these complexities we want to to provide a thorough mapping of occupations and tasks. This involves triangulating the data across all cases and attempting to represent tasks as accurately and reliably as possible (Willis et al., 2019).

To do this, we first gather and record tasks through observations and follow-up interviews. At the end of each fieldwork period the field researcher holds a focus group or follow-up interviews with staff to go over the task data, ask any clarifying questions, and in general use member checking to determine that the tasks observed accurately reflect the work of that occupation. This process occurs at each practice, and as the data set grows from additional observations, more staff are able to enrich the details of each task and provide further validation. At the end of the fieldwork a survey was sent to one additional practice. Survey participants selected their occupational role and were shown 15 tasks selected at random from the entire data set for their defined occupation. Participants were asked to verify "yes" or "no" if that was a task they perform in their occupation, and provide any additional comments about the task.

Figure 2 below summarises this process of using multiple research methods to understand the tasks performed in primary care, and to harness member checking as a validation technique.

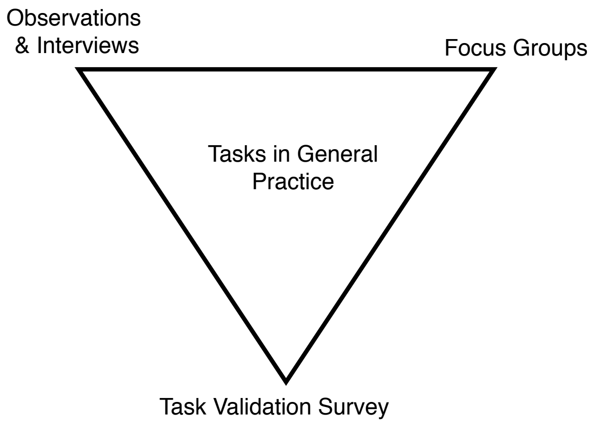


Figure 2: Triangulation approach for validating tasks through the application of three research methods.

4.2 Clinical Work

Following the description of how we categorized tasks and some of the main challenges we faced, we now turn to a discussion of how we conceptualized the distinction between clinical and administrative tasks in our analysis. We consider clinical tasks as those that mostly deal with specific diseases or conditions, and generally relate to the specifics of a patient's diagnosis and treatment, including interactions between healthcare provider and patient. They relate to the occupations of General Practitioners, Healthcare Assistants, Nurse Practitioners, Pharmacy Technicians, Phlebotomists, Practice Nurses, and Practice Pharmacists. This is not to say that these clinically active occupations do not perform administrative tasks - they certainly do! Vice versa, some administrative staff also have clinical experience and considerable medical knowledge. For example, if a member of staff has moved between occupations, or has semi-retired to work part-time on administrative work.

4.3 Administrative Work

Administrative tasks form the overwhelming majority of tasks performed in primary care. They are not simply "back office" tasks or general administration, but routine tasks, paperwork, and crucial workflows that permeate all aspects of primary care. They are tasks that most all occupations must perform (at some point) to various degrees, accounting for a different percentage of each occupation's overall work. That is, they form a crucial element of patient care, insofar as we observed GPs coming in on their day off to carry out administrative tasks. The emphasis here is that there is a very real amount of time-consuming administrative work that every occupation in primary care invests a considerable amount of time in performing. The potential to automate some of these tasks – important work that is rarely patient-facing – could make a significant positive impact on the availability of time, quality of care, and job satisfaction in primary care.

The quantitative analysis of this research, and the resulting insights, focus on administrative work (i.e. administrative tasks undertaken by *all* occupations, including by clinicians). We justify this by two facts: 1) A majority of the total tasks we observed in primary care are administrative tasks; of the 136 unique tasks observed, only 29 were considered clinical tasks (as defined above). 2) The O*NET database was developed and tested across a majority of non-healthcare occupations in the US. While we think one of the great strengths of the present study is that it is empirically grounded in NHS primary care, the database we use to support our empirical observations contains more "language" (capturing skills, knowledge, and abilities) to describe administrative tasks than clinical tasks.

4.4 Characterising Primary Care Work

Throughout the process of observing what tasks are performed by who, we identified 15 distinct occupations that constitute the main workforce of primary care. Each of these occupations may have alternate names or different ranks or steps. For example, we have combined all general practitioners, whether partners, salaried associates, or trainees, under the same occupation: "GP". Likewise, job titles such as receptionist manager, senior receptionist, or receptionist administrator are all grouped as "receptionist". In the following section we describe the main responsibilities of the following occupations (listed alphabetical order): Administrator, Deputy Practice Manager, General Practitioner, Healthcare Assistant, Nurse Practitioner, Pharmacy Technician, Phlebotomist, Practice Manager, Practice Nurse, Practice Pharmacist, Prescription Clerk, Receptionist, Secretary, Summariser, and Scanning Clerk.

Administrator

Administrators carry out many of the core administrative tasks of a practice. They are responsible for generating, completing, mailing, filing and retrieving essential documents to keep the practice running and for communicating with external bodies, the NHS, and others. Most practices in our dataset stored information in spreadsheets. From staff rotas and holiday hours to tracking certain patient contacts, excel spreadsheets are routinely found in this line of work and it is usually an administrator who maintains these documents.

Administrators may also help with phone calls or scanning documents, or assist with medical coding. It is important to note that the work of coding documents, creating much of the patient data in the health record, is carried out by a variety of staff. In some practices only GPs or clinically trained summarisers do the coding, while in other practices coding is carried out by receptionists and many other staff members, depending on the practice configuration. In general, medical coding is complex and relies on multiple bodies of knowledge such as diagnoses, classification systems, the organisation of healthcare, health policy, and other factors (Swinglehurst and Greenhalgh, 2015).

Administrators who work in primarily paper-based environments have a lot to keep up with compared to those working with digital workflows. This affects administrators more than the other administrative occupations we observed because less of their work can be performed in the electronic medical record, for example filling out audit forms, documenting significant events, providing information for GP meetings, and helping with e-referrals. Whereas a receptionist may create tasks and enter data into the health record or work with documents in document management software such as DocMan, the administrator has to spend more time with specialised documents and forms that may not be digital.

Deputy Practice Manager

Deputy Practice Managers assist the Practice Managers in the day to day operations of running a general practice, sometimes assisting the Practice Manager and sometimes developing a scope of work of their own. Examples of tasks include human resource interactions, Disclosure and Barring Service (DBS) checks, scheduling locums, scheduling staff work hours, accounting and finance tasks, buildings and facilities, patient management, text and phone communications, and dealing with specialist web portals like Open Exeter or CQRS.

General Practitioner

General Practitioners manage patient lists and provide consultation, physical examination, diagnosis, and treatment of patients. GPs are one of several clinical occupations that are able to prescribe medications and change treatments. During fieldwork the field researcher noted different styles of interpersonal communication with the patient during the consultation. Some GPs worked on the computer throughout the consultation, others maintained eye contact with the patient and only interacted with the computer after the patient had left. At times, the computer served as a reference tool to show the patient information or to refer to additional resources.

The outcome of consultations greatly impacts the quantity of GP's documentation. Routine checkups and conditions that GPs see day in and day out can equate to a code or two and only a few sentences in the patient's notes in the electronic record. Mental health consultations and other complex conditions may require follow-up letters and phone calls, as well as multiple paragraphs of notes and several minutes spent researching information to include in the documentation. Documentation presents numerous challenges and opportunities for automation, which we discuss later in this report.

Many GPs offer telephone consultations and all have to do a certain amount of work on the phone. They typically write letters, or dictate letters, that must be sent to other parts of NHS or to specialists. Again, there is variety in whether they choose to write their own letters or dictate them and have them transcribed by an in-house secretary or by a specialist outside the practice. Particularly difficult conditions that patients present with may require the GP to consult over the phone with a specialist, or with their colleagues in practice-based "huddles" where all GPs meet to discuss important cases. GPs also spend time mentoring and advising students or trainees gaining experience in general practice, as well as discussing and advising peers.

All practices in our sample designated a "duty doctor", which can be described as an alternating role whereby they deal with the day's emergencies and any other odds and ends and standalone tasks that would otherwise add friction to the other GPs' workflows (while also managing clinics and patients on their own list). The nature of the duty doctor means they must rely on their skills of adaptability and creative thinking as they deal with wide-ranging, urgent matters that arise throughout the day. These tasks can range from last-minute phone calls that must be made, signatures and documents to review, to walk-in patients with urgent conditions.

Healthcare Assistant

Healthcare assistants support the practice through running blood collection clinics, which can be described as blocks of clinical time devoted to taking blood from patients for various tests. While drawing blood, healthcare assistants provide essential emotional support and reassurance. Additionally, this involves using a laboratory test ordering and labelling system called ICE. After the blood has been collected, healthcare assistants must label each vial using this system and package the blood to send on to a designated laboratory. Healthcare assistants may also perform tasks such as ear canal irrigations and electrocardiograms (EKG), delivering the EKG result to the responsible GP either through a paper print-out or a digital data entry assigned to the GP. Healthcare assistants also gather patient vitals such as weight and height. Their main role is to support the clinical journey of the patient. After each patient encounter, much of their labour is spent cleaning and prepping the room for the next patient and entering data in the patient's electronic record.

Nurse Practitioner

Nurse practitioners assist with all clinical tasks as needed; they have the ability to conduct diagnostic tests, to give medical and lifestyle advice, and prescribe certain medications and treatments. Typically, nurse practitioners will run their own clinics for chronic diseases, including care planning and lifestyle reviews. Nurse practitioners provide emotional support and medical advice to new and returning patients undergoing a treatment or therapy. They also extensively document interactions with patients and typically study lab results and other documentation and data in the patient's electronic record before the clinical consultation.

Pharmacy Technician

From our observations, pharmacy technicians were typically dual roles shared with administrators or deputy practice managers. Pharmacy technicians usually have a pharmacy-focused background, and although not always trained pharmacists, they often have some medical knowledge allowing them to assist with processing prescriptions and tasks that require more responsibility than simply assigning scripts to a GP for signature. Pharmacy technicians must be registered with the General Pharmaceutical Council. They typically provide advice to patients on their medication, order medications for the practice, help with repeat dispensing, provide Nomad packs (i.e. blister packs of preorganised medication), and deal with any problems or concerns that arise while processing prescriptions. It should be noted that these pharmacy technicians were more administratively focused, as they were in dual roles, but typically a pharmacy technician and pharmacy assistant can support the dispensing and inventory control of medications. Our observations here are a good example of how occupational tasks and scope of work will change depending on the needs of the practice, existing configuration of occupational roles, and resources available.

Phlebotomist

Some practices in our sample employed phlebotomists. Their tasks focus on drawing blood and the associated tasks of cleaning the examination room, and properly packaging and labelling vials. In some cases, the phlebotomist took a dual role, sometimes also including some receptionist duties.

Practice Manager

Practice managers are general managers, responsible for managing day to day operations and human resources. Typically, practice managers control most of the finances, including paying bills, staff payroll and book-keeping. They are involved in staffing decisions, training, and developing and maintaining the policies of the practice. They interact with many systems and web portals designed to communicate and transmit information and data to different parts of the NHS. Practice managers are involved in making decisions and implementing any enhanced services the practice may decide to adopt. Their work is mostly administrative, including such tasks as completing risk assessments, managing pension schemes, and coordinating staff meetings.

Practice Nurse

Practice nurses typically run a chronic disease clinic, such as those for diabetes patients. During these clinics they meet with patients to discuss their recent test results and care plans, including lifestyle and prevention. They also carry out immunisation appointments and travel injections, advise patients, give information, and complete relevant documents including entering data into the patient's electronic medical record.

Practice nurses may also perform other tasks that healthcare assistants, phlebotomists and nurse practitioners also perform, such as drawing blood, performing ear canal irrigations, documenting patient exams, changing wound dressings, taking blood pressure, and filling out administrative documents for internal purposes.

Practice Pharmacist

Not all practices we visited employed a practice pharmacist. This occupational role was usually found only in locations that treated large numbers of patients, with most other practices relying on community pharmacies. The work of the practice pharmacist is to carry out medication reviews looking for drug interactions, suggesting alternative medication regimes where necessary, providing clinical staff with blood monitoring suggestions, monitoring drug shortages, and other tasks requiring specialist pharmaceutical knowledge.

Prescription Clerk

Prescription clerks were not employed at every practices we visited either, but mainly at those that process large numbers of prescriptions. The core work of a prescription clerk is to collate prescriptions arriving from other NHS sources, either through the post, fax transmission, or by various digital means. They review the prescriptions, and direct them to the appropriate GP for signing, either on paper or using a digital pin.

Receptionist

Receptionists are, more often than not, the first point of contact for patients entering the practice and attending appointments. While every location we visited had a touch screen in the waiting room for patients to check in, there was usually a queue of people at the reception desk with a variety of queries. Other tasks receptionists may perform include transcribing audio recordings, registering new patients, organising and scheduling clinics, checking prescription renewals, sorting and delivering post, and senior receptionists often help with creating and managing the administrative staff rota.

Aside from attending the front desk, answering phones, and helping to book appointments, receptionists perform information work in the form of gathering relevant information in letters or coding letters, getting various signatures on paperwork, writing notes and tasks for other staff members in the health record, managing the practice email inbox, and answering patients' queries.

As with many occupations in primary care, receptionists face frequent interruptions as they carry out their tasks. Although work can be configured in different ways for a receptionist, a typical workflow observed during fieldwork is for them to monitor the phones while they code letters or process prescriptions. The receptionist works at their computer until they receive a phone call, stopping their work as they deal with it, then going back to their previous computer work. They may also have to respond to questions from other colleagues in the practice. Some of the larger practices used a "phone bank", where a number of receptionists solely answer telephones for a period of time, allowing others to focus on tasks without phone interruptions.

Receptionists, secretaries, and some GPs (depending on the work style of the practice) spend significant amounts of time producing letters. The work involves opening letters, triaging, scanning and redirecting them to relevant staff members. These letters then need to be responded to in different ways. Sometimes the letter is just a notification to be archived. Other times the recipient of the letter must gather information from the medical record and write a new letter, usually addressed to someone working in secondary care. This kind of work is ubiquitous across all observed practices. The letters are highly structured and often a template is available to help the process. However, our observations found that these templates are often highly reconfigured or rarely used as intended, adding to the amount of time it takes to produce some letters. Fulfilling and transcribing letters, writing and sending them is a very time-consuming task.

In connection with letter work, there is significant time spent on the management of digital documents. These workflows often employ custom software or multiple software packages to organize, search, and archive the massive amount of letters and documents that the practice is required to maintain.

Scanning Clerk

The scanning clerks observed in our research were typically part-time employees who focused almost entirely on one task, that is, scanning paper documents and creating digital version of them. Some practices receive hundreds of pieces of mail over the course of a business week, and a part-time scanning clerk may perform scanning for 6-8 hours a day, several days a week. They are required to take sorted mail, scan it (potentially supplying additional metadata where required, or missing), then file the hard copy, attach the digital object to a patient record, archive the digital version on a hard drive, or store it in DocMan (the typical software used for the management of letters and other documents).

Secretary

The core set of tasks for secretaries is working with letters. This includes transcription of letters saved as an audio file if/when a GP prefers to dictate their letters or notes. Once transcribed, the secretary might edit the text and format it as a letter; then email it or post the letter to recipients, who may include secondary care specialists, employers, private insurers, or schools. Secretaries are also responsible for the management of GP letters and other documents they produce using custom document management software. We observed that secretaries sometimes work part-time as receptionists during certain hours. Secretaries also sometimes produce invoices, scan letters, perform DBS checks, place phone calls on behalf of others, assist with payroll, and look for errors in paperwork and administrative reports.

Summariser

On occasion, we observed practices that employ the help of a summariser to produce summary patient records. The summarisers we observed were part-time employees with some clinical and medical knowledge from past experience, for instance retired nurses or healthcare assistants. They often get involved in medical coding also, especially if they have clinical backgrounds. When new patients join the practice their records may not always be available in digital form, so the summariser will be responsible for entering the data into the electronic medical record. On occasion the summariser may be asked to “tidy up” information in a patient’s health record. They also help with phone calls and registering new patients.

Dataset Formulation

Synthesizing qualitative and quantitative data

5 Dataset Formulation: Synthesizing qualitative and quantitative data

We use the data from the empirical observations of job roles and administrative tasks performed in primary care as the basis of our analysis into the potential for increasing the use of automation within primary care. One of the great strengths of this work is that we rely on the empirical ground truth observations, and thoroughly validate them, so that the data is reliable for our quantitative analysis and insights.

We acknowledge that this data will not capture absolutely every task performed by a specific occupational group, but we are confident it covers a representative set of typical tasks. A dataset was formulated, where each row represents a unique occupation-task pair. An example is shown in Figure 2, showing seven observed tasks, performed by three identified occupations, an "Administrator", "Deputy Practice Manager" and "General Practitioner".

Table 2: Dataset exert

Observed Occupation	**Task**
Administrator	Medical Coding of letters and other documents
Administrator	Register new Patients
Deputy Practice Manager	Staff rotas (clinical or practice)
Deputy Practice Manager	Use accounting system, online or desktop software, for finances
General Practitioner	Indicating what is to be coded by a medical coder
General Practitioner	Review letters
General Practitioner	Respond to tasks in electronic medical record

The primary dataset is (referred to as the *dataset* hereafter), comprised of the 15 unique occupational roles described above. These 15 occupational roles perform the 137 unique tasks. We have documented different characteristics of these tasks, such as the type of technology used, or who else performs the task, whether the task is clinical (or not), structured (or not), plus other categories.

The formulated dataset consists of 333 unique occupation-task pairs, given that most tasks are performed by more than one occupation. The complete list of observed tasks is shown in Appendix 11.1.

5.1 Accessing the skills, knowledge, and ability requirements of work tasks

The dataset was then augmented with numeric attributes available from a publicly available occupational survey produced for the US Department of Labor called the Occupational Network (O*NET) 2016 database (National Center for O*NET Development, 2018). O*NET provides key features of an occupation as a standardised and measurable set of numeric variables, and provides open-ended descriptions of the tasks each occupation performs. The features are measured quantitatively on a 1 to 5 scale by dozens of employees and experts in the O*NET database; the strengths and weaknesses of ONET are reviewed in (Handel, 2016).

The O*NET database is a valuable resource, containing 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 (DWAs) and nearly 20,000 individual occupation-specific tasks arranged in a hierarchical structure.

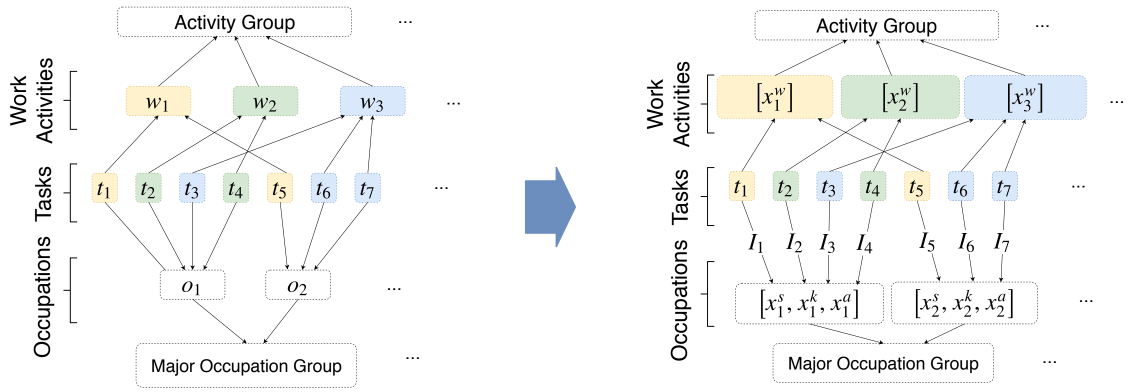


Figure 3: Overview of O*NET database architecture identifying how occupational characteristics are propagated to the work activities using the task's importance I_t relative to its occupation.

Shown in Figure 3, and described in detail in (Duckworth et al., 2019), the idea here is to use occupational-level characteristics $\boldsymbol{x}_o$ relating to the skills, knowledge and abilities required to perform an occupation $\left[\boldsymbol{x}_o^s, \boldsymbol{x}_o^k, \boldsymbol{x}_o^a\right] \in \mathbb{R}^{+120}$ in order to produce a feature vector for each of 2,067 Detailed Work Activities $\boldsymbol{x}_w \in \mathbb{R}^{+120}$ in the database. Where each Detailed Work Activity in O*NET is a collection of (child) tasks in the hierarchy. Since the dataset is developed in the US, our assumption is that the skills, knowledge and abilities required to perform an activity are representative between the UK and the US.

The O*NET occupation characteristic variables include: 35 *skill* attributes, 33 *knowledge* attributes, and 52 *abilities*, examples of each set are described in more detail next. For a complete list of variables, see Appendix 11.2.

The set of nine O*NET variables used in previous work to describe "bottlenecks to computerisation" in the popular work on automating occupations by Frey and Osborne (2017), are a subset of the 120 variables used in our study.

Skills: The 35 skills comprise of 10 Basic skills, such as: "Active Learning" - Understanding the implications of new information for both current and future problem-solving and decision-making. "Reading Comprehension" - Understanding written sentences and paragraphs in work related documents.

Six Social skills, such as: "Coordination" - Adjusting actions in relation to others' actions. "Negotiation" - Bringing others together and trying to reconcile difference one Complex Problem solving skill. 11 Technical Skills, three System skills and four Resource Management skills.

Abilities: The 52 abilities comprise of 21 Cognitive abilities, such as: "Category Flexibility" - The ability to generate or use different sets of rules for combining or grouping things in different ways. "Deductive Reasoning" - The ability to apply general rules to specific problems to produce answers that make sense.

