

Published on User Model User-Adap Inter (2017) 27:159–213:
<https://link.springer.com/article/10.1007/s11257-017-9190-5>
DOI: 10.1007/s11257-017-9190-5

MobiGuide: a personalized and patient-centric decision-support system and its evaluation in the atrial fibrillation and gestational diabetes domains

Mor Peleg^{a*}, Yuval Shahar^b, Silvana Quaglini^c, Adi Fux^a, Gema García-Sáez^d, Ayelet Goldstein^b, M. Elena Hernando^d, Denis Klimov^b, Iñaki Martínez-Sarriegui^d, Carlo Napolitano^e, Enea Parimbelli^c, Mercedes Rigla^f, Lucia Sacchi^c, Erez Shalom^b, and Pnina Soffer^a

^a *Department of Information Systems, University of Haifa, Haifa, Israel, 3498838*

^b *Department of Information Systems Engineering, Ben Gurion University of the Negev, Beer-Sheva, Israel, 84105*

^c *Dipartimento di Ingegneria Industriale e dell'Informazione, University of Pavia, Italy, 27100*

^d *Centre for Biomedical Technology (CTB), ETSIT,, Universidad Politécnica de Madrid, and Networking Research Centre for Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Spain, 28040*

^e *IRCCS Foundation “Salvatore Maugeri”, Pavia, Italy, 27100*

^f *Endocrinology and Nutrition Department, Hospital de Sabadell, Sabadell, Spain, 08208*

Acknowledgement

This study was part of the MobiGuide project partially funded by the European Commission 7th Framework Program, grant #287811.

Abstract

MobiGuide is a ubiquitous, distributed and personalized evidence-based decision-support system (DSS) used by patients and their care providers. Its central DSS applies computer-interpretable clinical guidelines (CIGs) to provide real-time patient-specific and personalized recommendations by matching CIG knowledge with a highly-adaptive patient model, the parameters of which are stored in a personal health record (PHR). The PHR integrates data from hospital medical records, mobile biosensors, data entered by patients, and recommendations and abstractions output by the DSS. CIGs are customized to consider the patients' psycho-social context and their preferences; shared decision making is supported via decision trees instantiated with patient utilities. The central DSS "projects" personalized CIG-knowledge to a mobile DSS operating on the patients' smart phones that applies that knowledge locally.

In this paper we explain the knowledge elicitation and specification methodologies that we have developed for making CIGs patient-centered and enabling their personalization. We then demonstrate feasibility, in two very different clinical domains, and two different geographic sites, as part of a multi-national feasibility study, of the full architecture that we have designed and implemented. We analyze usage patterns and opinions collected via questionnaires of the 10 atrial fibrillation (AF) and 20 gestational diabetes mellitus (GDM) patients and their care providers. The analysis is guided by three hypotheses concerning the effect of the personal patient model on patients and clinicians' behavior and on patients' satisfaction. The results demonstrate the sustainable usage of the system by patients and their care providers and patients' satisfaction, which stems mostly from their increased sense of safety. The system has affected the behavior of clinicians, which have inspected the patients' models between scheduled visits, resulting in change of diagnosis for two of the ten AF patients and anticipated change in therapy for eleven of the twenty GDM patients.

Keywords: Computer-interpretable guidelines, decision-support system, clinical guidelines, patient centrality, personalization, mobile health

1. Introduction

A key concept that has greatly influenced modern medical care has been the development of *evidence-based medicine*. A major aspect of evidence-based medicine includes the development of *clinical guidelines* for the diagnosis and/or treatment of common disorders. Such guidelines are developed by professional medical organizations, based on evidence accumulating from clinical research, often through randomized, controlled clinical trials. Over the past two decades, several types of clinical-guideline based *decision-support systems* (DSSs) have been developed to support clinicians in managing patients who have specific diagnoses, such as diabetes or cardiovascular diseases. Such DSSs represent evidence-based clinical guideline knowledge in computable specifications known as *computer-interpretable guidelines* (CIGs) (Peleg 2013). CIG formalisms provide detailed ontologies that capture data enquiry and abstraction, action management, and decision-making tasks. They provide semantics that go beyond those of existing telemonitoring systems (Lasierra et al. 2013) that use collections of rules to react to data values that are outside normal ranges.

A variety of CIG formalisms exist, each having its own application engine (Peleg 2013). CIG application engines match the CIG specification to a patient's data set, ideally imported from medical institutions' electronic health records. The collection of relevant patient data, to which each CIG refers, instantiates a *patient model* for that CIG. The behavior of the DSS is thus adapted to the patient model. Currently, most of these emerging CIG-based DSSs are used only by clinicians during the patient visit, to obtain patient-specific recommendations.

The MobiGuide system, described in this paper, goes beyond the current state of the art of CIG-based DSSs in two main ways. First, it is patient-centered, meaning that the users of the DSS are not only the clinicians, but are first and foremost, the *patients*. To accommodate this novelty, we have developed and used new knowledge-acquisition and specification methods for capturing in CIGs the parallel workflows of the patients who are following the CIG's recommendations, and of the clinicians caring for these patients (Sacchi et al. 2013).

The second way in which MobiGuide goes beyond traditional CIG-based DSS is personalization of decision support; unlike other emerging CIG-based DSSs, the *patient model of MobiGuide* includes not only the patient's clinical data but also the patient's *personal preferences and psychosocial context*. Preferences could be values that the patient assigns to his/her present health state and to future possible health states to which she might transition, based on the therapy choices decided together with his/her clinician, but also daily life preferences, such as meal times or self-monitoring schedule. The patient model includes also a set of *personal events*. These personal events *induce* (with certain temporal constraints) one or more of a set of predetermined *psychosocial contexts* that appear in the [customized] CIG, which can impact medical decisions. Thus, following the initial formal specification of clinical practice guidelines as CIGs (with optionally parallel workflows), the CIGs are *customized* so as to include additional contexts that are not mentioned by the standard clinical practice guidelines, but which

could occur for some periods of time when managing a patient at home through a mobile device. These customized contexts determine, for example, which sub-plans of the CIG will be activated or inactivated, and which knowledge will be applied to interpret the patient's data. For example, a context of "semi-routine schedule" may be associated with care plans that require more frequent monitoring of blood pressure. The customized CIGs are then *personalized* to specific patients by specifying the individual, personal events that *induce* each of the predetermined customized contexts (e.g., a "going on vacation" event may induce the predetermined, customized "semi-routine schedule" context). Hence, the patient model is both *dynamic*, reacting to current data from real-time monitoring, and it is also highly *adaptive* to the personal preferences and contexts of individual patients, rather than to stereotypical patient profiles.

