Digital Therapeutics: Virtual Coaching Powered by Artificial Intelligence on Real-World Data

Harm op den Akker, Miriam Cabrita and Aristodemos Pnevmatikakis*

Innovation Sprint, Brussels, Belgium

An ever-increasing number of people need to cope with one or more chronic conditions for a significant portion of their life. Digital Therapeutics (DTx) focused on the prevention, management, or treatment of chronic diseases are promising in alleviating the personal socio-economic burden caused. In this paper we describe a proposed DTx methodology covering three main components: observation (which data is collected), understanding (how to acquire knowledge based on the data collected), and coaching (how to communicate the acquired knowledge to the user). We focus on an emerging form of automated virtual coaching, delivered through conversational agents allowing interaction with end-users using natural language. Our methodology will be applied in the new generation of the Healthentia platform, an eClinical solution that captures clinical outcomes from mobile, medical and Internet of Things (IoT) devices, using a patient-centric mobile application and offers Artificial Intelligence (AI) driven smart services. While we are unable to provide data to prove its effectiveness, we illustrate the potential of the proposed architecture to deliver DTx by describing how the methodology can be applied to a use-case consisting of a clinical trial for treatment of a chronic condition, combining testing of a new medication and a lifestyle intervention, which will be partly implemented and evaluated in the context of the European research project RE-SAMPLE (REal-time data monitoring for Shared, Adaptive, Multi-domain and Personalised prediction, and decision making for Long-term Pulmonary care Ecosystems).

Keywords: digital therapeutics, virtual coaching, artificial intelligence, real world data, digital biomarkers, phenotypes

INTRODUCTION

More than 30% of the population above 15 years of age in 27 OECD countries (Organisation for Economic Co-operation and Development) lives with at least two chronic conditions and this percentage is expected to increase in the upcoming decades [Organisation for Economic Cooperation and Development (OECD), 2019]. Traditional healthcare delivery cannot respond to the resulting personal and socio-economic needs, requiring the uptake of digital health solutions. Digital Therapeutics (DTx), are a specific form of digital health solutions that "deliver evidence-based therapeutic interventions that are driven by high quality software programs to prevent, manage, or treat a medical disorder or disease. They are used independently or in concert with medications, devices, or other therapies to optimize patient care and health outcomes" (Digital Therapeutics Alliance, 2021). In other words, DTx can leverage the use of novel sensing devices combined with state-of-the-art artificial intelligence techniques to help individuals in the prevention, management,

or treatment of medical conditions. In the specific case of chronic conditions, management and treatment often go beyond management of symptoms in daily life and require the individual to change health-related behaviours, such as food intake, physical activity, and alcohol consumption. However, changing behaviour is difficult as it often concerns habits that are ingrained over years or even decades in the life of individuals. Ubiquitous technology such as smartphones and wearables design facilitate personalized DTx interventions that observe, understand, and coach the individual in their daily life. The first step—observing—consists of collection of data that is, meaningful for the management or treatment of the disease. By applying artificial intelligence techniques to the data collected, we reach the second step—understanding what is going on—for example, by finding interactions between clinical outcomes. The third and last step—coaching—concerns providing information to the user that is, understandable and actionable in a way that supports individuals with chronic conditions in reaching the desired behaviours. In DTx interventions, coaching combines theories of behaviour change with preferences and needs from users and healthcare professionals to define personalized strategies that help individuals breaking old patterns of behaviours or creating new healthy habits, in a way that is, informative, persuasive, and engaging. In this article we describe a novel approach to coaching in DTx powered by artificial intelligence on data collected in the daily life of the individual.

Virtual Coaching is the technology-mediated process of engaging in coaching conversations with an individual. For example, a real-life coach that assists the individual over a phone call or a teleconference service is seen as a form of virtual coaching, as the process is mediated by technology. This form of virtual coaching facilitating human-human interaction (e.g., healthcare professional and therapist) is already established in the healthcare domain and its market in the United States is expected to grow in 2021 at a 2.7% rate and 5.4% in 2022 (WebWire, 2021). However, in this paper we are focusing on an uprising specific form of virtual coaching in which technology plays a more central role—coaching that is, delivered through mobile or web apps, with no, or very little, direct intervention of a human being. This innovative way of virtual coaching is emerging in the healthcare domain, mainly targeting the promotion of healthy lifestyles (Tsiouris et al., 2020). Although not engaged in the interaction between machine and patient, in this form of virtual coaching, the healthcare professional is often available remotely as a safe-net to supervise the decisions of the automatic system. The virtual coaching can be delivered in a myriad of different ways—from short text messages [e.g., (Proctor et al., 2020)] to website-like content [e.g., (Murray et al., 2017)], from instructional videos [e.g., (Dekker-van Weering et al., 2017)] to realistic three-dimensional (3D) embodied conversational agents [e.g., (Poggi et al., 2005)]. Our approach focuses on applying virtual coaching to DTx interventions delivered through conversational agents allowing interaction with end-users using natural language. Conversational agents, also referred to as virtual agents, are receiving growing interest in the healthcare domain (Provoost
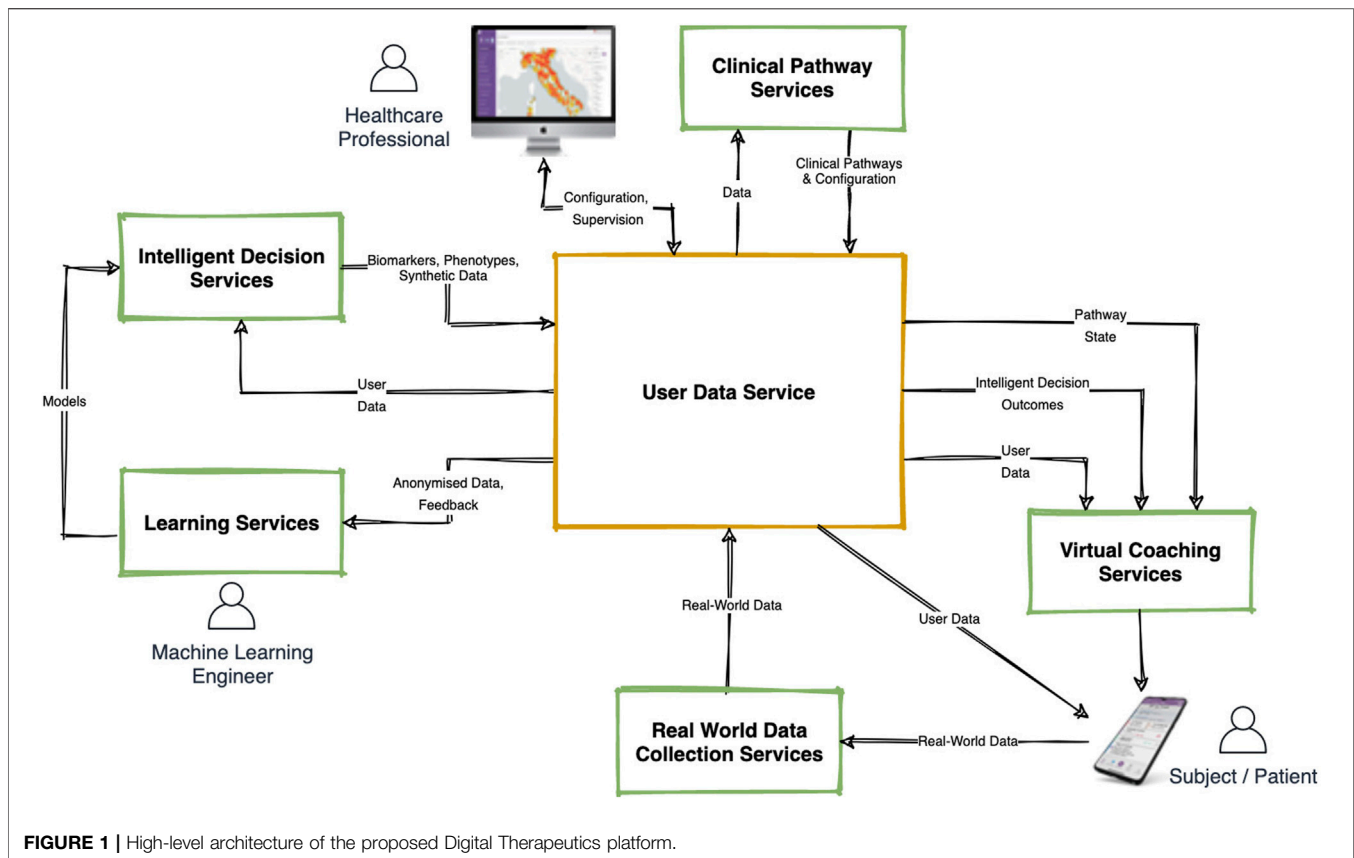
et al., 2017; Laranjo et al., 2018; Kamali et al., 2020). However, specifically when applied to the field of chronic conditions, the literature is still scarce (Schachner et al., 2020). Reasons for applying conversational agents in DTx are the demonstrated benefits to help building user trust (Bickmore and Cassell, 2001), support long-term engagement to interventions (Bickmore et al., 2010), or even short-term engagement when applying a bit of humour (Olafsson et al., 2020).

In this article, we present our approach to virtual coaching empowered by Real-World Data (RWD) capturing and processing in the Healthentia platform. Healthentia is an eClinical platform that captures clinical outcomes from mobile-, medical- and IoT devices, using a patient-centric mobile application and offers AI-driven smart services. Healthentia is a Class I Medical Device with a CE mark (Conformité Européenne) currently being used in more than 20 clinical studies. In this article, we introduce the new generation of the platform—Healthentia DTx—that will leverage on the knowledge obtained by the AI-driven services to deliver personalized virtual coaching in digital therapeutics. In the section "*Methods and Materials*" we start by giving a general overview of the Healthentia DTx platform, describing its main components. Next, we take a closer look to the services behind the three core purposes of the platform: observe, understand, and coach. The section ends with a short introduction of a fictitious study inspired in a current research and innovation project that serves as use-case to explore the potential of the proposed architecture to intervention delivery. In *Results* we provide a detailed walkthrough of the system and its DTx features following the described use-case in the Healthentia DTx platform. Finally, in the Discussion section we highlight some of the challenges associated with bringing this type of DTx to the market.

