CHAPTER 10

Case Studies

Learn from your failures or learn from other people’s failures. Whichever you choose, learn.

—Kirpa Gayatri Kaur

AI is improving the healthcare experience, bringing success to those who can leverage and adapt to a new health and care delivery paradigm. This chapter contains real-world case studies where applications of machine learning are being used to solve real-world problems in digital health.

Each case study provides a unique and engaging perspective of the use of big data, AI, and machine learning within health and social care. Real-life descriptions of organizational approaches to data-identified healthcare problems demonstrate the instant value within data and AI.

The purpose behind the case studies is to showcase real-world demonstrations of AI, machine learning, data, and IoT usage that are novel and innovative.

Each case study shares unique learnings developed from creation, training, and deployment of the respective intelligent system.

Real-World Inspiration

A call for case studies was placed on Twitter, and 108 case studies were submitted. Case studies were accepted if they were used in the real world, whether in pilot or otherwise, and were supported with evidence to demonstrate value from a health benefit or cost-saving perspective.
The criteria used to evaluate case studies for inclusion involved the following:

- There is focus on delivering precision medicine and big data-generating technologies.
- Machine learning or data-driven models must be deployed into the real world.
- Authors must highlight learnings from the experiences the project provided.
- Authors must share adequate information on the project scope and conclude on the organizational, business, clinical, or financial impact.

Real-World Application and Learnings

Case studies selected showcase a tremendous vision, focus, and drive to improve patient engagement, outcomes, and clinical success. Case studies serve as stimulating examples on the following:

- The use of big data to drive decision-making and provide patient- and population-scale precision medicine
- Advantages and pitfalls of developing machine learning models and their application in the real world.
- Data collection, analysis, management, and governance
- Developing artificially intelligent agents and blending digital and face-to-face healthcare
- The use of wearables, sensors, and IoT health devices to provide care within primary and secondary care systems
- Ethics of AI and morality
- The future of digital technology, healthcare, and evidence-based medicine in the pursuit of improving quality of life
Further details about case studies, their respective authors, and their organizations are available in the supplementary reading area available online.

The case studies chosen for inclusion all detail real-world applications of machine learning and AI that exist in healthcare today and include examples of digital health technologies provided by DDM, the organization I lead today.

Case studies comprise

1. AI for Imaging of Diabetic Foot Concerns and Prioritization of Referral for Improvements in Morbidity and Mortality
2. Outcomes of a Digitally Delivered, Low-Carbohydrate, Type 2 Diabetes Self-Management Program: 1-Year Results of a Single-Arm Longitudinal Study
3. Delivering a Scalable and Engaging Digital Therapy for Epilepsy
4. Improving Learning Outcomes for Junior Doctors Through the Novel Use of Augmented and Virtual Reality
5. Do Wearable Apps Have Any Effect on Health Outcomes? A Real-World Service Evaluation of the Impact on Activity
6. Big Data, Big Impact, Big Ethics: Diagnosing Disease Risk from Patient Data
7. Assessment of a Predictive AI Model for Personalised Care and Evaluation of Accuracy
8. Can Voice-Activated Assistants Support Adults to Remain Autonomous, a Real-World Service Evaluation of the Impact of a Voice-Activated Smart Speaker Application on Activity

**Case Study 1: AI for Imaging of Diabetic Foot Concerns and Prioritization of Referral for Improvements in Morbidity and Mortality**

Harkrishan Panesar, Gro Health, UK
Background

Globally, 463 million people have diabetes, 4 million in the United Kingdom [237]. Seventy-seven thousand diabetics in England have foot ulcers today [238]. The longer ulcers progress, the longer they take to heal. Untreated ulcers can lead to diabetic foot disease (DFD) and amputation [239]. Twenty-five lower limb (toe/foot/leg) amputations occur in England daily; 85% are preventable [240, 241].

From 2015 to 2018, 27,465 lower limb diabetes-related amputations happened in England with 147,067 DFD-related hospital admissions and an average 8-day stay totaling 1,826,734 bed days costing >£686 million [242, 243].

DFD accounts for £1 in every £100 the NHS spends [244]. Reduction in ulceration prevalence by 33% would save NHS England £250 million/year [9]. Foot checks are a chance for potential problems to be identified and assessed and action taken. NICE NG19 states people with diabetes should have diabetes foot checks annually and patients with active foot problems referred to a multidisciplinary team within one working day and triaged within another [245].

However, "poor symptom recognition by patients, inaccurate healthcare assessment, and difficulties in accessing specialist services" [246] exist due to

- Delays to examination: The National Diabetes Foot Care Audit found 39% of people waited 14+ days for their first foot ulceration specialist examination [247].
- Lack of testing: 800,000 patients in England do not receive an annual foot check. Eighty percent of practice nurses are not confident performing foot checks, causing variances in care and outcomes [248].
- Poor assessment: 33% of NHS England commissioners do not provide healthcare professionals (HCPs) with foot care training. Where training is in place, quality of assessment is poor [249].

Innovation can democratize patient care and reduce the economic and health burden of ulceration/DFD.
Cognitive Vision

Cognitive vision or image recognition is a machine learning technique designed to mimic how the human brain functions. With this technique, machine learning models are taught how to recognize visual components that make up an image. Through learning on large datasets of images and identifying patterns, models are able to make sense of their input and determine relevant tags and classifications. Image recognition is not an easy task. To successfully perform image recognition, neural networks are the technique of choice.

However, even still, the computation resource cost is expensive in practice. For instance, a 20-pixel by 20-pixel image received by a neural network would receive over 400 inputs. Although this sounds achievable in a typical hardware setup, images of greater size, say 1,000 pixels by 1,000 pixels, would require a powerful computational resource to process the higher number of parameters and inputs. In a traditional neural network, each pixel would be linked to a single neuron. This is computationally expensive.

Proximity of pixels within images has a strong relationship with their similarity. Convolutional neural networks specifically make use of this feature. Rather than treating two nearby pixels as distinct, convolutional neural networks assume there is more likely a relationship between these pixels than two that are further apart.

Through bypassing less significant connections between pixels, convolution solves the computational and time problems faced by traditional neural networks.

Convolutional neural networks improve the computational resource required for image recognition through filtering relationships by proximity.

Clinical evidence suggests that diagnostic tests and risk stratification can predict the risk of ulceration and amputation, with early referral reducing amputation rates and time to heal [251].

Solutions for diabetic foot care must be broader than the standard healthcare worker-dependent service and be deliverable at a lower cost for the patient and healthcare system.

Together with my colleague Yang Wang, we applied state-of-the-art AI techniques to automatically detect signs of diabetic foot concerns, alert the patient, and support communication and assessment by the healthcare team. Image recognition technology for diabetic foot concerns would be revolutionary, particularly considering this machine learning experience could improve ulceration healing time, cost, and risk of amputations—as this is related to how early ulceration is detected. The model has
achieved high accuracy in detecting subtle areas of bruising and cracked heels, which can soon develop into areas of foot ulceration.

A cloud-based system was developed to allow remote use of the application such that any device with an Internet browser was able to upload an image alongside a standardized protocol. The output would be a list of any concerns detected, the referral pathway most suited to the concern, and education, if relevant.

The model is constantly refined through a machine learning feedback loop that optimizes and retrain the model periodically with random subsets of labeled data, including the data received from the web-based system.

**Project Aims**

The aims of our project are to further develop and validate this machine learning technology, including the deep learning neural network used to predict the risk of foot concerns and provide precise medicine to patients. The convolutional neural networks used for multiple predictions can become more precise through their use and feedback loop from users.

As the user base is largely healthcare professionals and carers, the learning can be scaled up and validated by the healthcare professional for appropriate triage. The project also aims to deliver a service to patients, which is only possible when the model has reached maximum, and ideally near-absolute, accuracy. Reducing error is a key priority for the project.

Grading schemes and standardized protocols support the model to develop greater precision and recall. The ability to identify ulceration and other foot concerns earlier would accelerate diagnosis and treatment and save the lives of many. A cloud-based system would enable easier sharing of the application and its outcomes. A catchall event log was developed to identify all events taking place within the ecosystem. The platform includes behavior health mentoring and integrated tracking for blood glucose, blood pressure, mood, food, weight, and sleep through Bluetooth-enabled devices, wearables, and self-inputted data.

AI models provide boundless opportunity in predictive analytics. By analyzing and generalizing on thousands of images, medical images are being turned into science. There are few ways as engaging and immersive as digital, which enables healthcare to collect enormous amounts of data; analyze and provide useful feedback to inform patients; and prevent, predict, diagnose, and treat disease.
By the very nature of imagery, de-identification for aggregated learning is technically simple. However, the requirement to be able to explain the mechanisms of deep neural network prediction is technically difficult. A more general understanding of the application alleviates most stakeholder concerns, which can be explained through demonstration.

As a novel innovation, there has been a requirement to explain how models are generated, validated, and continually evaluated. Regulatory approval has proved more laborious than previous projects due to the loose definitions that exist around product categorization.

Error minimization has been a priority, and regulatory approval has focused on ensuring patient safety. The risk of false positives was determined to be more likely: early ulceration detection means quicker healing and less risk of amputation.

Thresholds of concerns were developed within the model, with any doubt referred to a human.

Outputs of the project are being documented through clinical papers to highlight the value of the project to patients, healthcare providers, and global healthcare providers. Perhaps of most significance, the outcomes will demonstrate how people with diabetes can achieve significant health improvements by preventing and detecting diabetic foot problems.

Challenges

There were a number of inertia points that were observed throughout the duration of the project:

Quality of Data

Any model can only be as good as the data that is used to create it. Data completeness and correctness are paramount for a robust machine learning algorithm. There were several significant challenges with the datasets collected that facilitated the development of robust data governance processes. First, lack of labeling meant that data was unlabeled and often unclear on receipt.

Second, handwritten notes that supported image scans were often illegible. Incorrect spelling or incomplete doctor notes complicated the data validation process. The cleaning and vetting of data advanced a dataset that was established as a source of truth for the model.
Data Governance

The requirement for national and international regulatory approval required appropriate, transparent data governance processes to be in place. Governance-covered areas included data architecture, archiving, management, metadata management, privacy, security, and validation. Standards, best practices, and procedures were developed for data acquisition, verification, and validation.

A central repository held all relevant data and metadata, with access restricted by credentials to ensure appropriate user management. A governance trail for all decisions was noted as per the requirements for regulation, which provided an opportunity for reflective learning as the project ended. Stringent data governance ensured the integrity of data used within the project.

Performance vs. “Explainability”

Neural networks are notorious for being difficult to explain, particularly as models become more complex. A key requirement of this project was transparency in algorithmic decisions.

As the model developed, iterations of learning and backpropagation tailored nodes within the network that became unexplainable. Model performance and accuracy was secondary to model transparency.

Ethical Governance

Strict data ethics and information ethics governance procedures were developed alongside data governance processes to ensure ethical and moral use of data. The machine learning model was used by healthcare professionals to predict and diagnose risk of foot ulceration.