Fig. 1 shows the MobiGuide patient model for AF and GDM patients. In addition to the clinical part (i.e., main diagnosis, comorbidities, and prescriptions), which is found in other emerging CIG-based DSSs, the MobiGuide patient model shows many **personalized elements** that are not found in other CIG-based DSSs and include (Fig. 1): (customized) context, patient-reported events that invoke DSS recommendations (symptoms, risky events for bleeding, reports of eating extra carbohydrates), and patients' personal preferences. These features, which change dynamically, influence clinical acts, including (a) the types and frequency of recommendations and reminders, which are sensitive to context and to reporting of personal events, and (b) the timing of measurement reminders, which depends on context and personal preferences. In addition to these personalized elements, the patient model also uses other unique features, including clinical abstractions inferred from raw data, and patient-specific DSS recommendations that were delivered to the patient or to his care providers.

The entire patient model resides in the dynamic PHR (see subsection 2.2) .

The following scenario describes the patient experience of using MobiGuide.

AF-specific	Common features	GDM-specific
One out of 4 AF types	Main Diagnosis	GDM
Tyroid diseases , IMA, Previous cerebrovascular disease	Comorbidities: Hypertension	
	Treatment Prescriptions: Drugs and Exercise prescription	Nutritional prescriptions
ECG, INR, quality of life	Monitoring Prescriptions: Blood Pressure, Heart Rate, Weight, Exercise	Blood Glucose, Ketonuria
	Clinical abstractions: Abnormal/normal Blood Pressure, Weight, Patient's compliance	Abnormal/normal Blood Glucose, Ketonuria trend
Increased physical activity	Contexts: Routine (e.g., working days), semiroutine (e.g., vacation)	
Symptoms, Events with risk of bleeding (e.g. dental surgery)	Patient's reported events: context-inducing events (e.g., vacation)	extra carbohydrates
Utility coefficients for the health state involved in the shared decision making process	Patient's preferences <i>Global</i> <i>Local (a set for each context) :</i> Meal times Willingness to receive reminders for drugs and measurements Preferred days/times for reminders	
At hospital, Presence of a caregiver	Patient's environment: At home	
Daily drug reminders, pill-in-the-pocket recommendation,	Recommendations and reminders Monitoring reminders, Contact your doctor	Nutritional recommendation
Patient's prognostic models: Decision model for anticoagulants Decision model for surgical ablation		
Patient's summary Comorbidities and current drugs		

Fig. 1. MobiGuide patient models. Features reported in the middle column of the figure are common to both the AF and GDM patient model, while the disease-specific ones are reported on the left and right column, respectively.

Michael is a 70-year old AF patient with paroxysmal AF and hypertension. He has been experiencing more palpitations recently, which has been making him anxious. His doctor, Carlo, suggested that he may feel safer to be monitored and advised by the MobiGuide system while he continues to function in his normal environment.

After signing the informed consent, Carlo activated the MobiGuide AF CIG for Michael, and his relevant hospital records were exported into the PHR. Then, additional information, such as current therapies, were added to the PHR. At this point, his patient model contains a relatively small number of relevant clinical data (shown in Fig. 1 and in Fig. 2(b) that presents the Summary (of clinical history) and Therapy prescriptions.

Using the **shared decision making** module of MobiGuide, Carlo and Michael elicited Michael's preferences in the form of utility coefficients and selected Warfarin as the most suitable anticoagulant that would bring Michael the highest value of quality-adjusted life years according to his personal utilities attributed to his actual state and other potential health states he could experience in the future. This **global preference** resulted in the selection of a specific arm in the CIG.

Next, the MobiGuide DSS guided Carlo in selecting from a relevant set of monitoring plans. He then set Michael's **local preferences regarding reminder days and times for different potential contexts, including routine schedule, semi-routine, increased exercise** (e.g., periods when he bikes more than usual), which like the global preferences, became part of his patient model (Fig. 1). Based on them, the system will remind Michael to perform his monitoring plans at the frequency appropriate to his current context.

On the following day at 6:55am, at his preferred reminder time, Michael received the reminder to take his medications. However, he decided to **report that he's not going to take the medication**, indicating the reason "because of side effects" (Fig. 2c)

He also received a reminder to measure his ECG. He accepted and activated his mobile sensor for 30 minutes (Fig. 2d). The signals were communicated via Bluetooth connection to the phone, analyzed by the AF detection algorithm and an ECG summary was saved in his PHR. It will be examined by nurse Roxana later that day.

Michael's BP monitor does not communicate via Bluetooth so he entered his measured values manually. He clicked on the logbook and scrolled to the BP column (Fig. 2e). But instead of reporting DBP of 90 mmHg he entered 900. The system detected the low quality data and asked him to re-enter (Fig. 2f). While he was on the logbook, he also decided to visualize his data collected so far (e.g., Fig. 2g shows all BP data).

Michael felt some palpitations that afternoon. So he clicked on the green face with a grimace (top of Fig. 2a) and **reported unacceptable palpitations** from the pull-down menu (Fig. 2h). He then received an automatic clinical recommendation from the MobiGuide DSS to monitor his ECG (Fig. 2i). Michael also remembered that he has a dental appointment two weeks from now. So he informed the system of this **risky event for bleeding** (Fig. 2j). The system recommended to call his doctor to appropriately manage anticoagulant therapy.



Fig. 2. MobiGuide patient UI. (a) regular first screen for Atrial Fibrillation (AF) (the Gestational Diabetes Mellitus [GDM] screen does not include the AF Monitoring icon); (b) active prescriptions; (c) - medication reminders with a place to indicate non-compliance with reason; (d) ElectroCardioGram (ECG) monitoring screen; (e) logbook with Blood Pressure (BP) data entry; (f) - bad quality of data message; (g) visualizing BP records; (h) selecting palpitations level; (i) recommendation to measure ECG due to symptom reporting; (j) informing the system of risky events (in this case risk of bleeding).