Nine Physical abilities, such as "Dynamic Flexibility" - The ability to quickly and repeatedly bend, stretch, twist, or reach out with your body, arms, and/or legs. "Stamina" - The ability to exert yourself physically over long periods of time without getting winded or out of breath. Ten psychomotor abilities and 12 sensory abilities.

Knowledges: 33 Knowledges, such as: "Mathematics" - Knowledge of arithmetic, algebra, geometry, calculus, statistics, and their applications. "Psychology" - Knowledge of human behavior and performance; individual differences in ability, personality, and interests; learning and motivation; psychological research methods; and the assessment and treatment of behavioural and affective disorders.

5.2 Knowledge Representation of tasks

The O*NET taxonomy is extremely granular at the task-level, and it's important not to collate a single task with a collection of tasks. Answering a phone, in this case, we judged to mostly mean beginning audio receipt on a phone system. We felt confident this was the case, as on the clinical side we noted other tasks such as "Use Telephone to consult with patients" and "Phone consultation with other medical professionals". In other words, each task in the O*NET database is a highly granular and highly literal component of a larger multi-task performance. Therefore, if a task is shared across occupations, it means the same thing in each, and other characteristics of the performance are differentiated by other tasks (as in the previous two examples). This is something we came to understand after working extensively with the O*NET database, examining tasks across all types of occupations.

The rationale for our high-level task classification is that the O*NET representation of a task (the skills, knowledge and abilities required to perform it) are aggregated features of the occupations who perform the task. As such, if a task was performed by a highly clinical occupation, it would contribute that occupation's clinical components (particularly in the "Science" skill, the "Biology" knowledge component, and the "Medicine and Dentistry" knowledge component). We took the position that variations between the medical training required to perform clinical tasks in the United States, and hence in the O*NET data, may not accurately represent the attributes of clinical staff performing the same task in the UK. We therefore excluded those tasks considered most at risk.

However, strictly ensuring we obtain accurate representations is difficult as the nature of the data is somewhat subjective. Therefore, we felt confident that our aggregation procedure reduced the likelihood of ignoring occupation-specific variation in automatability and clinical content. Of course, it may not always be the case that an entirely-clinical occupation does not perform that task. In that case, their feature vector would contribute a higher science/biology/medicine score in aggregation.

5.3 Measuring Automation Potential: A Scale of Automation

One of the main contributions of this work is the application of an "automation scale" to primary healthcare tasks, which we use to predict the automability of each task. In the section below we discuss the construction of an expert survey used to create an ordinal scale of task automation through eliciting expert knowledge of automating technologies.

This scale is designed to judge what tasks are automatable with technology that is available *right now*, and not technology 5 or 10 years in the future (or technology "on the horizon"). This survey provides data on automation, and an element of practicality of the outcomes of this report: in that the suggestions and findings in this report are technologically available to implement right now, and not within some hypothetical amount of time.

5.3.1 Expert Survey of Automation

We elicited expert knowledge of state-of-the-art automation technologies substituting for human work-place tasks across a huge number of experienced experts *in the field*. In order to obtain current estimates of the automation potential of tasks, we conducted the largest task-based survey of its kind: 156 experts in the machine learning, robotics and AI communities [2]. Participants were reached through relevant academic mailing listservs and our own volunteer lists we collected at multiple machine learning conferences (in particular the Conference on Neural Information Processing Systems, 2015 and 2017). Experts verified their academic and industrial experience and (optionally) their contact information.

Each survey participant was presented with five O*NET occupations (one at a time) and the five tasks with the largest task-importance score for that occupation. Participants were asked the following question:

"*Do you believe that technology exists today that could automate these tasks?*"

Then rated each task as either: *Not automatable today* (score of 1.0), *Mostly not automatable today (human does most of it)* (2.0), *Could be mostly automated today (human still needed)* (3.0), *Completely automatable today.* (4.0), or *Unsure*. The scale is displayed graphically in Figure 4. Respondents also reported overall confidence in their answers.

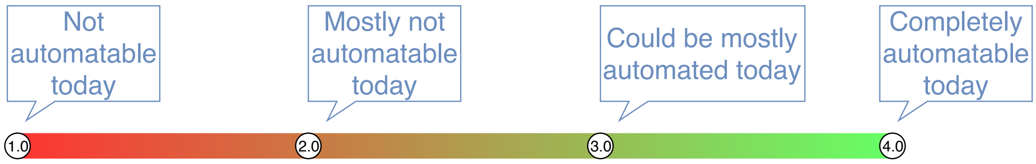


Figure 4: Linear Scale of Automatability.

Our dataset contains 4,599 task level responses from 156 academic and industrial experts from around the world, and across various scientific and industrial fields. The distribution of responses across each answer category is shown in Figure 5 (red).

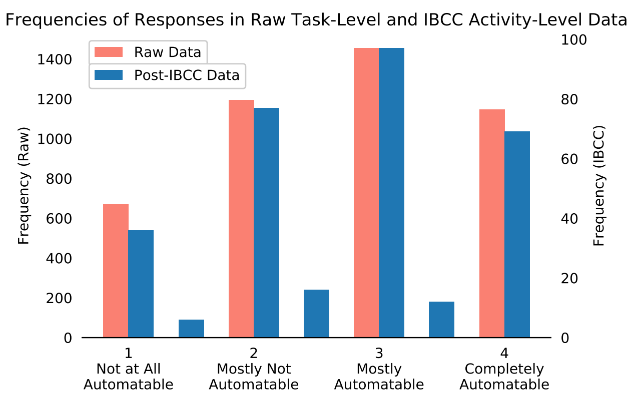


Figure 5: Expert survey response statistics. Distribution of expert task-level responses, and the IBCC combined activity labels.

The distribution of geographic location (left) and academic experiences of respondents are shown in Figure 6 (right). 97% of participants had field-relevant academic experience, with most participants

[2]Survey: `http://www.robots.ox.ac.uk/~survey/`

coming from computer science, ML, robotic or AI backgrounds. Most responses (52%) came from the US, UK and Germany, while the rest came from 30 other countries.

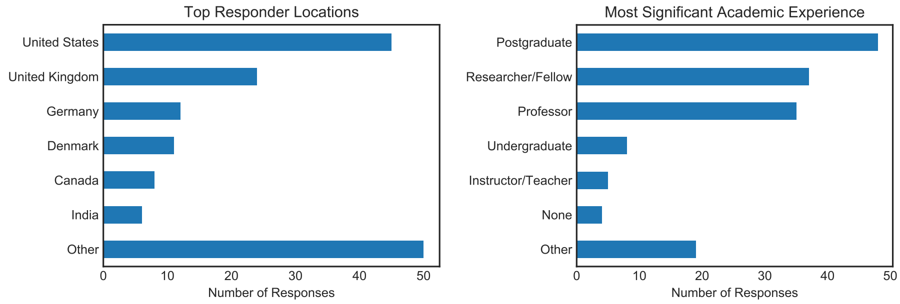


Figure 6: Expert survey response statistics. (left:) Number of responses by country. (right:) Number of responses by academic experience.

We combined each task's multiple expert labels using Independent Bayesian Classifier Combination (IBCC), a principled Bayesian approach to combine multiple classifications (Kim and Ghahramani, 2012; Simpson et al., 2013). IBCC creates a posterior prediction over labels that reflects the individual labellers' tendencies to agree with other labellers over ultimately chosen label values. We averaged IBCC task scores into their task's work activity, and so activity-labels concentrate around whole and half values and we round the final values to the closest 0.5 (a half-class). A score of four represents a "fully automatable" work activity, and one represents a work activity that "cannot be automated" using currently available technology. The distribution of IBCC activity labels is shown in Figure 5 (blue).

The survey demographic was specifically targeted towards technology experts, as opposed to healthcare experts, because we believe that classifying how automatable a specific task is requires little-to-no subject matter knowledge. We believe a survey of many experts is a transparent and authoritative method of obtaining data about the current state of automation at the task-level. This methodology requires no forecasting or prediction by the participant, reducing the risk of subjective responses.

We believe this expert survey provides a reliable reference 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.

Task Automatability

Distributions of Primary Care Automatability

6 Task Automatability: Distributions of Primary Care Automatability

We combine the following three data sources: 1) our observed healthcare task dataset, 2) augmented by the O*NET skills, knowledge, and abilities required to perform such tasks, and 3) the automatability scores of work activities (DWAs) contained in the automation survey. Then, we use a machine learning framework to learn a functional mapping between the skills, knowledge and ability characteristics of work activities and the ground truth automatability elicited from the expert survey.

We seek a flexible function estimation capable of modeling complex, non-linear relationships between the features (skills, abilities, knowledge) and automatability in high-dimensional space. Given the social scientific nature of the study, we also desire a measure of model uncertainty. Gaussian processes (Rasmussen and Williams, 2006) 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. We specifically use the ordinal likelihood function introduced in (Chu and Ghahramani, 2005) to reflect the nature of having discrete labels but with an ordinal interpretation (*not at all* to *completely* automatable). Other methods are compared in Duckworth et al. (2019).

The algorithm uses the trends and patterns it has learned from labelled data (from the expert survey) to provide a smoothly varying, probabilistic assessment of automatability as a function of the input variables on the unlabelled data. For the Gaussian process, this function is non-linear, meaning that it flexibly adapts to the patterns inherent in the training data. This approach has been successfully applied to *occupation*-based data in (Frey and Osborne, 2017) and (Bakhshi et al., 2017).

We use the (N=313) work activities included in the automation survey as a *training dataset* in order to train a machine learning model using Gaussian Process Ordinal Regression model. Once trained, the model allows us to estimate the automatability of "unlabelled" work activities, i.e. the observed healthcare dataset tasks where an expert surveyed label of automation potential would have been prohibitively expensive to generate from human annotators. These predictions provide an automation score for each of the observed tasks in primary healthcare, where one represents "not at all automatable" and a 4 represents a task that is "completely automatable".

We employ a Kernel Density Estimation (KDE) analysis to show the distribution of the automatability of the observed healthcare tasks. KDE is a technique that allows us to visualize a smooth curve,

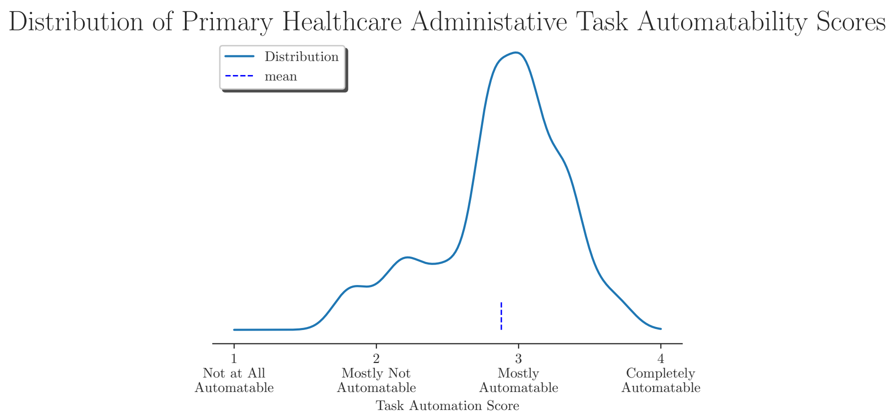


Figure 7: Distribution of Automatability Scores showing that most tasks in our Primary care dataset are "mostly automatable". Plot shows the Kernal Density Estimate of the Scores.

given a set of data, that represents the "shape" of the data distribution as a replacement for a discrete histogram. We use a Gaussian Kernel to generate smooth distributions.

Figures 7 shows the distribution of the potential for automation in primary care of all the administrative tasks performed by all practice staff. We can see that the machine learning model predicts that over 40% of administrative tasks are "mostly" or "completely" automatable using current technology. The mean is shown as a vertical dashed (blue) line at 2.9.

One of the ways we can interpret this number is that 44% of the administrative work performed in primary care presents opportunities to be altered or entirely eliminated from workloads. We believe this translates into significant room for rethinking how practices operate, how work is distributed, and how occupations (and the tasks they perform) are designed in the future. We discuss these issues further in Section 9.

The limitation of Figure 7 does not tell us which occupations perform which tasks or which occupational tasks are *more* automatable. It is simply a high-level view of what proportion of tasks are possible to automate today, given the assumptions in our automation scale. We tease out occupational-level data in Figure 8 and 9 below.

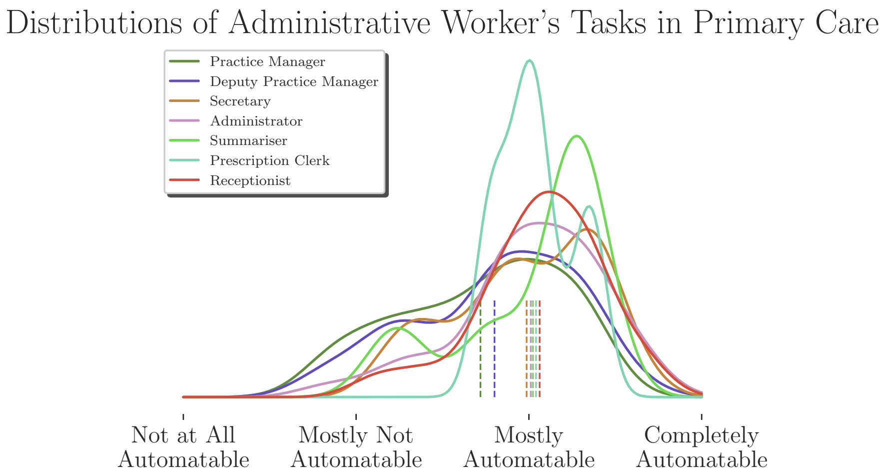


Figure 8: (KDE) Distribution of Task Automatability for Primary Care Administrative Occupations.

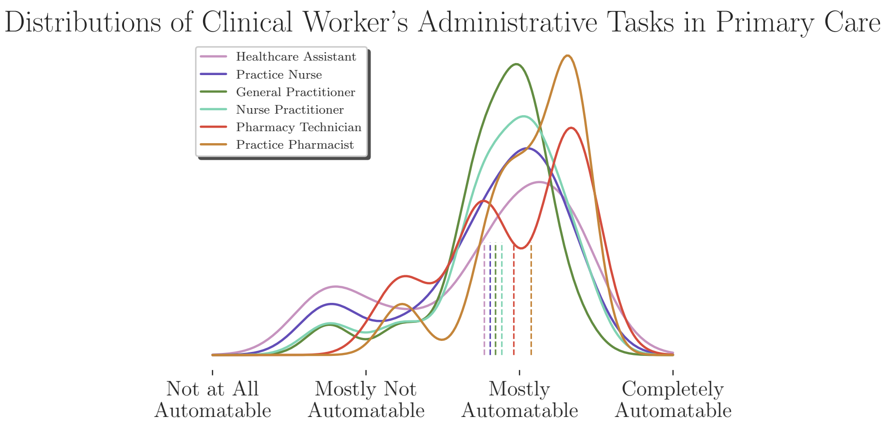


Figure 9: (KDE) Distribution of Task Automatability for Primary Care Clinical Occupations.

To accurately represent the administrative data across different occupations, we present the distributions of their task automatability scores for those primary care occupations whose main responsibility are administrative functions. We do not include clinical tasks in this analysis for reasons previously stated in Section 4. However, Figure 9 does present the automatability distribution of clinical occupations performing administrative tasks. As discussed above, tasks overlap and the automation potential of a task is occupation-agnostic, i.e. a task's automation score is fixed no matter which occupation performs it. For example, medical coding has an inferred automation score of 3.2. Thus, automating this task, which is mostly performed by receptionists, summarisers, and secretaries, would also eliminate that part of a GP's workload if they also perform medical coding at their practice.

An example of the most highly automatable tasks in our data include: payroll and managing finances, checking and sorting post, printing letters, managing patients through texting, management of paper archives (onsite or offsite), transcription, practice email account management, letter scanning, checking for errors in paperwork, and relaying information between personnel (for example messages or new employee inductions).

To further understand the automatability predictions made for each healthcare occupation, we generate three "automation categories" for healthcare tasks. They are: *automatable* (where the predicted task automatable score is > 3.0), *not*-automatable (predicted scores < 2.0), and *partly*-automatable (scores in-between). Figure 11 present the predicted automatability scores of administrative occupation's tasks classified into these categories.

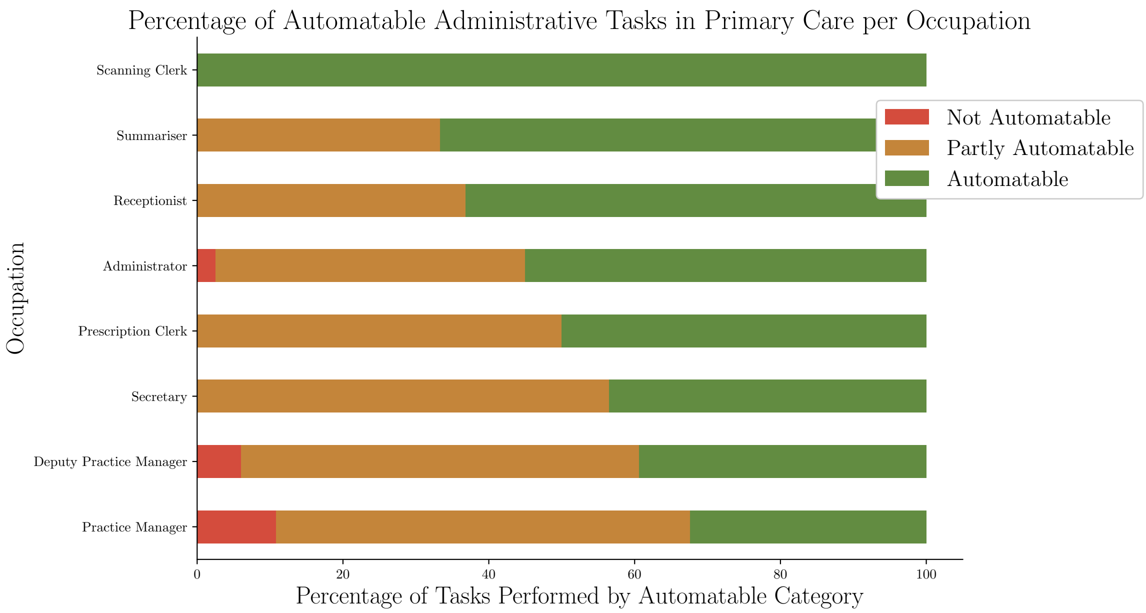


Figure 10: Percentage of Administrative Tasks performed in Primary Care by Automatability Categories.

Rather than present the entire list of tasks that are predicted to be automatable (this is available in the Appendix), we situate these numbers and the arising possibilities in the qualitative fieldwork we have conducted. This is critical for two reasons: First, it demonstrates how some technologies could be implemented straight away, automating what are considered as "easy wins" for technology. Whilst other important considerations are also highlighted: We work to clarify the complexities that emerge from the data because it is important for system designers, developers, and other stakeholders to be aware of our fieldwork and analysis before implementing technologies.

Second, it is important to be aware of the current state of primary care and how work (and collections of tasks creating workflows) are performed. While certain tasks may not be automatable, the following accounts may provoke other ideas and configurations that could improve quality and

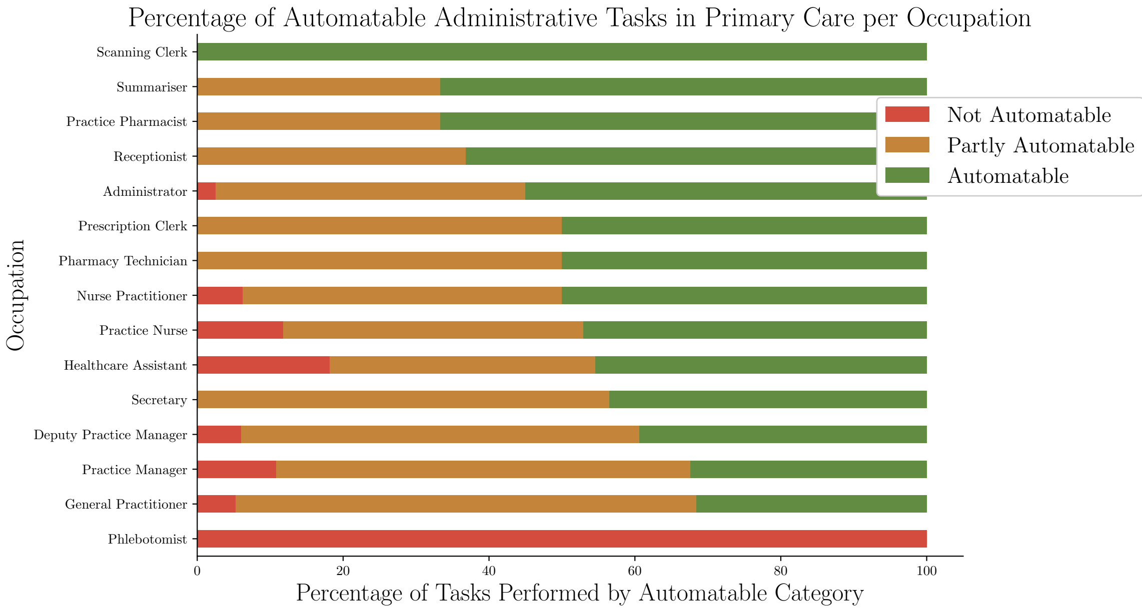


Figure 11: Percentage of Administrative Tasks performed in Primary Care by Automatability Categories.

workflow efficiency.

