# The Quantified Patient in the Doctor's Office: Challenges & Opportunities

**Peter West, Richard Giordano**
Faculty of Health Sciences
University of Southampton, UK
{p.west,r.giordano}@soton.ac.uk

**Max Van Kleek, Nigel Shadbolt**
Department of Computer Science
University of Oxford, UK
{max.van.kleek,nigel.shadbolt}@cs.ox.ac.uk

## ABSTRACT

While the *Quantified Self* and *personal informatics* fields have focused on the individual's use of self-logged data about themselves, the same kinds of data could, in theory, be used to improve diagnosis and care planning. In this paper, we seek to understand both the opportunities and bottlenecks in the use of self-logged data for differential diagnosis and care planning during patient visits to both primary and secondary care. We first conducted a literature review to identify potential factors influencing the use of self-logged data in clinical settings. This informed the design of our experiment, in which we applied a vignette-based role-play approach with general practitioners and hospital specialists in the US and UK, to elicit reflections on and insights about using patient self-logged data. Our analysis reveals multiple opportunities for the use of self-logged data in the differential diagnosis workflow, identifying capture, representational, and interpretational challenges that are potentially preventing self-logged data from being effectively interpreted and applied by clinicians to derive a patient's prognosis and plan of care.

## Author Keywords

Quantified self; clinical decision making; self-tracking

## ACM Classification Keywords

H.5.2 Information interfaces and presentation (e.g., HCI): User-centered design.; J.3 Life and medical sciences: Health

## INTRODUCTION

Empowering patients to "take charge" of their health is an idea frequently championed by politicians [22, 16], technologists [27], journalists [19] and healthcare experts alike [37]. Yet, despite both government and industry-led initiatives across both Europe and North America to encourage this "patient-led healthcare revolution," widespread adoption has been slow [41].

One area, however, where individuals have been taking the lead in understanding their own health is the *Quantified Self* movement. This primarily comprises of non-expert ordinary people who use technological tools to record and interrogate the minutiae of their physical and mental states over time [37]. As the population of those interested in self-logging has grown, industry has responded with a vast collection of wearable and embeddable sensors which enable people to keep an accurate record of their health with low effort and high fidelity. For these reasons, it has been proposed that the Quantified Self movement can contribute to health care, helping clinicians diagnose and treat illnesses [36]. However, some clinicians outright reject the use of self-logged data, citing concerns about data quality, time constraints, and insufficient resources [35].

What are the barriers to the use of self-logged data in critical clinical decision making settings? This is a delicate question to approach for several reasons: first, during the course of a single patient visit, there are many kinds of decisions made by a different clinicians in different roles in different settings. Paramedics in an ambulance, triage nurses within an emergency room, specialists in acute care units or hospital wards, to general practitioners (GPs) in their offices, all make decisions regarding patients under distinct situational and informational constraints [11]. Second, even if focus is centred around a single setting by a single class of medical professionals, such as GPs, there may be significant differences in day-to-day work practices between individuals. For example, there can be variation in the degree to which GPs use electronic medical records (EMRs) to organise patient data, how particular tests or treatments are prescribed, and the mechanisms that they use to maintain good patient relationships [34].

We focused on two clinical roles: primary care physicians on the "frontline" of the medical service, and secondary care specialists who work in hospitals. Filling a gap in empirical research, we sought to understand the nature of evaluation and use of 'self-logged' data voluntarily recorded by the patient without advice from clinicians, particularly of the most common types facilitated by consumer health monitoring tools. We wished to identify factors behind the underuse of self-logged data by clinical professionals, including *what* was captured, *how* they were captured, the *representation* made available during a patient consultation, and other, yet unidentified, issues. From this, we wished to extrapolate how such problems might be addressed through HCI research, such as by re-thinking tools people use to monitor themselves, the kinds of data they capture, or the ways that

physicians and medical professionals might access and use self-logged data.

Since these dimensions encompassed a broad set of possible factors, including data-oriented problems (pertaining to subject, quality and sampling), situational constraints, and practice constraints, we wished to understand which, if any, of these dimensions had support from previous studies. This led us to conduct a broad survey of medical literature, from which we identified a set of themes. This was followed by an empirical investigation from which we compared our findings to the identified themes and drew up a number of design implications for self-logging tools. Finally, we discussed the limitations of our study and future work.

## BACKGROUND

We contextualised our investigation against two closely related fields: the first, *evidence based medicine*, seeks to apply empirical methods (such as used in epidemiology) for evaluating and improving the effectiveness of clinical practice and clinical decisions. The second, *clinical decision making*, examines the cognitive, interactional, and situational processes which influence how practitioners arrive at decisions under the practical constraints necessary for conducting their practices. In this section, we introduce a high level view of how technology-enabled patient data-driven healthcare might look in the future and discuss the roles of evidence-based medicine and clinical decision making in our investigation.

### Visions for Data-Driven Healthcare

Notions of "big data" and "data driven healthcare" have inspired popular scenarios of data-informed healthcare designed for the individual or stratified for groups of individuals. In the popular press, Thomas Goetz's *The Decision Tree* outlined a vision in which every person will be DNA-tested at birth, and tracked with sensors throughout their lives [20]. The resulting data would be used to classify and compute optimal treatments and actions to support personalised medical treatments. Policy makers have set out national agendas towards such a goal. For example, the UK's Personalised Health and Care 2020 framework set out a vision in which health and well-being data, sensed from wearable and environmental sensors, would seamlessly integrate with patient health records by 2018 [25]. The framework proposes that these data would "fill in the gaps" between visits with their GP or specialist, enabling clinicians to perform more personalised differential diagnoses at point of care. It is anticipated that the introduction of such technologies will enable early onset detection of chronic conditions so that they can be controlled at their early stages, increase the quality of life for patients, reduce morbidity, and decrease national health-related costs [36]. Initial enthusiasm for this vision is also evident in the US, with the US Food and Drug Administration's approving consumer tracking devices for clinical trials, citing the importance of quantifiable analysis of physical activity to physiological monitoring [40].

This interest in better and more accurate physiological sensing has resulted in a number of new growing companies, such as Withings, Fitbit and Jawbone. Meanwhile, established sport device makers (such as Garmin), sport brands (such as Nike), and technology brands (such as Apple) have developed wearable tracking devices and mobile apps for tracking health. A few companies, such as iHealth and Withings, have had their consumer products certified for clinical use under the US FDA Class I and II medical device classifications [26].

### Evidence Based Medicine

The term *evidence-based medicine* refers to the use of epidemiological methods in both patient-level decision making and formulation of population-level clinical care guidelines. In the first use of the term, David M. Eddy pointed out that while outcome analysis may make it seem that there is a clear path towards achieving more accurate, or at least consistent, diagnoses, things are never that simple in practice:

> Decisions might be variable but they are not whimsical or flippant. The variability occurs because physicians must make decisions about phenomenally complex problems, under very difficult circumstances, with very little support. They are in the impossible position of not knowing the outcomes of different actions, but having to act anyway. [14]

Eddy, thus, argues that, due to the complexities of assessing individual patients' conditions under the constraints in place, care targets and practice recommendations alone will not achieve better health outcomes. Instead, working under extreme uncertainty is a necessity and should be embraced with the acknowledgement of the effects of various decision-making constraints and biases that are known to exist.