The paper is organized as follows. We start with a section that describes CIG formalisms and their patient models. This section covers the work of other research groups and also summarizes our previously published early work on the architecture of MobiGuide and on the process of customizing its formal knowledge base to the patients' psycho-social and demographic characteristics (Peleg 2013). In the Methods section, we explain in detail the knowledge elicitation and specification methodologies that we have

developed (as part of the MobiGuide project) for making CIGs patient-centered and enabling their personalization via the patient model. We then present the evaluation methods and results of a feasibility study, in two different clinical domains, and two geographic sites, as part of a multi-national feasibility study, of the personalized self-management MobiGuide architecture that we have designed and implemented, which applies complex CIGs to support patients and their care providers. We conclude with a Discussion section.

2. CIG formalisms and their patient models

In this section, we provide the necessary background regarding CIG formalisms in general, and the methods that we had previously developed to implement the MobiGuide system (Peleg 2013) in particular; we then complete the necessary background and paint the complete picture, by reviewing other existing approaches for CIG customization, pointing out their differences with respect to the MobiGuide model.

This background, describing the principles of our knowledge engineering process, will be useful in order to better understand the rest of our methodology as listed in the Section 3 (Methods), as well as to better appreciate the results that we describe in Section 5 (Results).

2.1. CIG formalisms: their patient models and application to patient data

Most of the CIG formalisms, such as Asbru (Miksch et al. 1997)— the CIG formalism used in the MobiGuide project— and also EON, GLIF3, GUIDE, and *PROforma* (Peleg et al. 2003), include *procedural* and *declarative* representations. In addition, all of them have application engines that can be used to run the model with a patient’s data. In MobiGuide, the Picard Asbru application engine (Shalom et al. 2016) is used.

The *procedural* representation specifies care processes as networks of task classes, such as medical actions, decisions, and compound plans. The control flow of task application is governed by scheduling constraints, task entry and exit criteria, and decision criteria. The decision criteria refer to the state of the care process and to abstractions of patient’s state, which are defined as declarative knowledge. Hence, the declarative CIG knowledge defines the patient model.

The various CIG ontologies make different assumptions regarding the available knowledge and patient data; but none of them explicitly refer to the patient. In addition, different CIG frameworks use various means to extract meaningful patterns (abstraction) from the raw, time-stamped clinical data. The definitions of clinical concepts and abstractions constitute the declarative patient model and they make assumptions about the patient information model used (Peleg and Gonzalez-Ferrer 2014). The declarative temporal-abstraction knowledge in MobiGuide, which is used to abstract raw time-stamped clinical data into clinically meaningful higher-level concepts and patterns, in a

context-sensitive fashion, is represented using the *knowledge-based temporal-abstraction (KBTA)* ontology (Shahar and Musen 1996), which is applied to the patient's longitudinal record by the KBTA problem-solving method (Shahar 1997), implemented by the IDAN *temporal-abstraction mediator* (Boaz and Shahar 2005) and thoroughly evaluated in several clinical domains (Shahar and Musen 1996)(Martins et al. 2008). The abstractions of the patient's data are often complex (e.g., a "repeating dietary non-compliance" pattern, based on two diet-associated non-compliance events -too much or not enough carbohydrates- during the past week).

Some CIG models include not just a patient model but a model of the clinician user. The Dementia Management and Support System (Lindgren 2011) provides advice to healthcare professionals, tailored to individual and often exceptional patient cases. The system is unique in providing the user with a learning environment that promotes the development of skills while assessing a patient case. This strategy acknowledges the different professional backgrounds, preferences, and different needs for individually tailored support. The system alerts the user when there is an ambiguity or information missing, or there is certain pre-defined types of ignorance detected in the user.

2.2. The MobiGuide architecture: Applying decision support to the dynamic patient model

Unlike commercial efforts such as Microsoft's HealthVault (<https://www.healthvault.com/us/en>), which focus on patient-centered storage and sharing of health information online, MobiGuide's centrality is in its knowledge-based decision-support services. These decision-support services provide evidence-based clinical recommendations that are specific to a patient's model that addresses the patient's clinical, psychosocial and demographic state. The architecture that we have developed for MobiGuide (Peleg 2013) is generic and was reused in different medical domains with different clinical guidelines, various sensors for collecting patient data and several electronic health records (EHRs). To enhance efficiency and self-sufficiency, much of the DSS's computation is performed locally for each patient. Thus, MobiGuide's DSS incorporates a novel distributed architecture (See Fig. 3). Within this architecture, a fully-fledged backend-DSS (**BE-DSS**), that operates for all patients, projects (projection & callback link) relevant parts of the **CIG knowledge** to the patient's local, mobile-based DSS (**mDSS**). The mDSS runs on the patient's Smartphone and has access to data collected from mobile sensors that measure the patient's biosignals (e.g., a mobile ECG monitor for AF patients and mobile blood glucose (BG) and blood pressure (BP) monitors for GDM patients). The mDSS can then act independently until the patient's state changes substantially, necessitating a *callback* to the BE-DSS, which projects new knowledge to the mDSS. The BE-DSS has access to the complete **CIG knowledge** (which is customized to allow individualized personalization as explained below) and to the patient model (i.e., the full historical and monitoring data, current context, clinical state, current therapy and individual preferences). The BE-DSS can also send evidence-

based recommendations to the patient’s care provider, through a dedicated interface. In addition, the BE-DSS interacts with a **shared decision module** that uses decision trees to select a CIG branch that best fits the patient's utility function.

Another central difference from efforts such as HealthVault is that MobiGuide’s personal model, which is called the *personal health record (PHR)* (Marcos et al. 2015), conveys to medical domain semantics and is based on a standard patient information models (HL7 vMR - http://www.hl7.org/implement/standards/product_brief.cfm?product_id=338). The PHR refers to controlled clinical vocabulary standardized terms (Bodenreider 2004), such as UMLS and SNOMED-CT which were used in MobiGuide. The PHR *semantically* integrates data from hospital **EHRs**, data collected from the mobile monitoring devices (e.g., electrocardiograms, BG), symptoms reported by patients, patient-specific preferences, and patient-specific recommendations and abstractions delivered by the DSS.