Processes around the effect of false positives and true negatives in particular were noted, as misdiagnosis was established to be the greatest ethical concern. The effects of ulceration and healing times can be greatly reduced through swift diagnosis and action.

As the application was used to support the actions of healthcare professionals, users were informed as to the limitations of the model, and it was explained how model accuracy was based on validation datasets.
Closing the Loop

The developed model predicted the likelihood of foot ulceration and other diabetic foot concerns. As this was to be used in the stages of prevention and diagnosis, a feature that was requested was the ability to validate and confirm model predictions. To do so, the healthcare professional using the application would either confirm diagnoses or refer the patient to a specialist for diagnosis.

The process of feeding back patient outcomes into the system enabled the system to learn in near real time. Data feedback from patients once referred to the hospital was difficult and cumbersome to gather.

Stakeholder Understanding

An unforeseen challenge was the lack of stakeholder understanding and sporadic resistance to technology-driven healthcare solutions.

Key stakeholders were typically biased by previous experiences or negative press or threatened by concerns of an age of robots taking jobs from humans. This was overcome with staff stakeholder training days; third-party experts and internal stakeholders explained how AI and machine learning could be used to optimize the tasks conducted by healthcare professionals, improve the resource burden, and mitigate concerns of a near-future robot race.

The project scope was developed with the input of a diverse, multifaceted stakeholder group to ensure all relevant parties were vested in the project.

Adoption Strategy

As a project that started through R&D funding, there was little incorporation of the pilot project into the wider organizational direction. Stakeholder comfortability and adoption of the novel innovation was progressed through workshops and hands-on experience. Workshops focused on mitigating stakeholder concerns and on the ease, efficacy, and resource-light environment of the predictive system.

Conclusions

Cloud-based intelligence that is sensitive and accommodating to differences in photography hardware and formats can be combined with patient clinical, demographic, behavioral, and genomic data to develop a precision medicine framework for detection of diabetic foot concerns.
The prototype was welcomed by healthcare professionals, with several remarking on elements that improved their own knowledge on the topic. Healthcare professionals felt more confident in diagnosing foot concerns when partnered with the digital tool. Digital assistance can confirm and support decisions and challenge healthcare professionals on how they arrive at their decisions. Applied across the spectrum of medical images and data collected by healthcare and research organizations, the technology will affordably enable large-scale prognostic, risk stratification, and best treatment decisions.

From our experience, having more data is almost always more important than the algorithm itself.

Overcoming the concerns of text digitalization and label recognition alongside medical imagery proves to be a challenging area within this field. Handwriting and human error have been the biggest problems to mitigate to reduce false recall. Conducting a large-scale clinical trial is the next phase of the project. A pilot project assessing the feasibility of the project is pending funding approval.

Case Study 2: Outcomes of a Digitally Delivered, Low-Carbohydrate, Type 2 Diabetes Self-Management Program: 1-Year Results of a Single-Arm Longitudinal Study

L. R. Saslow, PhD, University of Michigan, Ann Arbor, USA
Charlotte Summers, Diabetes Digital Media, UK
J. E. Aikens, PhD, University of Michigan, Ann Arbor, USA
D. J. Unwin, FRCGP, Principal in General Practice, The Norwood Surgery, Southport, UK

Background

Type 2 diabetes has serious health consequences including blindness, amputation, stroke, and dementia; and its annual global costs are more than $800 billion.

Although typically considered a progressive, nonreversible disease, some researchers and clinicians now argue that type 2 diabetes may be effectively treated with a carbohydrate-reduced diet, which could improve type 2 diabetes management and potentially even lead to remission [252]. Indeed, previous research with carbohydrate-reduced diets for type 2 diabetes does show improved outcomes (such as glycemic control, weight loss, and reductions in the use of hypoglycemic medications), for both
very-low-carbohydrate diets (roughly 20% or less of total dietary calories derived from carbohydrates) [253, 254, 255] and lower-carbohydrate diets (roughly 40% or less of total dietary calories derived from carbohydrates) [256, 257]. Although dietary interventions have historically been in person, online programs can be just as effective for some participants, as suggested by research that has examined diet and lifestyle interventions in adults with prediabetes [258]. Therefore, it is perhaps not surprising that the beneficial results of carbohydrate-reduced diets for people with type 2 diabetes (glycemic control, weight loss, and reductions in the use of hypoglycemic medications) have been replicated using online programs [259, 260].

Objectives
Our objective was to evaluate the 1-year outcomes of a digitally delivered Low Carb Program (LCP), a nutritionally focused, ten-session educational intervention for glycemic control and weight loss for adults with type 2 diabetes. The program reinforces carbohydrate restriction using behavioral techniques including goal setting, peer support, and behavioral self-monitoring.

Methods
The study used a quasi-experimental research design comprised of an open-label, single-arm, pre- and post-intervention using a sample of convenience.

From adults with type 2 diabetes who had joined the program and had a complete baseline dataset, we randomly selected participants to be followed for 1 year (N = 1,000; mean age 56.1, SD 15.7, years; 59% [593/1,000] women; mean HbA1c 7.8, SD 2.1, %; mean body weight 89.6, SD 23.1, kg; taking an average of 1.2 diabetes medications).

The Low Carb Program is a completely automated, structured, 10-week health intervention for adults with type 2 diabetes. Participants are given access to nutrition-focused modules, with a new module available each week over the course of 10 weeks. The modules are designed to help participants gradually reduce their total carbohydrate intake to < 130 g per day to meet their self-selected goals. The program encourages participants to make behavior changes based on “action points” or behavior change goals at the end of each module. These goals are supported with resources that are available to download including information sheets, recipes, and suggested food substitution ideas. The Low Carb Program online platform also includes digital
tools for submitting self-monitoring and device-driven data on a number of different variables including blood glucose levels, blood pressure, mood, sleep, food intake, and body weight. Weekly automated feedback is provided to users based on their use of the program through email notifications, and participants are notified when the next week’s module has been opened. Lessons are taught through videos, written content, or podcasts of varying lengths.

The program stresses the importance of regular contact with the participants’ healthcare providers for adjustments in medications in weeks 1, 2, and 10. After the 10 weeks’ worth of modules have been opened, participants continue to have access to the education content as well as the ability to continue to track their health (glycemic control, weight).

The content and strategies used in the program build off of prior research and theory. For example, evidence suggests that goal setting can act as an effective behavior change strategy used to improve adherence to lifestyle intervention programs in obesity management programs. The program therefore encourages participants to select a goal at the beginning of the program (such as to lose weight, reduce medication dependency, or make healthier choices for their whole family). Participants are also prompted to consider how their health would benefit from attaining their goal.

Throughout the program, participants are periodically prompted to consider how close they are to attaining their goal. The program further reinforces behavior change through integrated tracking, whereby program users are encouraged to track their health data including mood, food intake, blood glucose levels, weight, sleep, and HbA1c.

According to the control theory of behavior change, monitoring goal progress—that is, evaluating one’s ongoing performance relative to the standard—and responding accordingly is critical to goal attainment. Recent findings suggest that program interventions that elevate the frequency of progress monitoring are likely to induce behavior change.

**Results**

Of the 1,000 study participants, 708 (70.8%) individuals reported outcomes at 12 months, 672 (67.2%) completed at least 40% of the lessons, and 528 (52.8%) completed all lessons of the program.
Of the 743 participants with a starting HbA1c at or above the type 2 diabetes threshold of 6.5%, 195 (26.2%) reduced their HbA1c to below the threshold while taking no glucose-lowering medications or just metformin.

Of the participants who were taking at least one hypoglycemic medication at baseline, 40.4% (289/714) reduced one or more of these medications. Almost half (46.4%, 464/1,000) of all participants lost at least 5% of their body weight. Overall, glycemic control and weight loss improved, especially for participants who completed all ten modules of the program.

For example, participants with elevated baseline HbA1c (≥ 7.5%) who engaged with all ten weekly modules reduced their HbA1c from 9.2% to 7.1% (P < .001) and lost an average of 6.9% of their body weight (P < .001).

**Observations**

- The engagement platform used by patients for this study was only available on the Web and not as a mobile app. It would be expected that a mobile app would improve engagement.

- The criteria for inclusion within the project were the requirement to be over 18 and the ability to speak English.

- The percentage of individuals with an HbA1c level of < 6.5% increased from 26% (257/1,000) to 50% (503/1,000). This degree of control, when achieved through pharmacotherapy, is often accompanied by weight gain and risk of hypoglycemic events [261]. As the now famous Action to Control Cardiovascular Risk in Diabetes (ACCORD) study reported, intensive hypoglycemic medical therapy “increased mortality and did not significantly reduce major cardiovascular events” [262].

- A limitation was the rate of delivering the entire intervention, as only 528 (52.8%) completed all modules. However, a high rate (70.8%) reported 12-month outcomes. This could be due to the program being launched in November, with Christmas and seasonal activities affecting the rate of completion. On the other hand, given that this program was entirely automated and had a wide reach, a large number of individuals were able to complete the program.
Conclusions

Especially for participants who fully engage, an online program that teaches a carbohydrate-reduced diet to adults with type 2 diabetes can be effective for glycemic control, weight loss, and reducing hypoglycemic medications.

Case Study 3: Delivering a Scalable and Engaging Digital Therapy for Epilepsy

Charlotte Summers, Diabetes Digital Media, UK

Background

Epilepsy is a neurological condition of the brain. Different epilepsies are due to many different underlying causes [263]. The causes can be complex and sometimes hard to identify. A person might start having seizures because they have one or more of the following: a genetic tendency, structural changes in the brain, or genetic conditions [263].

Epilepsy is sometimes referred to as a long-term condition, as people often live with it for many years or for life [264]. Although generally epilepsy cannot be “cured,” for most people, seizures can be “controlled” (stopped) so that epilepsy has little or no impact on their lives. Treatment is often about managing seizures in the long term.

Most people with epilepsy take antiepileptic drugs (AEDs) to stop their seizures from happening [264]. However, there can be side effects with such medications, and there are other treatment options for people whose seizures are not controlled by AEDs, including the ketogenic diet (KD) [265].

This case study documents how the organization elucidates machine learning project ideas within teams and how an evidence-based approach is used to support decision-making to solve real-world problems.

Implementing the Evidence Base

The positive effects and therapeutic mechanisms of dietary ketosis and those of a ketogenic diet (KD) on human physiology have been well documented in literature. The KD is a high-fat, low-carbohydrate (usually less than 50 g/day), adequate-protein diet [266], whose metabolic effects originate back in the 1960s; however, the therapeutic
effects of KD can be traced back to the early 1920s when it was successfully used in the treatment of epilepsy [267].

Ever since, the potential clinical utility of KDs has been investigated by several studies, resulting in an accumulation of scientific evidence on the therapeutic role of KDs on various physiological disorders. Several studies have shown that the KD does reduce or prevent seizures in many children whose seizures could not be controlled by medications.