The following section presents themes and scenarios taken directly from fieldwork, including interviews, observations, and focus groups. These are included to help contextualise the findings of our research. We provide automation scores where appropriate, and relevant technologies that exist which may be applicable in a primary care context. The final section presents a discussion of the primary care practice of the future, disseminating what occupations *could* look like if or when the most susceptible tasks to automation are removed from their workload. We highlight the opportunities, challenges and current realities, and how these might evolve.

Critical Incidents

Workflows and Task Exemplars

7 Critical Incidents: Workflows and Task Exemplars

The following 10 examples are presented as "critical incidents": observed episodes that encapsulate and help ground our quantitative data. These incidents focus on the following examples: use of templates in the health record, GP record-keeping, what happens when patients do not attend clinics, letters and relevant administrative tasks, the use of paper in primary care, single and super-practices, stock and inventory management, staff scheduling, and making/managing phone calls.

We focus exclusively on administrative tasks because as we have seen in Figure 1: Administrative work is the majority of the work undertaken in primary care.

7.1 Templates

The use of templates is an important activity to observe and consider, because this is a task that has already been partially automated. Repetitive letters and highly structured documents have typically had templates created for them. These templates consist of preformatted documents or fields with information already populated by a program based on a patient's information or based on a previously run search query. The two most commonly used types of templates in general practices are: 1) letter writing templates, for the many different types of letters that a practice will generate; and 2) electronic medical record templates.

Word document templates are used to assist in the writing of letters to secondary care specialists, to employers with the results of medical exams, to patients inviting them to attend a clinic, and a range of other standard letters. If the practice needs to communicate with an external entity, it is almost always done by letter. Even telephone calls often require a follow-up letter. Communications from GPs to patients or specialists often involve the GP dictating text into an audio recorder, which is later transcribed and copied into a Word template.

The Word template always consists of the practice letterhead and other professional formatting. It will often be based on a specific template among many options, depending on the context of the letter. The template will have "boilerplate" text for the introduction and ending of the letter, with the body of the letter copy/pasted in from an earlier dictation. The boilerplate text is edited as the drafter of the letter sees fit. In some cases, the modification and editing of a letter can take as much time as it might require writing from scratch. While templates are intended to partially automate and therefore speed up the process of letter production, it remains a time-consuming task that requires administrative workers (or clinical workers when they write their own letters) to spend a large percentage of their workload each day working with these "automated" templates.

The second type of template is features in the electronic medical record designed to reduce keyboard strokes from repetitive or structured writing. These templates function like auto-complete on a smartphone keyboard. The idea is that you start typing and as you enter certain words, or key medical terms, a suggestion will pop up to format or complete the text.

The field researcher often observed GPs inputting data into patients' records. When starting a new entry, the system provides suggestions on relevant templates and pops up automatically with pre-populated documentation fields, although they were often ignored by the GP in favour of their personal writing style. This is one example of an application that might be more helpful if it could be tailored to an individual's preferences.

Data to tailor auto-suggestions could be generated by asking questions to help understand why a particular template was used in a certain way, for example: Why was a template ignored? Under what conditions are they used? What types of work does the template need to support? And, is the template too generic?

7.2 General Practitioner Documentation Work

Automation technologies can be expensive investments and take time to develop and implement in a healthcare organization. It is not economical to automate infrequent tasks that are easily/quickly performed. Automation provides the greatest advantages when it is applied to a task that is frequent or time consuming, is important, and repetitive. Our observations suggest that clinical documentation could benefit from automation for the following three main reasons.

First, these are tasks that every clinician must complete with a degree of urgency. It is considered good medical practice to document the patient consultation as soon after the end of a consultation as possible. This task cannot be left to the end of the day. Second, documentation requires attention to detail and sufficient time: Even brief encounters require accurate documentation (Ammenwerth and Spötl, 2009). Third, this work must somehow be squeezed into a busy schedule, typically performed during the time it takes a patient to leave the examination room, before the next patient arrives.

Smart technologies exist that are able to accurately capture a natural language discussion between two people. High quality speech recognition and transcription has been available on consumer smart phones for years. Recently the use of AI systems to streamline and automate documentation tasks in healthcare has been recommended (Lin et al., 2018; Verghese et al., 2018), for example, by using natural language processing to dissect patient-doctor conversations and create notes (Klann and Szolovits, 2009).Incorporation of smart technologies into clinical documentation practices has been a controversial topic, since research presents them as a web of complex work practices with institutional, social and situated dimensions (Bansler et al., 2016). However, it should be feasible to use a speech recognition device to convert the patient and provider discussion into a text transcript. With natural language processing this text could be further reduced and annotated to provide clinically relevant information that matches the way the provider writes, categorizing it to appropriate sections of the medical record. If connected to the electronic medical record, this device could utilize data already in the patient's record as well as that of similar patients at the practice, supporting the practitioner's decision making and partially automating the documentation responsibilities. After the clinical encounter the clinician could review a summary for editing and saving to the patient's record. Almost the entire process could be automated, at least in theory.

The above scenario was presented at each of the focus groups, provoking lengthy discussions. It was seen as potentially useful by some, with many "when can we have that" comments. While technically possible, some participants felt it might have negative consequences on the clinical encounter and could even undermine the skills and problem-solving ability of the clinician. One GP argued that generating an entire transcript of the conversation would be a waste of time, generating a great deal of unnecessary information: "I don't need the entire transcript, it's not useful or clinically relevant to me."

It was not uncommon for GPs to re-write and/or edit their notes, particularly notes for new cases or complex circumstances; for example, in mental health encounters, or when they were previously written by trainees. What may look like a relatively simple administrative task in fact involves a considerable amount of clinical skill and experience. Removing this task from the clinical workflow would remove an opportunity for the clinician to think and reflect critically about an interaction.

Research shows that writing engages the brain, allows GPs to be better observers, supports empathy, and engages critical thinking (Wald et al., 2012).

7.3 Did Not Attend

The field researcher observed several notices on display in the various health centres announcing the number of people who did not attend (DNA) their appointments, sometimes comparing the local rate to a regional or national average and giving an indication of the cost to the practice of these "no shows". Reasons for a DNA can range from clerical errors, communication issues, patient anxiety, absence of transportation, or other unforeseen events, including failure to cancel the appointment, or spontaneous resolution of the problem making the appointment unnecessary. A DNA leads to a set of tasks that may include making a telephone call to rebook the appointment, sending a letter or text message to the patient, or making an entry on the patient's electronic medical record. Interviews with staff revealed a range of reactions and frustrations with DNAs.

During focus groups and interviews a scenario was discussed where smart scheduling technology might be used to predict when a patient is likely to DNA based on their prior attendance history, the particular time of year, and relevant medical and transportation information available. This technology might help to reduce the DNA rate, perhaps by double booking likely DNAs, or inviting patients from the waiting list who could attend at short notice.

Interestingly, the subsequent discussions revealed a lack of incentives for clinicians to reduce patient DNAs. This is because DNAs provide valuable opportunities for nurses, GPs, and other clinicians to "catch up" on their workload. This was also observed during the fieldwork. One example observed included a practice nurse who was running behind on documenting patient notes, and was further behind because a travel vaccination consultation had run over time: Travel vaccination and immunisation appointments can often take longer than the allotted time depending on the questions the patient asks and their travel destination. The practice nurse was preparing to work into her scheduled break to catch up on her tasks, but one of her scheduled patients did not turn up. She was relieved about this because it afforded her more time to complete her tasks and take her scheduled break.

7.4 Letter Work

Of all the ways that primary care as an institution can communicate with patients and other NHS entities – including phone calls, emails, letters, text messages, and fax – letters are the most widely used media. We refer to this as letter work. Letter work is the process of reading and writing letters, or more precisely, deconstructing incoming letters to understand what they are requesting or what actions they create, and crafting new letters that contain all the relevant information that the letter needs to address.

The exact workflow and who does what with letters can vary from practice to practice. The workflow can look remarkably different in a single-site practice working independently, to that in a super-practice, where services are shared across multiple sites. An individual practice usually has a receptionist processing every part of the letter, while a super-practice might outsource all medical coding, audio transcription, and typing to a central location.

As an example, the field researcher observed the following workflow: Receptionists at a small independent practice sort incoming letters into a red, amber, or green status depending on the urgency of the required response. They scan the letter into the system and use DocMan to sort them using the colour coding system. Then, one of the receptionists will have the task of "dealing" with the letters. This involves first opening a letter, which has been briefly coded by a GP using the highlighter tool in the document management software, reviewing the highlighted medical terms or phrases, and allocating appropriate Read codes (a standard clinical terminology system). The letter is then archived on the system, the task is marked complete, and the receptionist continues with the next letter. If the receptionist has a question about the appropriate Read code, they add a text comment to the document, save it, and assign it as a task on the system, i.e. sending it back for the GP's attention. Other letters, for example, from a medical supplier requesting information about a patient who requires a specific device, must find the required information in the patient's health record, then use a template letter in order to respond to the supplier.

This is the nature of letters: reading them and understanding what must be responded to, gathering the required information and then packaging all that information together into a response, or coding the letter and putting it in an archive. The length of letters ranges from precise and short to long and verbose. Letters have a different utility based on who is writing them, who is reading them, and the context. This can be illustrated in the following example: The field researcher observed a GP reading a nine-page letter from secondary care. The GP skimmed the letter on his monitor and after a few minutes said "Ah hah, you see here it is", highlighting a paragraph with the mouse cursor, saying "This is all I need, out of this whole document this is what is relevant to me." He continued: "I can tell the rest of this letter is for their [the specialist's office] purposes, it's a form of documentation, I don't need all this." The GP did not require most of the content of the letter, only a single paragraph. His prior experience of the specialist who wrote the letter enabled him to extract the relevant information efficiently and discount the rest.

7.5 Miscellaneous Administrative Work

While much administrative work takes the form of the letter work discussed above, there are plenty of other tasks that are general administrative actions. These are tasks that occur in all offices, including completing spreadsheets, writing text documents, generating reports, providing information to schedule and support staff meetings, writing emails, and answering telephone calls. One technology expanding in use in many other industries, which could be applied in primary care, is Robotic Process Automation (RPA). RPA is a way to automate repetitive tasks and other processes through a series of digital scripts and application programming interfaces (APIs). This type of system is widely used for financial management and bookkeeping, but no use of RPA was observed at any of the practices in our sample.

One example of administrative processes observed during the fieldwork illustrates inefficiencies due to lack of appropriate automation. This concerned the task of adding digital documents, namely PDF files, to software for managing, manipulating, and archiving. Practices receive many digital documents via email. While these could be directly imported into the document management software, this was not done in one practice we observed. Instead, the documents were printed and scanned to a different computer, and then that file was uploaded to the document management software. It was unclear if the reason for this was to produce a paper document that, along with the digital document, would be archived and stored on the premises or whether there was some other reason for this inefficient procedure, which went unchallenged by those responsible for carrying it out.

7.6 The Use and Reliance on Paper for Particular Workflows

Our quantitative analysis shows that more than 40% of administrative work are mostly automatable, and others completely automatable. However, this prediction must be seen against the function of other tasks and tools that, while they may be automated, do not always lead to greater efficiency. Even in the most digitally transformed practice, paper was still found. One example of the beneficial use of paper is people's personal task lists: i.e. their external memory devices. It was common to observe nurses in busy clinics using a scrap of paper to write down all kinds of notes, tasks, numbers to remember, questions to ask, names, and other pieces of information. The use of paper and pen allows the nurse, GP, or healthcare assistant to quickly work with information, remember something, establish a quick task list, or otherwise work in a manner that is personal to the clinician. At the end of the clinic these scraps of paper are shredded or otherwise destroyed in secure bins. Maintaining this type of information on a computer would require several clicks and would be less accessible.

Another role of paper in practices is represented by the many printed letters. These physical documents can be demarcated, and information attached to them as they are transferred through the practice. An interesting use of paper was to apply a rubber stamp to letters and other paper documents to establish a common task list and add extra metadata to the document. Figure 12 provides an photograph of one of these stamps; The stamp shows the date the document was received, along with a series of check boxes, in effect adding extra metadata to the document. The stamp enables the recording of tasks that need to be performed to the document. The tick in the document pictured indicates that a drug update was requested to the pharmacy technician to process. Other uses of the stamp include summarising information from the document to the health record, or processing information from specific treatments and archiving. Stamps also often show to whom the document has been assigned. These stamps are available for a variety of different types of documents. The example here is from a more clinically oriented document, while other stamps are used for administrative and financial purposes.

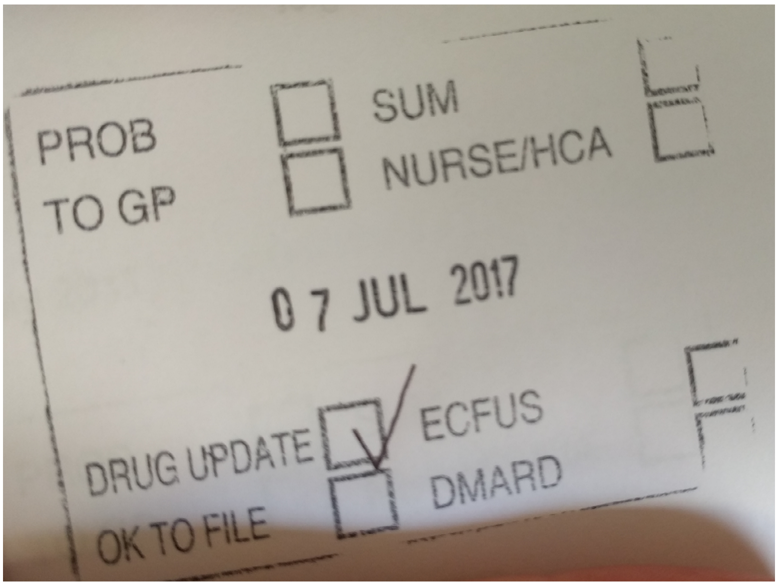
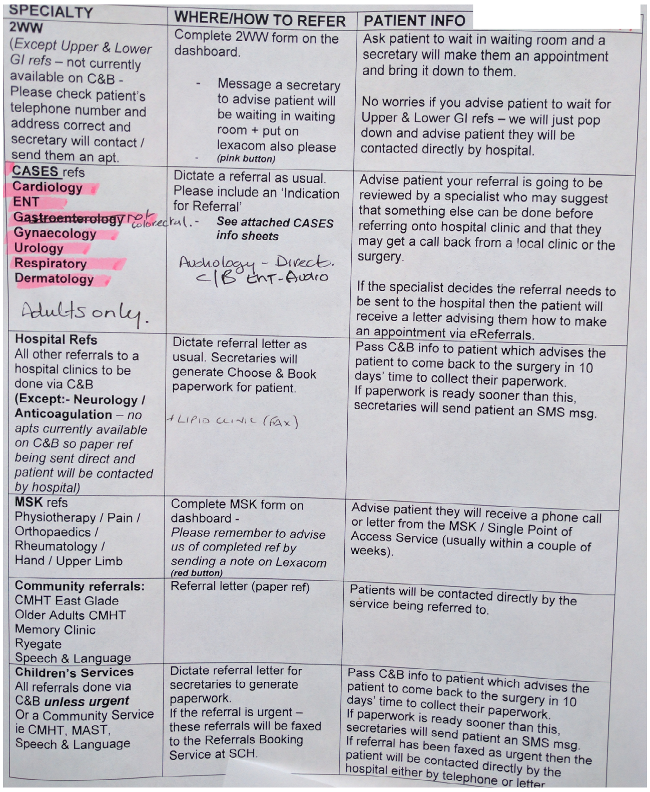


Figure 12: Paper stamps example

Another use for paper in primary care is for reference purposes. Clerical and administrative staff have to navigate a vast range of organisations, institutions, local groups, specialists, policy and legal frameworks, patient requests, and practice-specific protocols. The walls of admin offices, phone banks, practice manager offices, and even clinicians' offices are well covered – and in some cases layered – with different paper documents and reference materials. The theme that unites many of these paper documents is instructions on what to do or where to go in particular eventualities.

Figure 13 is an example spreadsheet showing the complex process of making e-referrals to specialty clinics. Each specialty service has its own processes and each referral demands a specific set of tasks and workflows. Some clinics require paper letters, some require a fax, and others use an online web portal. The process itself can evolve, requiring the addition of new information to be added to the document, as shown in the handwritten addenda.



SPECIALTY	WHERE/HOW TO REFER	PATIENT INFO
2WW *(Except Upper & Lower GI refs* – not currently available on C&B - Please check patient's telephone number and address correct and secretary will contact / send them an apt.	Complete 2WW form on the dashboard. - Message a secretary to advise patient will be waiting in waiting room + put on lexacom also please - *(pink button)*	Ask patient to wait in waiting room and a secretary will make them an appointment and bring it down to them. No worries if you advise patient to wait for Upper & Lower GI refs – we will just pop down and advise patient they will be contacted directly by hospital.
CASES refs **Cardiology** **ENT** ~~**Gastroenterology**~~ not colorectal. **Gynaecology** **Urology** **Respiratory** **Dermatology** Adults only.	Dictate a referral as usual. Please include an 'Indication for Referral' - ***See attached CASES info sheets*** Audiology - Direct. C&B ENT-Audio	Advise patient your referral is going to be reviewed by a specialist who may suggest that something else can be done before referring onto hospital clinic and that they may get a call back from a local clinic or the surgery. If the specialist decides the referral needs to be sent to the hospital then the patient will receive a letter advising them how to make an appointment via eReferrals.
Hospital Refs All other referrals to a hospital clinics to be done via C&B **(Except:- Neurology / Anticoagulation** – *no apts currently available on C&B so paper ref being sent direct and patient will be contacted by hospital)*	Dictate referral letter as usual. Secretaries will generate Choose & Book paperwork for patient. + LIPID CLINIC (fax)	Pass C&B info to patient which advises the patient to come back to the surgery in 10 days' time to collect their paperwork. If paperwork is ready sooner than this, secretaries will send patient an SMS msg.
MSK refs Physiotherapy / Pain / Orthopaedics / Rheumatology / Hand / Upper Limb	Complete MSK form on dashboard - *Please remember to advise us of completed ref by sending a note on Lexacom* *(red button)*	Advise patient they will receive a phone call or letter from the MSK / Single Point of Access Service (usually within a couple of weeks).
Community referrals: CMHT East Glade Older Adults CMHT Memory Clinic Ryegate Speech & Language	Referral letter (paper ref)	Patients will be contacted directly by the service being referred to.
Children's Services All referrals done via C&B ***unless urgent*** Or a Community Service ie CMHT, MAST, Speech & Language	Dictate referral letter for secretaries to generate paperwork. If the referral is urgent – these referrals will be faxed to the Referrals Booking Service at SCH.	Pass C&B info to patient which advises the patient to come back to the surgery in 10 days' time to collect their paperwork. If paperwork is ready sooner than this, secretaries will send patient an SMS msg. If referral has been faxed as urgent then the patient will be contacted directly by the hospital either by telephone or letter

Figure 13: Complex e-referral spreadsheet example

7.7 Single Site Independent Practices and Multi-Site Super Practices

A major differentiating factor in how practices perform tasks and who does what is the size of the organisation - whether it is a single-site practice operating independently or a multi-site "super practice". The latter have a conglomerate structure with a centralised administrative function and governing body. Typically, these super-practices will have a single "brand identity" and shared services.

The field researcher observed more sharing of tasks at single site practices than at the larger organisations. An example that emerged from the fieldwork concerns finances: At single-site practices the practice manager, deputy practice manager or senior administrator would be responsible for employee payroll, bookkeeping, and other essential practice finances, whereas these functions were more likely to be outsourced at the super-practices.

7.8 Stock and Inventory Management

Our researcher observed that some practices work more closely with community pharmacies than others. These provide the clearest empirical example of robotic automation. Figure 14 below shows a community pharmacy stock room curated by a robot. It is a narrow space that people are rarely required to enter except for maintenance. The robot stocks, organises, searches and removes expired drugs, and fulfils prescriptions autonomously.

Prescriptions are dispensed by the robot placing the items on a conveyor belt that delivers each item to the pharmacist's bench. The items slide down a spiral slide from the ceiling and into a container next to the human pharmacist. The pharmacist then checks the order, packages it for the customer and makes themselves available for consultation. This robotic pharmacist assistant has removed several tasks from the community pharmacist. Prior to the use of this robot the pharmacist would be responsible for inventory control, stocking, organising drugs, checking for expired stock, and dispensing. Much of the manual labour in maintaining stocks has been eliminated. When the field researcher asked pharmacists working with the robot if they had noticed any mistakes or errors made by the robot they said no, apart from an occasional barcode error or missing barcodes.

Figure 14: Pharmacy stock robot example

Stocking items and maintaining inventory is, in general, a highly automatable task. For example, international retailer Wal-Mart has been using robots in 50 stores across its US locations since early 2018 (Vanian, 2018). Once a day the robot travels down every aisle, scanning it three times, to count the quantity of items and checking if each item is located correctly. The robot does this mostly autonomously. Another example of automated stock management has been developed by the start-up company Pointy (Danziger, 2018). It uses a device that connects to point of sale barcode scanners. As items are scanned it adds those items to a searchable online database, efficiently creating an inventory log for the store managers. A similar system could be used in general practice to maintain the stock of items in each clinical room and to prompt reordering.

7.9 Scheduling Staff

Creating rotas, scheduling staff, and creating schedules for clinics are all tasks that require complex coordination, discussion, and revision. Responsibility for rotas and the methods used to schedule events such as annual leave, holiday times, and sick leave varied greatly between field sites. One strategy that we observed was for senior staff in each occupational group to take responsibility for managing the rotas of their colleagues. For example, the field researcher observed a receptionist manager who arranged the schedules of all the other receptionists. Figure 15 below shows a large wall calendar on the back wall of the receptionist's area used to orchestrate these tasks. Each receptionist was assigned a colour and different symbols denoted what type of "time" was being assigned. Our research observed nurse practitioners and practice nurses using a similar system.