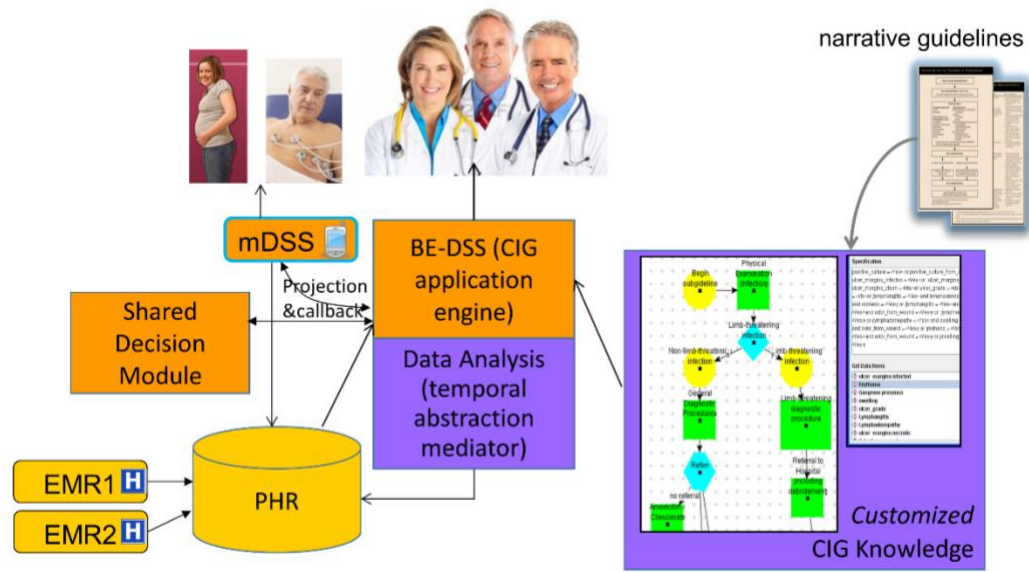


Fig 3. The overall architecture of the MobiGuide system

2.3. CIG formal specification, customization, personalization, and application in MobiGuide

The patient-centered care process in MobiGuide introduces two new steps into the CIG development and application process, thus creating a process that we refer to as the CIG *formal specification, customization, personalization, and application (SCPA)* process (Peleg et al., 2013) (Fig. 4). As in the non patient-centered case, the process of CIG development starts with manual **formal specification** of the evidence-based guideline as a CIG by a modeling team of knowledge engineers and clinical experts (Shalom et al. 2008).

The two next steps are unique to the patient-centered focus of MobiGuide. In the **customization** phase, the modeling team extends the CIG, so that it will now include also various new CIG-customized contexts (CCCs) and corresponding CIG branches that were not taken into consideration in the original narrative guideline, which was typically intended to be applied in a hospital or an ambulatory clinic context. These new contexts are crucial for applying the guideline through a remote-care, mobile-based architecture such as MobiGuide, and the new context address the patients' recurring psychosocial domains (e.g., whether patients have routine or semi-routine schedule) and technological domains (e.g., low Smartphone battery), in addition to clinical context (examples provided in Fig. 4).

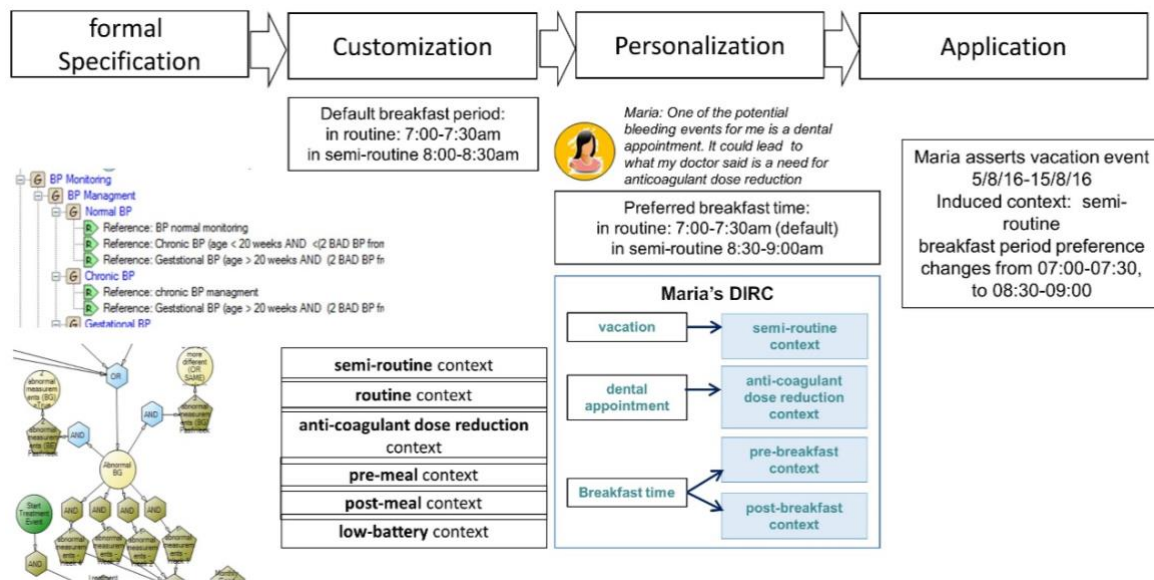


Fig 4. The Computer Interpretable Guidelines (CIGs) formal specification, customization, personalization, and application (SCPA) process in the MobiGuide project.

CCCs, like clinical contexts introduced into the CIG at the formal specification phase, can be *induced* by specified events, within predefined temporal constraints. Thus, an event of the type “a potential risk of bleeding,” induces the “anti-coagulant dose reduction” context, which starts five days before the scheduled bleeding risk event and finishes one day after the event. The methodology of *Dynamic Induction Relations of Contexts (DIRCs)*, was originally introduced by Shahar (1998), as one component of the KBTA methodology (Shahar, 1997). It can handle also *composite* (overlapping) interacting contexts, as well as multiple types of temporal relations and constraints between the inducing event and the induced context.

In addition, during the CIG customization phase, the modeling team prepares the CIG for the inclusion of patient personal preferences:

(a) *Global clinical preferences*: patient preferences that determine the selection of a major branch in the guideline over another. For example, preferring Dabigatran over

Warfarin when both are reasonable options as anticoagulants, in the case of AF. This preparation involves adding the appropriate shared decision model that enables a choice of the CIG branch that best fits the patient's utility function (Quaglini 2013);

(b) *Local clinical preferences*: local modifications within a guideline branch, or context. For example, preferring a particular morning hour for a reminder to measure a GDM patient's fasting BG. This preparation involves making explicit options such as choosing the preferred time to get a reminder, etc. Once a CIG is customized to refer to meal times, specific patients may set their actual meal times during the following personalization phase.