Over half of children who go on the diet have at least a 50% reduction in the number of their seizures. Some children, usually 10–15%, even become seizure-free [268]. The Epilepsy Society supports the KD as a treatment option for patients over 12 months old. Research highlights the therapeutic effects of the KD on patients with epilepsy as primarily reduced medication and reduced seizures. In a recent Cochrane systematic review of the evidence regarding the effects of KDs, Levy and Cooper found no randomized controlled trials (RCTs) [269].

This demonstrates that epilepsy can be managed; and in some cases, patients can live seizure-free lives through the sustained application of a ketogenic way of eating [270]. Through reducing sugar in the diet and at the same time improving blood glucose control, people with epilepsy can achieve significant health benefits such as reducing their risk of seizures and number of medications [271].

Sensor-Driven Digital Program

The Ketogenic Program for Epilepsy (KPE) is a structured education behavior change program for people with epilepsy. The KPE empowers users to sustainably adopt a lower-carbohydrate lifestyle with the appropriate personalized education, health tracking facilities, support, health mentoring, and resources to maintain a safe and sustainable lifestyle. As a result, people with epilepsy could expect to improve glycemic control and as a result reduce the incidence of seizures, reduce medication dependency, and improve confidence in managing their condition.

Health data is collected in real time through blood glucose monitoring, medication monitoring, and seizure tracking modules within the application. Wearables are used to collect data to sense seizures and falling. Unstructured data is particularly useful in understanding sentiment and the patient’s psychology.
Research activity takes place over the course of 12 months, reporting on 3-month, 6-month, 9-month, and then 1-year epidemiological health outcomes and engagement data followed by an 18-month and 2-year follow-up to demonstrate efficacy and adherence to the program over the short to medium term.

This is the world’s largest and longest study on patients engaged with a digital platform for epilepsy.

**Research**

As part of this project, we seek to learn the most efficient and effective methods to implement and accelerate at-scale delivery of our technology to the United Kingdom’s NHS population and international bill payers. The project is led by an award-winning and experienced team, recognized for innovating digital health, who are determined to revolutionize the health and well-being of people across the globe.

The technological challenge is to create a scalable, engaging, and effective solution for global implementation. This challenge is mitigated through the project extending on the infrastructure of DDM’s Low Carb Program, with clinically validated outcomes and 71% retention at 1 year. At the time of writing, out of 300,000 patients globally registered within the Low Carb Program for type 2 diabetes, there are 3,112 people with epilepsy (0.5% of the UK population)—showing the readiness for an online, nutrition-focused intervention for this population.

The potential cost-saving impact to health bill payers is significant. Patients who are empowered and have improved health and well-being are also more active members of their communities and have lower social care needs.

**Project Impact**

The greatest impact will be felt by patients with epilepsy and their families through empowering patients to manage their epilepsy. Evidence demonstrates that a KD in people with epilepsy can be maintained by infants and adults with improvements in number of seizures and medication consumed. Evidence also demonstrates positive effects on cardiovascular disease risk [272]. The health improvements and potential cost savings are sizable.
In addition to health improvements, the project is of tremendous value to the UK economy and the NHS, with reduced complications, improved mental and emotional health, reduced hospital admissions, and reduced medication having a direct impact on the budget of the UK Treasury.

Positive social impacts experienced by the patient have a repercussion for wider communities, businesses, and the UK economy.

As the result of improved patient and population health, organizations will benefit from fewer sick days and reduced absenteeism, improved mental health, and improved perceived quality of life, as well as a reduction in the perceived burden of managing epilepsy. The positive impact is also received by the clinical healthcare community—through reducing physician and healthcare professional burden and enhancing the evidence base.

**Preliminary Analysis**

The objective of the research arm of this project is to evaluate the 1-year outcomes of a digitally delivered Ketogenic Program: a nutritionally focused, 16-session educational intervention for epilepsy control through provision of a KD.

The program assists patients to achieve ketosis using behavioral techniques including goal setting, peer support, and behavioral self-monitoring. Interesting correlations between elevated blood glucose and mood enabled predictive algorithms to be developed to identify symptoms of impending seizures, alerting patients through notifications to take action.

Especially for participants who fully engage, an online program that teaches a carbohydrate-reduced diet to adults with epilepsy can be effective for glycemic control, weight loss, and reducing medication and the number of patient seizures.

A randomized controlled trial is required to understand the clinical impact of such a digital therapeutic.

**Case Study 4: Improving Learning Outcomes for Junior Doctors Through the Novel Use of Augmented and Virtual Reality**

Paul Duval, University of Liverpool, UK
Vidhi Taylor-Jones, University of Liverpool, UK
Background

Junior doctors arrive at a university eager to help mankind, yet often they are unaware of how to deal with high-pressure, critical scenarios. Dealing with medical concerns appropriately—whether emergencies or not—requires years of experience and learning.

Over time, exposure in doctor surgeries, hospitals, and clinics develops the self-assurance, readiness, and experience required for junior doctors to realize their vocation. High-pressure scenarios, such as a patient having a cardiac arrest, require pinpoint precision accuracy in decision-making and actions.

For some time after the first arrival to a university, junior doctors are not ready to make such decisions. This is where augmented and virtual reality provides an opportunity to engage students in learning that readies them for life as a doctor. When you are a doctor, every decision has consequences—and that can be burdensome for some.

What’s more, with 1,500 students in any one year, group trips to hospitals are expensive and logistically difficult to manage.

The project was recognized for its pioneering nature, winning an Association for the Study of Medical Education (ASME) Education Innovation Award.

Aims

The aims of the project were manifold:

- Increase the exposure to simulation and training environments for junior doctors without the use of simulation centers.

- Give inexperienced doctors the experience they require in a noncritical environment.

- Demonstrate effective and ineffective clinical practice.

- Encourage students to consider their own practices and thought processes.

- Immerse students in a clinical learning environment and identify whether the use of virtual reality enabled them to become familiar with the healthcare setting and patient unpredictability and provided the confidence to cope with high-pressure scenarios.
Project Description

It was agreed that the project has as wide a reach as possible. This eliminated virtual reality hardware such as the HoloLens or Oculus and focused the delivery onto mobile devices. The Samsung Gear and Google Pixel were able to facilitate the delivery of the content with the use of a headset to extend the view into virtual reality. This was logical considering almost every student owns a phone, whereas the use of HoloLens and Oculus is still relatively niche.

Simulations of a variety of clinical experiences were filmed with 360-degree videography using both doctors and patients as actors. These simulations were short 5- to 10-minute videos that featured decision moments where students needed to make an active choice—including a patient suffering from cardiac arrest and a patient with schizophrenia becoming increasingly agitated.

These videos were integrated into a virtual reality application that enabled students to explore a clinical setting in virtual reality and engage in the environment and encouraged them to make a variety of decisions during the process. The application encouraged users to “wear” the headset for a full, immersive experience.

The use of this augmented reality application to educate junior doctors was then evaluated with students.

What Is 360° Video?

360° video is a recent technology that uses omnidirectional cameras to capture a spherical video space rather than the landscape dimensions of traditional photography. These videos are slotted together to create an immersive viewing experience, placing the viewer in a scene within an environment rather than an observer or fly on the wall. The viewer has the ability to explore the scene, turning a full 360°. Capturing 360° video requires specialist cameras.

Playback of 360° video is most often experienced on mobile devices, where the smartphone display acts as the window to the 360° environment. Viewers can physically move their phone to view other areas of the 360° video—controlling both the orientation and viewing.

Physical devices which allow the viewer’s device to be worn as a headset can be obtained to provide a fully immersive experience. Desktop computers can also play 360° video, in a similar manner to mobile devices. Interaction with 360° video on desktop usually involves a mouse to move through the environment. Hardware-driven headsets
can also be used to engage with 360° videos; however, this was deemed to be prohibitive given the objectives of the project.

**Conclusions**

VR-driven education was warmly welcomed by team members, students, and staff alike. After an initial “wow” factor of using 360° videos and/or virtual reality, participants using the headsets engaged in the surrounding and participated in VR-delivered training.

- Virtual reality simulation provided a prime and safe clinical environment for junior doctors to explore their practice.
- In particular, the technology allowed students to focus on mistakes—unifying approaches to medicine across the cohort.
- The technology enabled students to learn anywhere, anytime: at home, school, or library.
- Virtual reality simulations were accessible by all medical students with the appropriate hardware.
- Students were able to become more active in decision-making in lectures.
- The virtual reality experience encouraged discussion and learning.
- It was cheaper and logistically easier to manage than student trips to hospital wards.
- The use of this innovative technology prepares junior doctors for the demands of the vocation.
- Research and exploration into the appropriate pedagogies to support virtual reality learning is required as the tools become more freely available.

Immersive education is a route of delivery that should be further explored to improve students’ learning experiences, reduce costs, and further innovate healthcare.
Case Study 5: Do Wearable Apps Have Any Effect on Health Outcomes? A Real-World Service Evaluation of the Impact on Activity

Charlotte Summers, Diabetes Digital Media, UK
Arjun Panesar, Diabetes Digital Media, UK

Background

The United Kingdom, like many developed countries, is experiencing increasing rates of chronic conditions associated with diet and lifestyle choices such as obesity, type 2 diabetes, and cardiovascular disease [273]. Obesity has been shown to be a COVID-19 risk factor [274]. Evidence to support adoption of healthy lifestyles in the prevention and management of these and other long-term conditions is strong [275]. Less than one in ten patients newly diagnosed with type 2 diabetes attends structured education in the United Kingdom [276]. There are a number of reasons for this, including the fact they are in the working week and often inconveniently timed [277]. Over and above this, learned behaviors diminish after 6 months of attending an offline education program [278]. Improvements in management of type 2 diabetes can be achieved by reduction in carbohydrates, and this approach has been demonstrated to be the most effective approach to achieving normal blood glucose control and sustainable weight loss. Weight loss is a key factor in improving self-management of obesity and metabolic health conditions such as prediabetes [279–281].

Patient-facing digital health applications (apps) provide the opportunity to change the way individuals take responsibility for their own health by enabling more effective delivery of health information, allowing better monitoring of health markers, and encouraging lifestyle behavior change. Wearables (and wearable apps) have been demonstrated to increase activity in older adults and can contribute to weight loss [282–284].

Enthusiasm about the potential of health apps has grown rapidly over the last few years with many NHS services now commissioning digital health apps. While health apps offer the potential to augment care for a breadth of health conditions from mental health to medical conditions like hypertension, cardiovascular disease, and type 2 diabetes, in comparison with uptake, surprisingly little is known about their functionality or impact on health outcomes.
From the data that is available, there is a suggestion that some factors mainly predictive of uptake are not necessarily the same as those predicting completion or health outcome and very few health apps demonstrate long-term adherence [285] or report long-term health outcomes. This may, in part, be due to the lack of standards for collecting data across health apps [286].