METHODS AND MATERIALS

The Healthentia DTx platform that we describe in this article consists of several interconnected components as shown in **Figure 1** below. This architecture shows the two user-interface components: the Healthentia Mobile App (bottom right), which is the main point of entry to the system for the primary end users (Subjects or Patients) and the Healthentia Web Portal (top left) which is the main entry point for Healthcare Professionals. The central component—the User Data Service—is a cloud service that handles storage of- and access to all the different types of data that is, collected or generated by this Digital Therapeutics platform. The main digital therapeutic functionalities are provided by the five surrounding services: Real World Data Collection Service, Learning- and Intelligent Decision Services, Clinical Pathway Services and the Virtual Coaching Services. The inner workings of each of these services will be discussed in more detail below.

We begin by describing in short how Digital Therapeutics is provided by the Healthentia platform, and the roles that each of these services play in enabling personalized virtual coaching for the Patients. Healthentia is a Software-as-a-Service platform that is, built to support various types of clinical studies and digital

FIGURE 1 | High-level architecture of the proposed Digital Therapeutics platform.

therapeutic interventions. For simplicity's sake, we will discuss the functionality of Healthentia from the view point of Digital Therapeutic Interventions with their end users being Patients (even though they might be better classified as Citizens or Subjects in different contexts). The starting point of the high-level walkthrough of Healthentia's functionality is the configuration of a specific intervention.

The healthcare professional can securely login to the Healthentia Web Portal in his role as an intervention administrator, to allow setting up new interventions. In this process, the professional starts by providing basic information such as the name, description, and logo of the intervention. Next, they can enable or disable various modules based on the needs of the Patients—for example, should they be asked to log their liquid intakes, or measure their daily physical activity? Enabling such modules will enable the corresponding functionalities on the Healthentia mobile app.

Next, the healthcare professional uses the Clinical Pathway Services to define the pathway (or care plan) that patients traverse. A clinical pathway is a set of states and transition rules between those states that allows the system to automatically track the progress of each patient in their care plan. This pathway can be defined at a high level—for example, pre-operation, in hospital, post-operation would be a three-state pathway describing the patients journey from before, during and after a hospital visit. The Clinical Pathway Services allows certain functionalities of the application to be linked to the

different states. Continuing the example, in the pre-operation state the patient's mobile app may be configured to measure baseline health status and coach towards improving physical condition [e.g., to facilitate better-in-better-out (Bongers et al., 2021)], during the in-hospital state, application functions may be fully disabled, while during the post-operation state, the virtual coach in the mobile application may be configured to provide post-surgery information and coaching. But the definition of a pathway is not limited to such high-level states, and states may be defined to represent anything, and transitions can be defined based on a large number of measurable parameters (time, questionnaire results, health outcomes, etc.). A behavior change intervention may include a pathway that defines the various stages of change according to Prochaska's model (Prochaska and DiClemente, 1986), a fixed weekly schedule, or a complex model based on many different RWD parameters. After having been configured, the Clinical Pathway is stored in the User Data Service and the intervention is ready to start registering Patients.

Once having been invited to participate in the intervention, Patients may start by downloading the Healthentia mobile app, registering their personal account and providing their electronic consent to participate in the intervention. Patients may then connect various devices to their Healthentia app (e.g., Fitbit or Garmin activity trackers) if enabled in the configuration stage to start collecting e.g., daily physical activity and sleep data. Patients will also start receiving scheduled questionnaires, and may report

various health symptoms or events at their own initiative. This process of observing the Patient through real-world data collection is described in more detail in *Observing: Real-World Data Collection* below.

The Real-World Data being collected is used in several ways. First, the Patient can view his own information from within the mobile app in order to increase awareness of his/her own behaviour (an essential part of the virtual coaching part as described in *Coaching: Virtual Coaching in Digital Therapeutics*). Second, healthcare professionals (mostly in case of clinical studies) are able to view the Patient's data, at the individual level, or through various dashboards that provide overviews over groups of patients. This way of using real-world data is not the focus of this article, as it plays a more significant role in the context of clinical studies, compared to the digital therapeutic scenario. The third way in which RWD is used is through deriving higher level of understanding of the Patient's behaviour, health status and progress through the course of their clinical pathway. Using various machine learning methods, the raw "first-order information" is used to derive valuable knowledge and understanding of the patient. This process is described in more detail in *Understanding: Real-World Data Processing*, and provides an important input to the process of virtual coaching described in *Coaching: Virtual Coaching in Digital Therapeutics*.

Observing: Real-World Data Collection

The Food and Drug Administration (FDA) defines Real-World Data (RWD) as "data related to patient health status and/or the delivery of health care routinely collected from Electronic Health Records (EHRs), claims and billing data, data from product and disease registries, patient-generated data including home-use settings, and data gathered from other sources that can inform on health status, such as mobile devices" (US Food and Drug Administration, 2017). The Healthentia DTx employs two of these broad categories of RWD, namely measurements from mobile devices, manufactured either for medical or wellness purposes, and patient-generated reports.

The measurements are automatically reported by the respective device to the data collection system, without the intervention of the patient (other than an initial configuration to connect the device to our platform). These measurements are so called objective RWD, since their quality only depends on the devices' measurement accuracy.

The RWD we measure have to do with physical activity, the heart, and sleep. Regarding physical activity, we measure steps, distance, elevation, energy consumption, and time spent in different zones of activity intensity. Regarding the heart, we measure the resting heart rate and the time spent in different zones of heart activity. Regarding sleep, we measure the time to bed and waking up time (so also the sleep duration) and the time spent in different sleep stages.

The measurements are obtained by the Healthentia DTx either directly from the devices, or from the platforms of the device manufacturers. Direct access to devices is achieved *via* Software Development Kits (SDKs) provided by the manufacturers. The SDKs are used in our patient companion mobile application to collect the data directly from the device. The mobile application periodically communicates the data to the RWD Collection services. Indirect access from manufacturers' platforms is achieved utilizing endpoints made available in their Application Programming Interfaces (APIs). The API endpoints are used by components in the RWD Collection services running periodically. SDK access is preferred since the patients' data only reside in the Healthentia platform and the collection process is spared the delay of device synchronization with the external platform and that platform returning the data to Healthentia, as well as the constraints the manufacturers impose in accessing their platforms. But unfortunately, in many cases SDK access is not supported, and in most of those providing it, the extra cost of using the SDK is prohibitive. Some medical devices do provide SDK access though. API access is ubiquitous and almost always free. In some cases, the recipient party (the Healthentia DTx) needs to apply for access, but this appears to be most often granted for medical purposes.

Patient-generated reports are formally termed Patient-Reported Outcomes (PRO), defined by the FDA as "any report of the status of the patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else" (US Food and Drug Administration, 2009). Assessing patients' status with PRO facilitates understanding of their experience, especially as regards to symptoms being experienced (Grey, 2017). The reports collected by our measurement system usually refer to symptoms related to a disease, functional statuses, or answers to complex questionnaires. But they can also go beyond such self-assessments, actually conveying measurements taken by the patient using a non-integrated device, whereupon the patient enters the data into the system. Patient-generated reports constitute subjective RWD, since their accuracy depends on the patients' understanding of the assessment, on their objectiveness and on their accuracy in data entry.

The reports we ask the patients to generate have to do with self-assessments of physiological, psychological, and social aspects of their life. The physiological aspects include but are not limited to important symptoms for the particular therapeutic area, nutrition (foods and liquids consumption), mobility, self-care, usual activities, and discomfort. The psychological aspects include a generic five-level self-assessment of mood and any specialized questionnaires, e.g., on anxiety or depression, important for the therapeutic areas. The social aspects are collected using questionnaires that gauge the patients' involvement with their social circle, peers, or community.

The patients generate the reports by answering simple or complex questionnaires pushed to them by the Healthentia mobile application as a result to some daily, weekly, or monthly timing of each questionnaire. In some cases, especially for symptoms, spontaneous reports are allowed. Finally, reports that should be compiled throughout the day usually are aided by widgets the patients can use to modify the reported value. A typical example of this is the collection of water intake, where patients are offered a widget to update the water they have been consuming in near real-time. The collected
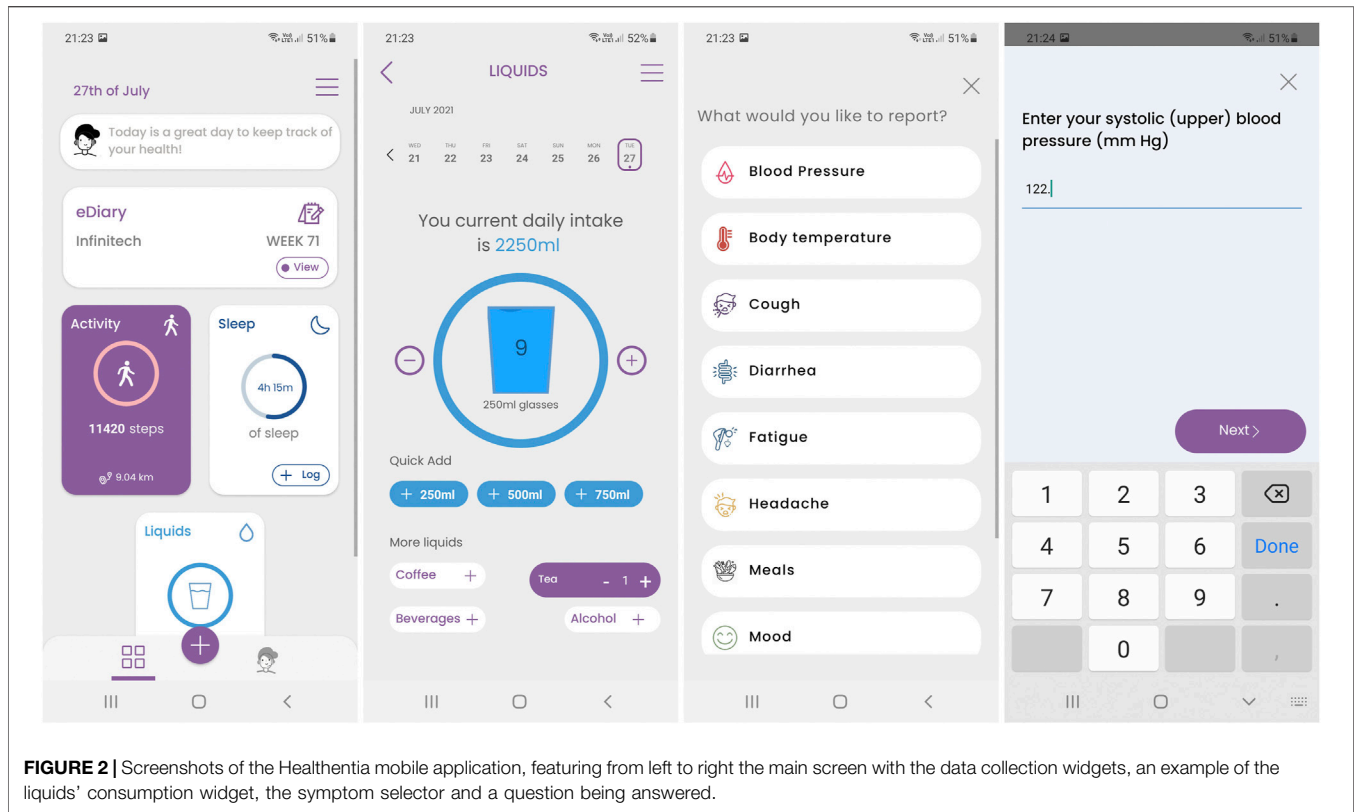
FIGURE 2 | Screenshots of the Healthentia mobile application, featuring from left to right the main screen with the data collection widgets, an example of the liquids' consumption widget, the symptom selector and a question being answered.