Figure 15: Staff schedule example

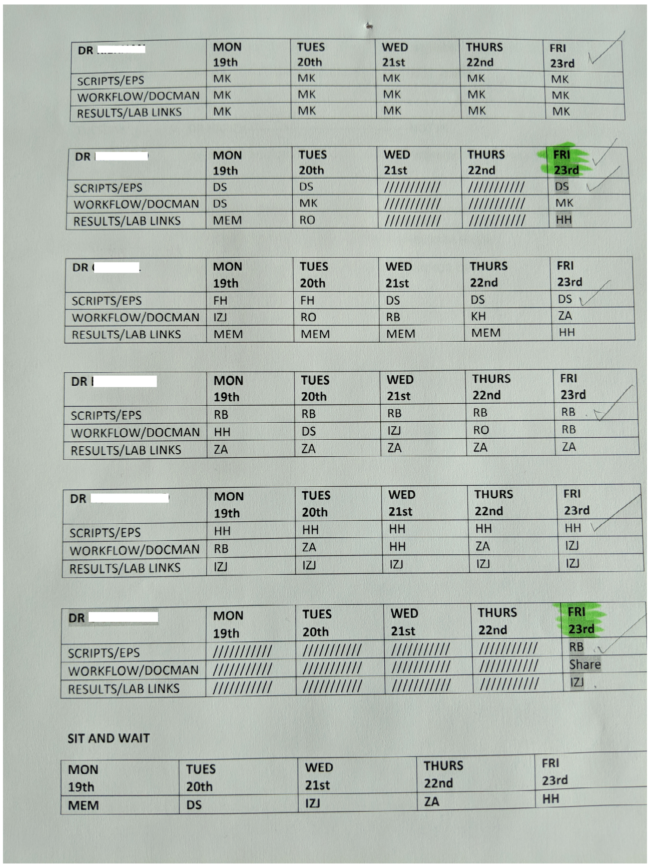
Other strategies involved the use of online software that allows each staff member to log-in and administer their own time management. Or, smaller practices may have the practice manager perform all of the scheduling, typically on paper or through email. Complex spreadsheets were used at larger practices. Whatever the method employed, staff scheduling could be complex and sometimes problematic, involving unforeseen disruptions requiring frequent consultation and negotiation between staff.

The same applies to booking GPs' holidays and other absences, while maintaining workflow. The GPs' core duties of reviewing and signing prescriptions, reviewing letters, addressing their workflow in the electronic medical record, and reviewing lab tests must be reallocated to other GPs or locums. The Figure 16 shows an example of such a system. Each of the responsibilities of a missing GP is allocated to one or more of the GPs on duty. Their initials show who is doing which allocated tasks. This document also shows which GP is scheduled to hold a "sit and wait" clinic. This is where GPs take turns to see patients who walk in without appointments.

Scheduling rotas and booking meetings in multiple calendars have been widely automated in recent years. Both Google Calendar and Microsoft Outlook have functionality to request meetings with other participants, where software will compare all involved parties' calendars and automatically schedule and suggest/book a meeting time that works best. Additionally, Google has employed machine learning to help users schedule "goals" and other important events automatically into their calendar (Shu, 2016). One could imagine an application of this technology in primary care for automatically scheduling both staff members and patient appointments, taking account of their given preferences. However, the reasons for scheduling and the context around what needs to be scheduled can vary depending on staff members involved and what purpose the scheduling is for.

7.10 Phone Calls

The telephone is a main channel of communication between the practice and the community it serves. Every practice experiences an ebb-and-flow of phone calls. The volume of calls varied greatly between our field sites. The smallest clinics managed a steady stream of calls throughout the day, a task



DR	MON 19th	TUES 20th	WED 21st	THURS 22nd	FRI 23rd
SCRIPTS/EPS	MK	MK	MK	MK	MK
WORKFLOW/DOCMAN	MK	MK	MK	MK	MK
RESULTS/LAB LINKS	MK	MK	MK	MK	MK

DR	MON 19th	TUES 20th	WED 21st	THURS 22nd	FRI 23rd
SCRIPTS/EPS	DS	DS	//////////	//////////	DS
WORKFLOW/DOCMAN	DS	MK	//////////	//////////	MK
RESULTS/LAB LINKS	MEM	RO	//////////	//////////	HH

DR	MON 19th	TUES 20th	WED 21st	THURS 22nd	FRI 23rd
SCRIPTS/EPS	FH	FH	DS	DS	DS
WORKFLOW/DOCMAN	IZJ	RO	RB	KH	ZA
RESULTS/LAB LINKS	MEM	MEM	MEM	MEM	HH

DR	MON 19th	TUES 20th	WED 21st	THURS 22nd	FRI 23rd
SCRIPTS/EPS	RB	RB	RB	RB	RB
WORKFLOW/DOCMAN	HH	DS	IZJ	RO	RB
RESULTS/LAB LINKS	ZA	ZA	ZA	ZA	ZA

DR	MON 19th	TUES 20th	WED 21st	THURS 22nd	FRI 23rd
SCRIPTS/EPS	HH	HH	HH	HH	HH
WORKFLOW/DOCMAN	RB	ZA	HH	ZA	IZJ
RESULTS/LAB LINKS	IZJ	IZJ	IZJ	IZJ	IZJ

DR	MON 19th	TUES 20th	WED 21st	THURS 22nd	FRI 23rd
SCRIPTS/EPS	//////////	//////////	//////////	//////////	RB
WORKFLOW/DOCMAN	//////////	//////////	//////////	//////////	Share
RESULTS/LAB LINKS	//////////	//////////	//////////	//////////	IZJ

SIT AND WAIT

MON 19th	TUES 20th	WED 21st	THURS 22nd	FRI 23rd
MEM	DS	IZJ	ZA	HH

Figure 16: GP cover schedule example

that is often manageable by a single receptionist. The largest and busiest clinics received around a thousand calls a day, with two to four receptionists working in a phone bank (a specialised room with multiple desks that each contain a phone and a desktop computer). The content of each phone call varied greatly and usually sent the receptionist on a hunt for information or consultation with other members of staff.

Phone communication has been in the spotlight recently with a remarkable automated and machine learning approach to phone calls and booking appointments. Google Duplex (Leviathan, 2018), which was unveiled in May 2018, is branded as an “AI system for accomplishing real-world tasks over the phone”. The Duplex technology has a human-like voice and allows humans to interact with it using natural language. Examples showcased included booking a hair salon appointment and booking a table at a restaurant for dinner, all automatically without much input from the user, or a receptionist. Google has another product dubbed “contact centre AI”: a similar technology to Duplex but fine-tuned to manage customer relations for businesses. These products represent a state-of-the-art advance in the technologies of natural language processing, machine learning, machine transcription, and an assortment of other tools used to perform these complex interactions.

It ought to be feasible for this kind of automated call centre technology to find applications throughout primary care and the NHS more generally. Given the volume of phone calls that occur each day across the country, this could become a financially advantageous investment. That is the reason that our machine-learning model predicted high automatability scores for tasks involving answering phone calls, ranging from 2.7 (for medical professional consultations) to 3.4 (reconciling information by phone call).

Moving Forwards

Impact, Skills, and Reconfiguration

8 Moving Forwards: Impact, Skills, and Reconfiguration

Our research shows great potential to aid decision making in a future where more automation is applied to administrative tasks in primary care. Our qualitative analysis tempers and contextualises our quantitative analysis, by showing the complexities and dependencies of some of the work in primary care. We would anticipate that by automating some administrative work, more time would be made available for other important tasks, including more patient contact and longer consultations. It might also allow more time for staff development and training, and it could help to create space for quality improvement and other service developments.

Our research suggests that simply focusing on automation to increase patient throughput would be a mistake. Employing automation to see more patients per hour would not address underlying problems of staff fatigue, communication issues, and healthy work environments. The aim should be to use automation to relieve pressure and workload stress, and to support over-worked practice staff. This means seizing the opportunity to rethink staff roles if/when certain tasks are automated.

In this concluding section we explore how staff workloads may change because of increased automation. First, we look at how task automation will impact on current roles in primary care. Next, we discuss certain skills, knowledge and abilities that are unlikely to be automated, showing that these are important for primary care to invest in. Finally, we discuss what the future practice may look like given our findings, how certain workloads will shift, and how the scope of work may change.

8.1 Potential Impact of Automation on Primary Care Workforce

First, we consider the extent of potentially automatable work performed by staff in the occupations we observed. We have used employment figures from NHS digital, visualizing each occupation as a collection of the tasks they routinely perform (from our observation dataset) and show what portion of these tasks are at risk of automation.

Figure 17 presents the automatability of work activities by the number of currently-employed individuals who perform them, classified into 10 administrative roles within primary care. It should be evident that most activities observed in our dataset lie between *mostly* and *completely* automatable.

This quantitative analysis shows that a large proportion of the administrative work performed in primary care (that is, the area under the curves in Figure 17) is "mostly" automatable. However, this figure also shows the distribution of occupations that are most likely to be affected by automation, for example Administrators (mauve), Scanning Clerks (sage), Receptionists (pink), and to a lesser extent Prescription Clerks (dark green). These occupations regularly perform tasks with high automatability scores.

The differences between occupations may be somewhat reduced by the overlapping nature of tasks within our healthcare dataset. For example, a Receptionist may also perform tasks such as "Review letters" (similarly to an Administrator), which has a high automatability score (3.22). Automating this task would remove it from the workload, and reduce the corresponding area under the curve in Figure 17, for both occupations (mauve and pink).

What is perhaps most interesting about this Figure is that it shows how much potential change and reconfiguration of occupations is theoretically possible. It shows the considerable amount of work

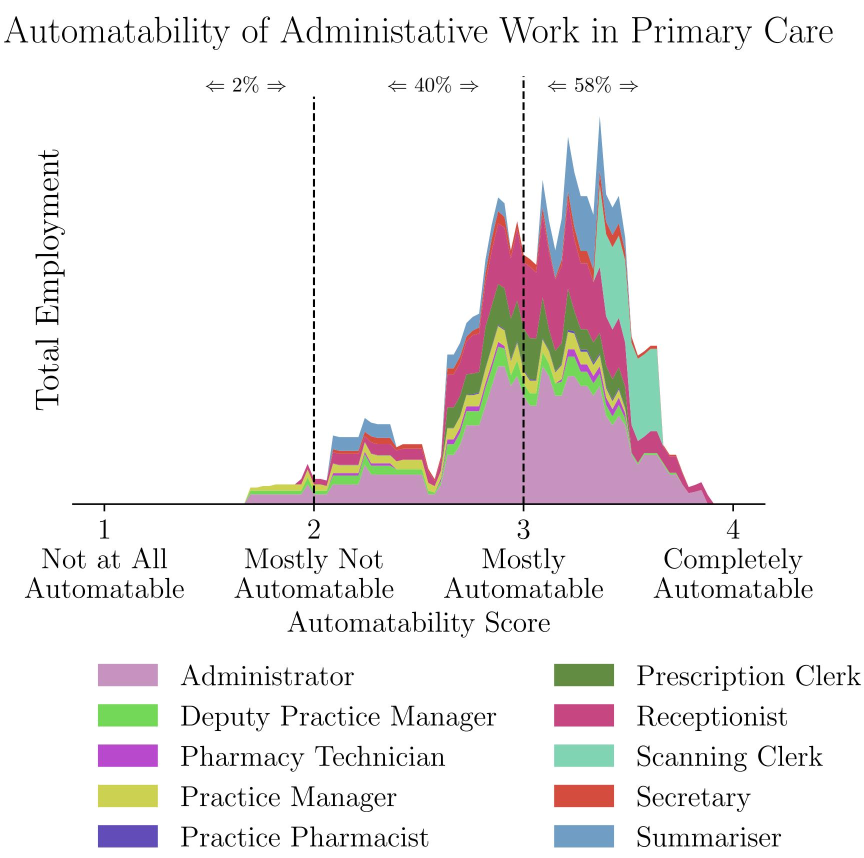


Figure 17: Amount of clerical workers affected by the automatability of their tasks.

shifted towards the right-hand side of the automatability scale, considered to be mostly or completely automatable. While we understand some of these tasks may continue to require human input, we would nevertheless expect them to require much less human attention once automated, freeing up valuable staff time.

8.2 Moving Forward: Skills for the Future of Primary Care

The work of primary care is made up of a variety of different tasks, some of which are difficult to automate, and others of which have been automated in sectors outside of healthcare for years. None of the core occupations in primary care is likely to disappear as a result of automation, but people whose jobs focus on a relatively small set of tasks with high probability of automation will be freed up for new roles based on their experience and knowledge acquired. For example, the demand for Scanning Clerks may decrease or even disappear in the future, but these people could probably be redeployed to add value elsewhere in the practice. We imagine that, as with any new technology implementation, staff would be required to manage the “gap” between what technology can and cannot do. New problems and opportunities arise when new technologies are adopted by organisations.

A key goal of our study was to understand the significance of each of the 120 O*NET occupation features (the skills, knowledges and ability variables required to perform a job) with respect to task automability within primary care work. Understanding important features can inform future policy decisions with regards to training or upskilling staff. We use the following criteria for a scheme to measure the importance of occupation features:

1. An *important* feature must be clearly predictive of automatability.

2. An increase in an *important* feature must lead to an increase in automatability.

Ideally, any method for identifying important occupational features would also be able to uncover complex non-linear interactions between features, and also capture complementarities between features (whose importance is contingent on the value of other features). To this end, we propose two complimentary methods, firstly to understand what O*NET characteristics are particularly automatability (or not) when performing tasks. We calculate this as the percentage difference in an occupation feature value between the average value in a subset of tasks in the same automatable category, when compared to the entire healthcare task dataset. This provides an interpretation about which occupational features have particularly high or low level of requirement within each of the automatability categories.

Secondly, for each occupational feature, we analyse the average gradient (derivatives) to understand its impact on a tasks' automatability score, as described in (Baehrens et al., 2010). For the nth O*NET feature, this is computed as $AG(n) := \mathbb{E}(^{\partial m(\boldsymbol{x})}/_{\partial \boldsymbol{x}_n})$, where $m(\boldsymbol{x})$ is the posterior mean distribution.

The average gradient of an occupational feature measures the expected increase in likelihood of task automatability for a unit increase in the required amount of the feature to perform a task, for instance, as a result of a policy intervention. The average gradient of a tasks' automatability gives an interpretable notion of sign (or direction), meaning positive relationships can clearly be distinguished from negative relationships using this method. Those occupational features with large positive gradients are complementary to the automatability of a task, i.e. increasing the feature increases the task automatability. Whereas those with large negative gradients are anti-complementary to automatability (increasing this feature decreases the task automatability).

Together, these analyses inform which features are predictive, and in what direction the automatability score will change given an increase in the feature.

Table 3 presents the five largest (positive) percentage differences in O*NET occupational features split into the three automatable categories, when compared to the entire healthcare dataset of tasks. This shows us that healthcare tasks inferred as "not-automatable" require much higher than average Installation skills, 26% higher Personnel and Human Resources knowledge, 17% higher Education and Training knowledge, and 17% higher Management of Resources skills. Whereas, "automatable" tasks require 24% more Clerical skills, and to a lesser extent higher than average Customer and Personal Service knowledge.

Table 4 presents the largest average gradients of each O*NET feature in our Gaussian Process model (averaged over all our healthcare tasks). We present the largest positive values indicating a positive "direction" or influence on task-automatability, given an increase of the particular feature. For example, Telecommunications and Clerical knowledge have the largest gradient. This means that if a task requires more of this knowledge, then its automatability score will increase, more than increasing any of the other O*NET occupation features.

Table 3 and 4 should be analysed concurrently. For example, in the "automatable" category, tasks require 24% more than average Clerical knowledge (Table 3). Clerical knowledge is also an occupational feature that, if the level of requirement increased, contributes to increasing the automatability score (+0.166 from Table 4). Another occupational feature, such as the amount of Customer and Personal Service knowledge required to perform a task, is on average 13% larger in the automatable risk group when compared to the population, however increasing it does not increase the automatability of a task. In fact, increasing the level of Customer and Personal Service knowledge required to perform

Table 3: Largest feature differences relative to the dataset by automatable category

Automatable Category	O*NET Feature	Feature Difference
not-automatable	Installation (skill)	+62.6 %
	Building and Construction (knowledge)	+27.2 %
	Personnel and Human Resources (knowledge)	+26.0 %
	Management of Financial Resources (skill)	+18.7 %
	Education and Training (knowledge)	+17.0 %
automatable	Clerical (knowledge)	+24.4 %
	Customer and Personal Service (knowledge)	+13.2 %
	Service Orientation (skill)	+5.0 %
	Economics and Accounting (knowledge)	+4.7 %
	Computers and Electronics (knowledge)	+4.5 %
partly-automatable	Medicine and Dentistry (knowledge)	+30.8 %
	Therapy and Counseling (knowledge)	+22.0 %
	Psychology (knowledge)	+13.4 %
	Clerical (knowledge)	+13.2 %
	Biology (knowledge)	+12.1 %

a task actually *decreases* the tasks' automatability, shown by its negative gradient of -0.023. The full list of global average gradients is provided in Appendix B in Table 11.2.

These results suggest that primary care occupations should focus on specific skills, knowledge and abilities in order to maintain an advantage in the face of automated systems that are already capable of substituting for many tasks. Skills such as Learning Strategies, Management of Personnel and Financial Resources and Instructing should be core values within training programmes (large negative gradients). Conversely, it would not be wise to make Telecommunications, Clerical and Economics and Accounting knowledge essential when hiring or investing in the future, as tasks requiring these characteristics are much more automatable.

Table 4: Global: Ten large O*NET occupational feature gradients.

O*NET Feature	Feature Gradient
Telecommunications (knowledge)	+0.167
Clerical (knowledge)	+0.166
Wrist-Finger Speed (ability)	+0.153
Number Facility (ability)	+0.118
Mathematics (skill)	+0.093
Depth Perception (ability)	+0.092
Building and Construction (knowledge)	+0.090
Mathematical Reasoning (ability)	+0.088
Economics and Accounting (knowledge)	+0.085
Control Precision (ability)	+0.082

Into the Future

The General Practice of the Future

9 Into the Future:
The General Practice of the Future

The future of occupations in primary care is not the removal of certain staff, but rather reconfiguration of job roles. How differently would a general practice function if they find that their administrators and receptionists have 20-50% less work to do? Would they fill that time by processing historical backlogs of documents that cannot be automated? Or would they start to take on new tasks and responsibilities? Would the practice choose to see more patients per hour, or would they see the same number of patients but offer them longer consultations? What changes can we envisage to the distribution and nature of primary care tasks?

In a previous section of the report we provided a description of the main occupational groups and their typical work characteristics. Below, we consider how these might change as a result of automation.

Administrator

As one of the core administrative roles in the practice, administrators are responsible for a large workload of office tasks, the length and complexity of which vary greatly. Our analysis suggests the following tasks are highly likely to be automated (score of 3.0 and above):

> Print letters; Use texting service or patient management service to contact patients; Use software or database to paper records that are in storage either onsite or offsite; Scan letters and other documents; Organise and manage main practice email account; Answer phone; Reconcile information over the phone; Work in patient management system for QOF, recalls, or other patient communication systems; Checking for errors in paperwork; Send messages in electronic health record to communicate to other staff; Connecting human resources/making introductions of new staff; Register new patients; Work in NHS CQRS online web portal; Medical Coding of letters and other documents; Review letters; Work in clinical query system to create reports; Address problems that arise with building; Process prescription renewals; Reviewing health records addressing any changes from health authority; Use software to convert printed letters into text on the computer; Patient recalls; work in spreadsheets; Mass mail letters; Type letters; Write letters for secondary care, other GPs, or for patients.

This would leave administrators with a workload that focuses mainly on internal communications. The types of tasks they might focus on could include developing the human resources functions, cleaning-up and optimising data in electronic medical records, and supporting procurement and inventory management for the practice.

Deputy Practice Manager

Deputy Practice Managers usually assist the practice manager with their scope of work, but depending on the practice size, they sometimes take on a more managerial role. The tasks that can be mostly automated from their workload include:

> Use texting service or patient management service to send reminders; Answer phone; Reconcile information with a phone call; Work in patient management system for QOF, recalls, or other patient communication systems; Checking for errors in paperwork; Send messages in electronic health record to communicate to other staff; Connecting human resources/making introductions of new staff; Medical coding of letters; Work in clinical query system to create reports; Address problems that arise with building; Create work or holiday schedules for practice staff; Staff rotas (clinical or practice); Patient recalls for QOF, Write letters or texts using software; working with spreadsheets; Mass mail letters; Use eReferral system to help patients book appointments and manage referrals to different specialists; Create time slots for clinical/bookable hours.

Here we see some overlap with several tasks that Administrators also perform, but removing the above tasks would still leave deputy practice managers with a significant workload of at least 21 other tasks. This level of automation might allow DPMs to focus on practice development, taking a more strategic and proactive approach to meeting population health needs. They might also have scope to extend their role in staff training and recruitment, financial management, and assisting in procurement and inventory management for the practice.

General Practitioner

Work in primary care, and indeed all of healthcare, is designed to shield GPs and other clinicians from spending too much time on paperwork and administrative tasks, to maximise patient contact time. However, even GPs must take on administrative responsibilities, and many of these tasks overlap with those of other occupations. The following tasks performed by GPs were classified as potentially automatable:

> Reconcile information with a phone call; Review letters; Respond to tasks in electronic medical record; Run search in health record for patients that need medication reviews; Look for drug interactions; Indicating what is to be coded by a medical coder; Maintain supplies and stock in clinical room; Study patient medical record; Type letters; Write letters for secondary care, other GPs, or for patients.