Aspects of digital health apps that are demonstrated to impact behaviors and health outcomes should be evaluated to improve health service design and delivery. Research is needed to extend previous health app studies and understand the effectiveness of particular app features on engagement and health outcomes. Previous publications on Low Carb Program, a structured education and behavior change app, demonstrate its clinical effectiveness for supporting obesity management including statistically significant weight loss [287].

Patients’ behavior directly contributes to their treatment success, with doctors relying on patients to take their prescribed medication alongside making and maintaining dietary and lifestyle changes. Many of the most significant challenges in healthcare, specifically in long-term or chronic conditions, such as obesity and type 2 diabetes, will only be resolved if we can influence behavior and support sustainable behavior change among other things such as the food environment.

There has been a detailed review of the features of Low Carb Program with an analysis of the theoretical underpinning for the presence of each app feature. The aim of this study was to assess some of these app features and understand how the engagement with these features impacts health outcomes, in particular, weight loss.

Obesity is associated with significantly worse cardiovascular risk factors, suggesting that more active interventions to control weight gain would be appropriate to help address the increasing burden of obesity on the NHS. Regardless of the interventions used to lose weight—pharmacological or behavioral—the weight is commonly regained [288–290].

Typically, half the weight lost is regained in the first year. Weight regain often continues up to 3–5 years after treatment; and, on average, 80% of people return to or exceed their pre-intervention weight [291].

Similarly, relapse rates are high for individuals who initiate attempts to stop smoking and those who try to reduce alcohol consumption [292–294].

Therefore, effective interventions that consider known factors associated not only with initial weight loss but also critically with weight loss maintenance such as building on internal motivations to lose weight, establishing social support mechanisms,
identifying coping strategies, or providing support for self-efficacy and autonomy can all enhance weight loss maintenance, which is crucial for the long-term success of any weight loss interventions [295].

Methods

Research Design

The study used a quasi-experimental research design consisting of an open-label, single-arm, pre- and post-intervention. Participants were not paid for their participation; members were referred by their NHS healthcare professional team to access the program and consented to analyses of de-identified data. Each R&D committee confirmed that as the research was a service evaluation, it would not require ethical approval.

The program was accessible as an app for smartphone devices running on iOS and Android with companion apps for Apple Watch 3 and above and Wear OS watches running Android 8.0 or later with Wear OS 2.7 or above running on the Android smartwatch.

Participants

Participants were recruited and referred into the platform through their NHS healthcare professional. A total of 200 users were found from people with type 2 diabetes referred to the Low Carb Program between January 1, 2019, and April 30, 2019. The population’s health and engagement data was analyzed for 6 months post-registration. All participants were from England, United Kingdom.

Low Carb Program

Low Carb Program is a behavior change app for patients with type 2 diabetes, prediabetes, and obesity. The app provides NHS-approved structured education for patients with type 2 diabetes, prediabetes, obesity, metabolic syndrome, polycystic ovarian syndrome (PCOS), and non-alcoholic fatty liver disease (NAFLD). The program educates patients to reduce the sugar in the diet to 130 g/carbs per day over a core 12-week implementation phase.

The platform provides users the ability to track their health (weight, HbA1c, blood glucose, food, mood, medication, ketones, cholesterol, blood pressure), track their nutrition with a food diary, and connect the app with a wearable (Apple Health, Google Fit, Fitbit, Nokia, Garmin) and prompts users to check in at regular intervals. The app is
used within the NHS and has been demonstrated to facilitate weight loss in patients with obesity, prediabetes, and type 2 diabetes [287].

The platform provides native-language education and culturally relevant support for South Asians (Punjabi) and is available on responsive Web, iOS, Android, Apple Watch, Android Watch, Alexa, and VR (Oculus).

The platform is built on a peer-reviewed behavior change architecture developed to support long-term behavior change within the app. Components include goal setting, tailored education, community support, integrated tracking, tailored resources, coaching, and AI-powered engagement based on behaviors and health status. Figure 10-1 visualizes the behavior change architecture.

**Figure 10-1.** Behavior change architecture used in the Low Carb Program
Wearable App: Low Carb Program for Apple Watch and Wear OS Watch

The Low Carb Program’s wearable app is available on Apple Watch and Android Watches, also known as Wear OS. The companion Low Carb Program watch app provides wearers with four functionalities.

- Health tracking: Companion app users can track their weight, blood glucose, mood, activity, steps, blood pressure, and heart rate.
- Relax: Participants can choose a set duration of time to focus on breathing. The smartwatch displays an animation intended to soothe and displays instructions on when to inhale and exhale.
- Lifestyle: Participants can read education and send articles to their phone from a daily-updated library of articles.
- Motivation: A screen within the companion app that displays random motivational quotes curated to prompt reflection and goal reappraisal.

Measures

At baseline, on registration to the smartphone app, participants were asked to report on their type of diabetes, year of diagnosis, weight, current medications, age, gender, dietary preference, socioeconomic status (based on weekly food budget), and presence of comorbid chronic illnesses.

At 3 and 6 months, participants were asked to report on their current weight and medications for reappraisal. Participant weight, activity (steps), and engagement data was extracted where available.

The research explored how many participants remained engaged at 6 months. It also explored the respective activity, weight loss, and engagement with the platform(s) for all users.

Statistical Analyses

Analyses were performed using the SPSS version 21.0 (SPSS Inc., Chicago, IL, USA). We conducted a multiple regression analysis to determine the predictive power of engagements with Low Carb Program app features on weight loss. The primary outcomes were engagement, body weight, and steps.
Outcomes were also analyzed within strata based on participant’s Low Carb Program completion (i.e., completers, engaged with at least 9 of 12 of the Low Carb Program weekly modules, n=117; partial completers, engaged with 4–8 modules, n=49; or non-completers, engaged with ≤3 modules, n=34.

Results took into account all of the sample, regardless of follow-up information or lesson completion. Any patient within the cohort who had activated their account was included.

For participants who did not report their outcomes at 6 months, we followed the highly conservative approach of assuming that they did not improve at all (last observation carried forward), by imputing their baseline values as their outcome values. For example, participants who did not comply with reporting a 6-month outcome were treated as having no change in the outcome variable and thus were not counted as having any weight change.

**Results**

**Participant Characteristics at Baseline**

At baseline, mean weight was 95.51 kg (SD 22.36), and mean age was 56.86 years (SD 10.75). Just over half were female (59%, 118/200), 88% (176/200) were white, and all participants were from England, United Kingdom.

**Platform Usage and Engagement**

Over half of the referred participants (53%, 106/200) registered on the Web. Over half of participants (121, 54%) accessed the program from an iOS device and 79 (39.5%) from an Android device. A minority of participants (1%, 2/200) downloaded the Alexa app.

Of the 200 participants tracking their activity through a wearable (Apple or Android Watch), 84% (168/200) were engaged and had outcomes at 6 months (defined as actively inputting a piece of data within the last 7 days). For the 32 people lost to follow-up at 6 months, the last recorded data point was carried forward.

**Wearable Engagement**

For members who completed the program, participants using a companion app who completed the program reported weight loss of 7.38% (7.34 kg, t_{116}=10.9247, p<0.001).
## Weight Change

| Status                      | 6-Month Weight (kg), Mean SD | Weight Change (kg), Mean (SD) | Weight Change (%), Mean (SD) | P Value |
|-----------------------------|------------------------------|-------------------------------|-------------------------------|---------|
| All participants (n=200)    | 90.43 (20.63)                | -5.07 (6.99)                  | -4.99 (6.31)                  | <0.001  |
| Completers (n=117)         | 89.61 (21.8)                 | -7.34 (7.27)                  | -7.38 (6.72)                  | <0.001  |
| Partial completers (n=49)  | 93.51 (19.13)                | -3.21 (6.35)                  | -2.78 (4.39)                  | .0009   |
| Non-completers (n=34)      | 88.81 (18.52)                | 0.02 (0.44)                   | 0.02 (0.52)                   | .8478   |

For the 117 participants using a companion wearable app who completed the program, the average number of steps taken a day increased by 54.2% (t116=11.57, P<.0001) by the end of 6 months. The number of average steps taken by users was significantly higher 6 months after registering for the app.
Discussion

This was not a randomized controlled experimental trial, so we cannot compare the 6-month results to a control or standard-of-care group. However, of 200 patients with obesity referred to the platform, 84% of those who used a wearable were still engaged at the 6-month follow-up which supports previously reported outcomes of the effectiveness of Low Carb Program.

Although the study design does not support inferences on the causality, there are statistically significant correlations between use of the wearable and reduction of body weight.

The study found that members who either complete (>9 modules) or partially complete (>4 modules) the program report greater weight loss than those who do not complete the program.

Members who complete the program lose a significant 7.38% of body weight at 6 months. Building on prior evidence for the use of the platform by patients with type 2 diabetes to achieve improvements in glycemic control and sustained weight loss, the results of this study go to demonstrate that for participants who fully engage, an automated online program delivered by smartphone and companion app teaching a carbohydrate-reduced diet and providing tailored on-demand workouts to adults with obesity can facilitate weight loss.

Main Findings

Our initial hypothesis that patient engagement with the wearable companion app would be predictive of uptake of lifestyle behavior change appears to hold true. In particular, participants completing the program and using a wearable companion app achieve significant weight loss.

Participants using the platform’s companion wearable app who completed the program lost more weight than all other strata of participants analyzed. Members using the companion app who completed the program lost more than 7% of their body weight at 6-month follow-up. These members also walked more—over 50% more at 6 months than they did at baseline. It is noteworthy that baseline data was provided to the patient by their healthcare professional on referral.

However, subsequent health outcomes (weight, steps) were input using self-report and wearable trackers, rather than measuring them through clinical health records.
However, previous research has found that these self-reported health outcomes can be quite close to actual values [296, 297].

A limitation of this study was that the analysis can only attribute active companion app feature engagements with health outcomes. There are a number of passive engagements that any given patient may undertake that would be deemed an engagement within the app but would not be explicitly tracked.

One such example would be peer support community; for “lurkers” who read posts but do not actively engage by asking or answering questions, this patient group can make up to 75% of a digital community population. There may be a significant impact on the likelihood of adopting lifestyle behavior change and subsequent health improvements, but these are a lot more difficult to monitor.

Conclusions

The aim of this study was to evaluate whether a companion wearable app contributes to weight loss. Participants tracking their health using a companion wearable app demonstrate significant weight loss at 6 months, as well as increased levels of activity.

The findings support previous studies that demonstrate wearable apps can improve health outcomes. Further analysis should be done to explore differences among populations, in particular if different ages, genders, ethnicities, and socioeconomic statuses impact health app engagements and subsequent health outcomes.

Case Study 6: Big Data, Big Impact, Big Ethics: Diagnosing Disease Risk from Patient Data

Dominic Otero, Diabetes.co.uk, UK

Background

The variety and velocity of patient data has exponentially increased over recent decades. This data has not progressed the robustness of decision-making. This leads to questions as to why decision-making is not becoming more data driven and what can be done to validate and expedite its adoption.