reports are forwarded by the mobile application to the RWD Collection services of Healthentia. The RWD Collection services abstract the actual information sources, and they offer the collected data to the User Data service for storage.

Evidently the reports, but in some cases also the measurements are facilitated by the Healthentia mobile application. The application serves as a patient companion, to facilitate RWD collection and offer RWD visualizations as shown in **Figure 2**. It also hosts the virtual coach, as discussed in *Coaching: Virtual Coaching in Digital Therapeutics.*

Understanding: Real-World Data Processing

The Healthentia DTx regards the collected RWD as the first-order information to be gained about the patients. The system then proceeds to uncover second-order information in a process carried out by the close collaboration of its learning and the intelligent decision services. The simpler forms of second-order information are obtained using signal processing. They involve long and short time averages, as well as trends of the different collected quantities. More elaborate processes employing Machine Learning are also involved in uncovering second-order information. First, decisions are reached on important clinical outcomes, in the form of predictions from composite biomarkers. Then, the outcome decisions are analysed to model the attributes of the patient mostly influencing the decisions. Finally, the patients are clustered into phenotypes.

Quantities that characterize a clinical outcome or a disease stage are enumerated using biomarkers (Coravos et al., 2019). The biomarkers are considered digital if they are collected using sensors and software or hardware computational tools. In their simple form, biomarkers are values of a single quantity. When the quantity is one of the factors or products of the disease that allows diagnosing a disease outcome or stage, the biomarker is termed direct. Direct biomarkers are usually measured in a clinical setting. Indirect biomarkers are values of quantities indirectly associated to, but highly correlated with, a factor or product of the disease. Indirect biomarkers are usually obtained using ubiquitous devices in an everyday life setting. Single quantities can be correlated with outcomes, but can better do so when combined into composite biomarkers (Kovalchick et al., 2017), the non-linear combinations of multiple quantities into a single value characterizing some outcome. In simple cases the combination can be done analytically, with the Body Mass Index (Garrow and Webster, 1985) being a typical example. There, the body mass and height are non-linearly combined in a simple equation, yielding a simple composite biomarker used in obesity.

The analytic combinations are the exception though. In most cases composite biomarkers enumerate models that perform the non-linear combinations (Guthrie et al., 2019). These models are learnt using Machine Learning algorithms (Geron, 2019) and they can be classifiers, if the biomarker values are finite and discrete, or regressors if they are continuous. E.g., in (Guthrie et al., 2019) Random Forests (Breiman, 2001) are employed to

learn a predictive model (binary classifier) that predicts if the systolic blood pressure of patients is expected to improve or worsen. The learning of the models whose outputs yield the composite biomarkers of interest is the first learning process facilitated by the Learning Services of the Healthentia DTx. It is an offline process involving different expertise. First, domain experts are consulted regarding the outcomes of interest and the direct or (mostly) indirect biomarkers (single quantities) that are expected to influence them, or whose influence on the outcomes needs investigation. These quantities are employed as attributes to be non-linearly combined by the predictive model to be learnt, forming the attributes input vector. The outcomes to be predicted form the output vector. A predictive model with multiple outputs can be used to evaluate multiple composite biomarkers, instead of multiple models, each with a single output. The model training involves the collection of the data via the user data services, and its processing for measurement error correction, missing values imputation, and anonymization (Jaidan et al., 2019). The processed and anonymized data are split into three sets, the training, validation, and testing ones (Pnevmatikakis and Polymenakos, 2009; Moghaddam, 2002; Manli Zhu and Martinez, 2006; Baudat and Anouar, 2000; Breiman, 2001; Schmidhuber, 2015). Candidate models are trained using the training set and their performance is evaluated using the validation set. Model parameters are tuned, and the process is repeated until the validation set errors are minimized. The final model is trained with the optimum parameters using the combination of the training and validation sets, while the performance is evaluated using the testing set. The final model is stored using the user data services.

The Intelligent Decision Services of the Healthentia DTx evaluate the learnt models to yield the composite biomarkers. There are timed components in the Intelligent Decision Services that periodically ask the user data services for new data regarding the patients, enumerate all the attributes of the predictive models and evaluate the composite biomarkers. The resulting composite biomarkers are stored as second-order data for the patients using the user data services. Unlike model learning, the data used to evaluate the composite biomarkers is not anonymized but is processed for measurement errors and is imputed for missing values.

The evaluated composite biomarkers (decisions of the predictive models) are analysed for the most important single quantities, i.e., predictive model attributes, influencing the predicted outcome. SHapley Additive exPlanations (SHAP) analysis (Lundberg et al., 2020) is employed in this analysis, enumerating as a SHAP coefficient the importance of each model attribute value in the particular decision for the given patient.

The SHAP coefficients are important in the Learning Services of the Healthentia DTx. Accumulated across patients, the SHAP coefficients can indicate attributes of small impact on the decisions overall. Such attributes can be ignored in subsequent offline model learning processes, allowing for the model complexity to decrease without any performance penalty, or in cases of small training sets, even increasing model performance. Individual SHAP coefficients on the other hand are employed in the second learning process facilitated by the Learning Services, this time an online one. The Healthentia DTx builds models for every attribute and every patient, yielding the overall importance of the attribute in the particular outcome for that patient. These models are continuously updated as new decisions are reached and SHAP coefficients are made available via their analysis.

Finally, there is a third type of models to be learnt for the patients: their phenotypes. Phenotypes are models of appearance, from the viewpoint of the therapeutic area, of similar groups of patients. The appearance is established using the most impactful first-order attributes and the second-order data, especially the time averages, trends, and composite biomarkers. These attributes form the phenotype clustering space, and their values form the data vectors to be clustered. These vectors provide a snapshot of a patient for a particular time interval (usually day or week, but actually adjustable to whatever the particular therapeutic area requires). Clustering is performed using a unsupervised Machine Learning algorithm to uncover structure in the data, like k-means or hierarchical clustering using trees (Theodoridis and Koutroumbas, 2008). As a result, every patient's snapshot is assigned to a cluster. Once a patient's snapshot is clustered, all their data attributes, not only those selected for the phenotype clustering space, are considered to belong to that cluster. Some of these attributes are then selected as representative of the patient in the particular therapeutic area and form the phenotype modelling space. Each cluster is described via a model for the attributes in the phenotype modelling space. Gaussian Mixture Models are used for modelling, trained by the Expectation-Maximization algorithm (Theodoridis and Koutroumbas, 2008).

The learnt phenotype models are utilized in the intelligent decision services for patient phenotyping. Each patient snapshot in time can be assigned to one of the phenotype models. Each patient is considered as exhibiting a particular phenotype for a time interval (typically much longer than the time interval of a snapshot), if the weight of that phenotype exceeds those of all others. These weights are updated every time a new snapshot of the patient is clustered. Transitions to different phenotypes are observed but are not frequent. Usually, they correspond to a transition to a new disease stage or new behavioural habits. No matter the reason, such transitions are considered important and are being monitored.

Coaching: Virtual Coaching in Digital Therapeutics

The final process in completing the Digital Therapeutics platform is the delivery of virtual coaching to patients. Introduced as the Virtual Coaching Services box in **Figure 1** in the beginning of this section, the virtual coaching service is actually a complex sub-system, consisting of many interconnected parts that together power a conversational agent that can hold natural language dialogue-based interactions with the patients on various coaching and support related topics.

