BMJ Open

Qualitative and quantitative approach to assess the potential for automating administrative tasks in general practice

Matthew Willis ,[1,2] Paul Duckworth,[3] Angela Coulter,[4] Eric T Meyer,[1,5] Michael Osborne[3]

To cite: Willis M, Duckworth P, Coulter A, *et al*. Qualitative and quantitative approach to assess the potential for automating administrative tasks in general practice. *BMJ Open* 2020;**10**:e032412. doi:10.1136/bmjopen-2019-032412

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2019-032412).

Received 19 June 2019
Revised 03 March 2020
Accepted 21 April 2020

Check for updates

© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

[1]Oxford Internet Institute, University of Oxford, Oxford, UK
[2]School of Information, University of Michigan, Ann Arbor, Michigan, USA
[3]Department of Engineering Science, University of Oxford, Oxford, UK
[4]Nuffield Department of Population Health, University of Oxford, Oxford, UK
[5]School of Information, University of Texas at Austin, Austin, Texas, USA

Correspondence to
Dr Matthew Willis;
matthew.willis@oii.ox.ac.uk

ABSTRACT

Objective To identify the extent to which administrative tasks carried out by primary care staff in general practice could be automated.

Design A mixed-method design including ethnographic case studies, focus groups, interviews and an online survey of automation experts.

Setting Three urban and three rural general practice health centres in England selected for differences in list size and organisational characteristics.

Participants Observation and interviews with 65 primary care staff in the following job roles: administrator, manager, general practitioner, healthcare assistant, nurse practitioner, pharmacy technician, phlebotomist, practice nurse, pharmacist, prescription clerk, receptionist, scanning clerk, secretary and medical summariser; together with a survey of 156 experts in automation technologies.

Methods 330 hours of ethnographic observation and documentation of administrative tasks carried out by staff in each of the above job roles, followed by coding and classification; semistructured interviews with 10 general practitioners and 6 staff focus groups. The online survey of machine learning, artificial intelligence and robotics experts was analysed using an ordinal Gaussian process prediction model to estimate the automatability of the observed tasks.

Results The model predicted that roughly 44% of administrative tasks carried out by staff in general practice are 'mostly' or 'completely' automatable using currently available technology. Discussions with practice staff underlined the need for a cautious approach to implementation.

Conclusions There is considerable potential to extend the use of automation in primary care, but this will require careful implementation and ongoing evaluation.

Strengths and limitations of this study

- ► This is currently the first use of a scalable measure of task-based automation applied to the work of healthcare grounded in empirical data from National Health Service England primary care.
- ► Using and Independent Bayesian Classifier Combination to achieve a single rating from multiple experts is advantageous both because it is fully Bayesian and reflects a higher chance of accurately recovering the true automatability label of a task in an environment of uncertainty and subjectivity.
- ► This study was only able to analyse administrative tasks; our use of the Occupational Information Network database in large part restricted us to looking at these administrative functions.
- ► Primary care looks different across all parts of the country, and unfortunately, we were unable to gather data on some of these nuances and cannot be certain that we captured all the diversity.
- ► The subjectiveness in the task simplification may have translated into our dataset providing very few entirely 'not automatable' tasks caused, in part, by the creativity of the responder's who would believe part of a rearranged task to be automatable.

INTRODUCTION

Primary care in England faces numerous challenges, including increased demand, workload pressures, recruitment difficulties and budgetary constraints.[1–4] All staff in general practice, including general practitioners (GPs), nurses, pharmacists and non-clinical staff, carry out many administrative tasks. There has been much discussion of the bureaucratic burden on general practice and how to streamline these tasks to free up more time for direct patient care. A report in 2015 identified several potential areas of concern, including time spent on reclaiming financial reimbursement for services rendered, entering data into various unintegrated practice-based information systems, processing information from hospitals and other external providers, keeping on top of changing requirements of commissioners and helping patients navigate a fragmented and poorly coordinated system.[5] The authors concluded that over 50% of practice time was spent on bureaucracy and much of this was potentially avoidable. A majority of GPs nowadays describe their workload as excessive and detrimental to the delivery of high-quality care.[6–8] This phenomenon of 'paperwork' growth is consistent with the general upward

BMJ Open: first published as 10.1136/bmjopen-2019-032412 on 8 June 2020. Downloaded from http://bmjopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

trend in bureaucratic work in public sector occupations.[9] In general practice in England, administrative, non-clinical staff now outnumber clinical staff by a factor of 2.5 to 1.[10] The Royal College of General Practitioners has recently called for significant investment in digital technology to transform care, including automation of routine administrative tasks.[11]