### Clinical Decision-Making

Modern texts for nursing and clinical evaluation have formalised the diagnostic process of determining a patient's disease or condition based on available evidence in the form of *differential diagnosis* [38]. Briefly, this is usually described as the following procedure: first, the physician gathers all available information about the patient, creating a list of symptoms. Then, the physician lists plausible candidate causes for the symptoms, prioritising the most urgently dangerous. Finally, plausible causes are ruled out through tests or further observations, and treated systematically.

## PRE-STUDY: LITERATURE REVIEW

The objective of our literature review was to establish a framework from which we could then identify areas of self-logging in clinical settings that require further investigation, as well as opportunities for HCI research to help.

### Literature Review: Method

We started with a set of search terms broad enough to encompass studies of clinical practice where patient-logged data (both paper-based and digital, manual and automatic) were introduced into a clinical setting. To do this we searched PubMed, Google Scholar and the ACM DL for keywords "patient diaries", "care diaries", "well-being diaries", "self-report diaries", "quantified-self", "self-tracking", "self-logging", "smartphone apps", and "wearable sensors".

Since we wanted to focus on the usage of data by medical experts, we excluded studies about use by patients themselves, such as for feedback, reflection, goal setting and self monitoring, including behaviour-change studies and studies of motivation to self-diaries, which were prevalent in the HCI community. Focusing only on existing practice, we omitted papers describing new interfaces and systems that have not had substantial adoption. We also excluded papers discussing the capture side of health diarising and life-logging by patients, except where aspects of capture affected its later use. We were careful to include papers that discussed any issues relating to the use of patient data in clinical settings, including those that discussed human factors issues specifically, to more broadly operating rooms and emergency rooms.

We then broadened our search to include studies that discussed the use of patient data in medical decision-making, including both patient-supplied and clinical data held by providers themselves. We also included "telemonitoring", and "electronic patient records" as our search terms. We expected to find a broad range of factors spanning human-factors issues to social, cultural, institutional, situational, among others.

For each paper, we identified factors that hinder the clinical use of data, which were first added to a spreadsheet and linked to their original source. After examining each paper, two researchers organised the list into themes.

**Literature Review: Results**
From an initial set of 2340 results, we identified 429 papers that contained at least one of the search terms among keywords and the abstract. We identified 22 relevant papers according to the criteria defined above. This set allowed us to derive the 11 themes contained in Table 1. We grouped these into four categories: data capture, data access, clinical practice, and situational constraints.

*Data capture*
This category pertains to how individuals conduct self-logging, and the devices used for doing so. Three themes were identified in this category:

- *Relevance*. Individuals may decide for themselves what information is important to self-log. This decision is influenced by their own knowledge of their condition. Without guidance from clinicians, the types of self-logged information presented by a patient may not align with what the clinician sees as useful and actionable.

- *Quality*. Most consumer devices for self-logging are not approved for medical use. Clinicians often perceive the data quality and sampling from such devices to be poor and perceive the devices to be unreliable. They are therefore unwilling to use them. Studies have, however, demonstrated that some devices have good reliability for particular purposes.

- *Completeness*. Individuals who self-log often do so sporadically, leaving out readings which may be important. Individuals attributed this to cumbersome logging tools that take time and effort to use. Such spotty data are suspect to

clinicians, who look for detailed information patterns and anomalies in data.

*Data access*
This category pertains to factors contributing to how individuals access and use self-logged data. Three themes were identified in this category:

- *Selective disclosure*. Some patients do not disclose data or information because they think it is irrelevant. Furthermore, some patients have privacy concerns with sharing information with health providers because they believe they have little control over how their data might be used outside a clinical setting. Some patients therefore conceal certain information from their health providers.

- *Representation*. Health providers use standardised forms for certain types of data, including blood pressure, symptom history, and glucose level. However, due to the quantity and variety of self-logging apps and devices, there is a lack of standardisation of how self-logged data is represented. Exercise logs, for example, often have distinct data structures and different meanings of 'activity level'. This heterogeneity in data representations results in ambiguity, and adds difficulty to interpreting self-logged data. However, applying standards to data may oversimplify it and remove important parts.

- *Interoperability*. A lack of standardisation of how data are represented creates problems when trying to share between colleagues and other health providers. There are concerns that self-logging tools will be poorly integrated into existing health-provider tools. In particular, there are challenges regarding the integration of web-based services (which many self-logged devices and apps are) into legacy systems. Many self-logging apps do not provide a means of exporting data, or they only do so in proprietary APIs which are not intuitive for patients or clinicians to use.

*Clinical practice*
This category pertains to how clinicians' training affects how effectively self-logging data can be used, and the potential barriers to introducing self-logged data to clinical practice. Three themes were identified in this category:

- *Data literacy*. Clinicians are concerned that they do not have appropriate expertise or training to effectively use or validate self-logged data. Contributing factors include the lack of standards for data representations, the lack of access to appropriate electronic tools for analysis, not being familiar with new tools for self-quantisation, and the wide variety of data. Clinicians may rely on specialists to interpret the information, but specialists may not always be available when needed.

- *Doctor-patient relationship*. The use of mobile devices in a clinical setting can draw the physician's attention away from the patient and toward the device. User interface and physical form factor affect how doctors and patients

| | Self-logged data | | | | | | | | | | | | | | Patient data | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | [8] | [3] | [5] | [35] | [1] | [28] | [36] | [24] | [4] | [2] | [17] | [43] | [6] | [29] | [33] | [42] | [21] | [12] | [39] | [18] | [23] | [13] |
| **Data capture** | | | | | | | | | | | | | | | | | | | | | | |
| Relevance | ■ | ■ | | | | | | | | | | | | | | | ■ | | ■ | | | |
| Quality | | | | ■ | | | | | ■ | ■ | | | | ■ | | | | | ■ | | | ■ |
| Completeness | | ■ | ■ | | | ■ | | | ■ | | ■ | ■ | | | | | | | ■ | | ■ | |
| **Data access** | | | | | | | | | | | | | | | | | | | | | | |
| Selective disclosure | ■ | | | | | | | ■ | | | | | ■ | | | | | | | | | |
| Representation | ■ | ■ | | | | | | | | | ■ | | | | | | | | | | | |
| Interoperability | ■ | ■ | | | | | | ■ | | | | | | | ■ | | | ■ | | | | |
| **Clinical practice** | | | | | | | | | | | | | | | | | | | | | | |
| Data literacy | ■ | | | ■ | | ■ | | | | | | | | | | | | | ■ | | ■ | ■ |
| Doctor-patient relationship | | | | ■ | | | | ■ | ■ | | | | | | ■ | ■ | | ■ | | | | |
| Legal issues | | | ■ | ■ | | | ■ | | | | | | | | | | | | | | | |
| **Situational constraints** | | | | | | | | | | | | | | | | | | | | | | |
| Time | ■ | | | | | | | | | | | | | | | ■ | ■ | | ■ | | ■ | |
| Information overload | | | ■ | ■ | | | | | | | | | | | | ■ | | | | | | |

Table 1. Major themes identified within the literature review, separating papers by whether they discussed self-logged data, or patient data such as telemonitoring or EMRs.

conduct face-to-face discussion and nonverbal communication. These aspects are crucial because they allow clinicians to better understand the patient's condition, and contribute to patient satisfaction. Exam rooms are often laid out such that clinicians can quickly switch between looking at a computer and the patient, but poor user interface can cause clinicians to spend much longer looking at the computer and disengage from discussion with the patient. Furthermore, some clinicians are concerned that patient self-logging may threaten professional autonomy, with concerns that patients may begin to make decisions which would be better made with clinical advice.