The conversational agent may be accessed at the initiative of the patient, or the virtual coach may call the attention of the patient in case there is something important to discuss. When engaging with the virtual coach, either in response to a notification or at the patient's own initiative, the patient is presented with the virtual coach user interface in which he/she
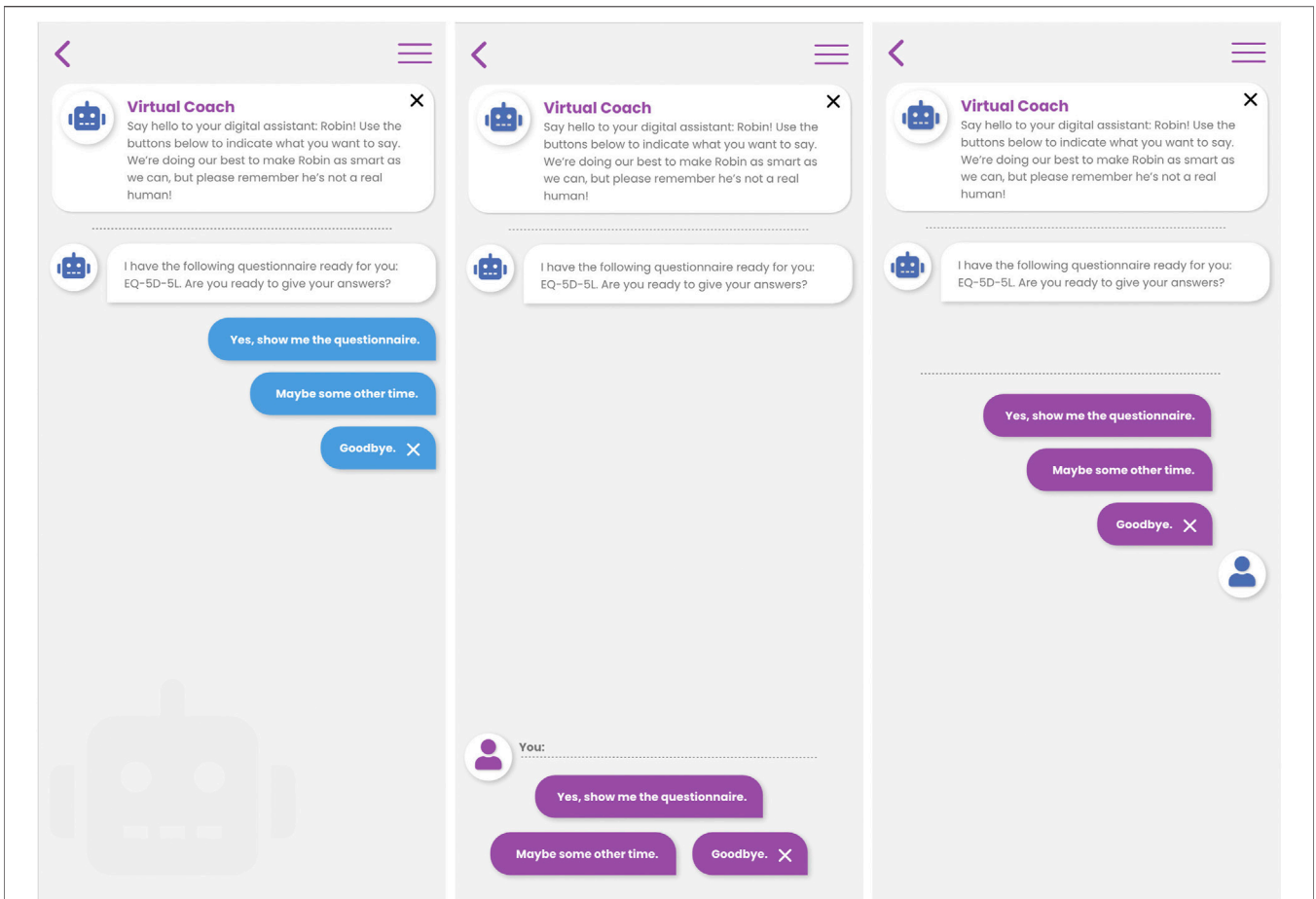
FIGURE 3 | Three different interaction designs for choice-based natural language dialogue interactions between patient and virtual coach in the Healthentia mobile application.

can enter into a dialogue with the coach using a choice-based input paradigm. **Figure 3** below shows three different variations of the choice-based dialogue interaction screen designs. The virtual coach statements are presented as chat-bubbles, and the user is given a number of options on how to respond to the coach's prompt. Based on the user's selected reply, the dialogue will advance to the next step.

This choice-based interaction metaphor may seem restrictive to the patient's ability to drive the conversation in the desired direction, but it is a fundamental design choice of our virtual coaching system that allows us to construct high quality engaging dialogue where the advice, suggestions or information that is, provided to the patient can be verified at every step of the process (from RWD collection, to processing, to the delivery of coaching).

When interacting with the virtual coach, patients are restricted to providing pre-defined replies at every dialogue step, and the flow of conversation within one such Dialogue (a conversation about a specific topic) follows a predefined hand-crafted script. These dialogue scripts are created using the WOOL Platform, an open-source platform for authoring and executing dialogues between users and agents, that was specifically designed for the creation of digital health applications. The fact that
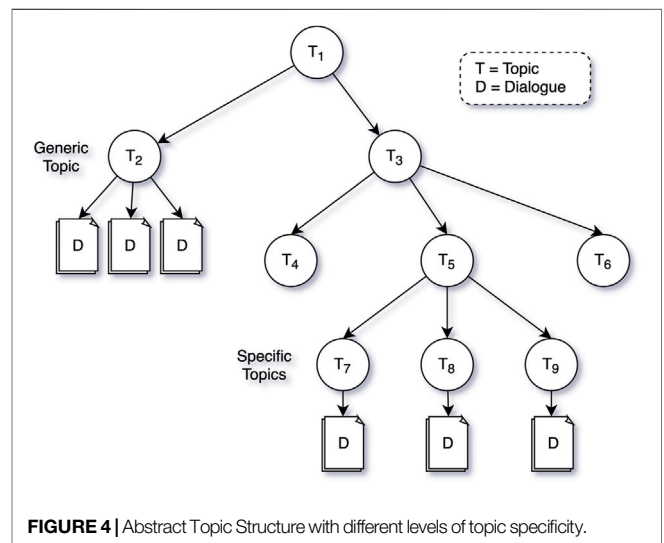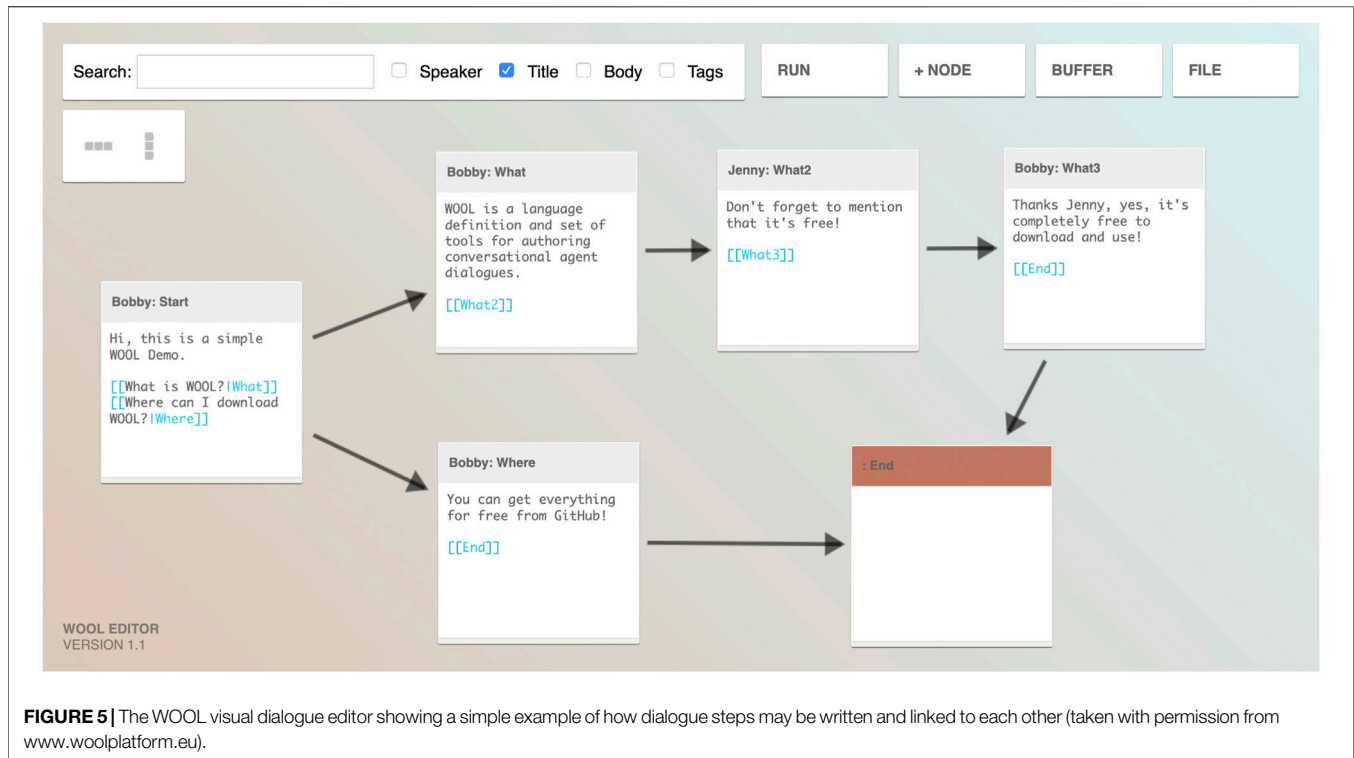

FIGURE 4 | Abstract Topic Structure with different levels of topic specificity.

interactions are scripted, and thus highly controlled in terms of which information the coach conveys to the user, is a solid foundation for creating a verifiable and reliable virtual coaching

FIGURE 5 | The WOOL visual dialogue editor showing a simple example of how dialogue steps may be written and linked to each other (taken with permission from www.woolplatform.eu).

platform. But almost as important as being correct, the interactions with the virtual coach should be personal and engaging. The process from authoring health coaching dialogues using the WOOL Platform to the triggering of the right dialogue at the right moment in time for the patient is outlined below.

Topics and Topic Structure

As a first step, domain experts define the type of information, or coaching that should be provided by the virtual coach by defining a set of topics—or topic structure (see **Figure 4**). A topic can be defined at a high level (e.g., "information about COPD[1]"), or at a more detailed level (e.g. "Physical Activity during low outside air quality"), and topics of different levels of specificity can co-exist in a single topic structure. The virtual coach can make decisions on which topic to discuss, meaning that by increasing the specificity of the topics in the topic structure, the virtual coach can make more precise suggestions (see *Triggering and Suggesting Dialogues*) to create a better personalized coaching experience.