Removing much of this administrative work should allow more time for patients, as well as reducing the stress of heavy workloads. The majority of GP tasks that hover just below "mostly automatable" with scores from 2.0 to just under 3, include clinical tasks such as diagnosing advising and treating patients, reviewing medical data and information, providing emotional support, and conducting medical exams. Some of these tasks with low automation scores may be relieved with a degree of automational support in future. As automation continues to advance, the future role of GPs may focus more on relationship building and social interactions with patients, and the interpretation and application of diagnostic data and decision support that can be provided by automation technologies.

Healthcare Assistant

Healthcare Assistants, as their title implies, assist with certain clinical tasks. They are often in training roles on the path towards greater clinical function and responsibility. While their tasks are mainly clinically focused, they also have various administrative duties, including the following:

> Answer phone; reconcile information with a phone call; run search in health record for patients that need medication reviews; maintain supplies and stock in clinical room; respond to tasks in electronic medical record; Send messages in electronic health record to communicate to other staff.

When these six standard administrative tasks are removed from the HCAs workload, the entirety of the remaining tasks (21) are all supporting the clinical workflow. We would expect automation to enable HCAs to focus entirely on clinical work. This would allow for more emphasis on upskilling and technical developments, patient consultations, research support, and medical knowledge.

Nurse Practitioner

Nurse practitioners face similar administrative workloads to GPs and other clinicians. The following nine administrative tasks typically found in a nurse practitioner's workflow could mostly be automated:

> Answer phone; Reconcile information with a phone call; Send messages in electronic health record to communicate to other staff; Respond to tasks in electronic medical record; Run search in health record for patients that need medication reviews; Indicating what is to be coded by a medical coder; Conduct diabetes reviews; Order immunizations; Maintain supplies and stock in clinical room.

In common with tasks carried out by other clinical occupations, these may vary according to seniority and responsibility. Similar to other clinical occupations, the extra time saved from automating these administrative tasks could enable nurses to see more patients, offer longer consultations, or take on new tasks such as health promotion, group clinics, social prescribing, safety monitoring, or quality improvement. The key opportunity with automating administrative tasks is that it makes time available to consider new possibilities. As a minimum, it allows for the same operations and process to continue as before, without the stress of facing a backlog of work each day.

Phlebotomist

Phlebotomists are one of several occupations that have a very focused task list. Usually employed in areas where large amounts of bloodwork need to be processed, there would appear to be only limited scope to automate the phlebotomist's workload. However, some HCAs are trained in phlebotomy, so automating their administrative tasks could free up time to take on more of this role.

Pharmacy Technician

Like the phlebotomists, pharmacy technicians are also responsible for a limited range of tasks. They spend most of their time consulting with other clinicians and sorting out drug and prescription issues that require medical knowledge. As a result, much of this work is exception based and its potential for automation is relatively low. The two exceptions are organising repeat prescriptions and answering the telephone. The role of a pharmacy technician is likely to continue to evolve as automation technologies are further developed to support their role; a continuation of the process that began in the 1960s with the advent of the digital pill counter, and numerous waves of inventory control and digital prescription information systems.

Practice Manager

Depending on the size of the practice, the practice manager usually has responsibility for financial management, staff management and recruitment, often supported by a Deputy Practice Manager and/or Administrators. Tasks which can be mostly automated from a practice manager's core responsibilities include:

Answer phone; Reconcile information with a phone call; Work in patient management system for QOF, recalls, or other patient communication systems; Checking for errors in paperwork; Send messages in electronic health record to communicate to other staff; Making introductions for new staff; Work in NHS online web portals; Work in clinical query system to create reports; Address problems that arise with building; Staff rotas (clinical or practice); Manage agenda and record minutes for meetings; Ordering immunizations; Significant event reporting.

Practice managers have a sizeable workload that includes at least 24 separate tasks. Greater automation could substantially transform their role, allowing more time for service development, innovation and research, and integrating the practice into the needs of the community. They may also take on tasks that involve upskilling and preparing all practice staff for a post-automation future.

Practice Nurse

Clinical occupations benefit from automation in different ways than administrative staff. This is because clinical staff, unsurprisingly, run clinics in addition to having to work on administrative tasks, often during clinical time. Administrative work can often lead to a tight schedule; a patient meeting running over time, or the documentation process taking more time than available can all lead to a stressful and rushed clinician. Thus, being able to remove the clinical tasks below would afford practice nurses and other clinicians the luxury of more clinical time, or the same clinical time at a more relaxed and gratifying pace. Tasks which can be mostly automated from a practice nurse's core responsibilities include:

Answer phone; Reconcile information with a phone call; Send messages in electronic health record to communicate to other staff; Staff rotas (clinical or practice); Respond to tasks in electronic medical record; Run search in health record for patients that need medication reviews; Indicating what is to be coded by a medical coder; Conduct diabetes reviews; Maintain supplies and stock in clinical room; Reorder stock of injections or immunizations.

While the field researcher was able to observe several practices in primary care. One practice featured nurses running a pilot group clinic for diabetes patients that was notably different than their usual one-on-one diabetes clinic. The future of primary care would allow for experimentation and exploration of new clinical styles and arrangements that best fit the community served by the practice.

Practice Pharmacist

While pharmacy technicians perform a specific and focused set of a few tasks, practice pharmacists perform a larger set of tasks that introduce more opportunities for automation. The following tasks below are mostly automatable, and largely concern repetitive tasks involving working with prescription information. They are:

Perform Medication reviews; Prescription queries; Answer phone; Control prescription refills or authorizations; Send messages in electronic health record to communicate to other staff; De-prescribing of high cost drugs and other drug audits; Look for drug interactions; Conduct diabetes reviews.

As with the pharmacy technician, pharmacy occupations should focus on both their medical and therapeutic knowledge in conjunction with the community they serve. Pharmacy-based occupations will see, as they have before, automation around the control and inventory management of drugs and advancement of pharmaceutical decision support systems. One opportunity for growth in pharmacy-related occupations might be to work with both the practice and broader community on research initiatives, and with the community on education and outreach.

Prescription Clerk

The prescription clerk is in a similar category as the scanning clerk, the similarity being that the tasks are specific and focused. Prescription clerks are employed by practices that have a high volume of prescriptions that need support to process them. More administrative tasks were observed being performed by prescription clerks than scanning clerks. The most automatable work of this occupation is:

> Check prescriptions for medication renewals and send to GP; Process prescription renewals; Answer queries from patients or from GPs about prescriptions.

We classify prescription clerks as an occupation that is unlikely to see new calls for employment in a primary care environment that has implemented the above mostly automated tasks. However, in the short term, current prescription clerks retain knowledge of the practice dynamics that allow them to transition into other roles. The knowledge they have of processing prescriptions and working with clinical systems may translate into an occupation that supports the organizational integration and maintenance of newly implemented automation systems, and to assure quality and accuracy of newly automated processes.

Receptionist

Receptionists can spend all day processing prescriptions, even if those prescriptions are received digitally. One receptionist we observed had to process around 80 prescriptions a day. This work entails double checking the prescriptions and assigning them to a GP for sign off. This work could be transformed by automation technologies that reduce paperwork and avoid the interruptions that are a typical fact of the receptionist's role. Receptionists have a pile of work that is constant, including coding letters and reviewing prescriptions, writing letters and answering phones, or performing other clerical work. Usually all of these tasks are done simultaneously, for example, answering phones while they code a document or assign prescriptions. The workflow is not optimal, as they end up going from one sudden interruption to the next and then back to their main set of tasks.

Most of the receptionists' tasks we observed (29 tasks) are potentially automatable, leaving only nine tasks classified as not-automatable. Automatable tasks in the receptionist workload include:

> Check and sort mail; Print letters; Use texting service or patient management service to contact patients for reminders; Scan letters and other documents; Check incoming emails to main practice email address, forward emails to other employees; Organise and manage main practice email account; Answer phone; Reconcile information with a phone call; Check prescriptions for medication renewals and send to GP; Transcribe audio-recorded letters; Send messages in electronic health record to communicate to other staff; Ordering of goods, stationary, office supplies; Register new patients; Work in NHS online web portals; Review letters; Help organise and schedule patients for clinics; Staff rotas (clinical or practice); Create and assign tasks for others in the electronic medical record or other software; Run search in health record for patients that need medication reviews; Run searches to look for repeat prescriptions; Process prescription renewals; Schedule/book patients; Assign prescriptions to doctors after reviewing to sign; Send patient texts to attend a chronic condition clinic; Type letters; Write letters for secondary care, other GPs, and patients; Use NHS eReferral system to help patients book appointments and manage referrals to different specialists; Assign letters to be reviewed by employees; Attend to the front desk.

The remaining tasks are the more complicated aspects of patient management, including booking, referring, and interactions with patients and other staff. We imagine that automation may cause the receptionist profession to shift towards more face-to-face interaction with patients. Our fieldwork showed that, even though kiosks and self-check-in services are regularly available, patients still value the ability to check-in and discuss their questions and concerns with a receptionist. Receptionists often become regular (and important) parts of the patient's care experience. Some have taken on the role of care navigators or patient advocates. We envision a future practice where the experience is reminiscent of a modern technology store, such as an Apple or Microsoft store. These have few barriers between staff and customers and more one-to-one time with a staff member to guide people through the experience. The point is that automation of tasks could allow these staff to spend more time with patients and reinvigorate communications, leading to improvements in patients' experience.

Scanning Clerk

The scanning clerk carries out a narrow and focused range of tasks. Document scanning received a score of 3.5 in our index, suggesting it is mostly automatable. Scanning clerks are among the occupations that could be shifted to existing or new roles that would add value to the practice. Scanning clerks may be well suited to transition into a role of support and maintenance of automation systems and organisational tasks.

Secretary

Another core administrative occupation is the secretary. While secretaries work primarily with writing and assisting in the production of letters, they perform a variety of other tasks that are all focused on documenting and reporting. The tasks below can be mostly automated:

> Print letters; Scan letters and other documents; Organise and manage main practice email account; Use software to manage transcription of letters or other documents; Answer phone; Reconcile information with a phone call; Checking for errors in paperwork; Transcribe audio-recorded letters; Send messages in electronic health record to communicate to other staff; Work in NHS online web portals; Type letters; Write letters for secondary care, other GPs, or for patients; Use eReferral system to help patients book appointments and manage referrals to different specialists; Type two-week wait letters; Runs DBS checks.

The remaining work of a secretary is about eight tasks that concern writing reports that are longer than a letter, supporting the process of auditing, and the maintenance and upkeep of practice documentation. Secretaries are a document-focused occupation and their role may shift to focus on documentation work that is resistant to automation, such as the creation and upkeep of policy documents and evolving the record keeping and documentation functions of secretaries. Furthermore, the practice of the future might see the combination of roles such as the secretary and administrator, or secretary and receptionist, to combine responsibilities when routine administrative work is mostly automated.

Summariser

Summarisers are typically employed at larger practices that can maintain the help of a summariser to tend to the vast amount of patient records that practices must maintain. Although they mostly focus on checking, editing, and revising data in the electronic medical record, summarisers may perform few other clerical tasks. Tasks that can mostly be automated include:

> Answer phone; Send messages in electronic health record to communicate to other staff; Register new patients; Medical coding of letters and other documents.

This leaves the core mission of summarisers' work just under the score of 3: cleaning up and amending information in the patient's record. Automation of administrative work may free up the time of summarisers to focus on just tending to the medical record. The summariser is an interesting occupation because they are responsible for “tidying up” or otherwise dealing with exceptions, odd problems, and errors in the record. While certain workflows and processes of this work may become automated, new issues and errors will arise that require a summariser to sort out medical data. The practice of the future would entail summarisers shifting to a role of “data experts” in which they work with data models and medical ontologies to ensure that electronic health data remains high quality, while also having the task of knowing how to support and maintain the automated processes that will be found throughout this type of work.

Concluding Remarks

10 Concluding Remarks

Our analysis of the potential for automation in primary care shows an exciting possibility and potential for change, reconfiguration, growth, and most importantly support for current and future staff. The prospect of occupational change is further enhanced because many tasks are shared or similarly performed by staff in different occupations. Removing answering the phone or processing prescriptions or medical coding from one occupation removes it from several other staff members at that practice. This creates a domino effect where modifying one task may affect the workloads and responsibilities of several occupations. In this section, we discuss the limitations and strengths of the project, and end with concluding remarks and discussion of what we have learned.

10.1 Strengths and Limitations

No study is without its limitations, and it is important to be aware of the limitations of this study. First, our study focused on administrative tasks in the analysis, although as we have emphasised, this is because these tasks constitute the majority of a practice's workload. Our use of the O*NET database in large part restricted us to looking at these administrative functions. We might assume that many purely clinical tasks are also potentially automatable, if not now then in the future, but we lack the data to responsibly make these claims. A further limitation of the O*NET database used to represent healthcare tasks is that it itself is survey data captured by the United States Department of Labor, and its population is US occupations. It is a large and useful dataset, however it does not specifically pertain to UK occupations, and more specifically the tasks performed by NHS medical occupations might lose nuance or fidelity when matched to US occupation / activity data.

Second, it is important to be mindful of the diversity of primary care practices, and the wide variations in how these are managed, organised, funded, and staffed. Primary care looks different across all parts of the country, and we were unfortunately not able to gather data on some of these nuances. We made extensive efforts to understand the tasks performed by each occupation and to ensure they were as representative as possible, but we cannot be certain that we captured all the diversity.

It is also good practice to take a critical lens on the data and methodology used, whose predictions provided quantitative validation of the detailed qualitative fieldwork. With regards to the machine learning expert survey, we believe using IBCC (Independend Bayesian Classifier Combination) to achieve a single rating from multiple experts is both appropriate and advantageous, because it is a fully Bayesian way of combining the data sources in order to reflect the highest chance of accurately recovering the true automatability label of a task in an environment of uncertainty and subjectivity. A side-effect of using this approach is that the ground-truth labels are more polarised at the extreme values (one and four), when compared with simply (mean) aggregating the task-level responses together. A complication of this is that when comparing two models based upon their Root Mean Squared Error (RMSE), lower absolute error is achieved by not using the (more polarized) IBCC combined labels. However, we believe they are more representative when combining multiple (semi)-reliable expert sources.

We also believe the expert survey could have captured data about *how* a task could be automated, for example, by simplifying or rearranging the task. This subjectiveness in the task simplification may have translated into our dataset, providing very few entirely "not automatable" tasks. Perhaps caused, in part, by the creativity of the respondents who may have believed part of a re-arranged task to be automatable. Perhaps our estimate ground-truth data would be more accurate if we could quantify how the expert labels were derived, i.e. by technology automating a task, or by the task being re-structured. To supplement more data in this area, we also need better and more clear definitions of automation.

This study also does not attempt to bridge the gap between what is technically possible to automate, and to what extent a task *will* be automated. This will be based on many socio-economic and policy making decisions outside the scope of this project.

Despite these limitations, we are proud to assert that this is currently the first use of a scalable measure of task-based automation applied to the work of healthcare. We hope it will inform future analyses of the likely impact of automation in healthcare and beyond. We recognize the importance of context to inform analysis and conclusions, so we consider it a strength that our data are grounded in the empirics of NHS primary care. Our approach was designed to look at what is technically possible to automate. We believe the successful transformation of primary care will depend, at least in part, on successful adoption of automation. We therefore hope our findings will be useful for policymakers and primary care stakeholders to inform their investment decisions. Judicious use of these technologies certainly offers potential benefits for staff deployment, leading to improvements in the quality and safety of patient care.

10.2 Discussion

It is important to note that neither our findings, nor the examples cited to demonstrate the potential of greater automation involves sensational claims about a futuristic AI that will diagnose and treat all patients and render swathes of the workforce redundant. On the contrary, many of the most needed and highest impact automation technologies are mundane. Automating and removing some of these tasks could free up countless hours for more rewarding tasks. Automation technologies that assist with tasks such as telephone answering, letter writing, document scanning, email monitoring, and information filtering would represent a remarkable change to the way all staff work in primary care. Primary care involves a great deal of work with information and data. Automation technologies will thus be impactful when they can present people with the right information they need at the right time, blocking out distracting information or information that can be read and processed later, and filtering the information by context and urgency. For instance, these technologies could reduce time wasted reading letters trying to locate the section most relevant to the GP, or introduce smart management of pop ups and alerts delivered by the electronic medical record.

Obviously, there are other tasks and roles which may emerge after the application of automation solutions that are hard to foresee. Once the tasks we have characterised as mostly automatable are removed from workloads, we would expect new tasks to emerge that are connected to or created indirectly by automation technologies. Indeed, there is already a heavy use of computers throughout primary care that have automated some form of work. Every single staff member at every practice we visited has access to a desktop computer. This allows for types of automation to occur in many ways, such as: automatic completion of certain terms and phrases when entering data, the use of templates, transforming printed text to digital text through optical character recognition, and pulling reports and data from electronic medical records for inclusion in letters or texts. Some automation functions are invisible to the user, such as automatic back-up, or pre-populating information using a common data source. The point is, that despite all this technological assistance there still exists a large amount of routine administrative work in primary care that is potentially automatable.

We have suggested four possible effects on the workloads of primary care staff in the short to medium term. These range from no modification of current tasks to major changes in people's new occupational responsibilities and identity. When automation influences an occupation one of the following possibilities will occur:

1. No change or little meaningful change, work as usual.
2. The staff member may find themselves doing more of their other tasks that cannot be automated, increasing their productivity in respect of these tasks.
3. Removing tasks prospectively could open new opportunities to deal with backlogs of document processing. When these are dealt with, staff members could move on to either #2 or #4.
4. Tasks are removed from day to day workflow and time is freed up to enable the development of new roles, responsibilities or training.

We want to focus on the organisational development and opportunity for change that automation may allow. Focusing just on the second point above: it may be easy for practices to "go with what they know" and see the automation of certain tasks as an opportunity to do more of an existing task. While this may be necessary for a period, a significantly automated workflow will have the potential to "disrupt" the day-to-day tasks of occupations. We echo the call of previous researchers in the field of healthcare automation for careful monitoring of the effects of these innovations (Steventon, 2018).

While most administrative tasks are broadly automatable, there is important context that is not fully captured in our method of characterising tasks. Our fieldwork and qualitative analysis has shown that the primary reason healthcare is, and will be, resistant to automation is that several key tasks are driven by exceptions and contextual specifics. There are many occasions when aspects of a task need to be discussed with other staff, or otherwise non-routine thinking must be applied to a patient's specific case. Interpersonal communication is an important feature of healthcare delivery. Applications of automation in the primary care setting need to consider the social and interpersonal aspects of the patient experience to a greater extent than in other fields outside of healthcare. Processes like letter scanning have little to do with the patient-practice relationship, but other tasks such as medical coding, letter writing, and other communication tasks may affect the relationship of the patient to the practice, for benefit or detriment.

This leads to advice and insights we have learned through our work that may be of benefit to organisations considering applying automation solutions and other digitisation strategies. One of the most important ways forward is to approach the prospect of automation strategically by asking direct questions that inform the use of the technologies: Which types of tasks would be most beneficial to the practice if they were automated? Are those tasks amenable to automation? If this results in a change in people's tasks, who will it affect most? If tasks are removed from the workload of a staff member, what kind of work (if any) will replace the task? Who will be responsible for supporting and maintaining the automated tasks while new ways of working are being developed? These questions should motivate deeper and critical thinking about the use and effects of modifying tasks with automation technologies. Rather than assume automating "as much as possible as soon as possible" is the best strategy, organisations should have a vision for how they plan to use the technology and how they will manage the change that automation will bring.

In general, we hope that careful introduction of automation will allow more mentorship, communication, and connection between clinical occupations, more professional growth and identity, more opportunities for multi-professional working and training, and increased opportunities to experiment and test new ideas for improving patient care. At a minimum we believe automation of the administrative tasks we have identified could contribute to more effective and efficient use of resources in primary care in the NHS.

About the Authors

Eric T. Meyer was at the time of this research Professor of Social Informatics at the Oxford Internet Institute, University of Oxford. Professor Meyer's research focuses on the transition from analogue to digital technologies in research and knowledge creation across disciplines in the sciences, social sciences, arts, and humanities. His research has included both qualitative and quantitative work with marine biologists, genetics researchers, physicists, digital humanities scholars, social scientists using big data, theatre artists, librarians, and organizations involved in computational approaches to research. Professor Meyer is currently Dean, Mary R. Boyvey Chair, and Louis T. Yule Regents Professor in the School of Information at the University of Texas at Austin. He also remains a Senior Fellow at the University of Oxford.

Michael A. Osborne Mike Osborne is the Dyson Associate Professor in Machine Learning at the University of Oxford. Back in 2013, he co-authored a paper (with economist Carl Frey) that used machine learning to estimate that 47% of US jobs are at risk of being automatable through advances in artificial intelligence and robotics. This work was born of Mike's interests in the practical use of machine learning to enable automation, while ensuring that such advances are made in sympathy with societal needs. Mike's technical expertise in Bayesian optimisation and probabilistic numerics underpins recent advances in automated and interpretable machine learning pipelines. His algorithms have been deployed in industrial and scientific applications ranging from battery monitoring, pigeon navigation and self-driving cars.

Angela Coulter is an Oxford-based health policy analyst and researcher. A social scientist by training, she has higher degrees in health services research from the University of London and the University of Oxford. Now freelance and still involved in research, Angela's previous roles include Chief Executive of Picker Institute Europe, Director of Policy and Development at the King's Fund, Director of the Health Services Research Unit at the University of Oxford and Director of Global Initiatives at the Informed Medical Decisions Foundation. She is an Honorary Professor at the University of Southern Denmark, an Honorary Fellow of the Royal College of General Practitioners and a Non-Executive Director of the National Institute for Health and Care Excellence (NICE).