- *Legal issues.* Clinicians raised concerns that the lack of regulation of self-logging apps and devices has implications for safety, security, data protection, and reliability of using their associated data. Health providers have also raised the issue of legal liability, particularly if data fell into the wrong hands and exposed them to privacy violations, or if decisions were made based on poor quality data.

*Situational constraints*
This category pertains to constraints that exist within healthcare situations. Two themes were identified in this category:

- *Time.* Clinicians often need to work within tight time constraints, within which they are already under pressure. Clinicians have been sceptical about how effectively they could use self-logged data within these time constraints. Some said that they could be more effective when working between visits, but that paid work is not recognised in those times, so clinicians are discouraged from doing so. Cumbersome user interfaces are frustrating to clinicians because of the time it takes to complete a task or to find a relevant piece of information. Moreover, clinicians expressed concern over the necessity to interpret and document patient-provided data, even when it is irrelevant to the current situation.

- *Information overload.* The sheer quantity of information generated by self-logging tools is seen by some to be a liability, reducing productivity, increasing levels of stress,

and reducing morale. The way in which information is presented contributes to the extent to which overload affects a user. That is, the simpler the display of complex information, the better.

**METHODOLOGY**
This section outlines the methodology behind our empirical study, which involved observing primary and secondary care clinicians from both the UK and US.

**Overview and Logic**
We conducted role-play interviews with clinicians from primary and secondary care within the United Kingdom and United States. The purpose of this study was to elicit reflections on and insights about using patient self-logged data in two distinct written vignettes based on real medical cases. Vignettes are frequently used to examine clinical judgements and decision-making among health professionals because these minimise confounding variables that may be introduced by using real patients or actors [15, 31]. For example, a clinician's response to an examination of a patient might be unpredictably influenced by the patient's physical appearance, ethnicity, body language, verbalisations, or eye contact. Furthermore, an actor would not respond believably to detailed questions from a clinician during a physical exam, and both actors and patients might change their responses between sessions. We also wanted clinicians to think out loud as they worked through a case, and this seemed neither realistic nor natural with a human subject. The use of paper-based vignettes based on real medical cases therefore permitted us to focus on initial presentation (before a care record and charts are synthesised) and observe what further data or history a clinician would seek. Because we focused on forms of self-logged data which are not in widespread use in clinical practice, we were unable to find medical cases in which these forms of data were used, hence those data are drawn from other subjects.

**Recruitment**
We recruited 10 participants (3 female) who were full time, highly experienced doctors. Three participants were general

| ID | Level of care | Gender | Country | Speciality |
|---|---|---|---|---|
| GP1 | Primary | Male | UK | – |
| GP2 | Primary | Female | UK | – |
| GP3 | Primary | Male | UK | – |
| Sp1 | Secondary | Male | USA | Nephrology |
| Sp2 | Secondary | Male | USA | Rheumatology |
| Sp3 | Secondary | Male | USA | Pulmonology |
| Sp4 | Secondary | Male | USA | Hepatology |
| Sp5 | Secondary | Male | USA | Cardiology |
| Sp6 | Secondary | Female | USA | Nephrology |
| Sp7 | Secondary | Female | USA | Pulmonology |

**Table 2. Level of care, gender, and country of practice for participants, and speciality for participants in secondary care. Participants comprised of general practitioners (GP) and specialists (Sp).**

practitioners in England, practising in three different urban settings; seven were board-certified specialists (Sps) from a variety of specialisms (such as rheumatology, nephrology, cardiology, and hepatology) at a single large urban hospital in the United States. We used a small snowball technique to build our population of participants. We recruited both specialist hospital doctors and general practitioners with the expectation that this would provide a variety of distinct approaches to selecting, judging, and using information and data. The participants are listed in Table 2.

**Data Collection**
We conducted semi-structured interviews using a protocol designed around the following questions:

- How would doctors judge patient-supplied data?

- Would doctors use patient-supplied data?

- How does patient-supplied data align with current work flows and work practices?

We drafted two written vignettes of patients presenting with a set of symptoms. These were actual cases of real patients selected from set of vignettes derived from the "Think Like a Doctor" column in *The New York Times*. These were edited so that they would be understandable to doctors in the US and UK, and we added further information on the nature of self-logged data provided by the patient. The vignettes were written as narratives that outline the patient's symptoms, some background and history, and the reasons why the patient decided to see the doctor.

In the first vignette, a middle aged man finds it difficult to sleep, feels nauseous, sweaty, out of breath, and finds, when lying in bed, that his legs disobey his mental orders to stop moving. There is some background history including that he has a *patent foramen ovale* (PFO, a small hole in the tissue separating the right and left atria), suffered a mild stroke years ago, is on blood thinners, and experiences occasional panic attacks where he feels his heart pounding in his chest. We added that his wife convinced him to become a strict vegetarian and she bought him a heart monitor that he regularly uses to check his pulse. When he visits the doctor, he brings a printed Excel chart of his resting pulse in the morning, afternoon, and late evening every day for the past month. Aside

from a few spikes of 130bpm, his pulse hovers around 85bpm. His pulse increased to about 100bpm in the three days prior to his doctor visit.

In the second vignette, a female university student presents lightheaded and dizzy. Her body feels heavy, she feels anxious, she has blurred vision and headaches, she is prone to fainting if she tries to stand, and her lips are bluish. She had surgery on her back twice about 18 months ago: once to fix a pinched nerve, and again to cut out an infection that rooted there. She had been on antibiotics after the surgery and the infection had cleared. She does not smoke, and her only alcohol consumption was an occasional glass of wine. She drinks copious amounts of coffee to keep her awake, especially during exam periods, and she was worried about this. She installed an app on her iPhone which she used to record her caffeine intake. She kept a log that she printed and brought with her when she visited her doctor. The chart showed that she was consuming in excess of 1000mg of caffeine a day, until the day prior to visiting her doctor, when she had no coffee. *The New York Times* vignette does not include the patient's caffeine intake, nor that she measures it; this is information that we added.

None of the participants had seen these vignettes in advance, and none knew in advance the true diagnoses: the man had a vitamin B12 deficiency, and the student was diagnosed with postural tachycardia syndrome. We chose these vignettes because they are relatively complicated, such patients could be seen by both GPs and specialists, a diagnosis is not obvious, and they would provoke a chain of thoughts and judgements by the participants.

We distributed a vignette and chart to each participant, and asked them to think out loud as they were reading them. We asked for general comments on the case, and how they would use the information supplied by the patient. GPs were interviewed individually; specialists read and commented on the vignettes one-by-one in a group. To overcome group influence, we asked specialists to write down their thoughts on the printed vignettes, and to hand them in at the end of the session. Interviews lasted for about an hour with specialists and 30 minutes with GPs, and were recorded and transcribed.