A topic is an abstract entity, that doesn't contain any content itself. This content is defined in Dialogues that can be attached to any topic node in the topic structure. A Dialogue is thus a specific conversation between the virtual coach and the patient on a specific topic.

Dialogues and Authoring

The actual content of the conversations between the virtual coach and patient is contained within Dialogue scripts. Dialogue scripts define a series of dialogue steps, or Nodes. Each node contains any number of

Statements (things the coach says), and any number of Reply Options (things the user can say). These dialogues are defined using the WOOL language. WOOL is an open-source platform that defines the WOOL language, and provides various tools for authoring, testing and integrating WOOL dialogues into applications.

WOOL has spun off from an existing open-source dialogue platform, called Yarn, which was developed for creating dialogue-based video games. But WOOL's focus is specifically on creating dialogue content for digital health applications. As such, an important aspect is that non-technical domain experts (i.e., health experts) should be able to easily author dialogues. The next step in the process is that for at least each leaf topic defined in the topic structure (see *Topics and Topic Structure*), one or many dialogues should be authored. This authoring is done using the WOOL Editor, which is shown in **Figure 5**.

The WOOL Editor visually shows the flow of the dialogue and has built-in functionality for testing (parts of) the dialogue. The example shown in **Figure 5** lists a small number of simple nodes with basic reply options, or auto-forward reply options (advancing the dialogue automatically without user input). **Code Snippet 1** below shows example WOOL script for a simple dialogue node.

Code Snippet 1. Simple (WOOL) Dialogue node with statement, basic replies and an auto-forward reply.

```
1. Hello, I am Victor, your virtual coach, nice to meet you!
2.
3. [[Nice to meet you Victor.|IntroNode2]]
4. [[Goodbye.|End]]
5. [[IntroNode3]]
```

When executing this node in the Healthentia mobile application, the virtual coach will say "Hello, I am Victor,

---

[1]COPD stands for Chronic Obstructive Pulmonary Disease

your virtual coach, nice to meet you!" Then, the user will be presented with two options for replying to the message, "Nice to meet you Victor" (advancing the conversation to the Node called "IntroNode2"), or "Goodbye." (Ending the conversation). The third reply option (line 5) is an "auto-forward" reply—in this case, when the user doesn't select any reply option within some predefined timespan (e.g., 10 s), the dialogue will automatically advance to the node titled "IntroNode3."

These are the basics of the dialogue authoring system, but the platform supports various ways to make conversations more dynamic, engaging, and personalized. Without going into details, **Code Snippet 2** below shows an example on how to create a personalized greeting to users, depending on their age and gender, and how to allow users to provide their own (text) input to the system.

Code Snippet 2. Complex (WOOL) Dialogue node with Variables, Conditionals, and a text input reply option.

```
<<if $userAge > 50>>
<<if $userGender == "Female">>
  Nice to meet you Miss $userLastName!
<<elseif $userGender == "Male">>
  Nice to meet you Mr $userFirstName!
<<else>>
  Nice to meet you $userFirstName $userLastName!
<<endif>>
<<else>>
  Nice to meet you $userFirstName!
<<endif>>

I can address you like that, right?

[[Please call me<<input type="text" value="$userPreferredName" min="2"
max="30">>.|NodePreferredNameGiven]]
[[Nice to meet you too.|NextNode]]
[[Goodbye.|End]]
```

Triggering and Suggesting Dialogues

With a well-designed topic structure and connected set of dialogue content at hand, the next step in the process of providing automated virtual coaching deals with triggering and suggesting dialogues by the virtual coach. As mentioned in the introduction to this section, the initiative of who starts the conversation can lie with the patient as well as with the virtual coach. If there is a topic and set of dialogues available on e.g., daily dietary tips and tricks, the patient is able to start an interaction with the coach, and choose to "navigate" to this topic in a conversational manner:

1. Patient: <taps on the virtual coach>
2. Coach: Hello, how can I help you today?
3. Patient: Selects I want to talk about diet.
4. Coach: Okay, let's talk food! Want to review your status, set some goals, or just want some simple tips?
5. Patient: Some simple tips, please.
6. Coach: Okay, I have a nice trick on reducing your portion sizes, would you like to hear it.
7. Patient: Yes, sure!

This example dialogue between Patient and Coach shows how the patient manually traverses the topic structure to choose what he wants to talk about. After the patients indicates that he/she wants some simple tips (Step 5), we have reached the "daily dietary tips and tricks" topic, which might contain a large number

of small "daily tips" dialogues. At this step, we invoke the Dialogue Proposal Engine to make a suggestion on which specific tip to suggest now (Step 6), which the patient accepts in Step 7. We will shortly explain how the dialogue proposal engine might make this suggestion. For now it's good to note that in the example above there is a collaborative decision between the patient and the coach on which exact dialogue to hold. The Patient initiates the conversation and chooses the topic, while the coach suggests the exact dialogue. Three scenarios can be distinguished:

- Fully User-driven: The user of the application initiates a conversation with the virtual coach, picks the topic to discuss, and picks the exact dialogue to start.
- Fully Coach-driven: The virtual coach generates a notification suggesting a specific topic and dialogue to be held.
- Collaborative: Both user and virtual coach are involved in either [Organisation for Economic Cooperation and Development (OECD), 2019] initiating the conversation, (Digital Therapeutics Alliance, 2021), choosing the topic, or (WebWire, 2021) choosing the specific dialogue.

Although in all three scenario's, the user (patient) has the final say on accepting any suggested dialogue, in the fully coach-driven and collaborative scenarios, an intelligent component provides suggestions on the timing, suggested topic, or specific dialogue to initiate.

The first step in this sequence is the Dialogue Trigger Engine. This component continuously makes assessments for every patient, analysing their profile data that it takes from the User Data Service (see **Figure 1**). Of particular interest for the trigger engine are (a) changes in clinical pathway state as provided by the Clinical Pathway Service, and (b) changes in health outcome predictions as provided by the Intelligent Decision Services (see *Understanding: Real-World Data Processing*). But relevant data may also come from other external services, such as e.g., a weather data service, or an air quality data service that may trigger a dialogue on physical activity during low air quality in the area of the patient. Depending on the details of the intervention, decision rules may be built that define critical moments for the virtual coach to intervene. Even is all is well, the Dialogue Trigger Engine may decide that, since the patient hasn't interacted with the virtual coach for some time, now it's a good time to have a conversation.

The Dialogue Trigger Engine works in collaboration with the Dialogue Proposal Engine. The Dialogue Proposal Engine can answer the question "what would be the most suitable dialogue to hold with the patient," given (a) the current status of the patient, and (b) a particular subtree of the topic structure. When the trigger decides that the patient hasn't interacted with the virtual coach for a long time, and it would be nice to have a conversation, it's mostly up to the dialogue proposal engine to decide on the topic and exact dialogue to propose. Alternatively, the trigger engine may be very specific in the topic of the conversation to trigger (e.g., physical activity during low air quality), in which case the role of the proposal engine is small. **Figure 6** below shows
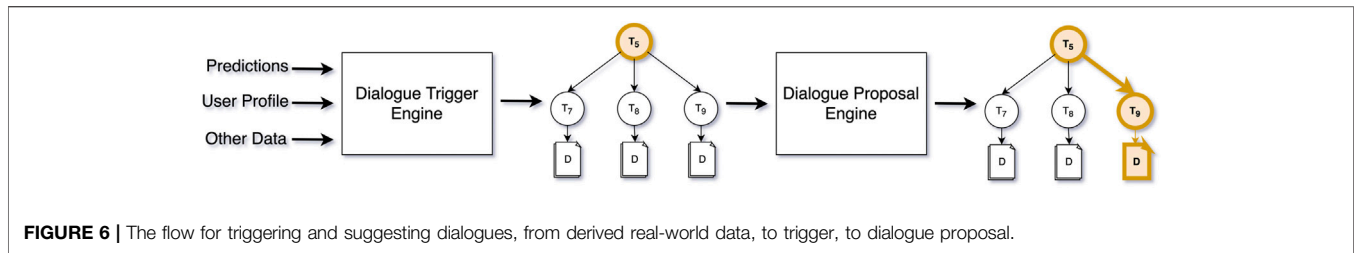
FIGURE 6 | The flow for triggering and suggesting dialogues, from derived real-world data, to trigger, to dialogue proposal.

the complete flow with an example where the trigger engine decides that a topic (T_5) should be triggered (a subtree of **Figure 4**), and the proposal engine finalising the decision process by proposing a concrete dialogue.

Finally, the proposed dialogue leads to a notification from the virtual coach on the patient's mobile phone—e.g., "Your symptoms are worsening, let's talk about how your sleep may influence this, if you have some time?" A simple text message, that is, made possible because:

1. Our real-world data collection services are measuring the patient's behaviour objectively (sleep) and subjectively (symptom tracking).
2. Advanced AI algorithms help understand the raw data to derive higher level, meaningful outcomes (predicted deterioration of health status, with sleep as major contributor).
3. The Virtual Coach is powered by the dialogue trigger engine, that combined with the dialogue proposal engine decide to start a conversation on the relevant topic (worsening symptoms and sleep).

Use-Case

To illustrate the potential of the Healthentia DTx platform to facilitate the deployment of digital therapeutics, we have created a fictitious use-case: a digital therapeutics trial for Chronic Obstructive Pulmonary Disease (COPD), combining the administration of a recently developed medication with a lifestyle intervention. In this digital therapeutic trial, Healthentia assists (i) healthcare professionals in performing and managing the trial *via* a web portal and (ii) study participants *via* a mobile application.

The trial consists of medication adherence, monitoring of disease-related symptoms, pre-scheduled standardized questionnaires, virtual coaching of physical activity and sleep.