Automation of administrative tasks may reduce the bureaucratic burden in primary care, but the nature and extent of its likely impact is currently unknown. In other sectors, automation is typically perceived as a threat to workers, conjuring visions of mass unemployment. Previous research by members of our team has classified 47% of occupations as being 'high risk' for computer automation within the next couple of decades.[12] That study based its estimates of the probability of automatability on the US Department of Labor's Occupational Information Network (O*NET) database of occupational characteristics and worker requirements.[13] O*NET is a large database that has been in development since 1997 and contains extensive information on a given job and the attributes of the worker, such as: personal requirements (skills and knowledge), personal characteristics (abilities and values), experience requirements (training, licencing and experience), requirements of the job (physical, social and organisational context of the work), occupation-specific information (title, tasks, technology skills and tools) and the outlook of the occupation in the labour market. O*NET enables the understanding of what is expected from a worker and the detailed tasks for which the worker is responsible. Later studies applied the same probability estimates to employment data from other countries, on the assumption that the probabilities are consistent globally. The proportion of jobs estimated as 'high risk' ranged from 45% to more than 60% across different European countries.[14–16] However, where there are workforce shortages as in general practice, automation presents not a threat but an opportunity, potentially increasing productivity and leading to more satisfying job roles. The assumption has been that healthcare occupations are less likely to be affected because they rely heavily on interpersonal interactions, but they also include tasks that require information acquisition, information analysis and decision making, all characteristics amenable to automation.[17] In primary care, automatable tasks include checking written documents for errors, reviewing and writing letters, creating referral letters, organising staff rotas, creating reports and maintaining records. Technologies such as natural language processing, voice recognition, text summarisation, robotic process automation and machine learning to support the manipulation of information are all potentially applicable.[18 19] It is unlikely that implementation of automation technologies in general practice will cause large-scale unemployment of healthcare staff, but it could enhance their capacity to keep on top of administrative work while freeing up more time for patient care.

Our aim in carrying out the present study was to quantify the likely impact of automation in general practice to inform the assessment of future resource needs. This was not a future-gazing project. We were interested in the potential for automation using currently available technologies, rather than those that might become available at some point in the future.

METHODS

This study deployed mixed qualitative and quantitative data collection and analysis. To help us make this assessment, author MW carried out 330 hours of ethnographic observations, interviews and focus groups in six general practices, conducted a survey of experts on automation technologies and used a predictive machine learning model to apply their assessments to primary care tasks, grouped according to the O*NET classification system. We present an overview of this process in figure 1. We used a similar quantitative framework to that developed by Frey and Osborne[12] where the authors estimated the probability of an occupation being affected by automation by classifying tasks according to the skills required to carry them out, such as dexterity, social perceptiveness, creativity, persuasion, negotiation and originality.

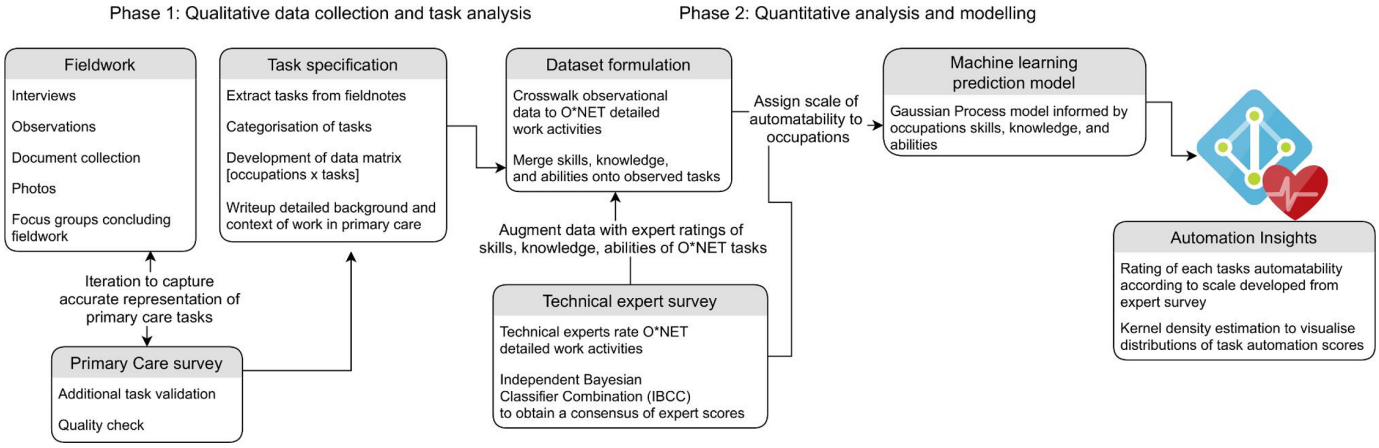

Figure 1 Overview of research method. O*NET, Occupational Information Network.

BMJ Open: first published as 10.1136/bmjopen-2019-032412 on 8 June 2020. Downloaded from http://bmjopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

However, our approach differs in two important ways. We have assumed that automation in primary care may not impact entire occupations, but specific tasks that are commonly carried out by a variety of occupational roles. Second, and most important, we have augmented our analysis with ethnographic observations in primary care to ensure that the complexity of primary care tasks is accurately captured.

Following Frey and Osborne's methods, we used the O*NET classification of occupational tasks to assess automatability. O*NET contains over 2000 detailed work activities (DWAs) for more than 1000 occupations across all sectors of the economy and nearly 20 000 individual occupation-specific tasks arranged in a hierarchical structure. DWAs are specific descriptions of tasks performed by different occupations. For example, some of the DWAs include: answer telephones to direct calls or provide information; diagnose medical conditions; document client health progress; and interact with patients to build rapport or provide emotional support. The O*NET system provides a useful basis for an analysis of automatability, but it had not been validated for use with British healthcare-specific occupation and task data. Our first task, therefore, was to carry out observation of primary care tasks to enable linkage to the O*NET system. The focus on task data is key, because it is individual tasks rather than entire occupations that are typically automated.[20–24] Recent research suggests that occupations are best analysed as evolving combinations of detailed tasks, skills and/or environments.[23 25–27] With more granular job data available than ever before, thousands of occupations can each be broken down into hundreds of numeric components or tasks, relating to the skills, knowledge and abilities required to perform them. Examples given in the appendix (online supplementary appendix images 1-3) show how O*NET classifies selected occupational tasks carried out by GPs.