It is important to emphasize that at the end of the customization phase we are left with a single CIG that could be applied to any patient. This CIG does not contain information to any specific patient.

Personalization is the second process that is unique to patient-centered CIGs. This process usually takes place during the first encounter of the patient with a care provider. Together, they define (a) which personal events might induce any of the pre-defined CCCs (and *when*, i.e., the relevant dynamic temporal constraints); and (b) the patient's global and local preferences regarding her treatment. These are stored in the PHR.

For example, in the case of the AF domain, the physician would always ask the patient, which events might lead to the predefined *semi-routine* context, and the patient may indicate that this CCC is induced by a *Vacation* event. The physician then adds "vacation" as a personal event that induces the semi-routine context for this patient. Similarly, the physician presents to the patient a list of potential bleeding events (risky events) and the patient can select some that are applicable, such as "visiting a dentist".

During the CIG **application phase**, patients will have the option to notify the system of their personal events. For example, the patient could report a vacation event or a dental appointment. These will be saved into the PHR. The BE-DSS will be notified by the PHR and the respective CCCs (*semi-routine* or *anti-coagulant dose-reduction*) would be induced (see Fig. 2j and Fig. 4).

With the two added phases, the CIG Specification, Customization, Personalization, and Application is termed **SCPA**.

2.4. Other existing approaches for CIG customization

The dynamic context-based personalization mechanism used in MobiGuide is novel and has not been used elsewhere. However, other researchers, whose studies are reviewed below, have customized CIG knowledge to infer patient-specific recommendations that address more than the patient's clinical data relating to a single disease. Such recommendations can address a patient's multi-morbidities, her social history, or considerations of the setting of a particular healthcare organization at which the patient is cared for. Unlike the SCPA approach, none of the existing studies have attempted to designate personal patient events as *inducers* of pre-customized CIG contexts (which might, for example, lead to activation of a new CIG plan).

Riaño et al. (2012) represent CIGs as State-Decision-Action (SDA) that are linked with clinical domain ontologies. They developed algorithms for adjusting patients' conditions based on disease profiles; these profiles are consulted to suggest additional signs and symptoms that the patient is likely to exhibit and which could be used to generate a more complete record. In turn, the complete record may help to establish the patient's diagnosis. For patients with multiple morbidities, they suggest a tool that provides a graphical interface to edit and merge SDA diagrams of individual intervention plans for comorbidities of the patient. Intervention plans and decision criteria may address the clinical as well as the social context of the patient.

Grandi and coauthors (Grandi et al., 2012)(Grandi 2016) suggest efficient management of multi-version CIGs collections by representing, in a knowledge base, multi-version clinical guidelines and domain ontologies in XML or in relational schemas. Personalized CIGs can be created by building on demand, from the knowledge base, which also contains historical versions of the guideline, a version that is tailored to the patient's current time (or desired temporal perspectives, e.g., what were the recommendations that would have been delivered by a previous CIG version) and to the patient's disease profile (set of comorbidities, e.g., hypertension in addition to AF). The versions could be tailored also to organizational settings in which the patient is cared for. While this approach allows creating different versions of CIGs over time, which are not supported in MobiGuide, these versions do not include individual preferences; they differ only in the composition of the different parts of the customized guidelines that are assembled together to create a guideline version for a particular patient.

Lanzola and co-authors (2014) developed guideline-based process indicators related to stroke care. The computation of those indicators is done within a stroke registry, which also stores the historical evolution of the resources available at the participating stroke units. While this is useful to justify low values of some indicators in some time periods, again, patients' preferences are not considered. Lasierra and co-authors (2013) describe the clinical concepts related to a monitoring task using two levels. The configuration level characterizes the clinical concept (e.g., systolic blood pressure is a vital body measurement) while the results data level characterizes the data used to acquire the concept and contains a more detailed patient model related to the current measurements. For example, whether the patient was sitting or lying down when the measurement was taken. Based on these characteristics, physicians can create personal versions of CIGs, for each patient. Note, that in their approach, there is no distinction between knowledge customization that fits a population of patients and personalization that links personal events to customized CIG contexts.

In their adaptive DSS called PRESYDIUM (Personalized Emergency System for Disabled Humans), Chittaro et al. (2011) developed a patient model that is partly based on the International Classification of Functioning, Disability and Health standard proposed by the World Health Organization. Important attributes from this standard include cognitive functions, body functions, mobility of joints, motor control, pain, and involuntary movements. The knowledge base of their system does not contain a procedural CIG model but rather a collection of frames defining rules. Given a patient

profile, frames with corresponding conditions are examined and are filtered based on the current emergency code and user category, to provide operating instructions.

3. Methods

This section presents the methods that we developed for knowledge acquisition. State of the art knowledge elicitation methods start from a narrative clinical guideline and define the semantics of the care process (Shalom et al. 2008). Our novel methods address aspects that are not present in the evidence-based clinical guidelines which view clinicians, rather than patients, as the target of the DSS recommendations. Our methods are used to detect opportunities that allow patients to use IT-technologies for making CIGs patient-centered and enabling their personalization. They include acquiring and specifying knowledge of parallel workflows and projection and acquiring CCCs and their effect on care plans. The parallel and customized CIGs could then be applied by the Asbru CIG engines mentioned in Section 2.1. The Methods section also discusses the personalization of the AF and GDM CIGs that were implemented and used in the evaluation of MobiGuide.

3.1. Methods for acquiring relevant CCCs and their effect on care plans

The customization step requires addition of CCCs that were not included in the Specification phase of the original clinical practice guideline, and corresponding care plans that should be applied for the CCCs. To support the teams of clinical experts and knowledge engineers who carry out the SCPA process in considering potentially relevant CCCs and in assessing their impact on clinical care, we apply the following strategies:

- (a) Examining all CIG decision points, and thinking whether different psycho-social and demographic characteristics of patients could impact decision-making at those points, changing the guideline's recommended plans and actions (Fux et al., 2012).
- (b) Thinking about scenarios that arise when patients are guided by a DSS such as MobiGuide, outside of clinically-controlled environments. For example, during travel.
- (c) Characterizing how care recommendations might change, by thinking about different types of modulations of the clinical goals and actions themselves.