Matthew Willis is a researcher at the Oxford Internet Institute. His interests include sociotechnical systems in healthcare settings, patient oriented infrastructures, and health information management. He has research experience in academic, government, and private institutional settings including Sandia National Laboratories, the U.S. Department of Veterans Affairs, and several university affiliated research centres where he was a contributor to multiple grants from the National Science Foundation (NSF), National Institutes of Health (NIH), Defence Advanced Research Projects Agency (DARPA), and Intelligence Advanced Research Projects Activity (IARPA).

Paul Duckworth is a Principle Researcher with the Machine Learning Research Group at University of Oxford. His research interests lie in probabilistic and unsupervised machine learning, information retrieval and mobile robotics. He is specifically interested in the impact machine learning technologies have on society and the future of human work, and recently, our understanding of algorithmic (creative?) intelligence. He has previous professional experience working in retail banking and performing clinical trail analysis for pharma companies.

Logan Graham is a PhD student with the Machine Learning Research Group at the University of Oxford, funded by the Rhodes Scholarship. Logan researchers how we can improve machine learning with causality and interpretability. Previously he held a residency at X (formerly Google[X]) and performed research in causality at Babylon Health.

Acknowledgements

We thank The Health Foundation (award #7559), the Oxford Martin Programme on Technology and Employment, NESTA, MIRI, the Rhodes Trust.

Advisory Board:

Trish Greenhalgh is an internationally recognised academic in primary health care and a practising GP. She joined the Nuffield Department of Primary Care in January 2015 after previously holding professorships at University College London and Queen Mary University of London. Her past research has covered the evaluation and improvement of clinical services at the primary-secondary care interface, particularly the use of narrative methods to illuminate the illness experience in 'hard to reach' groups; the challenges of implementing evidence-based practice (including the study of knowledge translation and research impact); the adoption and use of new technologies (including electronic patient records and assisted living technologies) by both clinicians and patients; and the application of philosophy to clinical practice. Trish is a fellow at Green Templeton College.

Roger Jones is the Founding President of the Primary Care Society for Gastroenterology and the European Society for Primary Care Gastroenterology. Educated at Oxford and St Thomas, he is Chair of the Royal Medical Benevolent Fund and Provost of the South London Faculty of the RCGP, as well as Editor of the British Journal of General Practice. He was Wolfson Professor and head of general practice and primary care at King's College London School of Medicine from 1993 – 2010, where he was also Dean for Teaching and Dean for External Affairs, and now holds an emeritus position there. He has published on mental health problems in primary care and on the relationships between health and ethnicity, and on health policy and medical ethics. He has been active in research and development in medical education since the introduction of the OSCE at Southampton in the 1980s. Professor Jones edited Family Practice, an international journal of primary care, between 1990– 2004 and was Editor-in-Chief of the prize-winning Oxford Textbook of Primary Medical Care, published in 2005.

Chris Jenkins is a Partnership Manager for the Thomas Pocklington Trust, a medium sized national charity for people with sight loss. He is a sport science and disability studies graduate, also with a disability equality training background. Chris has competed internationally in athletics for the visually impaired and has a variety of medals. In previous employment he worked at the Greater London Authority on culture, sport and volunteering initiatives in the build up to and delivery of the London 2012 Games. Chris is a keen advocate for stakeholder participation, working various service users and agencies across the health, social care, media, business, technology and third sectors; to work together to help people achieve their full potential.

David Clifton is an Associate Professor in the Department of Engineering Science of the University of Oxford, and a Governing Body fellow of Balliol College, Oxford. He is a Research Fellow of the Royal Academy of Engineering. His research focuses on the development of "big data" machine learning for tracking the health of complex systems. His previous research resulted in patented systems for jet-engine health monitoring, used with the engines of the Airbus A380, the Boeing 787 "Dreamliner", and the Eurofighter Typhoon. Since 2008, he has translated his work into the biomedical context for healthcare applications. He has worked on Visensia, the world's first FDA-approved multivariate patient monitoring system, and the SEND system, which is now used to monitor 20,000 patients each month in the NHS. His research has been commercialised via university spin-out companies OBS Medical, Oxehealth, and Medyc, in addition to collaboration with multinational industrial bodies.

Jens Rittscher holds a joint appointment with Department of Engineering Science and the Nuffield Department of Medicine. He aims to advance mechanistic understanding of disease and patient care through quantitative image analysis. He is a group leader at the Target Discovery Institute and is a adjunct member of the Ludwig Institute of Cancer Research. He is a senior research fellow at Harris Manchester College in Oxford. Prior to coming to Oxford Jens Rittscher was a senior research scientist and manager at GE Global Research in Niskayuna (NY, USA), one of the world's largest and most diversified industrial research laboratories. He held a position as an adjunct professor at the Rensselear Polytechnic Institute in Troy (NY, USA). In 2001 he was awarded a DPhil from the University of Oxford for his work on recognizing human motion under the supervision of Prof. Andrew Blake, FRS. To date he has been awarded 29 U.S. patents, 20 published patent applications, published more than 50 peer-reviewed publications, one edited book, and 5 book chapters.

Shaun Leamon works to build and share evidence and capability, and help others test and demonstrate innovations to improve quality in health service delivery. As part of this work, Shaun leads the Health Foundation's Insight programme, one of the Foundation's largest research programmes supporting novel research to improve health and care in the UK. Prior to joining the Health Foundation, Shaun led a team of research staff at the Royal National Institute of Blind People (RNIB). The group was responsible for positioning the RNIB as a key knowledge source on evidence and research in the sight loss sector. Shaun was also responsible for co-designing and delivering the RNIB's 5-year strategic plan for research. Shaun holds a BSc in Physiology and an MSc in Human and Applied Physiological Research.

References

General Practice Forward View, 2016. URL https://www.england.nhs.uk/gp/gpfv/about/.

Daron Acemoglu and Pascual Restrepo. Artificial intelligence, automation and work. Technical report, National Bureau of Economic Research, 2018.

Ajay Agrawal, Joshua Gans, and Avi Goldfarb. *Prediction Machines: The Simple Economics of Artificial Intelligence*. Harvard Business Review Press, 2018.

E Ammenwerth and H-P Spötl. The time needed for clinical documentation versus direct patient care. A work-sampling analysis of physicians' activities. *Methods of information in medicine*, 48(1):84–91, 2009. ISSN 0026-1270. URL http://www.ncbi.nlm.nih.gov/pubmed/19151888.

Melanie Arntz, Terry Gregory, and Ulrich Zierahn. The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis. *OECD Social, Employment and Migration Working Papers*, 2(189):47–54, 2016. ISSN 14777029. doi: 10.1787/5jlz9h56dvq7-en.

David H. Autor. Why Are There Still So Many Jobs? The History and Future of Workplace Automation. *Journal of Economic Perspectives*, 29(3):3–30, 2015. ISSN 0895-3309. doi: 10.1257/jep.29.3.3. URL http://pubs.aeaweb.org/doi/10.1257/jep.29.3.3.

David H. Autor and Michael J Handel. Putting Tasks to the Test : Human Capital, Job Tasks, and Wages. *Journal of Labor Economics*, 31(2):59–96, 2013. ISSN 0734306X. doi: 10.1086/669332.

David H. Autor, Frank Levy, and Richard J. Murnane. The Skill Content of Recent Technological Change: An Empirical Exploration*. *The Quarterly Journal of Economics*, 118(4):1279–1333, 11 2003. ISSN 0033-5533. doi: 10.1162/003355303322552801. URL https://doi.org/10.1162/003355303322552801.

Anand Avati, Kenneth Jung, Stephanie Harman, Lance Downing, Andrew Ng, and Nigam H. Shah. Improving Palliative Care with Deep Learning. In *IEEE International Conference on Bioinformatics and Biomedicine*, nov 2017. URL http://arxiv.org/abs/1711.06402.

David Baehrens, Timon Schroeter, Stefan Harmeling, Motoaki Kawanabe, Katja Hansen, and Klaus-Robert Muller. How to Explain Individual Classification Decisions. *Journal of Machine Learning Research*, 11:1803–1831, 2010. ISSN 1532-4435.

Beccy Baird, Anna Charles, Matthew Honeyman, David Maguire, and Preety Das. Understanding pressures in general practice. Technical report, The King's Fund, London, UK, 2016. URL https://www.kingsfund.org.uk/publications/pressures-in-general-practicehttps://www.kingsfund.org.uk/sites/files/kf/field/field{_}publication{_}file/Understanding-GP-pressures-Kings-Fund-May-2016.pdf.

Hasan Bakhshi, Jonathan M Downing, Michael A Osborne, and Philippe Schneider. *The future of skills: employment in 2030*. Pearson, 2017.

Jørgen P. Bansler, Erling C. Havn, Kjeld Schmidt, Troels Mønsted, Helen Høgh Petersen, and Jesper Hastrup Svendsen. Cooperative Epistemic Work in Medical Practice: An Analysis of Physicians' Clinical Notes. *Computer Supported Cooperative Work (CSCW)*, 25(6):503–546, dec 2016.

ISSN 0925-9724. doi: 10.1007/s10606-016-9261-x. URL http://link.springer.com/10.1007/s10606-016-9261-x.

Paul Beaudry, David A. Green, and Benjamin M. Sand. The Great Reversal in the Demand for Skill and Cognitive Tasks. *Journal of Labor Economics*, 34(S1):S199–S247, 2016. ISSN 0734-306X. doi: 10.1086/682347. URL http://www.journals.uchicago.edu/doi/10.1086/682347.

J. Beech, S. Bottery, A. Charlesworth, H. Evans, B. Gershlick, N. Hemmings, C. Imison, P. Kahtan, H. McKenna, R. Murray, and B. Palmer. Closing the gap: Key areas for action on the health and care workforce. Technical report, Nuffield Trust, The Health Foundation and The King's Fund, London, UK, 2019. URL https://www.nuffieldtrust.org.uk/research/closing-the-gap-key-areas-for-action-on-the-health-and-care-workforce.

James E. Bessen. How Computer Automation Affects Occupations: Technology, Jobs, and Skills. *SSRN Electronic Journal*, oct 2015. ISSN 1556-5068. doi: 10.2139/ssrn.2690435. URL http://www.ssrn.com/abstract=2690435.

Charlotte Blease, Michael H Bernstein, Jens Gaab, Ted J Kaptchuk, Joe Kossowsky, Kenneth D Mandl, Roger B Davis, and Catherine M Desroches. Computerization and the future of primary care: A survey of general practitioners in the UK. *PLoS*, 13(12):1–12, 2018. doi: 10.1371/journal.pone.0207418December. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0207418.

Jeremy Bowles. The computerisation of european jobs. *Bruegel, Brussels*, 2014.

Timothy F Bresnahan, Erik Brynjolfsson, and Lorin M Hitt. Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics*, 117(1):339–376, 2002.

Erik Brynjolfsson and Lorin M Hitt. Computing productivity: Firm-level evidence. *Review of economics and statistics*, 85(4):793–808, 2003.

Erik Brynjolfsson and Andrew McAfee. *Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy*. Brynjolfsson and McAfee, 2012.

Erik Brynjolfsson and Andrew McAfee. *The second machine age : work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company, 2014. ISBN 0393239357.

Erik Brynjolfsson, Lorin M Hitt, and Shinkyu Yang. Intangible assets: Computers and organizational capital. *Brookings papers on economic activity*, 2002(1):137–181, 2002.

Laura Byrne, Jennifer Bottomley, and Alex Turk. British Medical Association Survey of GPs in England. Technical report, ICM Unlimited, London, 2016.

Federico Cabitza, Raffaele Rasoini, and Gian Franco Gensini. Unintended Consequences of Machine Learning in Medicine. *JAMA*, 2015(8):573165, jul 2017. ISSN 0098-7484. doi: 10.1001/jama.2017.7797. URL http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2017.7797.

Robert Challen, Joshua Denny, Martin Pitt, Luke Gompels, Tom Edwards, and Krasimira Tsaneva-Atanasova. Artificial intelligence, bias and clinical safety. *BMJ Quality & Safety*, 28(3):231–237, mar 2019. ISSN 2044-5415. doi: 10.1136/bmjqs-2018-008370. URL http://www.ncbi.nlm.nih.gov/pubmed/30636200http://qualitysafety.bmj.com/lookup/doi/10.1136/bmjqs-2018-008370.

Wei Chu and Zoubin Ghahramani. Gaussian processes for ordinal regression. *Journal of machine learning research*, 6(Jul):1019–1041, 2005.

Henry Clay and Rick Stern. Making Time in General Practice, 2015.

A. Coulter and B. Mearns. Developing care for a changing population: Patient engagement and health information technology. *Nuffield Trust*, 2016. URL `\url{https://www.nuffieldtrust.org.uk/files/2017-01/patient-engagement-and-health-information-web-final.pdf}`. Discussion paper.

Pamela N. Danziger. Google And Pointy Partner To Help Local Retailers Compete With Amazon, sep 2018. URL `https://www.forbes.com/sites/pamdanziger/2018/09/12/google-and-pointy-partner-to-help-local-retailers-compete-with-amazon/`.

Ara Darzi, Harry Quilter-Pinner, and Tom Kibasi. Better health and care for all: A 10-point plan for the 2020s, 2018.

Thomas H. Davenport and Julia Kirby. Just How Smart Are Smart Machines? *MIT Sloan Management Review*, 2016. URL `https://sloanreview.mit.edu/article/just-how-smart-are-smart-machines/`.

Simon de Lusignan and Chris van Weel. The use of routinely collected computer data for research in primary care: opportunities and challenges. *Family Practice*, 23(2):253–263, apr 2006. ISSN 1460-2229. doi: 10.1093/fampra/cmi106. URL `http://academic.oup.com/fampra/article/23/2/253/527321/The-use-of-routinely-collected-computer-data-for`.

Paul Duckworth, Logan Graham, and Michael A Osborne. Inferring Work Task Automatability from AI Expert Evidence. In *AAAI/ACM conference on AI Ethics and Society*, 2019.

Thomas J S Durant, Eben M Olson, Wade L Schulz, and Richard Torres. Very Deep Convolutional Neural Networks for Morphologic Classification of Erythrocytes. *Clinical chemistry*, 63(12):1847–1855, dec 2017. ISSN 1530-8561. doi: 10.1373/clinchem.2017.276345. URL `http://www.ncbi.nlm.nih.gov/pubmed/28877918`.

Carl Benedikt Frey and Michael A. Osborne. The Future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114:254–280, 2017. ISSN 00401625. doi: 10.1016/j.techfore.2016.08.019. URL `http://www.oxfordmartin.ox.ac.uk/downloads/academic/The{\_}Future{\_}of{\_}Employment.pdfhttp://dx.doi.org/10.1016/j.techfore.2016.08.019`.

Jonathan Gibson, Kath Checkland, Anna Coleman, Mark Hann, Robbie McCall, Sharon Spooner, and Matt Sutton. Eighth National GP Worklife Survey. Technical report, Policy research unit in commussioning and the healthcare system, Manchester, 2016. URL `http://research.bmh.manchester.ac.uk/healtheconomics/research/Reports/EighthNationalGPWorklifeSurveyreport/`.

Kate Goddard, Abdul Roudsari, and Jeremy C Wyatt. Automation bias - a hidden issue for clinical decision support system use. *Studies in health technology and informatics*, 164:17–22, 2011. ISSN 0926-9630. URL `http://www.ncbi.nlm.nih.gov/pubmed/21335682`.

Maarten Goos, Alan Manning, and Anna Salomons. Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review*, 104(8):2509–2526, aug 2014. ISSN 00028282. doi: 10.1257/aer.104.8.2509.

Trisha Greenhalgh, Christopher Voisey, and Nadia Robb. Interpreted consultations as 'business as usual'? An analysis of organisational routines in general practices. *Sociology of Health and Illness*, 29(6):931–954, 2007. ISSN 01419889. doi: 10.1111/j.1467-9566.2007.01047.x.

Dhruv Grewal, Anne L Roggeveen, and Jens Nordf??lt. The Future of Retailing. *Journal of Retailing*, 93(1):1–6, 2017. ISSN 00224359. doi: 10.1016/j.jretai.2016.12.008.

Seung Seog Han, Gyeong Hun Park, Woohyung Lim, Myoung Shin Kim, Jung Im Na, Ilwoo Park, and Sung Eun Chang. Deep neural networks show an equivalent and often superior performance to dermatologists in onychomycosis diagnosis: Automatic construction of onychomycosis datasets by region-based convolutional deep neural network. *PLOS ONE*, 13(1):e0191493, jan 2018. ISSN 1932-6203. doi: 10.1371/journal.pone.0191493. URL `https://dx.plos.org/10.1371/journal.pone.0191493`.

Michael J Handel. The O* NET content model: strengths and limitations Stärken und Grenzen des O*NET-Models. *Journal for Labour Market Research*, 49(2):157–176, 2016.

Janet R. Hardy, Theodore R. Holford, Gillian C. Hall, and Michael B. Bracken. Strategies for identifying pregnancies in the automated medical records of the General Practice Research Database. *Pharmacoepidemiology and Drug Safety*, 13(11):749–759, nov 2004. ISSN 1053-8569. doi: 10.1002/pds.935. URL `http://www.ncbi.nlm.nih.gov/pubmed/15386720http://doi.wiley.com/10.1002/pds.935`.

Eleonora Harwich and Kate Laycock. Thinking on its own: AI in the NHS, 2018.

Alexander Laycock Kate Hitchcock and Emilie Sundorph. Work in Progress: Towards a leaner, smarter public-sector workforce. Technical report, Reform, 2017.

F D Richard Hobbs, Clare Bankhead, Toqir Mukhtar, Sarah Stevens, Rafael Perera-Salazar, Tim Holt, Chris Salisbury, and National Institute for Health Research School for Primary Care Research. Clinical workload in UK primary care: a retrospective analysis of 100 million consultations in England, 2007-14. *Lancet (London, England)*, 387(10035):2323–30, jun 2016. ISSN 1474-547X. doi: 10.1016/S0140-6736(16)00620-6. URL `http://www.ncbi.nlm.nih.gov/pubmed/27059888http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4899422`.

Chris Hopson. The State of the NHS provider sector. Technical report, NHS Providers, 2016. URL `https://www.nhsproviders.org/news-blogs/news/nhs-trust-leaders-warn-staff-shortages-now-outweigh-fears-over-funding`.

Blake Irving. AI and the Economy Without You: The Ethics of Automation in a World That's Unsure It Needs People. blog post, 2017. URL `https://blakesblog.com/2017/06/ai-and-the-economy-without-you/`.

James Manyika, Michael Chui, Mehdi Miremadi, Jacques Bughin, Katy George, Paul Willmott, and Martin Dewhurst. Harnessing automation for a future that works, 2017. URL `http://www.mckinsey.com/global-themes/digital-disruption/harnessing-automation-for-a-future-that-works`.

Fei Jiang, Yong Jiang, Hui Zhi, Yi Dong, Hao Li, Sufeng Ma, Yilong Wang, Qiang Dong, Haipeng Shen, and Yongjun Wang. Artificial intelligence in healthcare: past, present and future. *Stroke and Vascular Neurology*, 2(4):230–243, dec 2017. ISSN 2059-8688. doi: 10.1136/svn-2017-000101. URL `http://svn.bmj.com/cgi/doi/10.1136/svn-2017-000101`.

Alistair E. W. Johnson, Mohammad M. Ghassemi, Shamim Nemati, Katherine E. Niehaus, David Clifton, and Gari D. Clifford. Machine Learning and Decision Support in Critical Care. *Proceedings of the IEEE*, 104(2):444–466, feb 2016. ISSN 0018-9219. doi: 10.1109/JPROC.2015.2501978. URL `http://www.ncbi.nlm.nih.gov/pubmed/27765959http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5066876http://ieeexplore.ieee.org/document/7390351/`.

Hyun-Chul Kim and Zoubin Ghahramani. Bayesian classifier combination. In *Artificial Intelligence and Statistics*, pages 619–627, 2012.

Jeffrey G Klann and Peter Szolovits. An intelligent listening framework for capturing encounter notes from a doctor-patient dialog. *BMC medical informatics and decision making*, 9 Suppl 1(Suppl 1):S3, nov 2009. ISSN 1472-6947. doi: 10.1186/1472-6947-9-S1-S3. URL `http://www.ncbi.nlm.nih.gov/pubmed/19891797http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2773918`.

Angus Knowles-Cutler, C B Frey, and M A Osborne. Agile Town: The Relentless March of Technology and London's Response. *Deloitte*, 2014.

Clemens Scott Kruse and Amanda Beane. Health information technology continues to show positive effect on medical outcomes: Systematic review. *Journal of medical Internet research*, 20(2), 2018.

Wassily Leontief and Faye Duchin. Technological Change in Health Care. In *The Future Impact of Automation on Workers*, chapter 5, page 192. Oxford University Press, Oxford, UK, 1 edition, 1986. ISBN 0195036239.

Yaniv Leviathan. Google Duplex: An AI System for Accomplishing Real-World Tasks Over the Phone, may 2018. URL `https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html`.

Frank Levy and Richard J Murnane. *The new division of labor: How computers are creating the next job market*. Princeton University Press, 2005.

Steven Y Lin, Tait D Shanafelt, and Steven M Asch. Reimagining Clinical Documentation With Artificial Intelligence. *Mayo Clinic proceedings*, 93(5):563–565, may 2018. ISSN 1942-5546. doi: 10.1016/j.mayocp.2018.02.016. URL `http://www.ncbi.nlm.nih.gov/pubmed/29631808`.