Data collection and analysis

All fieldwork involving interactions with staff was undertaken by author MW, an experienced qualitative researcher. MW contacted each primary care centre using a telephone and email to explain the study, detail what is required from centres that participate in the research and answer any questions. These conversations involved the practice manager and at least one partner GP from the practice to agree on participating in the study. Practices were recruited to assure geographical diversity, variance in size, representation from multisite super practices as well as individual single site practices, and willingness to participate in the research project.

The focus of the fieldwork was to observe the work practices of each type of staff member (listed below) and to understand the typical scope of tasks they are responsible for and how those tasks are usually completed. When fieldwork concluded, there were a total of 138 unique tasks that describe work carried out by each of the following staff roles: administrator, deputy practice manager, GP, healthcare assistant, nurse practitioner, pharmacy technician, phlebotomist, practice manager, practice nurse, practice pharmacist, prescription clerk, receptionist, scanning clerk, secretary, and medical summariser.

Data collected during the fieldwork consisted of field notes, photographs and documents that elucidate tasks being observed and how different tasks are carried out at each practice. During observation of each occupational type, semistructured interviews were also conducted when appropriate (ie, the participant was not working with patients at the time or otherwise performing a task that did not allow them to respond to a question). The field researcher also attended staff coffee breaks where additional semi structured interviews occurred.

A focus group was held on the final day at each site to gather staff feedback on the researcher's observations and to check that tasks observed were representative of their occupations. All fieldwork and focus group sessions were conducted within the practice; focus groups were held in an available staff room or conference room. Data saturation was deemed to have been reached after observations at the last practice produced no new tasks. Next, a brief survey describing 15 random tasks per occupation was sent to one of the practices for further validation to check that tasks were accurately described and that the task descriptions were accurate.

A spreadsheet was developed by author MW detailing each observed occupation and a list of the tasks they perform. These were coded and categorised according to the frequency of task, any tools or technologies used, type of software required, other occupations in the practice that also perform the task, if the task is shared with other staff members, and any other features. Important themes were elicited and refined after a continuous process of conversation and collaborative sensemaking between the authors. All observed tasks were then matched to the equivalent task in the O*NET database. This involved reading each observed task and finding its equivalent representation of that task in O*NET, for which there was always an analogous DWA. Any differences in task representation were discussed by the research team, and the most accurate and representative task was selected. Crosswalking observed tasks to DWAs enabled us to use O*NETS rating of skills, knowledge and abilities required for each task and subsequence weighting of task importance per occupation. An example of the contents of O*NET using the GP occupation is provided in online supplementary appendix one. This shows the task list, top skills, knowledge and abilities as shown through the web portal interface.

The method conventionally used to identify and predict the likelihood of automation is to hypothesise the underlying dynamics that drive automation and extrapolate into the future an unspecified number of years.[12 25 28–30] This relies on subjective forecasts that may be unreliable. Instead, we decided to ask experts to base their estimates on *current* technology capabilities, rather than speculate on the future. To obtain current estimates of

BMJ Open: first published as 10.1136/bmjopen-2019-032412 on 8 June 2020. Downloaded from http://bmjopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

the automation potential of tasks, we contacted experts in the machine learning, robotics and artificial intelligence communities through academic mailing lists and lists acquired at two machine learning conferences (specifically 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 sequentially presented with five O*NET occupations and the five tasks with the largest task-importance score for that occupation. Task importance scores are rankings of their importance to the occupation that performs the task and are provided by O*NET. Participants were asked the following question: 'Do you believe that technology exists today that could automate these tasks?'

Each task was ranked on an ordinal scale from 1.0 to 4.0: not automatable today (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. Respondents were also asked to rate their overall confidence in their answers.

The multiple ratings or labels for each task were combined with Independent Bayesian Classifier Combination (IBCC), a principled Bayesian approach to combine multiple classifications.[31 32] IBCC creates a posterior prediction over labels that reflects the tendency of the individual labellers to concur with ultimately chosen label values. Crucially, it requires no forecasting or prediction by the expert participants. We believe it is the best method we have for reaching a consensus of experts' opinions. IBCC task scores were averaged into 313 work activities (note: an activity is a unique set of tasks in O*NET taxonomy). Therefore, activity automatability labels concentrate around whole and half values. The final values were rounded to the closest 0.5 (a half-class). A score of 4.0 indicates a 'fully automatable' work activity, while 1.0 indicates a work activity that 'cannot be automated' using currently available technology.[23]

Finally, all data sources were combined to form the basis of the analysis, including: (1) the validated list of primary care tasks obtained from the ethnographic observations, interviews, and focus groups which were qualitatively coded and classified allowing us to use; (2) the corresponding O*NET activity skills, knowledge and ability numerical attributes and (3) the inferred automatability scores of these activities informed by the expert survey. A machine learning framework was employed to infer a functional mapping between the skills, knowledge and ability characteristics of work activities and the ground truth automatability elicited from the expert survey. The anticipated output was a flexible function estimation capable of modelling complex, non-linear relationships between the features (skills, abilities and knowledge) and the target (automatability). Function mapping is routinely used to map the input (data you have) of a set into a different output (missing data) set. Just as we map skills, knowledge and abilities to the expert survey data, function mapping is also used to predict someone's weight based on their height or by using income to assess loan payments.