Diabetes.co.uk is the world’s largest diabetes community, based at the University of Warwick, and the platform welcomes over 45 million visitors a year. As the world’s
diabetes community, the organization provides a range of services that focus on the patient data ecosystem to empower patients to manage and control their health.

By collecting patient data, or real-world evidence, over time, the ecosystem is created such that it offers insights, education, and feedback on particular aspects of the patient’s health.

**Platform Services**

The platform collects a wealth of data. From the moment of platform access, behavioral and engagement metrics are collected, with patients progressing onto registering for the platform, opting in to part with over 120 variables that cover their diabetes and overall health status.

Once inside the platform, a variety of data types are collected—unstructured data in the form of conversations, interactions, engagements, and behavior and structured data in the form of self-reported, device-driven, wearable, IoT, and clinical health record data. The platform provides a number of services to patients and bill payers.

**Medication Adherence, Efficacy, and Burden**

Medication adherence, adverse effects, and wastage are concerns that pharmacology seeks to overcome. Patients with diabetes are typically on at least one medication on entry to the platform.

Within the platform, patients evaluate their medications at regular intervals, sharing their own opinions on the use of medicine in the real world, and report more structured data such as side effects, adherence, efficacy, and burden of using medications. This data is mapped geographically, by condition, medication, and other markers, and fed back in a de-identified, aggregate manner to pharmacological and academic partners for research purposes.

Data is used by partners to identify new drugs and understand interactions, usage, adherence, and real-world adverse effect prevalence. The same data is also used to improve patient safety. For instance, the platform alerts members on particular medications or using particular devices should there be an issue or recall.
Big Data: Pooling People to Empower Health Decisions

The Diabetes.co.uk community forum is the world’s largest support community for people with diabetes and provides an unrivaled insight into the global diabetes population. The platform welcomes people with all types of diabetes, carers, and healthcare professionals. The concept is simple: real people with diabetes talking about their own experiences and speaking to others who are in a similar position. Patient engagement is high; there is over 2 million years of cumulative experience among members, who spend over 17 minutes each visit on the platform.

The impact of participation in online health communities on well-being has been mainly studied in relation to the informational and emotional support that members receive from their community and the positive impact that such support has on their health condition [298, 299].

Community members provide and receive informational support by sharing practical information that can help them manage their own health condition. Emotional support involves community members seeking mutual understanding and comfort by sharing their story about how it can be frustrating and painful to live with a health condition [300].

Equally, not everyone is positive about health communities. There has been a long-standing debate among healthcare professionals regarding the quality of information shared online, as well as unintended consequences of health self-management, particularly among those with low health literacy [301, 302].

As a novel digital community, data and respective insights from the online forum and dissemination of academic findings are constantly shared within the Diabetes.co.uk patient support network, organizational medical advisory board, partners, and other interested networks.

A first-phase study by Bernardi and Wu from Royal Holloway, University of London, investigated the role of online health communities in patient self-management, assessing the following:

- How patients appraise knowledge from online health communities in managing diabetes
- How a patient’s healthcare professionals perceive and can potentially use such knowledge to innovate their clinical practice
Bernardi and Wu developed a model to investigate a) the factors that influence the level of engagement in online communities and b) the impact of engagement in online communities on members’ cognitive, behavioral, and emotional experiences of health self-management. The model was tested empirically through an online survey with the users of the Diabetes.co.uk forum [303].

The study aimed to provide a holistic perspective on the impact of a properly managed online health community on patients’ well-being.

The study concluded the following:

1. The Diabetes.co.uk forum is a catalyst of innovation.
2. The Diabetes.co.uk forum empowers patients, meaning
   - A good relationship with their healthcare professional
   - High confidence or self-efficacy in managing their condition
   - Feeling less emotionally burdened

This was supported with data from 13,000 patients who highlighted the patient benefit through platform usage:

- 53% improved blood glucose control
- 58% improved quality of life
- 59% improved confidence in managing diabetes
- 65% improved dietary choices
- 76% improved understanding of diabetes

### AI Prioritization of Patient Interactions

To meet regulatory requirements, any health platform must report patient health or safety concerns immediately on discovery. This would prove cumbersome for a developed health platform such as Diabetes.co.uk if completely human led for several reasons:

- Monitoring of medical discussion groups and interactions requires human-intensive staffing to spot concerns for regulation purposes.
- Lots of time would be wasted if humans were to do all the checking.
It was decided that to meet our obligations, the use of AI bots would speed up the process, enabling humans to spend time with the concerns that mattered. This was realized through the development of a neural network–based classifier to determine sentiment and concern based on contents, user profiles, and frequency of other key metrics.

The classification system went through several hundred rounds of validation, taking a subset of the huge data lake that was available from the millions of posts. The classifier was able to spot 99.8% of all concerns, with humans required to intervene in 0.2% of cases. There were a number of benefits to the AI prioritization:

- It saved a considerable amount of human time.
- Human time was spent on the problems that were of most concern.
- The flagging of discussions occurred in near real time rather than the 15–20 minutes it would take a human to do so.

This information also assisted the customer support teams to efficiently deal with user queries by knowing exactly what a conversation was about, typical responses to such a query, and what to say to a user to ensure a favorable outcome for all parties involved. Additionally, by gauging the sentiment of users and storing it, user data can be implemented to build additional systems in the future.

**Real-World Evidence**

Through the analysis of patient data, we can determine a discrepancy between official and real-world health statistics. For instance, 14% of people with diabetes in the United Kingdom are expected to have a mental or emotional health concern, whereby the platform sees 44% of users with this health condition [304]. This highlights that further exploration and understanding of the patient population is severely needed to ensure critical aspects of patient care are not neglected. Partners have become more interested in real-world evidence as the years have progressed, partially to do with concerns over research funding and conflicts of interest.
Ethical Implications of Predictive Analytics

Predictive analytics enables intelligent classifiers to predict the risk of disease. The collection of a varied patient dataset—including conversations, health markers, demographics, and behavior—enables the development of a sophisticated model that can be used to refine predictions further. One of the unintended consequences of the project was discovered when deploying an AI model that predicted the risk of a patient having pancreatic cancer.

Pancreatic cancer typically affects 1 in 10,000 people and is frequently misdiagnosed. Often, patients receive a diagnosis when it’s too late. A pancreatic cancer AI model was developed using a clinical dataset of health metrics, engagement, and behavior data and first predicted the likelihood of poor patient metabolic health. Should patients meet the threshold for poor metabolic health, the model evaluates the risk of the patient having pancreatic cancer.

Communicating this information to users was a first-time problem for the organization. All relevant stakeholders, including all members of the Medical Advisory Panel, were prompted for their input on this topic. Predictive analytics had previously been used to inform patients of possible hypertension or increased risk of conditions based on lifestyle.

However, informing patients of their risk of pancreatic cancer was considered to be unethical, particularly as we did not want to create additional mental health concerns or anxieties.

How should a user be told of a possible life-changing concern? The key action for this communication is to nudge the patient to see their doctor or physician as soon as possible. The factor for success in this case was determined to be user friction. The less friction there is between the user and communicating party, the more successful the engagement.

Stakeholders from behavioral economics were vital in testing and confirming an engaging communication to inform the user that their data fell outside expected norms without raising fear. To do this, the prediction was framed against patient and population demographic and medical data. Rather than being told they were the only patient to see anomalous data, patients were reassured that others have also seen similar patterns and been successful in improving their health.
Integration of the IoT

Sensors in wearables such as accelerometers, altimeters, and heart rate and skin temperature monitors provide a tremendous amount of useful information. For instance, motion, sweat molecules, and conversations are all useful and relevant data points. Such metrics can be used to take a snapshot of metabolic health, cardiovascular health, and mental health. For this to be conducted at scale, there must be little to no friction between the patient and providing parties.

The less friction there is between the user and whom they choose to share data with, the more successful the product. Wearables are becoming increasingly used in incentivized wellness propositions. Validating patient behaviors and general physiology facilitates the personalization of premiums and outcomes-based reimbursements.

Patients have the option to connect their devices and applications to the Diabetes.co.uk architecture. What is interesting is that younger users typically adhere to wearable usage and find it easy, while elderly users typically do not. This is of relevance to insurance and incentivized wellness propositions, as most claims come from older policyholders.

Conclusions

The realization of precision medicine is not without technical and ethical challenges. IoT and conversational and wearable data are just examples of a variety of data that can be collected and used to empower patients in real time, facilitating improvements in health and well-being.

With data and the application of AI come ethical concerns, particularly in the prediction of future disease risk. The lack of processes and governance on such topics is urgently required as the abilities of AI mature.
Case Study 7: Assessment of a Predictive AI Model for Personalised Care and Evaluation of Accuracy

Kelvin Summoogum, Founding CEO, MiiCare

Background

In order to provide personalized healthcare benefits and services to our users, our AI attempts to learn and interpret the user’s activity of daily living. This includes the behavioral pattern relating to their movement across the home, regular routines such as bathroom habits, time they wake up and go to sleep, and so on.

This forms the key foundation for all the predictive analysis that is carried out to create situational awareness and insights about the person living alone at home and the detection of abnormal activities that would warrant an alert to be raised.

In the previous trial where we achieved accuracies above 90% in identifying abnormal behavior based on the “states” and “transitions,” we had briefly discussed the potential of applying the Bayesian learning framework to automate user behavior prediction and enable proactive monitoring capabilities. In this paper, we define the problem statement for this application and outline the Bayesian approach taken to assess the effectiveness of a new predictive algorithm to add in our library.

Our initial assessment of the algorithm based on the data from a selected group of users shows a prediction of 82.2% accuracy, as an average. The accuracy relates to 10-minute windows within which the system is required to predict the location of the person. However, if the window is increased beyond 30 minutes, the accuracy reaches a maximum of 93.75%.

As the study progresses and more data about the subject is collected and the model trained on a wider dataset, we anticipate the accuracy of the 10-minute window to also increase. Some efforts are also being spent on the development of another predictive solution which uses state models for critical-time predictions without online connectivity.
Problem Statement

miicUBE network collects movement data in the form of sensor triggers. For every user’s home, the network comprises specialized sensors that detect motion installed in each room and passageways.

Every time the user passes by a sensor, the network registers the activity, with the timestamp of the event and the location where the sensor was triggered.

In our previous paper, we outlined how we had developed algorithms and formulae to visualize the daily routine of a user given sensor data collected over a period of time (days). However, that approach was static and didn’t allow for consideration of scenarios where the user changed their behavior or did something suddenly. Instead we required an approach capable of

1. Predicting where the user is supposed to be, at any point during the day
2. Highlighting instances where the user didn’t follow their routine
3. Continuously adapting to subtle differences in changes in behavior that happen over weeks and months as a result of ageing and cognitive decline

Expectation-Based Modeling Approach

This section outlines how we define the concept of a prediction window and train a Bayesian model for every such window to obtain a probability distribution of where the user is expected to be. This approach also allowed us to find events which are “unexpected” with respect to the model and continuously update our model with new data that is collected every day.