The use-case here presented is partly based on the RE-SAMPLE project[2], a 4-years project funded by the EU's Horizon2020 research programme (Grant Agreement #965315[3]). In RE-SAMPLE, a multi-disciplinary team of partners spread over 7 European countries, is working together to create beyond state-of-the-art tailor-made care for complex chronic diseases. RE-SAMPLE relies on the combination of three points of care. First, the patient with COPD and comorbidities is given a mobile application connected to wearable and other medical devices that collect relevant

behavioral, physiological, and clinical parameters to the specific diseases of the patient. This mobile application (the Healthentia mobile app) will also be used at later stages of the project to provide virtual coaching to the patient. Second, when deviation from normal symptoms is detected, the patient is encouraged to visit one of the shared-care facilities at the location to perform further examinations. The third and final component is directed to healthcare professionals who are given a platform to monitor the progress of the patients and take action whenever needed. The Healthentia DTx platform serves as the basis for the three components.

Noteworthy is that the RE-SAMPLE project does not include the testing of medication. This element of the use-case we describe in *Results* was added to show the full potential of the Healthentia DTx platform.

The principal investigator in one of the participating sites in the trial is Dr. Peter Smith, a pneumologist at the Health4All hospital. Our key representative of the study participants is Clara, a 54 years old woman diagnosed with COPD GOLD Stage II in 2013 after a diagnosis of diabetes mellitus type II in 2007. Clara experiences that her fitness level is deteriorating fast, a consequence of the symptoms of the disease and of a vicious cycle of avoiding exercise as she is constantly anxious about inducing an exacerbation. She hopes that she can learn how to listen to her body better. The biggest constraint caused by COPD in Clara's daily life is that she gets extremely tired and out of breath after a few minutes of movement. Simple tasks as doing the groceries or vacuum cleaning feel now like climbing a high mountain.

RESULTS

In *Methods and Materials* we have described how the Healthentia platform enables virtual coaching in clinical studies and digital therapeutic interventions in a three-step approach of observing, understanding, and coaching. Throughout the previous section, we have provided small examples from various application scenarios to exemplify certain technical aspects. In this section we will guide the reader through a complete use case in which the Healthentia virtual coaching process is applied.

The trial in this use case is a multi-centre study that will take place in three countries and aims to include 750 participants. Patients eligible for recruitment have been diagnosed with COPD GOLD Stage II or III and at least one other chronic disease (e.g., diabetes mellitus, chronic heart failure, or ischemic heart disease). Furthermore, potential participants are older than 40 years old, are able to understand and communicate in one of the languages

---

[2]https://www.re-sample.eu/
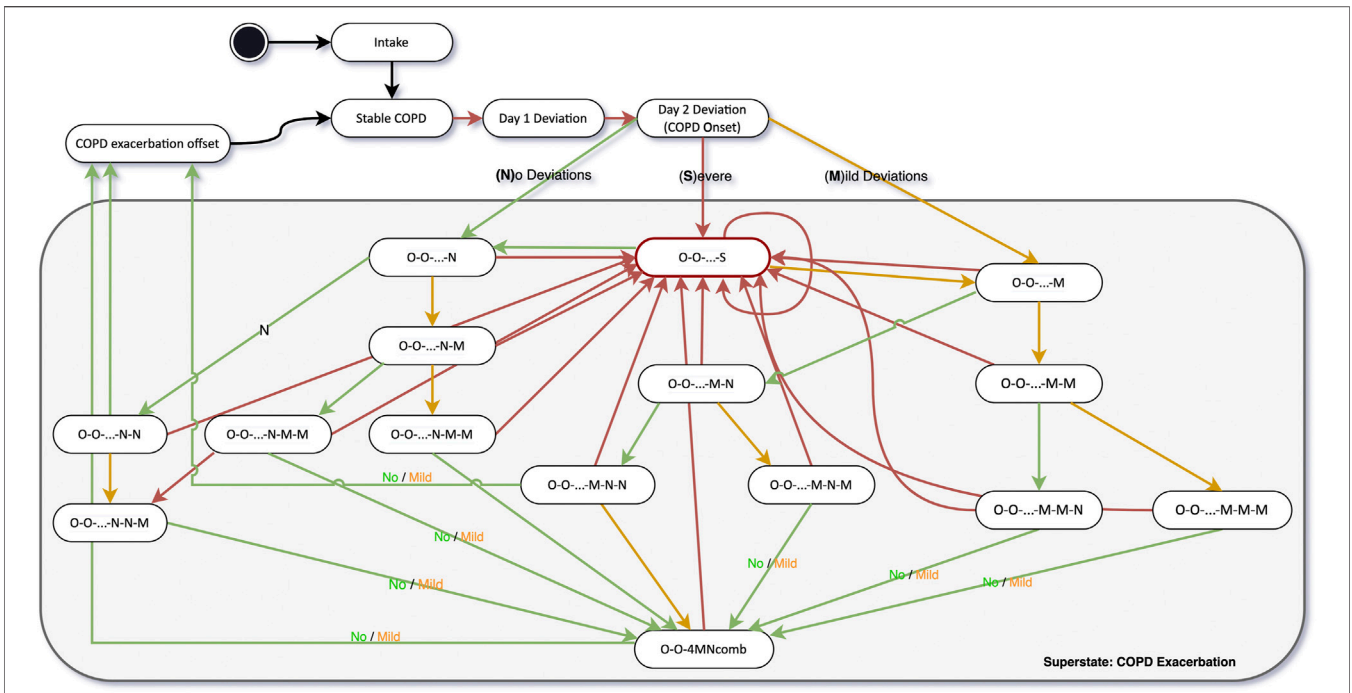[3]https://cordis.europa.eu/project/id/965315

FIGURE 7 | Simplified clinical pathway diagram of the COPD digital therapeutics trial.

spoken in the pilot site and have provided electronic or written informed consent prior to participation. In this section, we will explain how Healthentia DTx can facilitate the use-case in hands.

In preparation for the start of the study, Dr. Peter Smith, creates the study on the Healthentia portal, sets up the study procedure described in the protocol, including all questionnaires and functionalities available. Dr. Smith introduces four clinical pathway diagrams in the Healthentia platform, one related to COPD in general, and three other corresponding to each one of the comorbidities being investigated in the trial. For each state in the clinical pathway diagram, Dr. Smith defines the questionnaires that should be requested, their recurrency, additional actions—such as notifications and alerts, and the virtual coaching topics that should be enabled. This means that the digital therapeutic intervention is highly adaptable to the health status of each patient at any moment in time. A core section of the COPD clinical pathway is the daily monitoring of symptoms and adequate follow-up actions, based on the protocol of the COPE-III study introduced by Lenferink and colleagues (Lenferink et al., 2013). **Figure 7** illustrates the simplified version of the clinical pathway to detect the onset and offset of exacerbations based on the daily monitoring of symptoms. In a simplified and adapted manner from the original protocol, the start of an exacerbation is defined on the second day of deviations from normal symptoms. The offset of an exacerbation happens after three consecutive days of normal symptoms or 4 days of a combination of no or mild deviations from the symptoms. The original protocol defines exacerbation offset as 7 days of combination between no or mild deviation from symptoms.

After the consent form has been (electronically) signed by Clara and Dr. Smith, Clara is notified on the Healthentia mobile

application that she is now officially enrolled in the study. As soon as she starts the application, Clara is welcomed by a friendly human-like figure, who introduces herself as Anna, Clara's personal virtual coach. Anna informs Clara about her role in the Healthentia app, provides a tutorial on how to use the app and helps Clara setting up the accompanying devices.

The use-case introduced in this article involves the collection of varied RWD. First, connected to the aim of testing a new medication, the Healthentia mobile application allows monitoring of adherence to medication and reporting of deviations from the protocol. When the use of the smart medication dispenser given to each participant deviates from the expected treatment plan, Clara, the study participant, receives a reminder to take her medication and is given the option to report the reason why she skipped or took extra, medication. This information is visible by Dr. Smith, the principal investigator in one of the study sites, on the Healthentia web portal.

Second, physiological RWD is collected with a commercial fitness tracker given to the study participants. This wearable device tracks physical activity outcomes (e.g., steps, minutes spent in physical activity intensity category, sedentary bouts), sleep-related outcomes (e.g., sleep duration and time spent in each sleep phase), heart rate, and oxygen saturation. The observation and processing of physiological RWD serve as basis for the lifestyle intervention from the virtual coach described in *Coaching: Virtual Coaching in Digital Therapeutics*.

Third, the mobile application sends a daily reminder to report on deviations of normal symptoms related to COPD and eventual additional chronic conditions. The symptom diary is set up in a triage form to avoid burdening the participant: in cases of no

deviation from normality, the participant only needs to answer one question; otherwise, the participant is asked follow-up questions to understand exactly which symptoms have changed and potential triggers. The symptom diary is also the major trigger for transitioning the patient between various states in the Clinical Pathway (see **Figure 7**).

Fourth, standardized questionnaires (e.g., monitoring of the quality of life) are made available through the mobile application at recurrent moments or when deviations of symptoms from normality are detected.

Finally, psychological RWD is collected through conversations between the study participant and the virtual coach, based on simple questions mimicking real-life conversations, such as "Good morning Clara, how are you doing today?"

Understanding: Real-World Data Processing

Dr. Smith is using Healthentia as a tool to help him supervising his patients like Clara and in the long run help them *via* its virtual coaching features. Clara on the other hand is interested directly in Healthentia's coaching capabilities, seeing the tool as a companion in her patient journey. Both Dr. Smith and Clara utilize, at least indirectly an exacerbation biomarker discovered using Healthentia data. For this, the collected data are non-linearly combined by a function yielding the biomarker value, a binary value indicating whether an exacerbation is expected within the next week.

Exacerbation onset is declared deterministically given the answers of Clara in the daily COPD Symptom Questionnaire. Hence at the end of each day, Healthentia is setting a Boolean tag indicating the onset of an exacerbation for every patient that completed the questionnaire. The target of the biomarker is to predict this Boolean value, not using today's data (this is fully deterministic), but with data from the past. Thus, the attributes comprising the input to the model to be learnt are the physiological data collected by the activity tracker and the answers to the questionnaires of some input data, while the output is a Boolean value indicating if an exacerbation is detected in any of the next days, from 1 day after the input date, up to a week after the input date.