The collection of work activities (n=313) labelled in the expert survey of automatability was used as a training data set for the Gaussian Process Ordinal Regression model. Once trained, the model can estimate the automatability of 'unlabelled' work activities, that is, our observed healthcare tasks. This is preferable to asking experts to provide judgements on the full range of tasks observed, which would be time-consuming and prohibitively expensive. We used a Gaussian process model for two reasons. First, supervised learning problems in machine learning can be cast directly into the Gaussian process framework. Compared with traditional machine learning approaches, Gaussian processes are particularly well suited when large amounts of data—often referred to as big data—are not available or even possible, which is often required to prevent overfitting and to ensure a good generalisability of models to new data. Second, being a non-parametric model, Gaussian processes automatically control model complexity and provide a predictive distribution over the target function, providing formal bounds on uncertainty for classification and regression tasks. A Gaussian process can be defined as a distribution over functions, mapping from our input features X of work activities to the probability of automatability.

A Kernel Density Estimation (KDE) analysis was employed to show the distribution of the inferred automatability of the healthcare tasks. KDE allows visualisation of a smooth curve that represents the 'shape' of the data distribution as a replacement for a discrete histogram. Further technical details of the quantitative framework and analysis, as well as a detailed description of the observational methods, are documented in a previous publication.[33]

Patient and public involvement

No patients were involved in this study.

RESULTS

Administrative tasks

Of the 19 practices contacted, 6 agreed to participate, 6 expressed interest and intended to participate but never followed up and 7 declined to participate due to lack of time or financial burden. Ethnographic observations were carried out in six primary care centres in England, located in Oxfordshire, Yorkshire, Berkshire, Surrey and the West Midlands. The practice list sizes ranged from 5000 to 24 000 patients. The researcher spent 1 week in each location, observing work carried out by GPs, other primary care clinicians and non-clinical staff. Three hundred and thirty hours of observations resulted in descriptions of 65 unique sessions of work carried out by primary care staff, each ranging from 45 min to 90 min. These observations can be best illustrated by focusing on two large and significant types of work that account

BMJ Open: first published as 10.1136/bmjopen-2019-032412 on 8 June 2020. Downloaded from http://bmjopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

for many of the administrative tasks carried out—letter work and clinical documentation. It is important to note that although we classified documentation as administrative work, we appreciate the medical importance of the tasks and that they involve several subtasks, some of which require specialist medical knowledge.

Letters

This type of workload is the bundle of tasks and workflows that relate to the production of letters. Letter production flows in two directions: incoming and outgoing. Incoming letters contain multiple forms of implicit and explicit information. Processing this information results in tasks such as responding to requests, storing physical and digital copies of the letter, applying a medical coding scheme and using information contained in the letter to inform or modify patient treatment and/or other clinical decisions. Outgoing letters involved multiple processes as well, such as audio dictation, transcription of audio text, application of preformatted templates, appending directions and notes to audio files or digital documents, referencing information in the electronic medical record, archiving letters and posting the letter. The specific details of how this work is carried out varies from practice to practice. The workflow can look remarkably different in a single-site independent practice when compared with the workflow in a super-practice, sharing services across multiple sites. A single-site practice may allocate this work to receptionists who process letters, while a super-practice might outsource all medical coding, audio transcription and typing to a central location.

The following workflow was observed at a small single site practice. Receptionists sorted incoming letters into groups, giving them red, amber or green status dependent on response urgency. Then, they scanned the letters into the system and used a commercial software system called DocMan to sort them using the colour-coded system. One of the receptionists is assigned the task of 'dealing' with the letters. This involves opening a letter that has been briefly coded by a GP with the highlighter tool in the document management software, reviewing the highlighted medical terms or phrases and allocating appropriate Read codes. The letter is then archived on the system, the task is marked complete and the receptionist continues to the next letter. If the receptionist has a question about the appropriate Read code, text comments are added to the document, which is then saved, and assigned as a task on the system, that is, sent back to the GP. Other types of letters, for example, from a medical supplier requesting information about a patient who requires a specific device, must find the required information in the relevant health record, then use a template letter to respond to the supplier.

The length of letters ranges from precise and short to long and verbose. Letters have a different utility dependent on who is writing them, who is reading them and the context. For example, during our research, a GP was observed reading a 9-page letter from secondary care. The GP skim-read the letter on her monitor and after a few minutes declared, '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.' She continues: '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. Her prior experience of the specialist who wrote the letter enabled her to extract the relevant information efficiently and discount the rest.

Documentation of clinical consultations

A scenario in which a system could automate clinical documentation using natural language processing and text summarisation was described to staff in focus groups and interviews, provoking lengthy discussion. Some participants saw it as a potentially useful technology, with multiple 'when can we have that' comments, while others felt it might have negative consequences on the clinical encounter and could even undermine the skills and critical thinking of the clinician. One GP argued that having a transcript of the entire conversation would be a waste of time, producing a great deal of unnecessary information. It is not uncommon for GPs to rewrite or edit their notes, particularly notes for new cases or complex circumstances, or when the notes are 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 reflect critically about an interaction. Research shows that writing engages the brain, allows GPs to be better observers, supports empathy and engages critical thinking.[34]

Scope for automation

The online survey of experts in automation systems achieved 156 responses from academic and industrial experts around the world, providing a total of 4599 task level ratings. The distribution of respondents by country and qualification is shown in figures 2 and 3.