Prediction Windows

Earlier studies allowed us to deduct that a single predictive model cannot say where the user is in their home at a particular time of the day with a decent level of accuracy. This is because

- There are small events throughout the day which are impulsive or situational and are not indicative of true behavior.
Elderly people have a tendency to follow a fixed routine first thing in the morning, midafternoon, and late evening. These are task-based routine, like feeding the cat, taking medicine, watching the television, going out, and so on. But in between, they are more unpredictable mainly due to morbidities, frailty, or tiredness.

To overcome this issue, we divide the day into chunks or blocks of n minutes. The duration of this window defines our level of sensitivity to unusual behavior or events.

If the window is too big (40 minutes), we lose the ability to send early notifications. In contrast, if the window is too small (5 minutes), we compromise the accuracy of our predictive algorithm which would lead to a high rate of false positives.

Therefore, setting the window requires some level of pragmatism in terms of what is acceptable to a service provider or to an elderly person in the event of an emergency.

**Multinomial–Dirichlet Models**

Consider a prediction window of 15 minutes between 9 AM and 9:15 AM in the morning. If our user generally wakes up at around 9 AM, then we can expect sensor triggers in the Bedroom and the Bathroom. Perhaps, after the user has visited the Bathroom, they go to the Kitchen to make a cup of tea or coffee. So there can be Kitchen sensor triggers as well. If we are only monitoring for these three locations, we can visualize the distribution of the number of sensor triggers observed as shown in Figure 10-2.

![Figure 10-2. Sample visualization of sensor trigger counts](image-url)
This distribution of sensor trigger counts is an example of a multinomial distribution. Since the distribution values are sensor trigger counts, they directly translate into probability values when normalized.

For a prediction window of a day \( t \), we defined a multinomial distribution \((\text{location} = r)\) which tells us the probability of observing sensor triggers at a location \( r \). From Figure 10-2, we can say that

\[
(\text{location} = \text{Bedroom}) = \frac{20}{20+10+12} = \frac{10}{21} \approx 47.6\% \text{ between 9 AM and 9:15 AM.}
\]

The Dirichlet distribution is a powerful addition to a multinomial distribution: it serves to embed our prior personal knowledge about where we expect the user is at any point in the day. In the Bayesian framework, we always start with a prior knowledge of the probability of the event in question happening.

For instance, from our example in Figure 10-2, we know that the user will be in the Bedroom between 9 and 9:15 AM because they wake up at this time. We formalize this knowledge by defining an initial informative value for the parameter \( \alpha \) where \( \alpha_{\text{Bedroom}} > 1 \) and \( \alpha_{\text{other rooms}} = 1 \).

The value of \( \alpha \) is defined by the Dirichlet distribution \((\alpha)\) as

\[
\text{Dir}(\alpha) = \{ar\} \text{ for } r \in \{\text{all sensor locations}\}
\]

and it represents concentration power, alike to weightage. \( \alpha_r > 1 \) represents a higher concentration allocation to the probability of the user being in location \( r \), while \( \alpha_r = 1 \) represents an uninformative unbiased concentration, that is, the state of not knowing any special knowledge about that particular event happening.

Using the Dirichlet distribution allows us to specify certain known characteristics of the user and therefore add behavior personalization to the model.

For prediction window \( t \), our multinomial–Dirichlet distribution \((\alpha)(\text{location} = r)\) tells the expected probability of the user being in a location \( r \) by virtue of triggering the sensor at that location, given an initial prior knowledge. This definition forms the basis of our expectation-based predictive model \( ME \):

\[
ME = D\text{Dir}(\alpha t)(\text{location} = r)
\]

where

1. \( r \in \{\text{rooms where sensors are installed}\} \)
2. \( t \in \{\text{sequence of n-minute time blocks in a 24-hour day}\} \)
3. \( n \) = Duration of one time block in minutes

4. \( \alpha_t \) = Concentration parameter for the Dirichlet distribution for the predictive window \( t \)

**Conditioning the Multinomial–Dirichlet Models on Observed Data**

To compute \((\alpha_t)_{\text{location}=r}\), we compute the “observed” multinomial distribution for the prediction window \( t \) for every date in our “training” period. For example, if we set our training period to be between 1st and 7th of the current month and our predictive window between 9 AM and 9:15 AM in the morning, we must first compute the number of sensor triggers for every location for that window for every date in that week.

The result of this computation is a matrix where each row is the multinomial distribution of sensor triggers for a date in the training period while the column will represent the locations.

We then define a Bayesian model which consists of two distributions: \((\alpha)\) [initialized with a prior informative set of values] and \((\alpha)_{\text{location}=r}\). This model will condition itself with the observed data, that is, the distribution matrix. Once conditioned (“trained”), it can be used to generate matrices which have the same dimensions as the input matrix but now follows the true multinomial distribution of sensor triggers as learned from the input.

To compute the expectation for \((\alpha)_{\text{location}=r}\), we average the generated matrices column-wise and obtain an approximation of the true multinomial distribution.

**Evaluation of Predictive Accuracy**

To evaluate how performant the conditioned model \((\alpha_t)_{\text{location}=r}\) is

1. We compute the “observed” multinomial Distribution for the window \( t \) on a test date and find the first two locations \( LO1, LO2 \) with the highest observed sensor triggers.

2. We find the first two locations \( LE1, LE2 \) where the distribution \( Da_t \) has the highest expected probabilities.

3. We perform an equality between the observed and expected locations with highest probabilities as
   a. \( if \ LO1 = LE1 \ or \ LO2 = LE2, \ then \ we \ consider \ it \ a \ good \ prediction \ (“hit”) \).
b. if $LO_1 \neq LE_1$, then we consider it a bad prediction ("miss").

**Model Evaluation**

In this section, we evaluate the performance of our predictive model. We use six different durations of the prediction window so as to provide a comparison of variations in the predictive accuracy because of it.

Please do note that uniquely identifiable details about the clients have not been disclosed for privacy reasons.

Parameter values used are

1. Training period: March 2020 (31 days)
2. Testing period: 1st week of April 2020
3. $\alpha_t$ for all predictive windows $t$ and locations $r = 1$ (uninformative unbiased prior)
4. length of predictive window $n = 10, 30, 60,$ and $90$ minutes

For each of the evaluation table, 1) each row represents a day in the first week of April 2020, 2) each column heading represents a prediction window duration chosen for the trial, and 3) each cell value is the “hit rate,” which is computed as the number of times the model has a good prediction for a day.

User A is a 90+ elderly gentleman living in the United Kingdom. The user deployed activity sensors around ten locations in his home. Table 10-1 summarizes the hit rate for every day in the testing period with the different prediction window durations. We found that the model had a predictive accuracy of 80% when trained on one month of data and tested for 1 week of data.
Table 10-1. Evaluating predictive performance for User A

|     | 10 m | 30 m | 60 m | 90 m |
|-----|------|------|------|------|
| 1   | 75.9 | 81   | 81   | 90.4 |
| 2   | 80.6 | 83.5 | 81.4 | 88.1 |
| 3   | 77   | 81.4 | 77.3 | 88   |
| 4   | 70.5 | 76.5 | 76.5 | 82.8 |
| 5   | 76.9 | 77.5 | 81.4 | 82.2 |
| 6   | 72   | 79.4 | 77.5 | 82.1 |
| 7   | 71   | 71.7 | 74.4 | 77.4 |
| AVG | 74.8 | 78.7 | 78.5 | 84.4 |

We also expected a negative trend in the accuracy as we predicted further into the future as captured in Figure 10-3. The shaded region highlights the range of accuracy difference for using windows of 10 and 90 minutes.

User B is an 85+ elderly lady living in the United Kingdom. Her home has six activity sensors set up. Table 10-2 summarizes the hit rate for every day in the testing period with the different prediction window durations. For this user, we found that the model had a predictive accuracy of 84.4%.

Table 10-2. Evaluating predictive performance for User B

|     | 10 m | 30 m | 60 m | 90 m |
|-----|------|------|------|------|
| 1   | 74.2 | 89.17| 89.2 | 87.5 |
| 2   | 90.3 | 89.8 | 87.5 | 93.75|
| 3   | 82.8 | 87.5 | 79.2 | 93.75|
| 4   | 80.7 | 87.5 | 87.5 | 93.75|
| 5   | 82.7 | 83.3 | 87.5 | 87.5 |
| 6   | 77.4 | 85.4 | 83.3 | 87.5 |
| 7   | 76.4 | 77.1 | 80   | 82.5 |
| AVG | 80.6 | 84.3 | 83.4 | 89.4 |
Similar to User A, a negative trend is also observed for User B, as shown in Figure 10-4. The shaded region highlights the range of accuracy difference for using windows of 10 and 90 minutes.

**Figure 10-4. Prediction accuracy over time for User B**

**Improving the Solution**

In our trial for the two users, we set the value of $\alpha$ for all predictive windows as 1. This represents an uninformed unbiased state of the model before it has seen any data. Instead, we can manually define an informed biased value for $\alpha$ which reflects our
knowledge of the user’s presence during different points of the day. This can be a tedious manual task, but once defined can reduce the amount of data required for the model to fully condition itself to arrive at the real and precise expectation probabilities.

Increasing the duration of the predictive window boosts predictive accuracy; however, our objective is to enhance the model to bring the accuracy to exceed 90% prediction for 30-minute windows. Achieving such a high accuracy is very challenging given an elderly person may not necessarily stick to daily routine. However, with more data points, the predictive model will be able to learn about certain external factors which trigger an elderly person to deviate from their routines, for example, sickness, the effects of medications, and so on.

Put within the context of the intended use within a care home or a home setting, the model will predict where a person is “supposed” to be, and then the algorithm will compare this with where the person is for that particular window. In most cases, having a mismatch for one or two cycles of 30-minute duration is not going to be detrimental.

However, if, for example, we predict that the user will be in the kitchen in the next 30 minutes (first cycle) and then in the living room for another 30 minutes (second cycle), but the system detects that they are in the bathroom for more than 30 minutes, then an alarm will be triggered. At the same time, the system monitors their body vitals, and if something abnormal is observed, these will also be notified to a family member/carer.

Anything shorter than 30 minutes will very likely result in excessive false alerts given the behavioral pattern of an elderly tends to vary a lot. That said, anything abnormal with the body vitals is notified instantly, for example, low heart rate.

The two focus areas as we grow the fleet of miiCUBEs are user duration and data flows.

**User Duration**

Establish the duration a user has to be in the same location to trigger an alarm. This is currently a feature configurable by the user/family/carer; however, with enough data points, the machine learning model will be able to set those parameters automatically.