**Data Analysis**
The research team individually analysed each transcript, swapped analysed transcripts, and re-analysed them [30]. In the first iteration, we developed a set of keywords (such as 'Training,' 'Patient History,' 'Exam') that characterised what the participants were saying as they read through the vignettes and what they had written on the printed vignettes. At this stage, we were interested only in identifying and tagging what the participants said; this was a qualitative equivalent of performing simple descriptive statistics on a dataset. Using inductive qualitative methods [10], we iteratively developed a coding scheme related to participants' information acquisition practices, their judging and trusting of patient-provided data, and the patterns of interrogating the data. Subsequent analyses allowed us to discern how both data and information content and structure support or hinder their work practice and flow of thoughts. Our final analysis helped us focus on

the cyclic nature of their strategies, and helped us construct a model of this.

## RESULTS

In this section, we present results in the order as they unfolded during the clinicians' thinking out load as they read the vignette, the order of their notes on the printed vignettes, and in the followup discussion.

### Evaluating Data Quality and Completeness

GP1 stated the need to understand how the patient recorded their pulse rate. One important factor raised was the need to understand if any illness or medication had interacted which may contribute to a change in the patient's heart rate. GP2 and Sp2 mentioned they would need to understand what the patient was doing at the times of recording, particularly during the spikes. More specifically, Sp2 needed to know what other symptoms the patient was experiencing at those times.

The heart rate chart provoked comments on the reliability of the data, relating to both the accuracy of the recording equipment and how the patient recorded the data. GP1 could not assume that the data was "objective", and said that spot checking was insufficient – it would be necessary to take a trace using trusted calibrated equipment:

> I want to use my machine, which has been pre-calibrated, not off the shelf, because I don't know about this machine's calibration. Can I trust all the data? No. Can I assume all the data is correct? No. So I have to use a calibrated machine which I feel happy with, that I've used before, and which has had an electrician or some-one who said this machine is accurate. And then I apply that test – a very quick non-invasive test. I can do a heart trace in my office to see if I can spot anything. – *GP1*

All participants said that they could not truly understand the significance of the data without having the patient in front of them. In addition to conducting a physical examination, they would look for non-verbal and subtle cues. For example, GP1 explained that it is necessary to observe how the patient behaves by walking with them from the waiting room to the office.

> I observe the way they sit, or stomping gait, or something. You're eyes just think, 'oh!' And then your connection will be made consciously and unconsciously. – *GP1*

#### Understanding patient motivations

The patient's motivations for self-logging were a factor in determining whether the data should be used. There was general agreement among all participants that it is important to know *why* these patients took the time to record their pulse and caffeine intake. Pertaining to the caffeine chart, GP3 needed to understand why the patient recorded it, when it is not a normal thing to do:

> I would try and ask a little bit more about this caffeine chart and why she's done this anyway, just to have an understanding of the reasons. Because not everyone charts their caffeine intake. – *GP3*

In particular, Sp3 suggested that presentation of a caffeine chart suggested underlying psychological issues, such as being overwhelmed and struggling with their studies, job or relationship. Referring to both vignettes, Sp5 asserted that the patient must be obsessed to record this data, suggesting an underlying psychological issue. For Sp4, the mere existence of a heart rate plot provoked a jump to psychological issues:

> They're faking it. If someone brought this chart to me, there's a red flag that this guy's got psych issues. – *Sp4*

Sp3 commented that, among his patients, it was not uncommon for engineers to bring plots of self-logged data to consultations, and Sp7 proposed that patients bringing in data will, in time, become normal practice. Nevertheless, the overarching tone of the discussion questioned the motives of patients who present self-collected data, including the belief that collecting such data might mask a larger medical or psychological issue. Sp6 said that a "lot of patients come with a diagnosis that they put on", with Sp1 following this by saying:

> It's typical that patients like this that come in and they give you stuff, you get this whole story, and then they want you to focus on it. It takes your attention away. Or they're going to tell you "this is the reason why all of this is going on," and then you have to say "well, OK, but let's just put that aside". – *Sp1*

GP3 needed to understand why both patients had brought charts, expressing confusion at their presentation and questioning what they wanted out of their consultations.

> If she's still feeling unwell, I would then refer back to the caffeine chart, especially if she's a student and knowledgeable about that. That might be cause to listen. I'd say, "Hang on. OK, tell me more about this, and why? Why are you doing that?" – *GP3*

GP1 explained that they would question the presentation of self-collected data because in their experience "there is a lot of pressure for GPs to prescribe", and that it was necessary to resist this because patients push for prescriptions for "poorly evidence-based things."

#### *Deciding how to use self-logged data*

Sp1 explained that a decision needs to be made on how much time they would need to spend using the data:

> I think when you look at something like this, you do have to be able, in your own mind, to say, "OK am I going to spend 30 seconds on this, or am I really going to spend more time and give it more importance?" And I think that's what we face a lot. – *Sp1*

Sp1 explained that patient-provided data sources will add layers of data assessment to practice, and questioned whether this information may adversely affect efficient work flow.

> The layers of information, data assessment – it's ramping up and up, and all of these devices are certainly adding, or will add, yet more of this. Certainly here, we get a lot of things faxed to us. We get to know patients who come here. And then they go somewhere else, but then we are getting their lab work over and over again -

their X-rays, their visits. At some point you have to ask yourself, "what is efficient here and what is not?" – *Sp1*

Sp7 concurred, explaining that this adds "complexity to an already complex medical interface." Sp4 suggested that age of the clinician may be a large factor in acceptance of patient provided data, arguing that medics do not usually change their practice significantly over time.

> Younger doctors that are coming in are seeing patients for the first time and they're used to people bringing in this kind of stuff and will put more thought into it. Doctors are typically creatures of habit. You've been doing something for ten, fifteen years the same way, you're going to carry on doing it. There are minor changes that happen in terms of pathology and diagnostics, but in general you're used to doing it in a certain way... So I think it remains to be seen, but when I think what I'll be doing in 10 years time in terms of how I'll be managing a patient, it will be very similar. – *Sp4*

One GP raised the need for a physical exam, and that the scenario and data alone were insufficient for making a diagnosis.

> I haven't even touched him yet. After my history I'd do an examination, feeling the pulse myself, feeling if its strained, the quality, the character, and whether I can do anything to the pulse. And then listening to his heart and doing the various movements and motions – sit forward sit back – to see if he's got any murmurs associated with this. I'd see if he's been compliant with his anticoagulants – maybe he is maybe he isn't. – *GP1*

There were differences in how useful the clinicians found the self-logged information. GP3 chose to ignore the caffeine, stating that it was irrelevant, but conceded that it may become useful later. Despite also dismissing the heart rate plot, GP3 did observe that there were high peaks which motivated questions about what the patient was doing at those times.

> But certainly, I don't think this chart would influence me. The only influence it would have is to try and understand more about why he did this and what he wants out of the consultation. – *GP3*

The presentation of this information swayed Sp4 toward caffeine as a cause:

> I think if you see her at the office and she brings you this immediately you start thinking this is all from too much caffeine. – *Sp4*

GP1 said the caffeine chart suggested a coffee-withdrawal headache, and verbalised a unit conversion in trying to understand how much caffeine 400mg is.