Depending on the available number of input-output pairs, the most suitable Machine Learning algorithm is chosen to learn the model. Random Forests is a good choice for a moderate number of such pairs, i.e., from a few hundred onwards. Neural Networks with a couple of hidden layers would require a few thousand. The model is trained on the training set, while its hyperparameters are tuned using the validation set to evaluate performance. Once satisfied, the final model is trained on the combination of the two sets and its performance is assessed on the testing set. If the performance is adequate, then the model can be optimized using SHAP analysis across all the decisions to discard input attributes that are consistently of low impact to the decisions. The optimized model is then put into operation, predicting exacerbations in the coming days for every patient. If not, then more training is accumulated.

While in operation, SHAP analysis at the individual decision level yields the most important attributes pulling the outcome towards exacerbation or away from it for each patient. This information is used by the virtual coach to give actionable and personalized advice to Clara daily.

Finally, the most significant input attributes and the output variable are used to cluster the patients into phenotypes. It is possible to manually characterize these data-driven phenotypes and compare them to the COPD stage classification system, or to the behavioural groups Dr. Smith has classified his patients since the enrolment interview. Clara belongs to a phenotype. This allows the virtual coach to treat her as a member of a group, offering group advice and comparisons. But also Clara's phenotype is continuously updated, and changes in her phenotype are flagged by Healthentia for Dr. Smith's consideration.

Coaching: Virtual Coaching in Digital Therapeutics

The prediction of a possible future exacerbation for Clara, using the biomarker described above is picked up by the dialogue trigger engine. A predicted exacerbation is urgent, but not something to contact Clara about in the middle of the night. So, the dialogue trigger engine combines the information about the predicted exacerbation with Clara's sleep data to determine what the right moment would be (the next morning, an hour after Clara wakes up) to initiate a discussion on the topic of Preventing Exacerbation.

Preventing Exacerbation is a topic within the topic structure defined for this intervention that has several subtopics related to various aspects that can relate to the onset of exacerbation: physical activity, sleep, air quality, and medication intake. The dialogue proposal engine now uses the outcome of the SHAP analysis to decide on which of the sub-topics in the Exacerbation Prevention tree would be most suitable to suggest to Clara. In this case, the SHAP analysis indicated that various physical activity features were the main contributors to the machine learner's decision to predict an exacerbation, thus the dialogue proposal engine proposes "Exacerbation Prevention—Physical Activity" as the topic of discussion, and selects the one and only dialogue attached to this tree, e.g., copd-exacerbationprevention-pa.wool (see **Figure 8**).

Figure 8 shows the decision path through the topic structure that the dialogue trigger engine and dialogue proposal engine take in the case of a predicted exacerbation. It also shows a part of the coaching content that is, made available to Clara when engaging with the virtual coach in general. When Clara interacts with the virtual coach, she can choose to talk about Physical Activity, COPD, Sleep, or Medication. For each of these topics, a sub-tree of topics is available. For activity, sleep and medication, these subtrees follow the same structure (labelled Core Coaching Loop in **Figure 8**). The coaching loop is divided into six phases: relationship building, health education, insights, goal-setting, support, and reflection. The dialogues in these topics are part of a behaviour change coaching programme that aims to guide the user from unhealthy behaviour to a sustained positive change in behaviour.
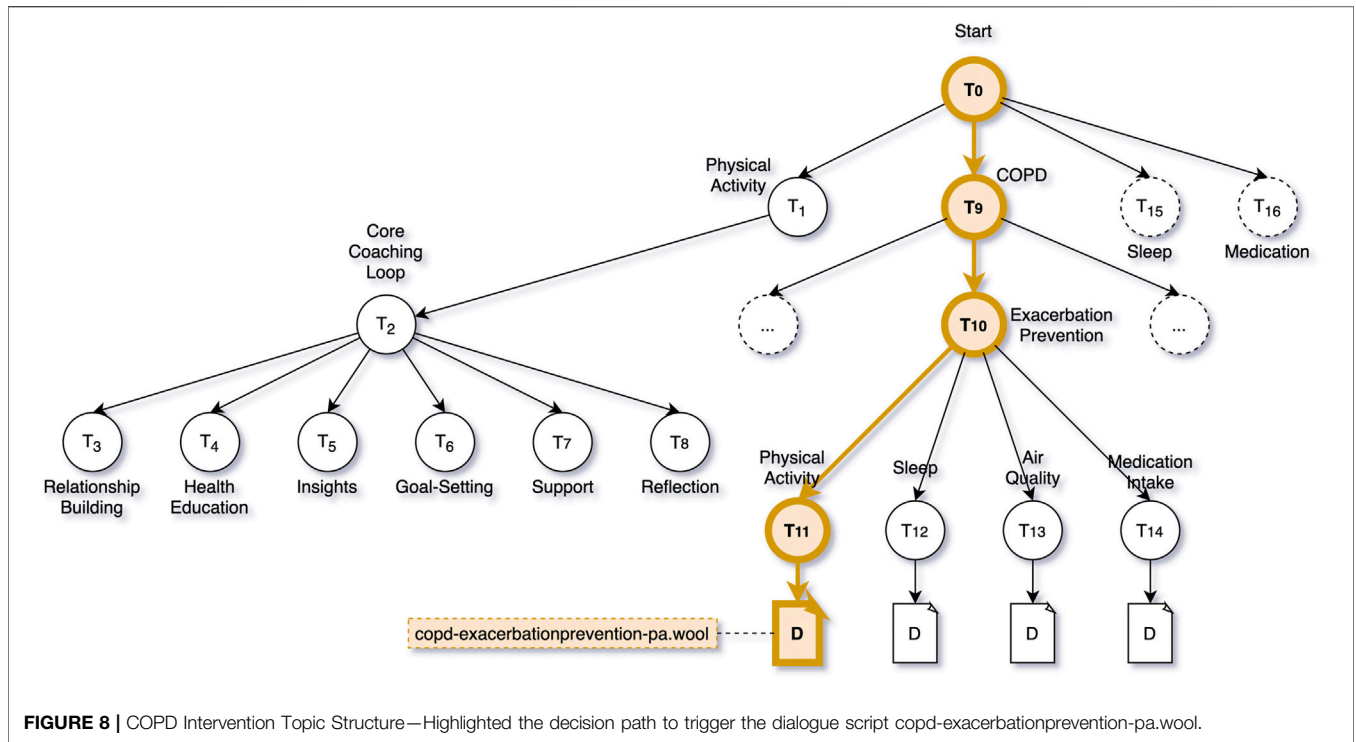
FIGURE 8 | COPD Intervention Topic Structure—Highlighted the decision path to trigger the dialogue script copd-exacerbationprevention-pa.wool.

As we expect the COPD intervention to last for quite some time (months or years), we "start" with relationship building which includes dialogues to simply get to know the user a bit better. In fact, we do not strictly linearly traverse through the six phases, but can jump back and forth. Once the virtual coach and Clara are better acquainted, we move to health education. Before changing any behaviour (e.g., physical activity), it is important to be informed about the benefits of positive behaviour, or risks of negative behaviour. Once Clara is aware of the benefits and risks, it is time to start getting insights into her own behaviour—is Clara already active enough, or could she be doing more? Assuming Clara could work on improving her levels of daily physical activity, the next step is goal-setting in which the virtual coach and Clara can collaborative decide on a long-term physical activity goal, and short-term goals to gradually reach it. In the support phase, Clara is offered support to reach her goals. This includes feedback, motivational messages, reinforcement, tips and tricks, or even community support where Clara's progress is compared to her peers. As Clara has been made aware of the importance of physical activity, has seen her own shortcomings, collaboratively set goals, and managed to reach those goals with support of the coach, the final phase is meant to help Clara sustain her now healthier behaviour by reflecting on e.g., positive impacts on Clara's life.

DISCUSSION

In this paper we describe a methodology to support virtual coaching empowered by artificial intelligence on real-world data and delivered through conversational agents in a DTx intervention

to manage or treat chronic conditions. We started by describing the main components that are needed to enable DTx, how they operate on a high level, and how they work together to form a platform that can be used to define DTx interventions for many different types of conditions. In order to provide a complete example, from observing to understanding to coaching, we present in *Results* a use case describing a COPD intervention and use this case to provide a walkthrough of all the relevant steps in the DTx process.

There are several promising DTx applications in the pipeline that use virtual coaching, but the number of solutions that arrives to the market is very scarce. A successful case is that of the company Lark with its programmes using fully automated conversational coaching for prevention and management of diabetes showing very positive results (Stein and Brooks, 2017). Unclear and non-harmonized regulatory considerations, concerns with security and data governance, and unclarity in reimbursement are some of the reasons for the scarcity of DTx solutions in the market (Patel and Butte, 2020). Especially when it comes to Artificial Intelligence (AI) features, and self-learning properties of DTx systems, passing regulatory clearance remains a challenging obstacle.

The potential issues that can arise from an artificially intelligent system that is, involved in the process of providing health or medical advice in an unsupervised manner are obvious. If a Google image search algorithm mistakenly classifies an image of a lynx as a common house cat, no real harm is done. When a self-driving Tesla car mistook the moon for a yellow flashing traffic light[4], no one was injured, but the potential risk becomes

---

[4]https://www.ndtv.com/offbeat/watch-tesla-autopilot-feature-mistakes-moon-for-yellow-traffic-light-2495804

clear. When self-learning AI systems in a DTx environment start saying the wrong things to patients in a coaching session, the problems are real.