James Manyika, Michael Chui, Me Miremadi, J Bughin, K George, P Willmott, and M Dewhurst. A Future that Works: Automation, Employment, and Productivity. *McKinsey Global Institute*, 2017.

Sara Martin, Edward Davies, and Ben Gershlick. Under pressure: What the Commonwealth Fund's 2015 international survey of general practitioners means for the UK. Technical report, The Health Foundation, 2016. URL `http://www.health.org.uk/publication/under-pressurehttp://www.health.org.uk/sites/health/files/UnderPressure.pdf`.

Andrew D. Maynard. Navigating the fourth industrial revolution. *Nature Nanotechnology*, 10(12): 1005–1006, 2015. ISSN 17483395. doi: 10.1038/nnano.2015.286. URL `http://dx.doi.org/10.1038/nnano.2015.286`.

Patrick Milner, Emily Greenwood, Beth Hooper, Thom Cheminais, and Vivienne Roach. The Long-term Sustainability of the NHS and Adult Social Care. Technical report, The House of Lords, London, UK, 2017. URL `https://www.parliament.uk/business/`

committees/committees-a-z/commons-select/health-committee/news-parliament-2017/long-term-sustainability-launch-17-19/.

Joel Mokyr. *The Lever of Riches: Technological Creativity and Economic Progress*. Oxford University Press, jun 1990. ISBN 9780199854981. doi: 10.1093/acprof:oso/9780195074772.001.0001.

Ian Morris. *Why the west rules-for now: The patterns of history and what they reveal about the future*. Profile books, 2010.

Vincent C. Müller and Nick Bostrom. Future progress in artificial intelligence: A Survey of Expert Opinion. In Vincent C. Müller, editor, *Fundamental Issues of Artificial Intelligence*. Springer, Berlin, 2014. URL http://www.nickbostrom.com/papers/survey.pdf.

National Center for O*NET Development. O*NET OnLine, 2018. URL https://www.onetonline.org/.

Andrew Ng. What Artificial Intelligence Can and Can't Do Right Now. blog post, 2016. URL https://hbr.org/2016/11/what-artificial-intelligence-can-and-cant-do-right-now.

Ziad Obermeyer and Ezekiel J. Emanuel. Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. *New England Journal of Medicine*, 375(13):1216–1219, sep 2016. ISSN 0028-4793. doi: 10.1056/NEJMp1606181. URL http://www.nejm.org/doi/10.1056/NEJMp1606181.

U.S. Department of Labor. Occupational Employment Statistics, 2017. URL www.bls.gov/oes/.

Kevin O'Rourke, Ahmed Rahman, and Alan Taylor. Luddites and the Demographic Transition. *Journal of Economic Growth*, 18(4):373–409, nov 2008. doi: 10.3386/w14484. URL http://www.nber.org/papers/w14484.pdf.

Dympna O'Sullivan, Paolo Fraccaro, Ewart Carson, and Peter Weller. Decision time for clinical decision support systems. *Clinical medicine (London, England)*, 14(4):338–41, aug 2014. doi: 10.7861/clinmedicine.14-4-338. URL http://www.ncbi.nlm.nih.gov/pubmed/25099829.

Mika Pajarinen, Petri Rouvinen, et al. Computerization threatens one third of Finnish employment. *ETLA Brief*, 22(13.1):2014, 2014.

Michael Pooler. Amazon robots bring a brave new world to the warehouse, 2017. URL https://www.ft.com/content/916b93fc-8716-11e7-8bb1-5ba57d47eff7.

Carl Edward Rasmussen and Christopher K I Williams. *Gaussian Processes for Machine Learning*, volume 1. MIT Press, 2006.

Wayne D Rasmussen. The Mechanization of Agriculture. *Scientific American*, 247(3):76–89, sep 1982. ISSN 00368733 (ISSN). doi: 10.1038/scientificamerican0982-76.

Mingjun Zhang Robin Felder, Majd Alwan. *Systems Engineering Approach to Medical Automation*. Artech House, 2008.

Martin Roland and Sam Everington. Tackling the crisis in general practice. *BMJ (Clinical research ed.)*, 352:i942, feb 2016. ISSN 1756-1833. doi: 10.1136/bmj.i942. URL http://www.ncbi.nlm.nih.gov/pubmed/26887896.

Jeffrey D Sachs and Laurence J Kotlikoff. Smart machines and long-term misery. Technical report, National Bureau of Economic Research, 2012.

Catherine Shu. Google Calendar's newest feature uses machine learning to help you actually accomplish your goals, apr 2016. URL https://techcrunch.com/2016/04/12/google-calendar-goals/.

Edwin Simpson, Stephen Roberts, Ioannis Psorakis, and Arfon Smith. Dynamic bayesian combination of multiple imperfect classifiers. In *Decision making and imperfection*, pages 1–35. Springer, 2013.

Adam Steventon. How to enable the uptake of technology in the NHS?, oct 2018. URL https://www.hsj.co.uk/technology-and-innovation/how-to-enable-the-uptake-of-technology-in-the-nhs/7023539.article.

Richard E. Susskind and Daniel Susskind. *The future of the professions: how technology will transform the work of human experts*. OUP Oxford, 2017. ISBN 0198713398.

Deborah Swinglehurst and Trisha Greenhalgh. Caring for the patient, caring for the record: An ethnographic study of 'back office' work in upholding quality of care in general practice. *BMC Health Services Research*, 15(1):1–12, 2015. ISSN 14726963. doi: 10.1186/s12913-015-0774-7. URL http://dx.doi.org/10.1186/s12913-015-0774-7.

TUC. Breaking point: The crisis in mental health funding. Technical report, Trade Union Congress, London, UK, 2018. URL https://www.tuc.org.uk/research-analysis/reports/breaking-point-crisis-mental-health-funding.

Jonathan Vanian. How Walmart Uses Robots at 50 Stores To Track Inventory, mar 2018. URL http://fortune.com/2018/03/26/walmart-robot-bossa-nova/.

Abraham Verghese, Nigam H. Shah, and Robert A. Harrington. What This Computer Needs Is a Physician: Humanism and Artificial Intelligence. *JAMA*, 319(1):19, jan 2018. ISSN 0098-7484. doi: 10.1001/jama.2017.19198. URL http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2017.19198.

Hedy S. Wald, Jeffrey M. Borkan, Julie Scott Taylor, David Anthony, and Shmuel P. Reis. Fostering and Evaluating Reflective Capacity in Medical Education: Developing the REFLECT Rubric for Assessing Reflective Writing. *Academic Medicine*, 87(1):41–50, jan 2012. ISSN 1040-2446. doi: 10.1097/ACM.0b013e31823b55fa.

Shijun Wang and Ronald M. Summers. Machine learning and radiology. *Medical Image Analysis*, 16(5):933–951, jul 2012. ISSN 13618415. doi: 10.1016/j.media.2012.02.005. URL http://www.ncbi.nlm.nih.gov/pubmed/22465077http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3372692http://linkinghub.elsevier.com/retrieve/pii/S1361841512000333.

Jonathan P Weiner, Susan Yeh, and David Blumenthal. The impact of health information technology and e-health on the future demand for physician services. *Health affairs (Project Hope)*, 32(11):1998–2004, nov 2013. ISSN 1544-5208. doi: 10.1377/hlthaff.2013.0680. URL http://www.ncbi.nlm.nih.gov/pubmed/24191092.

Matthew Willis, Paul Duckworth, Angela Coulter, Eric T Meyer, and Michael Osborne. The Future of Health Care: Protocol for Measuring the Potential of Task Automation Grounded in the National Health Service Primary Care System. *JMIR Research Protocols*, 8(4):e11232, apr 2019. ISSN 1929-0748. doi: 10.2196/11232. URL https://www.researchprotocols.org/2019/4/e11232/.

Jionglin Wu, Jason Roy, and Walter F. Stewart. Prediction Modeling Using EHR Data. *Medical Care*, 48(6 Suppl):S106–S113, jun 2010. ISSN 0025-7079. doi: 10.1097/MLR.0b013e3181de9e17. URL http://www.ncbi.nlm.nih.gov/pubmed/20473190https://insights.ovid.com/crossref?an=00005650-201006001-00017.

Supplemental Material

11 Supplemental Material

11.1 Appendix A: Task Automatability Predictions

Table 5: Observed Clerical Primary Care Tasks and Automatility Scores.

Observed Task	Automatility score
Print out lab test labels in ICE	1.760
Label blood vials	1.760
Staff recruitment	1.824
Manage finances for the practice including paying bills	1.833
Having practice staff take online training	1.908
Writing and updating policies	1.958
Manage petty cash	2.070
Manage pension schemes	2.108
Use accounting system, online or desktop software, for finances	2.108
Managing meeting diaries	2.137
Conducting training in-person	2.215
Generate QOF letters	2.235
Write notes on paper	2.235
Write medical report letters	2.235
Conduct risk assesment	2.236
Invoicing (private insurance)	2.378
Use software for invoicing	2.378
Maintain documentation for 2 week waits	2.399
Talk to other staff about incentive schemes	2.419
Work in Open Exeter online web portal	2.519
Helping with research participation or recruitment for research projects at the practice	2.526
Prepare for CQC regulation checks	2.538
Payroll in system called IRIS	2.553
Fill out/sign third party reports	2.590
Giving advice to colleagues/discuss medicine and other professional topics	2.687
Perform administrative tasks	2.743
Order ambulance for medical emergency in practice.	2.760
Request referral for outside clinic to hospital or specalist	2.760
Admit patient to hospital	2.760
Schedule locums	2.764
Use incentive scheme online system to manage commissioning of care, additional services, and other practice incentives.	2.781
Use system to see and interact with data and information from radiology, haematology, and distchange documents	2.781
Process patient deductions	2.783
Register new patients to the practice	2.783
Process audit paperwork	2.783
Process out of hours reports and create tasks to assign to GPs	2.783
Cleaning up information in the patients electronic health record if there are any problems or incorrect or incomplete information	2.783
Process prescription declaration forms	2.783

Fill out or process audit forms, conduct an audit or keep track of data incase of an audit.	2.783
Schedule clinics	2.793
Health and saftey checks	2.822
Ordering supplies for practice and office	2.824
Reorder stock of injections or immunizations	2.827
Attend to the front desk	2.838
Provide data, notes, and other information to GPs for pallative care meeting or other case discussion meetings.	2.839
Conduct business intelligence for future directions, business plans, practice management strategies	2.839
Meetings/human resources	2.855
Enter data for enhanced services	2.856
Look into drug shortages or other issues with drugs	2.880
Provide clinical staff with blood monitoring suggestions	2.897
Assign letters to be reviewed to other employees.	2.903
Runs DBS checks.	2.936
Create timeslots for clinical/bookable hours	2.950
Type two week wait letters	2.953
Use NHS eReferral system to help patients book appointments and manage referrals to different specialists	2.960
Answer queries from patients or from GPs about perscriptions	2.964
Deal with complaints	2.968
Type letters	2.983
Write letters for secondary care, other GPs, or for patients. Usually involves a template.	2.983
Significant event reporting	2.985
Maintain supplies and stock in clinical room	3.000
Patient recalls for QOF, Write letters or texts using software or spreadsheets (example: patientchase). Mass mail letters with Docmail	3.004
Send patient texts to attend a chronic condition clinic	3.016
Use software to convert printed letters into text on the computer	3.027
Reviewing health records addressing any changes from health authority	3.034
Assign prescriptions to doctors after reviewing and to sign.	3.045
Ordering immunizations	3.055
Schedule/book patients	3.064
Conduct Diabeies reviews	3.065
Process perscription renewals	3.067
Indicating what is to be coded by a medical coder	3.075
Manage agenda and record minutes for meetings	3.076
Look for drug interactions	3.079
Run search in health record for patients that need medication reviews	3.081
Run searches to look for repeat prescriptions	3.081
Respond to tasks in electronic medical record	3.091
Create and assign tasks for others in the electronic medical record or other software	3.091
Create work or holiday schedules for practice staff	3.100
Staff rotas (clinical or practice)	3.100
Address problems that arise with building	3.105
Work in clinical query system to create reports on patients data	3.158

Help organise and schedule patients for clinics, such as: diabetics, injections, and other chronic diseases.	3.167
Deprescribing of high cost drugs and other drug audits	3.209
Review letters	3.215
Medical Coding of letters and other documents	3.220
Work in NHS CQRS online web portal	3.232
Register new Patients	3.235
Connecting human resources/making introductions of new staff	3.268
Ordering of goods, stationary, office supplies	3.283
Send messages in electronic health record to communicate to other staff	3.304
Transcribe audio recorded letters	3.311
Checking for errors in paperwork	3.329
Work in patient management system for QOF, Recalls, or other patient communication systems	3.348
Check prescriptions for medication renewals and send to GP	3.351
Control prescription refills or authorizations.	3.351
Answer phone	3.362
Reconcile information with a phonecall	3.362
Use software to manage transcription of letters or other documents	3.366
Perscription queries	3.375
Perform Medication reviews	3.375
Organise and manage main practice email account.	3.387
Check incoming emails to main practice email address, forward emails to other employees.	3.399
Scan letters and other documents	3.514
Use software or database to paper records that are in storage either onsite or offsite	3.524
Use texting service or patient management service to contact patients for different clinics and to send reminders.	3.604
Print letters	3.609
Mass mail letters for checkups using DocMail	3.725
Check and sort mail.	3.737

11.2 Appendix B: O*NET Important Features

Table 6: Global: All O*NET feature derivatives.

O*NET Feature	Feature Gradient
Telecommunications (knowledge)	+0.167
Clerical (knowledge)	+0.166
Wrist-Finger Speed (ability)	+0.153
Number Facility (ability)	+0.118
Mathematics_x (skill)	+0.093
Depth Perception (ability)	+0.092
Building and Construction (knowledge)	+0.090
Mathematical Reasoning (ability)	+0.088
Economics and Accounting (knowledge)	+0.085
Control Precision (ability)	+0.082
Response Orientation (ability)	+0.081
Arm-Hand Steadiness (ability)	+0.077
Sales and Marketing (knowledge)	+0.076
Equipment Selection (skill)	+0.072
Finger Dexterity (ability)	+0.067
Perceptual Speed (ability)	+0.066
Static Strength (ability)	+0.059
Visual Color Discrimination (ability)	+0.058
Far Vision (ability)	+0.046
Spatial Orientation (ability)	+0.046
Multilimb Coordination (ability)	+0.045
Manual Dexterity (ability)	+0.042
Extent Flexibility (ability)	+0.041
Production and Processing (knowledge)	+0.038
Flexibility of Closure (ability)	+0.035
Night Vision (ability)	+0.034
Transportation (knowledge)	+0.025
Near Vision (ability)	+0.023
Dynamic Strength (ability)	+0.023
Writing (skill)	+0.023
Speed of Closure (ability)	+0.022
Computers and Electronics (knowledge)	+0.022
Operation Monitoring (skill)	+0.017
Visualization (ability)	+0.016
Medicine and Dentistry (knowledge)	+0.014
Equipment Maintenance (skill)	+0.013
Education and Training (knowledge)	+0.012
Category Flexibility (ability)	+0.012
Biology (knowledge)	+0.012
Written Comprehension (ability)	+0.011
Sound Localization (ability)	+0.011
Inductive Reasoning (ability)	+0.011
Selective Attention (ability)	+0.009
Mathematics_y (knowledge)	+0.009
Written Expression (ability)	+0.008

Peripheral Vision (ability)	+0.007
Reading Comprehension (skill)	+0.006
Operation and Control (skill)	+0.006
Rate Control (ability)	+0.003
Administration and Management (knowledge)	+0.002
Information Ordering (ability)	+0.002
Programming (skill)	-0.172
Installation (skill)	-0.162
Technology Design (skill)	-0.157
Fine Arts (knowledge)	-0.137
Dynamic Flexibility (ability)	-0.114
Psychology (knowledge)	-0.113
History and Archeology (knowledge)	-0.109
Science (skill)	-0.107
Sociology and Anthropology (knowledge)	-0.105
Food Production (knowledge)	-0.105
Personnel and Human Resources (knowledge)	-0.102
Speed of Limb Movement (ability)	-0.100
Management of Personnel Resources (skill)	-0.082
Gross Body Equilibrium (ability)	-0.080
Learning Strategies (skill)	-0.071
Negotiation (skill)	-0.069
Troubleshooting (skill)	-0.068
Coordination (skill)	-0.068
Glare Sensitivity (ability)	-0.067
Law and Government (knowledge)	-0.067
Chemistry (knowledge)	-0.065
Explosive Strength (ability)	-0.064
Social Perceptiveness (skill)	-0.063
Gross Body Coordination (ability)	-0.062
Originality (ability)	-0.058
Persuasion (skill)	-0.057
Management of Financial Resources (skill)	-0.057
Public Safety and Security (knowledge)	-0.047
English Language (knowledge)	-0.047
Instructing (skill)	-0.046
Operations Analysis (skill)	-0.044
Communications and Media (knowledge)	-0.042
Management of Material Resources (skill)	-0.041
Engineering and Technology (knowledge)	-0.041
Time Management (skill)	-0.038
Fluency of Ideas (ability)	-0.038
Physics (knowledge)	-0.037
Active Learning (skill)	-0.036
Service Orientation (skill)	-0.030
Mechanical (knowledge)	-0.030
Time Sharing (ability)	-0.029
Oral Expression (ability)	-0.029
Geography (knowledge)	-0.027
Monitoring (skill)	-0.025

Reaction Time (ability)	-0.023
Customer and Personal Service (knowledge)	-0.023
Speaking (skill)	-0.020
Active Listening (skill)	-0.020
Hearing Sensitivity (ability)	-0.019
Stamina (ability)	-0.018
Foreign Language (knowledge)	-0.018
Complex Problem Solving (skill)	-0.016
Speech Clarity (ability)	-0.016
Philosophy and Theology (knowledge)	-0.015
Speech Recognition (ability)	-0.014
Trunk Strength (ability)	-0.014
Problem Sensitivity (ability)	-0.013
Oral Comprehension (ability)	-0.012
Judgment and Decision Making (skill)	-0.012
Critical Thinking (skill)	-0.012
Auditory Attention (ability)	-0.011
Memorization (ability)	-0.010
Repairing (skill)	-0.008
Quality Control Analysis (skill)	-0.006
Deductive Reasoning (ability)	-0.005
Therapy and Counseling (knowledge)	-0.003
Systems Evaluation (skill)	-0.002
Systems Analysis (skill)	0.000
Design (knowledge)	0.001

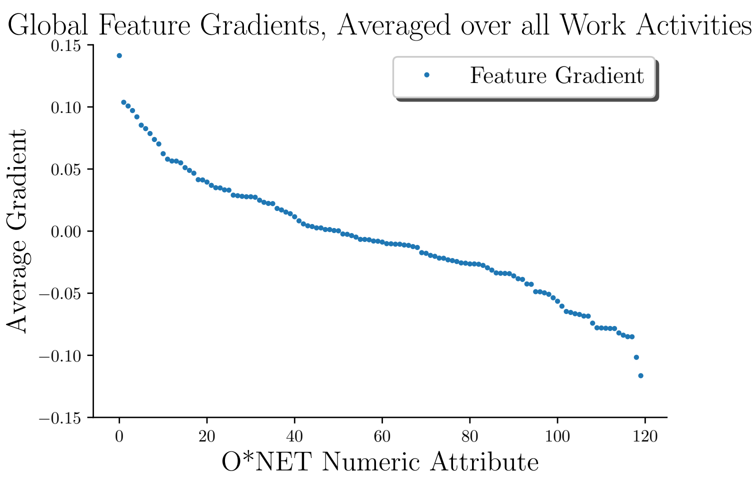


Figure 18: Importance of each of the 120 O*NET job characteristics are to Healthcare Task Automatability

11.3 Appendix C: Occupation-Level Analysis

Here, we give a detailed break down of each occupation's observed tasks, their most and least automatable tasks, the important O*NET features of each, and the largest positive and negative average derivatives of the features.

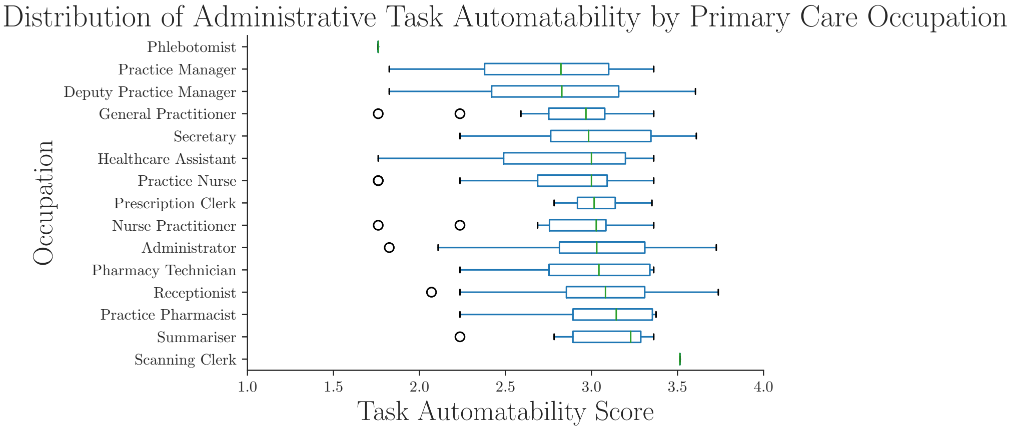


Figure 19: Distributions of Administrative Task Automatability Scores.