> Right! This could be a coffee headache. Well if you stop drinking coffee you get a headache. If you start drinking coffee you get a headache. Daily consumption – wow – above 400mg, 150mg per cup. Yeah, so this could be a coffee withdrawal headache. – *GP1*

Sp2, Sp3 and Sp5 did not know what a normal caffeine level was, and could not comment on whether the headaches were potentially caused by caffeine withdrawal. Similarly, Sp2 made an initial observation of the heart rate chart based on personal beliefs of how heart rate varies, but said it would be necessary to consult a cardiologist to fully understand it:

> Well one thing that struck me is how little variability there was in the heart rate during the time of the day. I would need to ask a cardiologist, but I thought there was greater variability in heart rate. – *Sp2*

**Prioritising Patient Safety**

In general, the clinicians would identify a set of possible conditions, and then order them from the most serious to the least. In addition, they would use information to construct a safe care pathway for the patient. One specialist said,

> I'm really just looking for things that are more serious, like trying to see what's the worst possible thing that she could have, approaching it that way. And hopefully it will turn out to be nothing... I'm looking for what's the worst possible thing the person could have and work backwards from there... – *Sp5*

In ruling out the worst cases, knowledge of caffeine consumption changed GP1's perspective on possible causes; although caffeine withdrawal initially seems likely, there could be a far more serious tumour. After reading the scenario and devising a plan, GP1 said:

> At the moment I've chopped, chopped, chopped, chopped, and we come to here. And now I think, "Right, we've pruned off all of that, now I've got the bare tree." It's deciduous, not perennial! And it's very easy to see, this is my path now. It's your heart, mate. And I need to do just one or two tests to show. Otherwise the trunk of this tree becomes thicker, and I will go that way. That's how I think. – *GP1*

**DISCUSSION**

Most of the major themes we identified in our literature review were found in our results; *Relevance*, *Quality*, *Completeness*, *Doctor-Patient Relationship*, *Information Overload*, and *Time Constraints* were especially prominent. Little related to *Legal Issues* and *Selective Disclosure* emerged. In this section we interpret an extended set of themes resulting from our analysis into how future Quantified Self and self-logging tools might be made to better support use in clinical diagnosis.

**Evidence Gathering in Support of Risk Mitigation**

We have observed six sets of interrelated activities: *Discovery*: Gather information and evidence; *Evaluation*: Evaluate the evidence for quality, reliability, validity, and completeness; *Form initial hypotheses*: Formulate a possible list of diagnoses based on the evidence presented; *Identify knowledge gaps* that are needed to test hypotheses (rule out possible diagnoses); *Refine hypotheses* based on further evidence and data; and *Construct safe care pathway*. These are not discrete activities, nor do they occur in sequence. Instead, doctors tend to work recursively, moving back and forth from one set of activities to another, and using the outcomes of one set to inform and guide another.

The guiding principle behind each of these sets of activities is to mitigate risk and maintain patient safety; this is made clear not only by the order of hypotheses and steps taken, but also by the self-reflection during the think-aloud process. The highest-risk, usually life-threatening, possible explanations for a set of presented symptoms were always considered first and systematically eliminated by gathering or identifying supporting evidence. This process was repeated for the next-highest risk hypotheses and so on, until either a single working hypothesis was devised or hypotheses could not be ruled out due to inadequate evidence. It was also clear that, during the evidence gathering phase, to find support or to eliminate any given hypothesis, clinicians carefully considered *all* evidence they encountered to see if there were potential connections to either new hypotheses not yet considered, or previous ones already excluded.

Given the importance and consistent application of this workflow in both primary and secondary care settings, one might ask how self-logged information might be better designed to support this process. In our experiment, it was clear that the self-logged data we included in our scenarios were often irrelevant to the most severe hypotheses, and therefore were not appropriate for use in hypothesis elimination until much later in the process. Yet, such data were always considered by the clinicians at least once, and usually quite early in the process, for the purposes of introducing new hypotheses not yet considered. Therefore, in the context of hypothesis creation and elimination, it is clear that self-logged data have two potential roles: both in *discovery* (for example, identifying potential causes not yet considered), and *refinement* (for example, eliminating hypotheses or supporting existing hypotheses).

A third, related use for self-logged data in this workflow is its role as a communications aid, for helping patients explain their symptoms, or recent history. The importance of clearly understanding what a patient was experiencing was emphasised several times during the experiment, with clinicians not only talking with the patient but also carefully studying the patient's body language to ascertain how the patient was feeling. This process was particularly important during the first few minutes of a patient consultation, to set the stage for the subsequent risk-mitigation workflow process.

In the evidence gathering phase, when inadequate supporting information was available to definitively rule out a hypothesis, clinicians resorted to adding actions to plans to gather more evidence. Such actions included conducting physical exams, simple diagnostics (such as taking blood pressure), to more time and resource extensive diagnostics, such as blood work, ECGs, and imaging. These actions were often prioritised by how quickly and easily they could be done, and the potential value of the data gleaned, against the potential inconvenience, discomfort, and risks posed to the patient. It was clear that if more appropriate and trustworthy evidence were available (such as through self-logged sources), fewer such tests might be required, saving potentially not only time and resources, but also affording a more expedient diagnosis. The potential for risk and discomfort in additional diagnostic tests also suggests that *not* having to do such tests could bring about additional direct benefit to patients, as well.

## Frameworks for Evidence: Form and Representation

Somewhat unsurprisingly, we observed a critical relationship between the aforementioned process of risk-mitigation and the ways that clinical evidence was ordered, structured, and represented. Our secondary care specialists, in particular, expressed having to 're-arrange' information that was being presented to them in order for them to be able to effectively think about the evidence, and expressed frustration when trying to ascertain how particular evidence fits within the frameworks to which they were accustomed.

The aspect deemed particularly crucial was the timeline of patient events. Clinicians mentioned mentally placing evidence (corresponding to patient events) along this timeline as they went through the scenarios. Both chronological ordering and duration between events were important for identifying potential relationships between symptoms, and establishing connections between symptoms and potential causes. Such relationships were germane to evidence admission; that is, for determining whether a particular piece of evidence was relevant for the current set of hypothesised causes. During the process of restructuring evidence, several participants noticed that crucial information was missing from the scenario about the timeline of patient events. This highlights an additional aspect of standardised representations, that they make it easier to identify when crucial information is missing.

Work identified within our literature review made similar observations regarding lack of standardisation in self-logging tools. In particular, Chung et al. [8] observed that, while standardised forms for some data do exist (for example, glucose level and blood pressure), consumer self-tracking tools rarely use them, and, where they are used, reduce it to a factor which is not useful to clinicians (for example, physical activity points). They suggest that standardised forms have been designed to facilitate quick and accurate review of data, and the use of non-standard data representations makes it more difficult for clinicians to use self-logged data effectively.

Given the importance of standard representations, we speculate whether digital tools might be capable of re-ordering and representing self-logged data into forms that clinicians are accustomed to. If successful, such representations could reduce the mental effort needed to transform presented evidence and facilitate reasoning, supporting external cognition [32] for the hypothesis refinement process. Additionally, such representations might make it easier to spot irregularities in the evidence, such as missing data, errors or hidden causes. Although not yet common in use, even for clinical data, our research suggests that visualisations of patient-event timelines, if designed in a clinician-centric way, might facilitate temporal reasoning across related events.