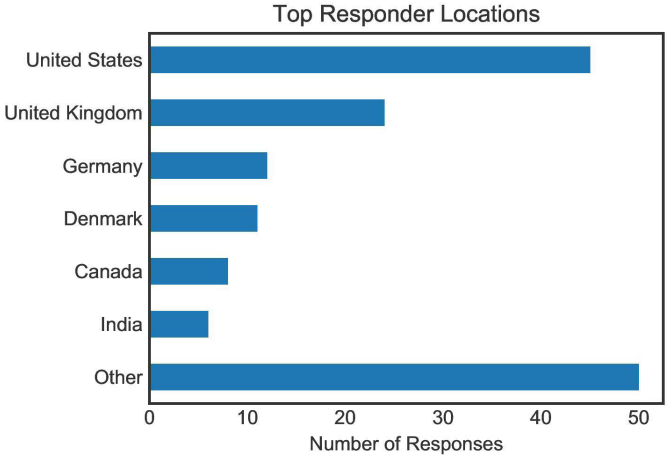

Figure 2 Number of survey respondents by country.

BMJ Open: first published as 10.1136/bmjopen-2019-032412 on 8 June 2020. Downloaded from http://bmjopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

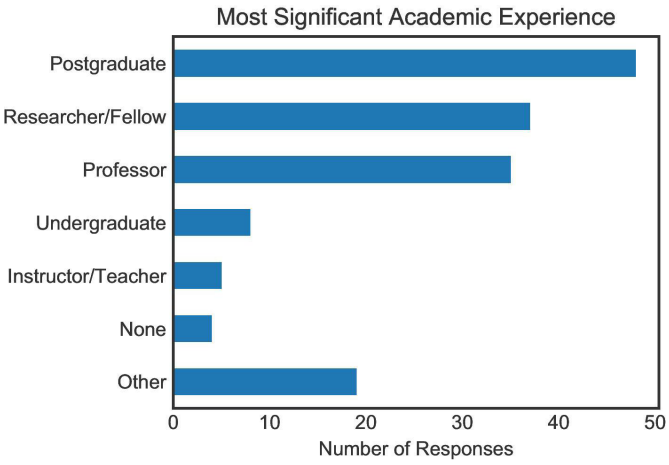

Figure 3 Number of survey respondents by qualification.

The Gaussian Process machine learning model that we applied to the data suggests that about 44% of the administrative tasks performed in primary care are 'mostly' or 'completely' automatable using current technology. Figure 4 shows the distribution of automatability for all administrative tasks performed by general practice staff.

As discussed above, tasks performed by various occupations overlap, but the automatability of those tasks are independent of the occupation; that is, an automatability score is not dependent on who performs the task. For example, medical coding has an estimated automatability of 3.2. Thus, automation of this task, which is mostly performed by receptionists, summarisers and secretaries, but also on occasion by GPs depending on the practice, could either remove this task from their workloads altogether, if the systems are considered sufficiently reliable, or reduce it to checking codes produced by an automated system. Some of the most highly automatable tasks in our data include: payroll and managing finances, checking and sorting post, printing letters, communicating with patients through texting, management of paper archives (onsite or offsite), transcription, email account management, letter scanning, checking for errors in paperwork and internal communications (eg, messages to staff or new employee inductions).

To further understand the automatability predictions, three categories were generated: automatable (predicted score is >3.0), not automatable (predicted score <2.0) and partly automatable (predicted score ≥2.0 to ≤3.0). Figure 5 presents the predicted automatability scores of administrative occupation tasks classified into these risk categories.

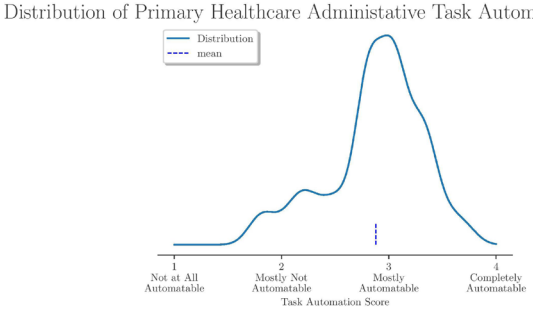

Figure 4 Distribution of automatability scores. Plot shows the kernel density estimate of the scores.

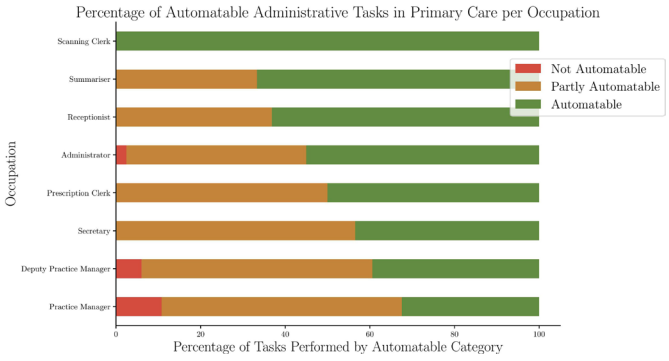

Figure 5 Percentage of administrative tasks performed in primary care by automatability categories.

Scanning clerks are typically, but not always, part-time employees that are responsible for scanning physical documents (mostly letters) and creating digital objects (usually a PDF file) and attach these files to the patients' health record or import files into other document management systems. Scanning clerks also add metadata to the PDF so the files are searchable and contain all contextual information. Results show that only the work of scanning clerks could be fully automated using currently available technology, thus eliminating the need for that occupation entirely. Many routine administrative tasks and typical office work (eg, answering phones, writing letters and managing email) are undertaken by all occupational groups, but the impact of automation will vary according to the number of tasks in each automatable category that they perform. Figures 6 and 7 below show the potential impact of automation technologies by occupation for clinical and non-clinical staff.