To implement these strategies, we used qualitative research methods, including structured questionnaires, interviews, and text analysis. We then used the elicited psycho-social contexts and the effect types to develop an elicitation instrument for thinking of relevant CCCs while trying to customize a particular clinical guideline.

We now outline our context-elicitation methodology in more detail.

A. Elicitation of generic psycho-social and demographic concepts and their potential effects on clinical goals and actions during guideline customization

To elicit generic psycho-social and demographic concepts and their potential effects on clinical goals and actions, we developed a set of questionnaires. The questionnaires were scenario-based and included parts that are increasingly structured. This was done in order to focus the interviewee on a top-down thinking process, thereby eliciting increasingly detailed responses, as it has been indicated (Pitts and Browne 2007) that a systematic thinking process during elicitation serves for avoiding cognitive biases (Kahneman and Tversky 1982). The questionnaires to clinicians and patients are shown in appendices A and B. The interviews were conducted using the questionnaires and were held with different stakeholders, including eight patients, twenty-one physicians, and nine other care providers (nurses, social workers, and nutritionists) from Israel, Italy, Spain, and USA.

Thus, the care providers were first asked to assess the main psycho-social and demographic contexts that might affect the clinical decision-making process. The second part of the first questionnaire lists several general situations. The questionnaire describes a patient who complains that he/she suffers from different symptoms. The care providers were asked to describe the general scenario with emphasis on psycho-social and demographic information, including which information items they collect during the medical interview and how they take them into consideration when they set clinical goals and make therapy choices. In the third part, clinicians were asked to think of patients of theirs that already have an established diagnosis (e.g., AF, hypertension) and describe the process of patient management, addressing major decision points and the role of psycho-social and demographic information items in their decision-making process. The fourth part considers the potential of the MobiGuide system and the goal of personalization. A scenario of a diabetic patient using MobiGuide is described in this part. The interviewees were asked to create an analogous scenario from their clinical domain.

Then, the interviewees were asked to think of the information that they expect to gain from the system in the process of patient follow-up, treatment and work up, considering decision points in the process, the potential use of ambulatory monitoring devices and the relevance of such data and of psycho-social and demographic information to decision-making. The interviewees were then asked more structured and technical questions regarding psycho-social and demographic data: their allowed range and valid time, who would be able to enter and modify those data items, and whether specific values of such data would require immediate notification to patients or care providers.

Similarly to the care providers' questionnaire, the patients' questionnaire referred to a series of treatment encounters that the patient had with his/her care provider. In the first part, the patients were asked to describe the course of the meeting with emphasis on treatments and decisions. They were asked how psycho-social and demographic considerations affected treatment on the one hand, and influenced their daily routine (e.g., work, family life, hobbies, habits, diet) on the other hand. As in the clinician's questionnaire, a second part considered the introduction of the MobiGuide system. The

patients were asked how the system and approach would change the way in which they handle their treatment and adherence.

After analyzing the interviews, we defined the emerging psycho-social and demographic context-inducing concepts: ability to comply with the treatment (i.e., fully by the patient herself, fully while assisted by care provider, partially by the patient herself, partially while assisted by an informal care giver, or no ability); communication level, which can be aggregated from cooperation level, desire to know the truth about the prognosis, education level, language level, and trust level; need for an accompanying person during visits to the care provider; degree of family support; whether the daily schedule and the daily diet are routine, semi-routine, or completely non-routine; distance from the nearby medical center; financial capability; living area accessibility, living area pollution, and family-support level.

These concepts were used in a second questionnaire to care providers (Appendix C), which focused on elicitation of effects on clinical goals and actions. Elicited effects included: target thresholds of concepts (e.g., pre-prandial BG \leq 100 mg/dL), appointment schedule, diet change, measurements change, medication change, physical activity change, and change in the mode of communication with the patient.

B. An instrument for elicitation of relevant CCCs and of their effects on treatment

Based on the generic psycho-social and demographic contexts and on the types of effects that they might have on clinical goals and actions, and on the experience gained from the first sets of interviews and on additional interviews held with clinicians, an elicitation instrument (Appendix D) was developed to facilitate guideline customization.

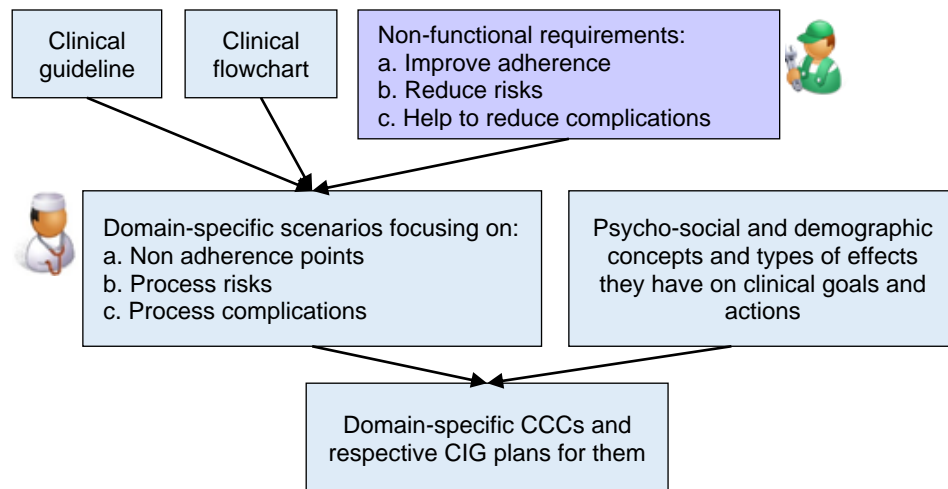


Fig 5. A methodology for elicitation of CCCs which customize clinical guidelines

The elicitation process is presented in Fig. 5 and is carried out by the SCPA modeling team. The elicitation instrument is used within the context of considering a specific

clinical guideline and the specific clinical decision points established within the guideline. Ideally, the clinical guidelines should contain flowcharts representing the structure of the guideline, as we would like to go over it and check which non-clinical aspects might change a specific treatment step. Therefore, if such a flowchart is not included in the guideline, then the first step would be to prepare this flowchart, focusing on the decision points related to the clinical state of the patient, while paying attention to non-adherence points and to process risks and complications. To establish the treatment steps in the guideline that are affected based on the generic psycho-social and demographic contexts, the instrument contains a set of questions that guide the thinking process needed in order to introduce CCCs and respective care plans into a clinical guideline.