## Unpacking "Expertise"

Our literature review revealed that clinicians are concerned with their lack of expertise in using self-logged data. Based on results from our experiment, we now have a clearer idea

of what "lack of expertise" might actually mean. We believe that there are at least two related factors.

The first pertains to what was captured, in particular reasoning about self-logged information that fell *outside the set of markers* normally used for diagnosis. For example, within our experiment, when presented with the caffeine chart, clinicians admitted being unable to effectively decide whether the information was worth considering. One remarked that they were not familiar with the normal caffeine level (displayed in milligrams per day), and attempted to convert the scale to *cups of coffee per day* to help them reason about it. Had this been a standard marker used regularly for diagnosis, it is much more likely that interpretation of this measure would have presented less difficulty.

The second pertained to (statistical) data interpretation. With the heart rate chart, a number of clinicians said they were not confident in interpreting the data, but when asked for clarification, it became clear what they meant was they had difficulty interpreting it *in the form that it was presented*. Several of our participants interrogated the salient features of the heart rate chart – comprising fluctuations, peaks and troughs of aggregate 3-per-day statistics – but could not discern easily if they were normal (uninteresting) or abnormal (possible evidence). This contrasts with the ways that outcomes of clinical diagnostic test results are typically reported: with statistical likelihoods pre-calculated, and deviation from expected means explicitly represented.

We believe that these observations have a number of implications for the design of self-logging and Quantified Self tools; first, the plurality (and popularity) of tools for self-measurement already demonstrates wide heterogeneity in what can be measured, many of which pertain to routine activities that may not correspond to standard measures used by clinicians. We suggest that tools might consider ways to make it easier for clinicians to reason about such unfamiliar kinds of data. For example, by comparing a person's measurements against a population (or demographic) average, useful comparisons could be directly made. Statistical pre-validation could be performed by tools to determine whether such measurements fall within the normal range for people within the patient's demographic, and, like lab results, this summary could be made explicitly visible. Finally, continuous information may be better initially presented as aggregated data, highlighting significant events.

### Psychology of Motivations for Self-Logging

One of our most unexpected findings was that clinicians found it important to try to understand the reasons that people chose to meticulously self-log data about themselves, and attempted to use this line of reasoning to ascertain potentially hidden physical or psychological problems in patients. In particular, several of our participants considered whether the patient's act of recording data indicated the presence of psychological disorders, ranging from obsession to depression, which could explain some of the symptoms presented. We note that earlier work observed doctors being concerned that patients who self-log appear obsessive and compulsive [4].

This kind of reasoning, however, would only be relevant to the subset of self-logged data that patients voluntarily captured themselves, without any clinical reason for doing so. Drawing from this, we propose three categories of self-logged data. The first comprises data which the patient recorded voluntarily, without being told to by a clinician, such as out of curiosity, personal goals, or obsession. This category would include, for example, data captured incidentally by an activity tracker purchased or acquired for the purpose of getting fit. The second describes data which have been recorded for the purpose of managing one or more chronic conditions [3]. Our literature review found studies documenting the often complex processes involved in keeping track of symptoms, medications and other aspects of having one or more chronic disease. The third describes data which have been recorded as a result of a clinicians instructions. This includes, for example, telemonitoring devices, such as implantable ECGs, as a result of a diagnosis of arrhythmia.

Both the caffeine chart and heart rate chart fell within the first category, and as such, prompted questions about why the patient felt the need to record these data, whether it was evidence that the patient had self-diagnosed, and if the presence of the data was a symptom of an underlying psychological issue, such as obsession, stress, or a co-morbid psychiatric disorder. Due to this possibility, clinicians were more sceptical of the accuracy of the data, suggesting that the patient may have been motivated to manipulate or fake the data to force a particular outcome from the consultation.

We suggest that, with the increasing prevalence of wearable technology and self-logging apps, the volume and variety of data falling in the first category will grow. While this may prompt more questions about why these data were captured, most of these data will be captured automatically by wearable sensors and devices will mean that the psychological reasons for doing so may become less significant.

### Data Believability: Veracity, Sampling and Context

Clinicians were unwilling to make decisions based on self-logged data without knowing details about *how* it was recorded. When contextualised within the risk mitigation workflow, it became clear why; for evidence to be able to be used to justify the elimination of a high-risk hypothesis (such as an undiagnosed, potentially fatal condition) it was imperative that this evidence was reliably obtained and interpreted correctly. A number of factors were mentioned around this, including what device had been used to perform the measurement, whether the device was calibrated, how the data were sampled. In order to ascertain whether there were potential sampling issues for the blood pressure charts, for example, clinicians asked questions about the context(s) at the times of capture: where the patient was when the data were recorded, what the patient was doing, among others. When satisfactory answers were not available, clinicians wanted to retake the readings using their own calibrated devices.

In order for future wearable devices and smartphone apps to capture data that can be interpreted believably, they might capture information about the calibration of their sensors and associate these with the data. More critically, we propose that

there may be opportunities to solve some of the sampling concerns that arose in our study through the use of digital context capture techniques. For example, physical activity sensing could enable annotation of heart rate readings with physical activities, while location sensing could provide physical location annotations as well. Cryptographic approaches such as proof of work [9] could further be used to prove that such readings were, in fact, taken at a particular time to eliminate the possibility of post-hoc data fabrication. The benefit of such techniques would be that they might substantially improve the trustworthiness of sensed data without adding additional burden to the process of capturing it.

**Implications for HCI**
The market for health tracking devices is currently being led by technology companies in fierce competition to sell new, technologically-advanced systems, rather than to best support patients' long-term needs. We feel that HCI, which has long examined personal health and well-being technologies, should continue to illuminate ways that such tools could better meet both patients' and clinicians' needs, through in-depth investigations, two examples of which we describe next.

The first is in resolving the disparity between kinds of data individuals think is important, and data that clinicians actually find useful for diagnosis. We feel that there is room for extending the current understanding of why and how people self-log, such as studied by Choe et al. [7], to explore ways that such practices might be aligned to better support the taking of clinically relevant, high-quality measures. Well-being diary "support systems" of the future, might, for example, teach (and nudge) individuals to capture more clinically-relevant data, and guide them on how to make more accurate measurements.

The second pertains to improving the clinicians' ability to explore and interrogate data brought in by patients during consultations. Although our study did not evaluate specific presentation or interaction techniques, it was clear that such particulars were important, as discussed earlier. But we also believe that further studies of clinicians' evidence-gathering and hypothesis-refinement workflows during consultations could greatly benefit the design of interfaces for performing *in-situ*, patient-provided data sensemaking and discovery.