Further detailed results can be found in the technical report for The Health Foundation, currently in preparation.[35]

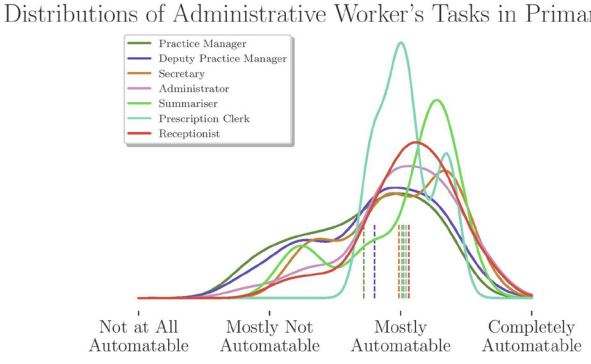

Figure 6 (KDE) Distribution of task automatability for primary care occupations (non-clinical staff). KDE, Kernel Density Estimation.

BMJ Open: first published as 10.1136/bmjopen-2019-032412 on 8 June 2020. Downloaded from http://bmjopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

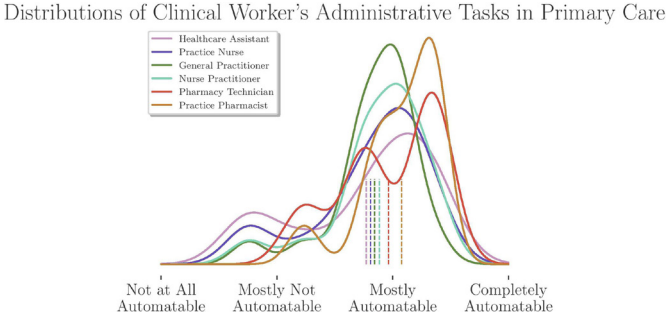

Figure 7 (KDE) Distribution of task automatability for primary care occupations (clinical staff). KDE, Kernel Density Estimation.

DISCUSSION

We believe this study is the first to combine detailed validated observations of workflows and a specially developed ordinal scale of automatability to assess the potential for automating various administrative tasks routinely carried out by primary care staff. Our findings indicate considerable scope for extending the use of automation in general practice. It is important to acknowledge that certain capabilities of automation already exist in primary care. In some practices, scanning clerks put paper documents into scanners that turn physical text into digital type using optical character recognition. Another form of automation is the use of templates and pregenerated text in desktop word processors or in electronic medical record notes. Automated report generation and custom search queries to input patient data to different systems is another form of automation that is sometimes used. Our analysis has looked at the scope for further automation of tasks, and we conclude that it is considerable.

A few caveats must be borne in mind. Observations took place in only six general practices. These varied in terms of list size and geographical location, but we cannot claim that these were representative of all practices in England. The survey of automation experts was, we believe, the largest of its kind to date and it provides a robust baseline estimate of the extent to which tasks are automatable using only currently available technology, but participants did not have specific expertise in the application of automated systems to healthcare. We made up for this deficiency by inviting discussion and feedback from primary care staff involved in interviews and focus groups, but they, in turn, lacked direct experience of using the automated systems that might eventually modify their working practices.

Automation technologies will be most useful when they provide clinicians with the right information at the right time; employing filters of context and urgency to remove details that are irrelevant or distracting. Automation of some mundane tasks, such as telephone answering, letter writing, document scanning, email monitoring and information filtering, could free up valuable time for patient-facing tasks. We find that many tasks involving data entry, sending information, formatting information, maintaining records, transcription and operating office equipment to name a few, have high potential for automation. Many practices have basic forms of automated telephone answering, but more sophisticated systems are now available that would enable a voice assistant to help triage the hundreds of phone calls that a practice can receive in a day. However, there are drawbacks to automated phone systems that would need careful consideration and evaluation. Receptionists help prevent unneeded appointments by screening phone calls and getting patients to the right area for their appointment; these skills rely on social perception and an understanding of the patient's history. Automated phone systems must be aware of this function and help guide patients through their case without scheduling unneeded appointments.

Clinical documentation could benefit from automation for 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 as possible; that is, it cannot be left to the end of the day. Second, documentation requires attention to detail and is time intensive. Even very brief encounters require accurate documentation. Third, this work is typically squeezed into a busy schedule, usually performed during the time it takes a patient to leave the examination room and the next patient arriving. Vendors of electronic health record systems have advocated the use of automation to facilitate record-keeping by use of structured data formats.[36–38] Natural language processing can be used to dissect patient–doctor conversations and create notes.[39] How to ensure the validity and reliability of such systems has been the subject of much research and remains a controversial topic.[18 40] However, a recent survey of UK GPs revealed considerable optimism about the likelihood that clinical documentation would be fully automated in the near future, with a large majority believing this would occur within 10 years.[41]

Documentation of clinical consultations is a good example of why automation should be approached with caution. Automating this type of work could have unintended consequences. Documentation of clinical encounters serves as a 'tool for thinking' for doctors and as an important means of coordination and communication.[40] Reading and writing clinical notes engage the medical professionals' critical faculties in a useful way. Allowing machines to take over would circumvent the way clinicians have been trained to document clinical encounters.[42] It would be important to identify the skills medical professionals would need in order to successfully work with automation technologies while maintaining a high level of care quality.

Automation technologies can be expensive and time-consuming to implement, so those contemplating the purchase of such systems should consider the following issues: which types of tasks would be most beneficial to the practice if automated? Are those tasks amenable to automation? If this results in a change in occupational roles, who will it affect most? How will staff be trained and supported to use the new technology? If tasks are

BMJ Open: first published as 10.1136/bmjopen-2019-032412 on 8 June 2020. Downloaded from http://bmjopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

removed from the workload of a staff member, what kind of work (if any) will replace the task? And who will be responsible for the maintenance of the tasks while new workflows are in development?