**Data Bandwidth**

The miiCUBE unit serves as an IoT hub to connect the network of sensors to the servers and also houses a small onboard computer for its other functions, such as having a conversation with the elderly person. Our predictive model, if implemented, will be housed in the server where it will be fed with data incoming from the miiCUBE for
its training and predictions. This workflow is dependent on the quality of Internet connection within the region where the user lives. If this connection is severed for a prolonged period of time, the network will lose its capability to proactively monitor for emergencies.

**Next-Generation Model**

The next version of the predictive model we will be exploring is state and transition models based on probabilistic graphical models (PGMs). This was mentioned in the previous white paper along with the Bayesian approach defined in this paper.

PGMs serve to represent, model, and infer probable courses of actions based on order-sensitive data. We define such a model using sensor locations as “states” and the connections between these locations as routes the user takes to move between them.

We can therefore isolate and detect emergencies like falls when the user is in a location (“state”) or when in transit (between “states”). We can also find probabilities for the user to be in a state or a transit between two states for a specific time of a day. Such a model can work locally in miiCUBE and will be beneficial when connectivity is not available, hence maintaining service continuity.

**Conclusion**

In this paper, we use the Bayesian framework as the key predictive framework of operation for miiCUBE’s sensor network. We define an expectation-based predictive approach where

- There is a multinomial–Dirichlet Bayesian model for every n-minute window of a day.
- The Bayesian models are conditioned on sensor trigger data for every location specific to the window of the day.
- The models state the top two locations where the user is expected to be at a specific window in the day.

This approach has three operating parameters: the training period, the duration of the predictive window, and prior values chosen for \( \alpha \) for every window of the day.
The predictive accuracy of the model is dependent primarily on the latter two where

- There is a trade-off between relevance of predictions and uncertainty in prediction for a model in a window. Increasing the window duration decreases prediction uncertainty while also decreasing how useful the prediction becomes in real-time monitoring.

- Defining strong informative prior values for $\alpha$ reduces the amount of data required to condition a model to arrive at the real precise expectation probabilities of the user’s location in the future.

The approach was able to achieve an average predictive accuracy of 82.6% in a trial involving two of our miiCUBE users.

For fair comparison, the training and testing periods along with the duration of the prediction window and values for $\alpha$ for all windows in a day were kept constant for both cases. The trial has been rolled out to all of our users.

Case Study 8: Can Voice-Activated Assistants Support Adults to Remain Autonomous, a Real-World Service Evaluation of the Impact of a Voice-Activated Smart Speaker Application on Weight and Activity

Charlotte Summers, Diabetes Digital Media, UK

Background

Regular and moderate physical activity practice provides many physiological benefits. It reduces the risk of disease outcomes and is the basis for proper rehabilitation after a severe disease [305]. Most importantly, regular activity can improve quality of life [306]. Physical activity has been demonstrated to increase fitness, physical function, cognitive function, and positive behavior in people with dementia and related cognitive impairments [307]. One study found that gaming systems are motivating and can have a positive effect on strength, balance, and overall fitness for elderly populations [308].
Patient-facing digital health applications (apps) provide the opportunity to change the way individuals, in particular the elderly, take responsibility and care for their own health by enabling more effective delivery of health information. New platforms and devices allow better monitoring of health markers and can act as tools to support and encourage lifestyle behavior change. Apps and wearables have been demonstrated to empower older patients to increase activity and can contribute to weight loss [309].

**Health Apps**

Enthusiasm about the potential of health apps has grown rapidly over the last few years with many global health services commissioning digital health apps. While health apps offer the potential to augment care for a breadth of health conditions from mental health to medical conditions like hypertension, cardiovascular disease, and obesity, surprisingly little is known about their functionality or impact on health outcomes. Only 15% of health apps are formally reviewed by a governing body in the United Kingdom [310].

Engaging in positive healthy lifestyle behaviors continues to be a public health challenge, requiring innovative solutions [311]. Voice-activated assistants are a form of conversational agent that include Amazon Alexa, Apple Siri, Microsoft Cortana, and Google Assistant. These agents are a form of human-centered computing which provides interaction between human and computer through sound-based input/output. This provides the opportunity to engage audiences in a novel way—through conversing with an agent that processes the speech of the user (invocation) and returning a statement, image, audio, or video as output to the user.

**Conversational Agents**

A conversational agent is a computer system intended to converse with a human [312]. Voice-activated assistants are connected to the Internet and listen for commands, interpret them, and take action.

**Smart Speakers**

Smart speakers are a form of speaker and voice command device, with an integrated virtual, voice-activated assistant that offers interactive actions and hands-free activation with the help of a wake word [313]. Smart speakers can run third-party apps. For Amazon Alexa, this is known as a Skill.
Voice-Activated Assistants

The use of smart speakers, or voice assistants, like Alexa, Siri, Cortana, and Google Assistant has soared with at least one in five homes in the United Kingdom estimated to be using them [314]. Amazon Alexa is the most popular smart speaker device [315]. Research has shown Alexa is primarily used for checking weather forecasts, playing music, and controlling other devices [316]. The use of devices such as Amazon Alexa provides a novel mechanism to engage new audiences and demonstrate potential to help older adults [317, 318]. Alexa uses over the weekends are more frequent than on weekdays, but overall usage tends to decrease over time [318].

Health and fitness apps for voice-activated assistants are considered a nascent area of smart care, shown by one of the first literature reviews being conducted only this year [321]. Indications suggest that voice-activated assistants are capable of reaching populations that are diverse, underserved, and hard to reach.

User satisfaction and receptivity to voice assistants are promising yet compared to smartphone apps, voice-enabled devices are very limited and lack evidence [321]. A review of evidence shows voice-activated assistants have been used to remind people to take medications, as pregnancy companions, to help find answers to questions, and to provide coaching [321]. Notably, evidence shows that the accuracy of apps could be improved [320, 321].

The accessibility, ubiquity, and conveniences of voice assistants have led voice-activated devices (e.g., smartphones, smart speakers) to be a significant component of a digital health ecosystem. Such an ecosystem has been envisioned especially for patients who have chronic conditions [322]. The anticipated benefits of voice assistants are also opined to be in elderly care and helping older adults live with more autonomy (“smart aging”) [323].

One study found use of an AI voice assist may be a practical method to deliver scalable individualized behavioral health coaching to address cardiovascular health in cancer survivors [324].

Cybersecurity

However, a key area of inertia for voice-enabled assistants, particularly if they are ever to be integrated within the EHR, is the real and perceived threat of privacy, cybersecurity, and exploitation [320, 325]. Others have raised questions about their longer-term disruptive impact on the consumption of information, user profiling, and people’s relationship with technology [325, 326].
It has been suggested that smart speakers can help older people stay more autonomous at home [327, 328]. As part of a “smart” future, smart healthcare could measure this in a number of ways—through wearables, connected devices, and smart speakers. There are a plethora of signals that could be used to measure activity—including active minutes, activities, calorie burn, and steps.

The aspects of digital health apps that are demonstrated to impact behaviors and health outcomes should be evaluated to improve health service design and delivery. It has been demonstrated that there is a low comfort level to new technologies, but with appropriate training, it could have significant roles in healthcare [329, 330].

Research is needed to extend previous health app studies and understand the effectiveness of particular app features on engagement and health outcomes.

Previous publications on Low Carb Program, a structured education and behavior change app, demonstrate its clinical effectiveness for supporting obesity and type 2 diabetes self-management including statistically significant weight loss, HbA1c improvements, and a reduced reliance on medication leading to type 2 diabetes remission [91].

There has been a detailed review of the features of Low Carb Program with an analysis of the theoretical underpinning for the presence of each app feature [16]. In 2019, a companion Low Carb Program Amazon Alexa Skill was launched for Alexa Echo and Echo Show. The skill enables users to converse with Alexa to watch all lessons of the program, update their weight, update their HbA1c, ask for the nutritional values of foods, and request on-demand exercise workouts, guided stretch workouts, and yogas [331].

Features included usage history (clicks/taps, date and time, device, screen size), age, gender, goal, health status, medication, dietary preference, allergies, and location. Features were transformed to reduce complexity, for example, by classifying age.

Real-World Digital Health Evaluation

As providers of the Low Carb Program as a multi-platform digital health intervention on the Web, mobile, smartwatch, and voice-activated assistants, there is an opportunity to assess voice-activated assistant engagement and outcomes and their real-world impact.

The aim of this study was to examine the profile characteristics and health outcomes for participants using the Low Carb Program voice-activated smart speaker Skill (app) as a health management tool to support sustainable health-promoting behaviors.
Objective

The aim of this study was to examine the profile characteristics and activity outcomes for participants using a conversational agent as a health management tool as part of their lifestyle.

Methods

A total of 27 patients with obesity, prediabetes, and type 2 diabetes were referred to Low Carb Program by their NHS healthcare professional in February 2020. Thirteen patients also downloaded the platform’s Amazon Alexa Skill.

All participants were from England, United Kingdom. Health conditions that would exclude healthcare professional referral were carnitine deficiency (primary), carnitine palmitoyltransferase (CPT) I or II deficiency, carnitine translocase deficiency, beta-oxidation defects, medium-chain acyl dehydrogenase deficiency (MCAD), long-chain acyl dehydrogenase deficiency (LCAD), short-chain acyl dehydrogenase deficiency (SCAD), long-chain 3-hydroxyacyl-CoA deficiency, medium-chain 3-hydroxyacyl-CoA deficiency, pyruvate carboxylase deficiency, acute porphyria, history of mental health disorders, or pregnancy.

All users were selected for analysis. Patient health and engagement data from use of the mobile and companion apps were tracked, collected, and analyzed over 4 months to determine if use of the Alexa Skill contributed to change in health-promoting behavior.

Data analyzed included demographic data (age, gender, ethnicity, location, employment status), weight, and amount of minutes spent in on-demand classes. Qualitative feedback on features was also assessed.

Low Carb Program

Low Carb Program is an evidence-based structured education and behavior change self-management platform. The platform has demonstrated clinical benefits including weight loss, reduction in medications, and supporting type 2 diabetes remission. The platform is built on an evidence-based architecture grounded in behavior change psychology [16].

The Low Carb Program provides NHS-approved structured education for patients with obesity, prediabetes, type 2 diabetes, metabolic syndrome, polycystic ovarian syndrome (PCOS), and non-alcoholic fatty liver disease (NAFLD) on general well-being
and aids patients to reduce the sugar in the diet to 130 g/carbs per day over a core 12-week implementation.

Participants engaging with the Low Carb Program show a directional increase in the number of active minutes as they progress through the intervention. To support this, on-demand activities were made available in-app. On-demand activities comprised guided exercise, yoga, and stretch classes led by qualified personal trainers.

On-demand activities were available on all app platforms, including Amazon Alexa where activities were presented in video or audio format to the participant in English only.

**Voice-Activated Assistant App: Low Carb Program for Alexa**

The Low Carb Program’s Amazon Alexa Skill is downloaded from the Alexa Skills store hosted by Amazon or by invoking the Skill by saying “Alexa, enable Low Carb Program” to an Amazon Alexa Echo or Amazon Echo Show smart speaker [331].