**LIMITATIONS**
The first, perhaps most obvious limitation was our sample size; the number of clinicians that we interviewed was small (3 GPs and 7 specialists). Since the practice of clinicians may differ substantially depending on specialism, training, background, and where they work, and so forth, we would like to follow this study with another featuring a different set of GPs, and clinicians with different specialities, as well as ward and triage nurses. Second, our interviews with GPs were conducted individually, whereas our interviews with specialists were conducted as a group. As a group, people may be more selective in what they say because of the presence of colleagues. Because of this, our observations cannot be used to directly compare differences in approaches to diagnosis between primary and secondary care. A third limitation pertains

to the way that setting and scenarios were presented. Instead of seeing real patients, or a professional actor, scenarios were written as textual descriptions on paper. As a result, clinicians did not have access to the patient (or an actor) or their medical file. This limits our ability to observe how self-logged data affects doctor-patient relationships. Fourth, since we presented data on a piece of paper and not on a device, we cannot infer the role of the user interface on the usefulness or introduction of barriers to use. Finally, we only showed them two different forms of self-logged data, in two representations within two different scenarios. We cannot necessarily understand the implications of the type of information, form of data representation on data usability. We wish to explore this further in follow-up work.

**CONCLUSION**
This paper investigated challenges to the adoption of self-logged data in clinical practice. By conducting a review of literature surrounding the use of self-logged data, we identified potential factors pertaining to data capture, data access, clinical practice, and situational constraints. Through our diagnostic role-play experiment with clinicians, we were able to get a deeper understanding of how such data would be used in practice, including factors that influenced whether these data were considered during the risk-mitigation decision-making process, barriers to effectively interrogating these data during clinical visits, and factors that influenced whether clinicians viewed data as trustworthy.

More specifically, we identified opportunities for how self-logged data could help with various phases of the diagnostic workflow process, including *communication* (of symptoms) with the patient, *discovery* of potential causes not yet considered, and *refinement* of hypothesis. We discovered that the motivation for self-logging activities was also of critical importance to the diagnostic process, by helping clinicians discover potential psychological disorders, or unspoken organic disorders. Finally, we identified reasons that clinicians distrusted self-logged data pertaining to accuracy of instruments, sampling methods, potentially missing data and patient activity or context, and identified potential ways that some of these issues might be addressed in the future. These initial findings suggest that, although there may be a significant number of design challenges remaining, the use of self-logged data may eventually significantly improve clinicians' abilities to effectively draw together evidence for clinical diagnosis.

## REFERENCES

1. Manal Almalki, Kathleen Gray, and Fernando Martin Sanchez. 2015. The use of self-quantification systems for personal health information: big data management activities and prospects. *Health Information Science and Systems* 3, Suppl 1 (24 Feb. 2015), S1. DOI: `http://dx.doi.org/10.1186/2047-2501-3-S1-S1`

2. Ole Andreas Alsos, Anita Das, and Dag Svanæs. 2012. Mobile health IT: The effect of user interface and form factor on doctor-patient communication. *International Journal of Medical Informatics* 81, 1 (14 Oct. 2012), 12–28. DOI: `http://dx.doi.org/10.1016/j.ijmedinf.2011.09.004`

3. Jessica S. Ancker, Holly O. Witteman, Baria Hafeez, Thierry Provencher, Mary Van de Graaf, and Esther Wei. 2015a. The Invisible Work of Personal Health Information Management Among People With Multiple Chronic Conditions: Qualitative Interview Study Among Patients and Providers. *Journal of Medical Internet Research* 17, 6 (14 June 2015), e137. DOI: `http://dx.doi.org/10.2196/jmir.4381`

4. Jessica S. Ancker, Holly O. Witteman, Baria Hafeez, Thierry Provencher, Mary Van de Graaf, and Esther Wei. 2015b. "You Get Reminded You're a Sick Person": Personal Data Tracking and Patients With Multiple Chronic Conditions. *Journal of Medical Internet Research* 17, 8 (19 Aug. 2015), e202. DOI: `http://dx.doi.org/10.2196/jmir.4209`

5. Stefan Becker, Talya Miron-Shatz, Nikolaus Schumacher, Johann Krocza, Clarissa Diamantidis, and Urs-Vito Albrecht. 2014. mHealth 2.0: Experiences, Possibilities, and Perspectives. *JMIR mHealth and uHealth* 2, 2 (16 May 2014). DOI: `http://dx.doi.org/10.2196/mhealth.3328`

6. Yunan Chen, Victor Ngo, Sidney Harrison, and Victoria Duong. 2011. Unpacking Exam-room Computing: Negotiating Computer-use in Patient-physician Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2011) *(CHI '11)*. ACM, New York, NY, USA, 3343–3352. DOI: `http://dx.doi.org/10.1145/1978942.1979438`

7. Eun Kyoung Choe, Nicole B. Lee, Bongshin Lee, Wanda Pratt, and Julie A. Kientz. 2014. Understanding Quantified-selfers' Practices in Collecting and Exploring Personal Data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2014) *(CHI '14)*. ACM, New York, NY, USA, 1143–1152. DOI: `http://dx.doi.org/10.1145/2556288.2557372`

8. Chia-Fang Chung, Jonathan Cook, Elizabeth Bales, Jasmine Zia, and Sean A. Munson. 2015. More Than Telemonitoring: Health Provider Use and Nonuse of Life-Log Data in Irritable Bowel Syndrome and Weight Management. *Journal of Medical Internet Research* 17, 8 (21 Aug. 2015), e203. DOI: `http://dx.doi.org/10.2196/jmir.4364`

9. Jeremy Clark and Aleksander Essex. 2012. CommitCoin: carbon dating commitments with Bitcoin. In *Financial Cryptography and Data Security*, Angelos D. Keromytis (Ed.). Springer Berlin Heidelberg, 390–398. DOI: `http://dx.doi.org/10.1007/978-3-642-32946-3_28`

10. Juliet Corbin and Anselm Strauss. 2008. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (3 ed.). Sage Publications, Inc. DOI: `http://dx.doi.org/10.4135/9781452230153`

11. Pat Croskerry. 2013. From Mindless to Mindful Practice – Cognitive Bias and Clinical Decision Making. *New England Journal of Medicine* 368, 26 (27 June 2013), 2445–2448. DOI: `http://dx.doi.org/10.1056/NEJMp1303712` PMID: 23802513.

12. M. Demonceau, A. Donneau, J. Croisier, E. Skawiniak, M. Boutaayamou, D. Maquet, and G. Garraux. 2015. Contribution of a Trunk Accelerometer System to the Characterization of Gait in Patients with Mild to Moderate Parkinson's disease. *IEEE Journal of Biomedical and Health Informatics* 19, 6 (19 Aug. 2015), 1803–1808. DOI: `http://dx.doi.org/10.1109/JBHI.2015.2469540`

13. Manon L. Dontje, Martijn de Groot, Remko R. Lengton, Cees P. van der Schans, and Wim P. Krijnen. 2015. Measuring steps with the Fitbit activity tracker: an inter-device reliability study. *Journal of Medical Engineering & Technology* 39, 5 (27 May 2015), 286–290. `http://www.tandfonline.com/doi/abs/10.3109/03091902.2015.1050125`

14. David M. Eddy. 1990. The challenge. *Journal of the American Medical Association* 263, 2 (12 Jan. 1990), 287–290. DOI: `http://dx.doi.org/10.1001/jama.1990.03440020121046`