3.2. Knowledge acquisition and specification of parallel workflows and projection

After achieving a consensus regarding the semantics of the care process based on the evidence-based clinical guideline, the elicitation task splits into the two subtasks of eliciting two parallel workflows (Sacchi et al. 2013). The first is a “traditional” workflow directed at the care professional; the second is a parallel process that focuses on the patients’ behavior and on their interaction with the MobiGuide system (i.e., both the Smartphone and the sensors).

For example, a traditional guideline may define a plan for monitoring the patient’s compliance to diet as a set of instructions to a nurse to check whether the patient reported in her diary of eating too many carbohydrates more than twice during the past two weeks (non-compliance). In the parallel workflow (García-Sáez et al. 2013), this recommendation is translated into an automatic evaluation of the patient’s non-compliance condition every week and delivery of an alert to the patient through the Smartphone. The automatic evaluation is carried out by retrieving data from the patient’s digital log book and checking for a temporal pattern that occurred during the past two weeks. This part of the parallel workflow is managed by the mDSS, hence the relevant CIG plans are indicated as projection points (García-Sáez et al. 2014) and allow at CIG application time the passing of control to the mDSS. The mDSS receives the relevant knowledge in the projection format, which consists of one or more procedural scripts running in parallel.

3.3. The modeled and implemented CIGs and the recommendations they provide

Using the elicitation methods described above, we formalized, customized, and personalized two CIGs in very different clinical domains: atrial fibrillation (AF) and gestational diabetes mellitus (GDM). We selected these clinical domains as representing chronic conditions that require monitoring with different sensors and with different patient populations so as to demonstrate that the same architecture, modeling languages,

and software components, could be used in different domains. The system architecture and patients' and care providers' user interfaces for the two clinical domains were the same (see Fig. 2 for the patients' user interface). However, the implementations varied (Table 1) in the types of data collected and patterns monitored, in the amount of interaction between the BE-DSS and mDSS, in the types of customized context, and in the amount of patient-specific recommendations and notifications that they provided to patient and clinician users.

Table 1 compares the knowledge of the two CIGs according to overall characteristics of the two domains. The number of concepts (e.g., fasting BG result), data patterns (e.g., two weeks of negative ketonuria), and criteria provide a general impression of the complexity of the modeled clinical care flows.

Table 1 also shows the different types of monitoring plans followed by patients. Plan activation was controlled by plan entry and exit conditions, which specify conditions that addressed clinical or customized context, which are part of the CIGs' knowledge. Upon CIG activation, these conditions are matched against each patient's personal model to yield patient-personal decision-support recommendations and reminders. In (MobiGuide Consortium 2016), we provide examples of the monitored patterns and conditions, and sample notifications and recommendations sent to patients and care providers, used in the two CIGs.

Table 1. Difference in CIG characteristics between the two clinical domains

	AF	GDM
Raw concepts	100	300
(Temporal) data patterns	71	124
Conditions/Criteria	20	69
# Customized context	4 (semi-routine, routine, 24h monitoring, increased physical activity)	2 (semi routine, routine)
Monitored measurements	ECG, BP, weight [§] , INR [#] , physical activity*	BG, ketonuria, BP, weight*, physical activity*
Notifications to patients	7	10
Notifications to care providers	20	2
Recommendations to patients	5	1
Recommendations to care providers	18	7
Callbacks from mDSS to BE-DSS	2	16
Periodic plan projections (for reminders)	18	22
Monitoring plan projections (for detecting data patterns that trigger plan changes)	2	17
Control mostly by	mDSS	BE-DSS

§ prescribed only to 2 patients; # prescribed only to one patient; * not prescribed during personalization.

CIG customization: CCCs and projection points

During customization, relevant CCCs are elicited using the methods described in Section 3.1, and added to the CIGs. Four CCCs were added to the AF CIG: semi-routine and routine schedule, 24 hour monitoring, and increased physical activity. Monitoring plans for these contexts vary (two rather than one daily BP measurements for hypertensive patients with BP monitoring prescriptions while in semi-routine context, in which patients may not be as well controlled; two rather than one daily 30-minute ECG monitoring sessions when in increased physical activity context; continuous ECG monitoring when in 24 hour monitoring context).

Two CCCs were added to the GDM CIG: routine and semi-routine context. However, CIG plans did not change; only preferred meal times and reminder preferences could be set by individual patients for these two contexts.

During individual personalization at enrollment, relevant CCCs were initialized and configured with meaningful names for the patients (e.g., “holiday” for semi-routine context). Patients could then activate relevant context using their UI. 24 hour monitoring had to be triggered by the clinicians.

Using the knowledge elicitation methods defined in Section 3.2, projections and call back points were defined in the CIGs. In the case of AF, most of the knowledge was controlled by the mDSS, as is reflected by the small number of callbacks defined in the knowledge base. On the other hand, most of the knowledge of the GDM guideline was defined in the parallel workflows for patients and care providers.

Monitored measurements

Table 2 presents the expected number of weekly measurements for AF and GDM patients.

All AF patients were prescribed monitoring their ECG for 30 minutes daily via mobile monitors. These monitors send the biosignals to the patient’s Smartphone using Bluetooth, from which the data is sent to the Backend server at the hospital. The nurse would look at their ECGs once a day to assess whether the MobiGuide AF detection algorithms correctly identified AF events, and would inform the doctor if the situation necessitated intervention, such as change of diagnosis or therapy. All patients received reminders for ECG monitoring at their preferred times, which were context-dependent.