The DTx platform described in this article is designed from the ground up to be explainable in every step of the way. Predictions made by the digital biomarkers are augmented, using SHAP analysis to provide information on what the most contributing factors to a decision are. The coaching components that display intelligent behaviour are simple components, making simple rule-based decisions in a step-wise manner. In this way, the overall "emerging behaviour" of the virtual coach can be very personalized, and seem intelligent, but the decision process can be analysed to catch mistakes. Perhaps more importantly, the virtual coaching is fundamentally based on expertly crafted coaching dialogues, instead of e.g., employing natural language generation techniques (Cawsey et al., 1997) to automatically construct coaching sentences.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

REFERENCES

Baudat, G., and Anouar, F. (2000). Generalized Discriminant Analysis Using a Kernel Approach. *Neural Comput.* 12, 2385–2404. doi:10.1162/089976600300014980

Bickmore, T., and Cassell, J. (2001). "Relational Agents: A Model and Implementation of Building User Trust," in CHI '01: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Seattle, Washington, March 1, 2000, 396–403.

Bickmore, T., Schulman, D., and Yin, L. (2010). Maintaining Engagement in Long-Term Interventions with Relational Agents. *Appl. Artif. Intelligence* 24 (6), 648–666. doi:10.1080/08839514.2010.492259

Bongers, B. C., Dejong, C. H. C., and den Dulk, M. (2021). Enhanced Recovery after Surgery Programmes in Older Patients Undergoing Hepatopancreatobiliary Surgery: What Benefits Might Prehabilitation Have? *Eur. J. Surg. Oncol.* 47 (3 Pt A), 551–559. doi:10.1016/j.ejso.2020.03.211

Breiman, L. (2001). Random Forests. *Machine Learn.* 45, 5–32. doi:10.1023/a:1010933404324

Cawsey, A. J., Webber, B. L., and Jones, R. B. (1997). Natural Language Generation in Health Care. *J. Am. Med. Inform. Assoc.* 4 (6), 473–482. doi:10.1136/jamia.1997.0040473

Coravos, A., Khozin, S., and Mandl, K. D. (2019). Developing and Adopting Safe and Effective Digital Biomarkers to Improve Patient Outcomes. *NPJ Digit Med.* 2 (14), 14. doi:10.1038/s41746-019-0090-4

Dekker-van Weering, M., Jansen-Kosterink, S., Frazer, S., and Vollenbroek-Hutten, M. (2017). User Experience, Actual Use, and Effectiveness of an Information Communication Technology-Supported Home Exercise Program for Pre-Frail Older Adults. *Front. Med.* 4 (4), 208. doi:10.3389/fmed.2017.00208

Digital Therapeutics Alliance (2021). Product Library - Digital Therapeutic Alliance. [Online]. Available from: HYPERLINK http://dtxalliance.org/understanding-dtx/product-library/ (cited July 29, 2021).

Garrow, J. S., and Webster, J. (1985). Quetelet's index (W/H2) as a Measure of Fatness. *Int. J. Obes.* 9 (2), 147–153.

Geron, A. (2019). *Hands-on Machine Learning with Scikit-Learn, Keras and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems.* 2nd Ed. Sebastopol, CA: O'Reilly.

AUTHOR CONTRIBUTIONS

All authors have contributed equally to the drafting of this manuscript. In particular, HA has written the sections on coaching, the discussion and has performed the editing. MC has provided the use-case, connecting all components discussed to user stories, and has written the introduction and abstract. AP has written the sections on collecting and understanding real-world data.

FUNDING

ACKNOWLEDGMENTS

Grey, M. (2017). Lifestyle Determinants of Health: Isn't it All about Genetics and Environment? *Nurs. Outlook* 65 (5), 501–505. doi:10.1016/j.outlook.2017.04.011

Guthrie, N. L., Carpenter, J., Edwards, K. J., Appelbaum, K. J., Dey, S., Eisenberg, D. M., et al. (2019). Emergence of Digital Biomarkers to Predict and Modify Treatment Efficacy: Machine Learning Study. *BMJ Open* 9, e030710. doi:10.1136/bmjopen-2019-030710

Jaidan, D. N., Carrere, M., Chemli, Z., and Poisvert, R. (2019). "Data Anonymization for Privacy Aware Machine Learning," in *Machine Learning, Optimization, and Data Science. LOD 2019. Lecture Notes in Computer Science.* Editors G. Nicosia, P. Pardalos, R. Umeton, G. Giuffrida, and V. Sciacca (Cham: Springer), 11943, 725–737. doi:10.1007/978-3-030-37599-7_60

Kamali, M. E., Angelini, L., Caon, M., Carrino, F., Rocke, C., Guye, S., et al. (2020). Virtual Coaches for Older Adults' Wellbeing: A Systematic Review. *IEEE Access* 8, 101884–101902. doi:10.1109/access.2020.2996404

Kovalchick, C., Sirkar, R., Regele, O. B., Kourtis, L. C., Schiller, M., Wolpert, H., et al. (2017). Can Composite Digital Monitoring Biomarkers Come of Age? A Framework for Utilization. *J. Clin. Trans. Sci.* 1 (6), 373–380. doi:10.1017/cts.2018.4

Laranjo, L., Dunn, A. G., Tong, H. L., Kocaballi, A. B., Chen, J., Bashir, R., et al. (2018). Conversational Agents in Healthcare: A Systematic Review. *J. Am. Med. Inform. Assoc. JAIMA.* 25 (9), 1248–1258. doi:10.1093/jamia/ocy072

Lenferink, A., Frith, P., van der Valk, P., Buckman, J., Sladek, R., Cafarella, P., et al. (2013). A Self-Management Approach Using Self-Initiated Action Plans for Symptoms with Ongoing Nurse Support in Patients with Chronic Obstructive Pulmonary Disease (COPD) and Comorbidities: The COPE-III Study Protocol. *Contemp. Clin. Trials* 36 (1), 81–89. doi:10.1016/j.cct.2013.06.003

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., et al. (2020). From Local Explanations to Global Understanding with Explainable AI for Trees. *Nat. Mach Intell.* 2, 56–67. doi:10.1038/s42256-019-0138-9

Manli Zhu, M., and Martinez, A. M. (2006). Subclass Discriminant Analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* 28, 1274–1286. doi:10.1109/tpami.2006.172

Moghaddam, B. (2002). Principal Manifolds and Probabilistic Subspaces for Visual Recognition. *IEEE Trans. Pattern Anal. Machine Intell.* 24, 780–788. doi:10.1109/tpami.2002.1008384

Murray, E., Sweeting, M., Dack, C., Pal, K., Modrow, K., Hudda, M., et al. (2017). Web-Based Self-Management Support for People with Type 2 Diabetes (HeLP-Diabetes): Randomised Controlled Trial in English Primary Care. *BMJ Open* 7, e016009. doi:10.1136/bmjopen-2017-016009

Olafsson, S., O'Leary, T. K., and Bickmore, T. W. (2020). "Motivating Health Behavior Change with Humorous Virtual Agents," in IVA '20: Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents, Scotland, UK, October 20–22, 2020, 1–8. doi:10.1145/3383652.3423915

Organisation for Economic Cooperation and Development (OECD) (2019). *Health at a Glance 2019*. Paris: OECD Indicators.

Patel, N. A., and Butte, A. J. (2020). Characteristics and Challenges of the Clinical Pipeline of Digital Therapeutics. *Npj Digit. Med.* 3, 159. doi:10.1038/s41746-020-00370-8

Pnevmatikakis, A., and Polymenakos, L. (2009). Subclass Linear Discriminant Analysis for Video-Based Face Recognition. *J. Vis. Commun. Image Representation* 20, 543–551. doi:10.1016/j.jvcir.2009.08.001

Poggi, I., Pelachaud, C., de Rosis, F., Carofiglio, V., and De Carolis, B. (2005). "Greta. A Believable Embodied Conversational Agent," in *Multimodal Intelligent Information Presentation*. Editors O Stock and M Zancanaro (Dordrecht: Springer Netherlands), 3–25. doi:10.1007/1-4020-3051-7_1

Prochaska, J. O., and DiClemente, C. C. (1986). Toward a Comprehensive Model of Change. *Appl. Clin. Psychol.* 13 (1), 3–27. doi:10.1007/978-1-4613-2191-0_1

Proctor, J., Naughton, F., Sloan, M., Hopewell, S., Brimicombe, J., Prevost, A. T., et al. (2020). Assessment of the Effectiveness and Cost-Effectiveness of Tailored Web- and Text-Based Smoking Cessation Support in Primary Care (iQuit in Practice II): Protocol for a Randomized Controlled Trial. *JMIR Res. Protoc.* 9 (7), e17160. doi:10.2196/17160

Provoost, S., Lau, H. M., Ruwaard, J., and Riper, H. (2017). Embodied Conversational Agents in Clinical Psychology: A Scoping Review. *J. Med. Internet Res.* 19 (19), e151. doi:10.2196/jmir.6553

Schachner, T., Keller, R., and v Wangenheim, F. (2020). Artificial Intelligence-Based Conversational Agents for Chronic Conditions: Systematic Literature Review. *J. Med. Internet Res.* 22 (9), e20701. doi:10.2196/20701

Schmidhuber, J. (2015). Deep Learning in Neural Networks: An Overview. *Neural Networks* 61, 85–117. doi:10.1016/j.neunet.2014.09.003

Stein, N., and Brooks, K. (2017). A Fully Automated Conversational Artificial Intelligence for Weight Loss: Longitudinal Observational Study Among Overweight and Obese Adults. *JMIR Diabetes* 2 (2), e28. doi:10.2196/diabetes.8590

Theodoridis, S., and Koutroumbas, K. (2008). *Pattern Recognition*. 4th ed. Orlando: Academic Press.

Tsiouris, K. M., Tsakanikas, V. D., Gatsios, D., and Fotiadis, D. I. (2020). A Review of Virtual Coaching Systems in Healthcare: Closing the Loop with Real-Time Feedback. *Front. Digital Health* 2, 21. doi:10.3389/fdgth.2020.567502

US Food and Drug Administration (2009). *Patient-Reported Outcome Measures: Use in Medical Products Development to Support Labelling Claims*. FDA Guidance Document UCM193282.

US Food and Drug Administration (2017). *Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices: Guidance for Industry and Food and Drug Administration Staff*.

WebWire (2021). Health Coaching Market Worth $7 Billion in the U.S. [Online]. Available from: HYPERLINK https://www.webwire.com/ViewPressRel.asp?aId=268817 (cited July 29, 2021).