The Amazon Echo provides output as audio only, whereas Echo Show outputs in video also. Participants are required to link their Amazon account with their Low Carb Program account in order to utilize the profile recall invocations available to participants using the Skill.

**Engaging with the Alexa Skill**

Once enabled, the Low Carb Program Amazon Alexa Skill is invoked by saying “Alexa, ask Low Carb Program…” followed by one of a range of commands to start an activity:

- Give me an exercise.
- I want to warm up.
- I want to cool down.
- I want to get active.
- Give me a yoga.
- Give me a workout.
- Give me a stretch.

Participants are informed of the commands available on the Amazon Alexa Skill listing page and also within the app. Commands allow the user to converse with Alexa...
CHAPTER 10  CASE STUDIES

to watch all lessons of the program, update their weight, update their HbA1c, ask for the nutritional values of foods, and request on-demand exercise workouts, guided stretch workouts, and yogas.

Deep neural network recommender systems are used to make predictions about other resources, activities, and education based on the current user’s profile and behaviors. The neural network was trained on the behaviors of 77,713 people. An initial dataset of 105,394 records was extracted. On exploration, it was identified that only fully completed attribute sets would be used and the remaining 26,681 semi-labeled datasets would be used to validate the model in the real world.

Exercises output by the Amazon Alexa Skill were available as guided audios or guided videos, respectively, for the Amazon Alexa Echo and Amazon Alexa Echo Show. A total of 21 activities were available—7 yoga workouts, 7 HIIT exercise workouts, and 7 stretch workouts. Workouts are tailored to a beginner’s level of fitness.

After 12 weeks of using the Amazon Alexa Skill, users were sent a survey to request opinions and feedback on the Amazon Alexa Skill. Users have 4 weeks to complete the feedback survey.

Measures

At baseline, on registration to the app, participants were asked to report on their type of diabetes, year of diagnosis, weight, current medications, age, gender, dietary preference, socioeconomic status (based on weekly food budget), and presence of comorbid chronic illnesses. On completion of registration, participants were notified by email that an Amazon Alexa Skill was available and prompted to download it.

At 4 months, participant weight, active minutes, and engagement data was extracted where available. Active minutes defined as the “number of minutes an on-demand exercise, stretch, or yoga class was opened and not backgrounded” was collected.

The total number of active minutes was calculated over the study period. Participants using the Amazon Alexa Skill were sent a survey to request feedback on the features, functionality, ease of use, and comfortability.

The research explored how many participants downloaded and used the Amazon Alexa Skill to engage with the on-demand classes. It also explored the respective active minutes, weight loss, and sustained engagement with the platform(s) for all users.
Statistical Analyses

Analyses were performed using the SPSS version 21.0 (SPSS Inc., Chicago, IL, USA). We conducted a multiple regression analysis to determine the predictive power of engagements with Low Carb Program app features on active minutes.

The primary outcomes were engagement and downloads of the Amazon Alexa Skill and active minutes. Secondary outcomes were health outcomes including weight loss. Outcomes were also analyzed within strata based on whether participants downloaded the Low Carb Program Amazon Alexa Skill comparing Skill users (ASU, downloaded Amazon Alexa Skill, n=13) and non-Skill users (NSU, did not download Amazon Alexa Skill, n=14).

Results took into account all of the sample, regardless of follow-up information or lesson completion.

For participants who did not report their outcomes at 4 months, we followed the highly conservative approach of assuming that they did not improve at all (last observation carried forward), by imputing their baseline values as their outcome values.

For example, participants who did not comply with reporting a 4-month outcome were treated as having no change in the outcome variable and thus were not counted as having any weight change.

Results

For the 13 participants syncing their account with the Alexa Skill (requiring an Amazon account), 92% (12/13) were engaged and had outcomes at 4 months defined as actively inputting an item of data within the last 7 days.

Participant Characteristics at Baseline

At baseline, mean weight was 95.1 kg (SD 22.9 kg) and mean age was 61.7 years (SD 5.7 years) across all participants. Participants using an Amazon Alexa Skill had a mean weight of 91.7 kg (SD 21.7 kg) at baseline.

Two-fifths of participants were female (40.7%, 11/27), over half were white (55.5%, 15/27), and all participants were from England, United Kingdom.
CHAPTER 10  CASE STUDIES

Platform Usage and Engagement

Almost half (48.1%) of the participants downloaded the Amazon Alexa Skill (13/27).

Of the 27 total participants, 24 (88.9%) were active at 4 months. 85.7% (12/14) of non-Amazon Alexa users were engaged at 4 months compared to 92.9% (12/13) of the Amazon Alexa users (defined as actively inputting a piece of data within the last 7 days).

Amazon Alexa Skill Usage

Participants using the Amazon Alexa Skill registered 85.82 active minutes, an increase of 68.9% compared to the 50.86 registered active minutes by participants not using the skill.

| Status                          | Total Active Minutes (mins/day), Mean SD |
|--------------------------------|------------------------------------------|
| All participants (n=27)         | 67.74 (28.36)                            |
| Skill users (n=13)              | 85.92 (12.84)                            |
| Non-Skill users (n=14)          | 50.86 (28.61)                            |

Weight

Participants using the Amazon Alexa Skill lost an average of 1.77 kg across the duration of the study.

| Status                      | Baseline Weight (kg), Mean SD | 4-Month Weight (kg), Mean SD | Weight Change (kg), Mean (SD) | Weight Change (%) Mean (SD) | P Value |
|-----------------------------|-------------------------------|-----------------------------|-------------------------------|------------------------------|---------|
| All participants (n=27)     | 95.05 (22.94)                | 93.66 (22.86)               | -1.39 (3.22)                 | 1.49 (3.43)                  | 0.557   |
| Skill users (n=13)          | 91.69 (21.79)                | 89.92 (21.68)               | -1.77 (2.80)                 | 2.03 (3.25)                  | 0.041   |
| Non-Skill users (n=14)      | 98.17 (24.45)                | 97.14 (24.17)               | -1.03 (3.63)                 | 0.99 (3.63)                  | 0.308   |
Qualitative Feedback

As part of the user experience, a feedback survey was sent to participants using the Amazon Alexa Skill. Of the 13 Alexa Skill users/participants, 7 (53.8%) completed the survey.

Discussion

This was not a randomized controlled experimental trial, so we cannot compare the 4-month results to a control or standard-of-care group. However, of 13 participants using the Low Carb Program Alexa Skill, 92.8% were still engaged at the 4-month follow-up which supports previously reported outcomes of the engagement of Low Carb Program [91].

Engagement

Surprisingly, almost half of participants downloaded the Alexa Skill, demonstrating the rise in popularity of smart speakers. Healthcare professionals did not give all participants an Amazon Alexa on referral, and therefore an unaccounted bias in this study is ownership of an Amazon Alexa Echo or Echo Show. The study is intrinsically biased toward those who already own the smart speaker. If smart speakers are still considered an emerging technology, it could be reasonably inferred that the population the intervention engaged with would be considered “early adopters.”

Although the study design does not support inferences on the causality, there are statistically significant differences between the total active minutes registered by participants using the Alexa Skill and those who did not.

Total Active Minutes

In this study, Alexa Skill users logged statistically significantly more total active minutes (85.92 minutes, SD 12.84 minutes) than non-Skill users (50.86 minutes, SD 28.61 minutes), \( t_{(25)} = -4.05, p<0.01 \).

Participants using the companion Alexa Skill registered more than 30 minutes of activity on average as defined by total active minutes across the duration of the study compared to participants who did not use the Alexa Skill. There are several possible explanations for this, including the smart speaker’s more optimal user interface and ease of information retrieval through conversation.

Alexa’s command-and-repeat nature seems to suit engagement in activities. Usage of the Alexa Skill could have been influenced by COVID-19, which saw the wider public
engage with and consume digital public health services such as remote monitoring while staying safe and well at home.

Total active minutes was defined as the total number of minutes spent engaged in an on-demand exercise, stretch, or yoga class while the user interface is opened and not backgrounded.

The veracity of outcomes could be improved through use of wearable data to improve the accuracy of total active minutes. An on-demand class playing is used as a proxy for minutes the user is active, and engagement is assumed based on the video continuing to play. Surprisingly, no users who paired with Alexa also used a smartwatch with the app.

The possibility of leaving on-demand classes playing was mitigated by the platform’s auto-logout security mechanisms, as per medical software best practice.

Further exploration could investigate the use of wearables and Alexa to improve the precision of positive health behaviours such as time active.

### Weight Loss

In this study, Alexa Skill users had a greater weight loss (-1.77 kg, SD 2.80 kg) than non-Skill users (-1.03 kg, SD 3.63 kg), $t_{(25)}=0.59, p=0.56$. Alexa Skill users lost more weight than non-Skill users. There are several possible explanations for this.

Given that Alexa Skill users also logged more total active minutes, one could hypothesize that the Alexa smart speaker may be biased in ownership to those already interested in fitness. Indeed, the use of the app during COVID-19 may also contribute to its engagement.

The average age of participants using the Low Carb Program Alexa Skill was almost 5 years older than non-Alexa users, supporting other research that shows voice-activated assistants and smart speakers can assist the elderly to be more active and autonomous and improve social bonds [318, 321, 332, 333].

### Feedback

Of 13 Skill-using participants to complete a feedback survey evaluating the Alexa Skill

- Exactly three-quarters of participants (75%) found it easy to set up and 75% found it enjoyable to use.
- Three-quarters of participants (75%) were interested in new technology.
• The majority (76.9%) used the Skill for on-demand classes or meditation and mindfulness.
• Over a third (38.4%) of participants found using the platform on a new technology a comfortable experience.
• The majority of participants, 84.6%, would recommend the Skill to a friend.
• Over half of participants (53.8%) felt accountable to the Low Carb Program Alexa Skill, and three-quarters (75%) felt more active because of using it.
• Almost all participants (92.3%) stated they would like training on how to use the Skill.

The outcomes support evidence that older populations are interested in new technology, yet feel uncomfortable using them without appropriate training. There was a clear signal that participants expect onboarding support for new technologies.

The study contributes to the evidence base around smart care. There are very few studies available evaluating use of digital health apps for smart speakers.

The study ignites other research questions based on app platforms, including whether the use of other companion apps for smartwatch and augmented reality or virtual reality may impact health or engagement with the platform.

**Conclusion**

The aim of this study was to assess whether use of the Low Carb Program Alexa Skill would contribute to improvements in positive health behaviors. Participants using the Alexa Skill demonstrated improvements in total active minutes.

The findings support previous studies that demonstrate voice-activated assistant apps can improve health outcomes. Participants who use the Alexa Skill were motivated to try innovative technologies, older in age, and desire training to overcome low comfort levels. Further analysis should be done to explore differences among populations, in particular if different ages, genders, ethnicities, and socioeconomic statuses impact engagement and subsequent health outcomes.