Some tasks with high automatability scores will be challenging to implement for social and organisational reasons. Previous research has shown that staff contribute to patient quality and safety through hidden ways that emerge in the moment and require tacit knowledge.[43] The tacit and contextual work, such as reading social cues, currently lies outside the capabilities of automated systems and machine learning technologies.[29] Healthcare is crucially dependent on these human skills, and patient care would suffer if automation is pursued to their detriment. Our observations revealed a type of work that is often driven by exceptions, interruptions, competing critical requests and time sensitivity. Work practices that may appear routine, inefficient or superfluous often have an intrinsic value for cooperative teamwork.[44] While our quantitative findings reveal considerable potential for automation, our in-depth fieldwork demonstrated the need for a cautious approach given the complexity of social systems in healthcare delivery environments.

Observing and gathering data on tasks in primary care presents at least three major challenges: task variance, task fit and clinical or administrative task categorisation.

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.

There are further complexities that emerge due to the task being performed in a healthcare context, when attention is given to patients' needs, backgrounds and particular problems. In this context, many administrative tasks require specialist medical knowledge. This is what makes work in healthcare different and exceptional, when compared with other fields with similar task descriptions.

The second challenge is that the same occupation does not always perform the same set of tasks across practices. We found that some tasks that receptionists routinely perform in one practice were performed by administrators at another practice. The practices we observed 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', that is, allocating tasks to the appropriate person, involves both matching the task to the most appropriate pay grade and making sure clinicians are shielded from unnecessary administrative work. The driving force behind the assigning of tasks to occupations in this study was to accurately represent the core identity and 'scope of work' of each observed occupation.

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 categorised nearly every clinical interaction that takes place in the domain of primary care, it was often hard to distinguish administrative tasks from those that are purely clinical. Also, the O*NET database that was used for our analysis provided more terminology and language to describe administrative tasks and office work than it did for nuanced clinical work.

CONCLUSION

Evidence from our study show there is great potential to impact workloads in primary care by automating certain administrative tasks (approximately 44%, broadly). We believe that careful introduction of currently available technologies has considerable potential to reform administrative workflows and increase productivity in primary care. However, this should not be seen as a magic bullet that will solve the workload crisis in the immediate future. Introduction of new forms of automation should proceed cautiously with ongoing evaluation of their wider impacts.

Correction notice This article has been corrected since it was published. The article title has been updated.

Twitter Angela Coulter @acpatient

Contributors MW: writing, figure 1, coordination of authors, literature search, citation management, study design, formatting, draft iteration, data collection, data interpretation, manuscript submission and revisions. PD: figures, draft writing, editing, data collection, data analysis, literature search, data interpretation and revisions. AC and ETM: reviewing and editing, writing, data interpretation, study design and revisions. MO: data analysis, data interpretation, technical advising, figures, data collection and study design.

Funding This work was supported by The Health Foundation, Award #7559.

Competing interests All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: all authors had financial support from The Health Foundation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Ethics approval This study received a favourable opinion from the North East Tyne & Wear South Research Ethics Committee and approval from the Health Research Authority under project number 212 952.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available. All data relevant to the study are included in the article.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD

Matthew Willis http://orcid.org/0000-0002-9120-7319

BMJ Open: first published as 10.1136/bmjopen-2019-032412 on 8 June 2020. Downloaded from http://bmjopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