Administrator

Administrator: Three most and least automatable tasks

	Task	Automation Score
Lowest	Staff recruitment	1.824
	Manage pension schemes	2.108
	Write notes on paper	2.235
Highest	Mass mail letters for checkups using DocMail	3.725
	Print letters	3.609
	Use texting service or patient management service to contact patients for different clinics and to send reminders.	3.604

Administrator: Largest feature differences relative to the population

O*NET Feature	Feature Difference
Clerical (knowledge)	+20.1 %
Customer and Personal Service (knowledge)	+10.0 %
Computers and Electronics (knowledge)	+6.9 %
Service Orientation (skill)	+6.3 %
English Language (knowledge)	+6.2 %
Mechanical (knowledge)	−27.2 %
Building and Construction (knowledge)	−26.1 %
Repairing (skill)	−25.5 %
Troubleshooting (skill)	−24.4 %
Equipment Maintenance (skill)	−24.2 %

Administrator: O*NET features with the largest (positive and negative) derivatives.

O*NET Feature	Feature Gradient
Programming (skill)	-0.178
Technology Design (skill)	-0.164
Fine Arts (knowledge)	-0.152
Installation (skill)	-0.148
Food Production (knowledge)	-0.124
Clerical (knowledge)	+0.174
Telecommunications (knowledge)	+0.163
Wrist-Finger Speed (ability)	+0.158
Number Facility (ability)	+0.125
Building and Construction (knowledge)	+0.103

Deputy Practice Manager

Deputy Practice Manager: Three most and least automatable tasks

	Task	Automation Score
Lowest	Staff recruitment	1.824
	Manage finances for the practice including paying bills	1.833
	Use accounting system, online or desktop software, for finances	2.108
Highest	Use texting service or patient management service to contact patients for different clinics and to send reminders.	3.604
	Reconcile information with a phonecall	3.362
	Answer phone	3.362

Deputy Practice Manager: Largest feature differences relative to the population

O*NET Feature	Feature Difference
Clerical (knowledge)	+16.5 %
Economics and Accounting (knowledge)	+9.1 %
Customer and Personal Service (knowledge)	+8.5 %
Communications and Media (knowledge)	+7.9 %
Service Orientation (skill)	+6.9 %
Repairing (skill)	−24.4 %
Reaction Time (ability)	−24.0 %
Multilimb Coordination (ability)	−23.2 %
Equipment Maintenance (skill)	−22.9 %
Mechanical (knowledge)	−22.5 %

Deputy Practice Manager: O*NET features with the largest (positive and negative) derivatives.

O*NET Feature	Feature Gradient
Programming (skill)	-0.172
Technology Design (skill)	-0.163
Installation (skill)	-0.144
Fine Arts (knowledge)	-0.142
Science (skill)	-0.117
Clerical (knowledge)	+0.169
Telecommunications (knowledge)	+0.164
Wrist-Finger Speed (ability)	+0.153
Number Facility (ability)	+0.123
Building and Construction (knowledge)	+0.100

General Practitioner

General Practitioner: Three most and least automatable tasks

	Task	Automation Score
Lowest	Print out lab test labels in ICE	1.760
	Write notes on paper	2.235
	Fill out/sign third party reports	2.590
Highest	Reconcile information with a phonecall	3.362
	Review letters	3.215
	Respond to tasks in electronic medical record	3.091

General Practitioner: Largest feature differences relative to the population

O*NET Feature	Feature Difference
Clerical (knowledge)	+19.1 %
Medicine and Dentistry (knowledge)	+16.4 %
Customer and Personal Service (knowledge)	+12.7 %
Therapy and Counseling (knowledge)	+11.6 %
English Language (knowledge)	+8.5 %
Repairing (skill)	−26.6 %
Equipment Maintenance (skill)	−26.0 %
Reaction Time (ability)	−25.1 %
Engineering and Technology (knowledge)	−24.6 %
Rate Control (ability)	−24.3 %

General Practitioner: O*NET features with the largest (positive and negative) derivatives.

O*NET Feature	Feature Gradient
Programming (skill)	-0.170
Technology Design (skill)	-0.156
Fine Arts (knowledge)	-0.149
Installation (skill)	-0.136
Food Production (knowledge)	-0.126
Clerical (knowledge)	+0.168
Telecommunications (knowledge)	+0.156
Wrist-Finger Speed (ability)	+0.151
Number Facility (ability)	+0.122
Mathematics_x (skill)	+0.097

Healthcare Assistant

Healthcare Assistant: Three most and least automatable tasks

	Task	Automation Score
Lowest	Label blood vials	1.760
	Print out lab test labels in ICE	1.760
	Write notes on paper	2.235
Highest	Reconcile information with a phonecall	3.362
	Answer phone	3.362
	Send messages in electronic health record to communicate to other staff	3.304

Healthcare Assistant: Largest feature differences relative to the population

O*NET Feature	Feature Difference
Installation (skill)	+33.0 %
Clerical (knowledge)	+12.3 %
Customer and Personal Service (knowledge)	+9.3 %
Foreign Language (knowledge)	+5.6 %
Telecommunications (knowledge)	+5.4 %
Science (skill)	−22.8 %
Chemistry (knowledge)	−19.3 %
Food Production (knowledge)	−17.6 %
Biology (knowledge)	−15.7 %
Management of Financial Resources (skill)	−12.8 %

Healthcare Assistant: O*NET features with the largest (positive and negative) derivatives.

O*NET Feature	Feature Gradient
Programming (skill)	-0.149
Installation (skill)	-0.142
Technology Design (skill)	-0.141
Fine Arts (knowledge)	-0.129
Dynamic Flexibility (ability)	-0.111
Telecommunications (knowledge)	+0.163
Clerical (knowledge)	+0.155
Wrist-Finger Speed (ability)	+0.141
Number Facility (ability)	+0.110
Mathematics_x (skill)	+0.088

Nurse Practitioner

Nurse Practitioner: Three most and least automatable tasks

	Task	Automation Score
Lowest	Print out lab test labels in ICE	1.760
	Write notes on paper	2.235
	Giving advice to colleagues/discuss medicine and other professional topics	2.687
Highest	Reconcile information with a phonecall	3.362
	Answer phone	3.362
	Send messages in electronic health record to communicate to other staff	3.304

Nurse Practitioner: Largest feature differences relative to the population

O*NET Feature	Feature Difference
Medicine and Dentistry (knowledge)	+16.2 %
Clerical (knowledge)	+12.7 %
Customer and Personal Service (knowledge)	+10.2 %
Therapy and Counseling (knowledge)	+8.9 %
Service Orientation (skill)	+6.0 %
Engineering and Technology (knowledge)	−20.4 %
Design (knowledge)	−18.9 %
Building and Construction (knowledge)	−17.0 %
Mechanical (knowledge)	−14.7 %
Physics (knowledge)	−14.6 %

Nurse Practitioner: O*NET features with the largest (positive and negative) derivatives.

O*NET Feature	Feature Gradient
Installation (skill)	-0.164
Programming (skill)	-0.157
Technology Design (skill)	-0.147
Fine Arts (knowledge)	-0.135
Food Production (knowledge)	-0.123
Telecommunications (knowledge)	+0.161
Clerical (knowledge)	+0.161
Wrist-Finger Speed (ability)	+0.145
Number Facility (ability)	+0.116
Mathematics_x (skill)	+0.092

Pharmacy Technician

Pharmacy Technician: Three most and least automatable tasks

	Task	Automation Score
Lowest	Write notes on paper	2.235
	Perform administrative tasks	2.743
	Cleaning up information in the patients electronic health record if there are any problems or incorrect or incomplete information	2.783
Highest	Answer phone	3.362
	Control prescription refills or authorizations.	3.351
	Send messages in electronic health record to communicate to other staff	3.304

Pharmacy Technician: Largest feature differences relative to the population

O*NET Feature	Feature Difference
Medicine and Dentistry (knowledge)	+21.5 %
Clerical (knowledge)	+11.5 %
Communications and Media (knowledge)	+6.9 %
Customer and Personal Service (knowledge)	+6.3 %
Therapy and Counseling (knowledge)	+5.7 %
Building and Construction (knowledge)	−29.0 %
Engineering and Technology (knowledge)	−25.7 %
Design (knowledge)	−22.6 %
Mechanical (knowledge)	−22.0 %
Peripheral Vision (ability)	−18.8 %

Pharmacy Technician: O*NET features with the largest (positive and negative) derivatives.

O*NET Feature	Feature Gradient
Installation (skill)	-0.176
Programming (skill)	-0.170
Technology Design (skill)	-0.155
Fine Arts (knowledge)	-0.141
Food Production (knowledge)	-0.137
Clerical (knowledge)	+0.168
Telecommunications (knowledge)	+0.164
Wrist-Finger Speed (ability)	+0.149
Number Facility (ability)	+0.121
Mathematics_x (skill)	+0.095

Phlebotomist

Phlebotomist: Three most and least automatable tasks

	Task	Automation Score
Lowest	Label blood vials	1.760
Highest	Label blood vials	1.760

Phlebotomist: Largest feature differences relative to the population

O*NET Feature	Feature Difference
Installation (skill)	+228.0 %
Building and Construction (knowledge)	+130.5 %
Mechanical (knowledge)	+82.6 %
Design (knowledge)	+72.4 %
Repairing (skill)	+65.8 %
Medicine and Dentistry (knowledge)	−41.6 %
Therapy and Counseling (knowledge)	−40.6 %
Biology (knowledge)	−33.1 %
Sociology and Anthropology (knowledge)	−33.0 %
Psychology (knowledge)	−31.9 %

Phlebotomist: O*NET features with the largest (positive and negative) derivatives.

O*NET Feature	Feature Gradient
Gross Body Equilibrium (ability)	-0.147
Speed of Limb Movement (ability)	-0.091
Medicine and Dentistry (knowledge)	-0.091
Glare Sensitivity (ability)	-0.089
Sociology and Anthropology (knowledge)	-0.077
Telecommunications (knowledge)	+0.149
Wrist-Finger Speed (ability)	+0.077
Clerical (knowledge)	+0.075
Design (knowledge)	+0.071
Number Facility (ability)	+0.056

Practice Manager

Practice Manager: Three most and least automatable tasks

	Task	Automation Score
Lowest	Staff recruitment	1.824
	Manage finances for the practice including paying bills	1.833
	Having practice staff take online training	1.908
Highest	Answer phone	3.362
	Reconcile information with a phonecall	3.362
	Work in patient management system for QOF, Recalls, or other patient communication systems	3.348

Practice Manager: Largest feature differences relative to the population

O*NET Feature	Feature Difference
Clerical (knowledge)	+16.2 %
Economics and Accounting (knowledge)	+8.8 %
Customer and Personal Service (knowledge)	+8.6 %
Personnel and Human Resources (knowledge)	+8.4 %
Service Orientation (skill)	+7.1 %
Repairing (skill)	−25.0 %
Equipment Maintenance (skill)	−24.3 %
Equipment Selection (skill)	−24.1 %
Multilimb Coordination (ability)	−23.8 %
Mechanical (knowledge)	−23.3 %

Practice Manager: O*NET features with the largest (positive and negative) derivatives.

O*NET Feature	Feature Gradient
Programming (skill)	-0.171
Technology Design (skill)	-0.162
Installation (skill)	-0.146
Fine Arts (knowledge)	-0.145
Science (skill)	-0.115
Clerical (knowledge)	+0.169
Telecommunications (knowledge)	+0.162
Wrist-Finger Speed (ability)	+0.151
Number Facility (ability)	+0.123
Mathematics_x (skill)	+0.099

Practice Nurse

Practice Nurse: Three most and least automatable tasks

	Task	Automation Score
Lowest	Print out lab test labels in ICE	1.760
	Label blood vials	1.760
	Write notes on paper	2.235
Highest	Answer phone	3.362
	Reconcile information with a phonecall	3.362
	Send messages in electronic health record to communicate to other staff	3.304

Practice Nurse: Largest feature differences relative to the population

O*NET Feature	Feature Difference
Installation (skill)	+16.1 %
Clerical (knowledge)	+13.8 %
Customer and Personal Service (knowledge)	+10.9 %
Medicine and Dentistry (knowledge)	+6.0 %
Service Orientation (skill)	+5.2 %
Science (skill)	−18.7 %
Engineering and Technology (knowledge)	−17.2 %
Chemistry (knowledge)	−17.0 %
Physics (knowledge)	−14.8 %
Food Production (knowledge)	−14.6 %

Practice Nurse: O*NET features with the largest (positive and negative) derivatives.

O*NET Feature	Feature Gradient
Programming (skill)	-0.155
Installation (skill)	-0.149
Technology Design (skill)	-0.146
Fine Arts (knowledge)	-0.134
Food Production (knowledge)	-0.116
Telecommunications (knowledge)	+0.161
Clerical (knowledge)	+0.159
Wrist-Finger Speed (ability)	+0.144
Number Facility (ability)	+0.115
Mathematics_x (skill)	+0.092

Practice Pharmacist

Practice Pharmacist: Three most and least automatable tasks

	Task	Automation Score
Lowest	Write notes on paper	2.235
	Cleaning up information in the patients electronic health record if there are any problems or incorrect or incomplete information	2.783
	Look into drug shortages or other issues with drugs	2.880
Highest	Perscription queries	3.375
	Perform Medication reviews	3.375
	Answer phone	3.362

Practice Pharmacist: Largest feature differences relative to the population

O*NET Feature	Feature Difference
Medicine and Dentistry (knowledge)	+18.4 %
Clerical (knowledge)	+14.9 %
Customer and Personal Service (knowledge)	+11.1 %
Sales and Marketing (knowledge)	+6.3 %
Service Orientation (skill)	+4.8 %
Engineering and Technology (knowledge)	−25.8 %
Design (knowledge)	−25.6 %
Building and Construction (knowledge)	−25.5 %
Spatial Orientation (ability)	−19.6 %
Glare Sensitivity (ability)	−19.5 %

Practice Pharmacist: O*NET features with the largest (positive and negative) derivatives.

O*NET Feature	Feature Gradient
Installation (skill)	-0.171
Programming (skill)	-0.166
Technology Design (skill)	-0.157
Fine Arts (knowledge)	-0.136
Food Production (knowledge)	-0.127
Clerical (knowledge)	+0.166
Telecommunications (knowledge)	+0.163
Wrist-Finger Speed (ability)	+0.149
Number Facility (ability)	+0.122
Mathematics_x (skill)	+0.096

Prescription Clerk

Prescription Clerk: Three most and least automatable tasks

	Task	Automation Score
Lowest	Process out of hours reports and create tasks to assign to GPs	2.783
	Answer queries from patients or from GPs about perscriptions	2.964
	Process perscription renewals	3.067
Highest	Check prescriptions for medication renewals and send to GP	3.351
	Process perscription renewals	3.067
	Answer queries from patients or from GPs about perscriptions	2.964

Prescription Clerk: Largest feature differences relative to the population

O*NET Feature	Feature Difference
Medicine and Dentistry (knowledge)	+64.6 %
Therapy and Counseling (knowledge)	+27.9 %
Biology (knowledge)	+26.4 %
Customer and Personal Service (knowledge)	+20.5 %
Clerical (knowledge)	+19.7 %
Design (knowledge)	−36.2 %
Engineering and Technology (knowledge)	−33.2 %
Building and Construction (knowledge)	−32.7 %
Mechanical (knowledge)	−30.7 %
Repairing (skill)	−29.9 %

Prescription Clerk: O*NET features with the largest (positive and negative) derivatives.

O*NET Feature	Feature Gradient
Programming (skill)	-0.176
Installation (skill)	-0.161
Technology Design (skill)	-0.158
Food Production (knowledge)	-0.152
Fine Arts (knowledge)	-0.150
Clerical (knowledge)	+0.175
Telecommunications (knowledge)	+0.152
Wrist-Finger Speed (ability)	+0.149
Number Facility (ability)	+0.132
Mathematics_x (skill)	+0.104

Receptionist

Receptionist: Three most and least automatable tasks

	Task	Automation Score
Lowest	Manage petty cash	2.070
	Write notes on paper	2.235
	Maintain documentation for 2 week waits	2.399
Highest	Check and sort mail.	3.737
	Print letters	3.609
	Use texting service or patient management service to contact patients for different clinics and to send reminders.	3.604

Receptionist: Largest feature differences relative to the population

O*NET Feature	Feature Difference
Clerical (knowledge)	+23.3 %
Medicine and Dentistry (knowledge)	+17.0 %
Customer and Personal Service (knowledge)	+11.2 %
English Language (knowledge)	+7.2 %
Service Orientation (skill)	+7.0 %
Mechanical (knowledge)	−31.1 %
Building and Construction (knowledge)	−31.0 %
Repairing (skill)	−28.6 %
Engineering and Technology (knowledge)	−28.4 %
Design (knowledge)	−27.6 %

Receptionist: O*NET features with the largest (positive and negative) derivatives.

O*NET Feature	Feature Gradient
Programming (skill)	-0.178
Technology Design (skill)	-0.163
Fine Arts (knowledge)	-0.159
Installation (skill)	-0.145
Food Production (knowledge)	-0.134
Clerical (knowledge)	+0.176
Telecommunications (knowledge)	+0.160
Wrist-Finger Speed (ability)	+0.158
Number Facility (ability)	+0.126
Building and Construction (knowledge)	+0.103

Scanning Clerk

Scanning Clerk: Three most and least automatable tasks

	Task	Automation Score
Lowest	Scan letters and other documents	3.514
Highest	Scan letters and other documents	3.514

Scanning Clerk: Largest feature differences relative to the population

O*NET Feature	Feature Difference
Clerical (knowledge)	+33.1 %
Computers and Electronics (knowledge)	+21.8 %
Programming (skill)	+20.5 %
Fine Arts (knowledge)	+9.9 %
Perceptual Speed (ability)	+9.4 %
Science (skill)	−41.9 %
Chemistry (knowledge)	−37.2 %
Physics (knowledge)	−33.5 %
Building and Construction (knowledge)	−32.6 %
Mechanical (knowledge)	−31.2 %

Scanning Clerk: O*NET features with the largest (positive and negative) derivatives.

O*NET Feature	Feature Gradient
Programming (skill)	-0.183
Technology Design (skill)	-0.172
Fine Arts (knowledge)	-0.156
Installation (skill)	-0.155
Science (skill)	-0.131
Clerical (knowledge)	+0.181
Telecommunications (knowledge)	+0.175
Wrist-Finger Speed (ability)	+0.172
Number Facility (ability)	+0.127
Building and Construction (knowledge)	+0.115

Secretary

Secretary: Three most and least automatable tasks

	Task	Automation Score
Lowest	Write notes on paper	2.235
	Write medical report letters	2.235
	Use software for invoicing	2.378
Highest	Print letters	3.609
	Scan letters and other documents	3.514
	Organise and manage main practice email account.	3.387

Secretary: Largest feature differences relative to the population

O*NET Feature	Feature Difference
Clerical (knowledge)	+19.7 %
Customer and Personal Service (knowledge)	+8.8 %
Communications and Media (knowledge)	+8.4 %
Medicine and Dentistry (knowledge)	+7.1 %
English Language (knowledge)	+6.7 %
Mechanical (knowledge)	−30.5 %
Building and Construction (knowledge)	−30.2 %
Repairing (skill)	−27.3 %
Troubleshooting (skill)	−26.7 %
Engineering and Technology (knowledge)	−26.7 %

Secretary: O*NET features with the largest (positive and negative) derivatives.

O*NET Feature	Feature Gradient
Programming (skill)	-0.179
Technology Design (skill)	-0.163
Fine Arts (knowledge)	-0.156
Installation (skill)	-0.146
Food Production (knowledge)	-0.136
Clerical (knowledge)	+0.175
Telecommunications (knowledge)	+0.163
Wrist-Finger Speed (ability)	+0.158
Number Facility (ability)	+0.126
Building and Construction (knowledge)	+0.104

Summariser

Summariser: Three most and least automatable tasks

	Task	Automation Score
Lowest	Write notes on paper	2.235
	Cleaning up information in the patients electronic health record if there are any problems or incorrect or incomplete information	2.783
	Medical Coding of letters and other documents	3.220
Highest	Answer phone	3.362
	Send messages in electronic health record to communicate to other staff	3.304
	Register new Patients	3.235

Summariser: Largest feature differences relative to the population

O*NET Feature	Feature Difference
Medicine and Dentistry (knowledge)	+19.0 %
Clerical (knowledge)	+15.6 %
Communications and Media (knowledge)	+10.6 %
Telecommunications (knowledge)	+7.5 %
Computers and Electronics (knowledge)	+7.4 %
Building and Construction (knowledge)	−27.8 %
Spatial Orientation (ability)	−22.0 %
Mechanical (knowledge)	−20.9 %
Design (knowledge)	−18.9 %
Engineering and Technology (knowledge)	−18.7 %

Summariser: O*NET features with the largest (positive and negative) derivatives.

O*NET Feature	Feature Gradient
Programming (skill)	-0.172
Installation (skill)	-0.162
Technology Design (skill)	-0.157
Fine Arts (knowledge)	-0.137
Dynamic Flexibility (ability)	-0.114
Telecommunications (knowledge)	+0.167
Clerical (knowledge)	+0.166
Wrist-Finger Speed (ability)	+0.153
Number Facility (ability)	+0.118
Mathematics_x (skill)	+0.093

11.4 Appendix D: NHS Digital Employment Figures

Table 22: NHS Digital Employment Figures (December 2017)

Occupation	**Employment (Dec 2017)**
Administrator	64,565
Deputy Practice Manager	9,585
General Practitioner	3,3947
Healthcare Assistant	6,580
Nurse Practitioner	4,670
Pharmacy Technician	2,418
Phlebotomist	740
Practice Manager	9,585
Practice Nurse	18,307
Practice Pharmacist	658
Prescription Clerk	13,486
Receptionist	34,213
Scanning Clerk	13,486
Secretary	5,777
Summariser	13,486