15. Spencer C. Evans, Michael C. Roberts, Jared W. Keeley, Jennifer B. Blossom, Christina M. Amaro, Andrea M. Garcia, Cathleen Odar Stough, Kimberly S. Canter, Rebeca Robles, and Geoffrey M. Reed. 2015. Vignette methodologies for studying clinicians decision-making: Validity, utility, and application in ICD-11 field studies. *International Journal of Clinical and Health Psychology* 15, 2 (29 Jan. 2015), 160–170. DOI: `http://dx.doi.org/10.1016/j.ijchp.2014.12.001`

16. Jeffrey L. Fox. 2015. Obama catapults patient-empowered Precision Medicine. *Nature Biotechnology* 33, 4 (7 April 2015), 325. DOI: `http://dx.doi.org/10.1038/nbt0415-325`

17. Richard Frankel, Andrea Altschuler, Sheba George, James Kinsman, Holly Jimison, Nan R. Robertson, and John Hsu. 2005. Effects of Exam-Room Computing on Clinician-Patient Communication. *Journal of General Internal Medicine* 20, 8 (1 July 2005), 677–682. DOI: `http://dx.doi.org/10.1111/j.1525-1497.2005.0163.x`

18. B. M. Frier, M. M. Jensen, and B. D. Chubb. 2015. Hypoglycaemia in adults with insulin-treated diabetes in the UK: self-reported frequency and effects. *Diabetic Medicine* (7 Sept. 2015). DOI: http://dx.doi.org/10.1111/dme.12878

19. Thomas Goetz. 2010. It's time to redesign medical data. TEDMED. (Oct. 2010). http://www.ted.com/talks/thomas_goetz_it_s_time_to_redesign_medical_data?language=en

20. Thomas Goetz. 2011. *The Decision Tree: How to make better choices and take control of your health* (1 ed.). Rodale Books.

21. Benjamin Hughes. 2008. To 2.0 or not to 2.0-have junior doctors already answered the question? *Medicine 2.0 Conference* (9 May 2008). http://www.medicine20congress.com/ocs/index.php/med/med2008/paper/view/92

22. Tim Kelsey. 2013. Personalised care plans will give patients control of their own health - Tim Kelsey. *NHS England* (25 Sept. 2013). https://www.england.nhs.uk/2013/09/25/tim-kelsey-2/

23. Samir S. Khariwala, Steven G. Carmella, Irina Stepanov, Dipankar Bandyopadhyay, Heather H. Nelson, Bevan Yueh, Dorothy K. Hatsukami, and Stephen S. Hecht. 2015. Self-reported Tobacco use does not correlate with carcinogen exposure in smokers with head and neck cancer. *The Laryngoscope* 125, 8 (5 April 2015), 1844–1848. DOI: http://dx.doi.org/10.1002/lary.25290

24. John Millar. 2004. A shared care model for complex chronic disease care: A community of practice. (2004). PHSA Navigation Workshop.

25. National Information Board. 2014. Personalised health and care 2020: Using Data and Technology to Transform Outcomes for Patients and Citizens. (Nov. 2014). https://www.gov.uk/government/publications/personalised-health-and-care-2020

26. Aditi Pai. 2013. Timeline: Smartphone-enabled health devices. *MobiHealthNews* (7 June 2013).

27. Aditi Pai. 2015. iHealth Core, a new weight scale for people with chronic conditions. *MobiHealthNews* (15 Sept. 2015).

28. Rupa A. Patel, Predrag Klasnja, Andrea Hartzler, Kenton T. Unruh, and Wanda Pratt. Probing the benefits of real-time tracking during cancer care. In *AMIA Annual Symposium Proceedings*. 1340–1349. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3540467/

29. Emily S. Patterson, Anh D. Nguyen, James P. Halloran, and Steven M. Asch. 2004. Human Factors Barriers to the Effective Use of Ten HIV Clinical Reminders. *Journal of the American Medical Informatics Association* 11, 1 (Jan. 2004), 50–59. DOI: http://dx.doi.org/10.1197/jamia.M1364

30. Catherine Pope, Sue Ziebland, and Nicholas Mays. 2000. Qualitative research in health care: Analysing qualitative data. *British Medical Journal* 320, 7227 (2000), 114–116. DOI: http://dx.doi.org/10.1136/bmj.320.7227.114

31. Donald A Redelmeier, Derek J Koehler, Varda Liberman, and Amos Tversky. 1995. Probability judgment in medicine discounting unspecified possibilities. *Medical Decision Making* 15, 3 (Aug. 1995), 227–230. DOI: http://dx.doi.org/10.1177/0272989X9501500305

32. Mike Scaife and Yvonne Rogers. 1996. External cognition: how do graphical representations work? *International journal of human-computer studies* 45, 2 (Aug. 1996), 185–213. DOI: http://dx.doi.org/10.1006/ijhc.1996.0048

33. R. Schoenberg, L. Nathanson, C. Safran, and D. Z. Sands. 2000. Weaving the Web into legacy information systems.. In *Proceedings of the AMIA Symposium* (2000). 769–773. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2243891/

34. W. S. A. Smellie, M. J. Galloway, D. Chinn, and P. Gedling. 2002. Is clinical practice variability the major reason for differences in pathology requesting patterns in general practice? 55, 4 (April 2002), 312–314. http://jcp.bmj.com/content/55/4/312

35. Mark Sullivan. 2014. Guess what? Doctors don't care about your Fitbit data. *VentureBeat* (2014).

36. Melanie Swan. 2009. Emerging Patient-Driven Health Care Models: An Examination of Health Social Networks, Consumer Personalized Medicine and Quantified Self-Tracking. *International Journal of Environmental Research and Public Health* 6, 2 (Feb. 2009), 492–525. DOI: http://dx.doi.org/10.3390/ijerph6020492

37. Melanie Swan. 2012. Health 2050: The Realization of Personalized Medicine through Crowdsourcing, the Quantified Self, and the Participatory Biocitizen. *Journal of Personalized Medicine* 2, 3 (Sept. 2012), 93–118. DOI: http://dx.doi.org/10.3390/jpm2030093

38. Carl Thompson and Dawn Dowding. 2002. *Clinical decision making and judgement in nursing*. John Wiley & Sons.

39. Athanasios Tsoukalas, Timothy Albertson, and Ilias Tagkopoulos. 2015. From Data to Optimal Decision Making: A Data-Driven, Probabilistic Machine Learning Approach to Decision Support for Patients With Sepsis. *JMIR Medical Informatics* 3, 1 (2015). DOI: http://dx.doi.org/10.2196/medinform.3445

40. U.S. Food and Drug Administration. 2014. *K132764 Approval Letter*. Technical Report. Department of Health & Human Services. http://www.accessdata.fda.gov/cdrh_docs/pdf13/K132764.pdf

41. Paraskevas Vezyridis and Stephen Timmons. 2015. On the adoption of personal health records: some problematic issues for patient empowerment. 17, 2 (2015), 113–124. DOI: http://dx.doi.org/10.1007/s10676-015-9365-x

42. Signe Vikkels. 2005. Subtle Redistribution of Work, Attention and Risks: Electronic Patient Records and Organisational Consequences. *Scandinavian Journal of Information Systems* 17, 1 (2005). http://aisel.aisnet.org/sjis/vol17/iss1/10

43. Nathan Zeldes and Neil Baum. 2011. Information overload in medical practice. *Journal of Medical Practice Management* 26, 5 (2011), 314–316.