REFERENCES

1 Hopson C. The state of the NHS provider sector, 2016. Available: https://www.nhsproviders.org/news-blogs/news/nhs-trust-leaders-warn-staff-shortages-now-outweigh-fears-over-funding
2 Baird B, Charles A, Honeyman M, *et al*. Understanding pressures in general practice, 2016. Available: https://www.kingsfund.org.uk/publications/pressures-in-general-practice
3 Roland M, Everington S. Tackling the crisis in general practice. *BMJ* 2016;352:i942.
4 NHS Engl. General practice forward view, 2016. Available: https://www.england.nhs.uk/gp/gpfv/about/
5 Clay H, Stern R. *Making time in general practice*. Primary Care Foundation, 2015: 1–83.
6 Byrne L, Bottomley J, Turk A. *British Medical association survey of GPs in England*. London, 2016.
7 Gibson J, Checkland K, Coleman A, *et al*. Eighth national GP Worklife survey, 2016. Available: http://research.bmh.manchester.ac.uk/healtheconomics/research/Reports/EighthNationalGPWorklifeSurveyreport/
8 Hobbs FDR, Bankhead C, Mukhtar T, *et al*. Clinical workload in UK primary care: a retrospective analysis of 100 million consultations in England, 2007-14. *Lancet* 2016;387:2323–30.
9 Hitchcock ALK, Sundorph E. *Work in progress: towards a leaner, smarter public-sector workforce*, 2017.
10 statistics NHS. Facts and figures. London, UK, 2017. Available: https://www.nhsconfed.org/resources/key-statistics-on-the-nhs
11 Royal College of General Practitioners. Fit for the future: a vision for general practice. London, UK, 2019. Available: https://www.rcgp.org.uk/policy/future-vision.aspx
12 Frey CB, Osborne MA. The future of employment: how susceptible are jobs to computerisation? *Technol Forecast Soc Change* 2017;114:254–80.
13 National Center for O*NET Development, 2016. Available: https://www.doleta.gov/programs/onet/eta_default.cfm
14 Bowles J. The computerisation of European jobs. Bruegel., 2014. Available: http://bruegel.org/2014/07/the-computerisation-of-european-jobs/
15 Pajarinen M, Rouvinen P. Computerization threatens one third of Finnish employment, 2014. Available: https://econpapers.repec.org/paper/rifbriefs/22.htm [Accessed 27 May 2019].
16 Knowles-Cutler A, Frey C, Osborne M. Agile Town: The Relentless March of Technology and London's Response, 2014. Available: https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/uk-futures/london-futures-agiletown.pdf
17 Jamieson GA, Skraaning G. Levels of automation in human factors models for automation design: why we might consider throwing the baby out with the bathwater. *J Cogn Eng Decis Mak* 2018;12:42–9.
18 Kreimeyer K, Foster M, Pandey A, *et al*. Natural language processing systems for capturing and standardizing unstructured clinical information: a systematic review. *J Biomed Inform* 2017;73:14–29.
19 Ng A. What artificial intelligence can and Cant do right now. Available: https://hbr.org/2016/11/%0Awhat-artificial-intelligence-can-and-cant-do-right-now
20 Bresnahan T, Brynjolfsson E, Hitt L. *Information technology, workplace organization and the demand for skilled labor: Firm-Level evidence*. Cambridge, MA, 1999.
21 Brynjolfsson E, McAfee A. *Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy*, 2012.
22 Brynjolfsson E, McAfee A. *The second machine age : work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company, 2014.
23 Duckworth P, Graham L, Osborne M. Inferring work task Automatability from AI expert evidence. *ACM conference on Artificial Intelligence Ethics, and Society*, Honolulu, Hawaii, 2019.
24 Susskind RE, Susskind D. *The future of the professions: how technology will transform the work of human experts*. OUP Oxford, 2017.
25 Acemoglu D, Robots RP. *Robots and jobs: evidence from US labor markets*. Massachusetts Ave, Cambridge: National Bureau of Economic Research, 2017.
26 Arntz M, Gregory T, Zierahn U. Revisiting the risk of automation. *Econ Lett* 2017;159:157–60.
27 Manyika J, Chui M, Miremadi M. *Harnessing automation for a future that works*. McKinsey Glob, 2017. http://www.mckinsey.com/global-themes/digital-disruption/harnessing-automation-for-a-future-that-works
28 Autor DH, Handel MJ. Putting tasks to the test: human capital, job tasks, and wages. *J Labor Econ* 2013;31:S59–96.
29 Bakhshi H, Downing JM, Osborne MA, *et al*. *The future of skills: employment in 2030*, 2017.
30 Goos M, Manning A, Salomons A. Explaining job polarization: Routine-Biased technological change and Offshoring. *American Economic Review* 2014;104:2509–26.
31 Kim H-C, Ghahramani Z. Bayesian Classifier Combination. In: *Proceedings of the fifteenth International Conference on artificial intelligence and statistics*. PMLR, 2012: 619–27. http://proceedings.mlr.press/v22/kim12.html
32 Simpson E, Roberts S, Psorakis I. *Dynamic Bayesian combination of multiple imperfect classifiers*, 2012.
33 Willis M, Duckworth P, Coulter A, *et al*. The future of health care: protocol for measuring the potential of task automation Grounded in the National health service primary care system. *JMIR Res Protoc* 2019;8:e11232.
34 Wald HS, Borkan JM, Taylor JS, *et al*. Fostering and evaluating reflective capacity in medical education: developing the reflect rubric for assessing reflective writing. *Acad Med* 2012;87:41–50.
35 Willis M, Duckworth P, Graham L, *et al*. *Future of healthcare: computerisation, automation, and general practice services*. London, UK, 2019.
36 Rule A, Rick S, Chiu M, *et al*. Validating free-text order entry for a note-centric EHR. *AMIA Annu Symp Proc* 2015;2015:1103–10.
37 Lin SY, Shanafelt TD, Asch SM. Reimagining clinical documentation with artificial intelligence. *Mayo Clin Proc* 2018;93:563–5.
38 Verghese A, Shah NH, Harrington RA. What this computer needs is a physician: Humanism and artificial intelligence. *JAMA* 2018;319:19.
39 Klann JG, Szolovits P. An intelligent listening framework for capturing encounter notes from a doctor-patient dialog. *BMC Med Inform Decis Mak* 2009;9 Suppl 1:S3.
40 Bansler JP, Havn EC, Schmidt K, *et al*. Cooperative Epistemic Work in Medical Practice: An Analysis of Physicians' Clinical Notes. *Comput Supported Coop Work* 2016;25:503–46.
41 Blease C, Bernstein MH, Gaab J, *et al*. Computerization and the future of primary care: a survey of general practitioners in the UK. *PLoS One* 2018;13:e0207418–12.
42 Willis M, Jarrahi MH. *Automating documentation: a critical perspective into the role of artificial intelligence in clinical documentation*. Washington, D.C: Springer, 2019.
43 Swinglehurst D, Greenhalgh T, Russell J, *et al*. Receptionist input to quality and safety in repeat prescribing in UK general practice: ethnographic case study. *BMJ* 2011;343:d6788.
44 Cabitza F, Ellingsen G, Locoro A, *et al*. Repetita luvant : Exploring and Supporting Redundancy in Hospital Practices. *Comput Support Coop Work* 2018.

BMJ Open: first published as 10.1136/bmjopen-2019-032412 on 8 June 2020. Downloaded from http://bmjopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.