Table 2. Expected number of weekly measurements for AF and GDM patients

Context / Entry condition (CCC shown in bold)	Measurement/ recommendation	# reminders/ recommendations	Total #weekly reminders
AF patients			
<i>Measurement and medication reminders</i>			
All AF patients	ECG ¹	2/day	30-42 depending on BP stability and routine/semi routine context; 30/week for typical patients
Hypertensive patients in routine context with stable BP	BP	2/week	
Hypertensive patients in routine context , unstable BP	BP	1/day	
Hypertensive AF patients in semi-routine context	BP	2/day	
Patients with prescriptions	Medication reminders ²	~2/day	
Patients taking anti-coagulants	INR ³	(1/2weeks	
Prescribed patients	Weight	1/week	
<i>Clinical recommendations</i>			
Patients on anticoagulants reporting bleeding risk event	Clinical recommendations relating to risk events	Sparse, maximum 1/month	Sparse
Patient-reported AF symptom	Measure ECG		
GDM patients			
<i>Measurement reminders</i>			
BG, ketonuria, and BP values controlled	Ketonuria	2/week	11
	BG	2/week; 4 times a day on those days	
	BP	1/week	
BG, ketonuria, or BP values <u>not</u> controlled	Ketonuria	1/day	42
	BG	4/day	
	BP	1/day	
<i>Clinical recommendations</i>			
Positive ketonuria value detected	If patient answered that ate enough carbs, advise carbs increase; if not first time: see doctor	sparse	Sparse, maximum # of recommendations on any given day: 2
Two high BG values detected within one week	If patient answered that ate enough carbs, ask to see doctor; Insulin recommended (to doctor)		
Negative ketonuria values for	Lower ketonuria	maximum 2/month	

two weeks (ketonuria controlled)	monitoring frequency		
Good BP values for two weeks (BP controlled)	Lower BP monitoring frequency	maximum 2/month	
Good BG values for one month (BG controlled)	Lower BG monitoring frequency	maximum 1/month	
Typical patients that are stable half the time			26.5

¹ AF patients were expected to measure their ECG once a day but received two reminders each day in order to increase their chances of seeing the recommendation

² Medication reminders arrived according to drug prescriptions for all morbidities of the patients, which typically ranged from 0-3/day; 2/day on average.

³ INR, an assay meant to assess blood coagulation was to be measured at the hospital every two weeks by all patients taking anti-coagulants.

AF patients were optionally prescribed to also measure their BP, weight, or INR and enter it manually into the system. The rate of recommended BP measurements depended on context (See Table 2).

The clinical recommendations and measurement frequency for GDM patients did not depend on the patients' psychosocial context. The frequency of measuring BG, ketonuria and BP depended on the clinical context (stable or unstable), as shown in Table 2. In case an entry condition became true (see lower part of Table 2), monitoring plans were switched automatically by the BE-DSS and the respective knowledge was projected to the mDSS.

All GDM patients were provided with portable glucose meters and BP meters that sent data to the Smartphone via Bluetooth. The data was downloaded by the patients to the backend via a menu button on the phone. Patients also had to measure ketones in their urine and manually record results into the system. The GDM system included an option for weight monitoring prescription, although none of the patients on the study were prescribed to do so. Physical activity monitoring plans could be prescribed, but were not prescribed by clinicians in this study. This monitoring was supported using the physical activity detector (PAD) integrated in the MobiGuide application, which provides information about physical activity intensity, energy expenditure and steps and is based on the Smartphone's accelerometer. Patients activated the PAD and ended sessions manually; continuous monitoring was not a feature of the system.

When patients enrolled into the MobiGuide system, they determined the setting of whether they wanted to receive reminders for monitoring and in the case of AF, also for medications (related to AF and to other comorbidities). They also set the timing of reminders, which could be different under different contexts.

Patients' self reporting of events and clinical recommendations

AF Patients could report symptoms related to AF by selecting them from a drop-down list and indicating their severity (unacceptable, acceptable, absent). Symptom reporting triggered a recommendation to measure their ECG. Additionally, patients could report personalized future events with risk of bleeding (e.g., dental visit) which could indicate change in anticoagulation therapy, resulting in a system notification to schedule an appointment with their doctor. Upon delivery of medication reminders, patients could report taking the medication or not taking it, indicating their reason for non-compliance.

GDM patients could self-report diet non-compliance events and could answer questions sent via the mDSS regarding following diet prescription. The bottom part of Table 2 presents the different clinical recommendations delivered by the system in response to detection of a patient's patterns of clinical data, depending in some cases on answers to questions posed by the system.

Recommendations to clinicians

In the case of the AF guideline, most of the recommendations to care providers concerned cardioversion decision support. However, in practice, only one patient was eligible for cardioversion. In GDM, recommendations to clinicians concerned starting insulin and changing diet prescriptions.

4. Evaluation

We tested the MobiGuide system in an extensive pre-pilot testing with healthy volunteers (described in Appendix E). This testing has shown that the system responded (adapted) correctly to changes in the dynamic patient model (see Fig. 1) and delivered appropriate recommendations.

The evaluation of the system was then performed in a multi-national feasibility study in the clinical domains of AF and GDM. This paper reports results that focus on the patient model and its impact on users' behavior. An analysis of compliance to specific recommendations and reminders is addressed in a separate paper that examines clinical outcomes.

In the case of AF, the system evaluation study was planned for a maximum of 9 months, considering that the study was done with clinicians' follow up that was in addition to normal patient care.

Gestational diabetes usually starts after month 5 of pregnancy; the GDM patients could use the system till delivery of the babies, upon which GDM resolves. Thus the evaluation study was planned for a duration of at least 2 months for each patient and a maximum of 5 months.

In addition to the variability of the CIGs for these two guidelines (see Section 3.3), the patient user population varied as well (Table 4): while AF patients were older Italian

men and women, chronic patients, had additional comorbidities and were much less experienced with technology, the GDM patients were Spanish women with an average age of 35, more experienced with technology, who are otherwise healthy and have complications of pregnancy related to diabetes with/without hypertension. We recruited ten AF patients and twenty GDM patients for the evaluation study. They used the system during April-December 2015.

4.1. Hypotheses

The technical evaluation of the system was successful and showed that the system reacts to the dynamic patient model. In order to design an analysis that would focus on the users' experience and how it was related to the personal patient model, we defined the following three hypotheses:

1. Sustainable usage by patients. Most patients will not drop out and will use the system for a long duration of time continuously. For AF patients, we considered that patients could be enrolled until 2 months before the end of the project, thus we expect follow-up lengths from 2 to 7 months. For GDM patients, the expected time period was constrained by the gestational age at which the patient was enrolled in MobiGuide and the gestational age at delivery.

We conjectured that this behavior could be attributed to a combination of two factors. First, patients should find the system useful and usable, and second, patients in high risk domains are intrinsically motivated.

One of corollaries of this hypothesis is that at first, patients might use the system more than later, but they would still perform at least about 60 actions a week for AF patients and 53 for GDM patients. This estimation is based on the numbers reported in Table 2. The number of reminders and recommendations that a typical patient should receive depends on the patient's context and is estimated as 30 per week for AF patients and 26.5 per week for GDM patients (see Table 2). In addition we expected that each patient should view (click) on such reminder/recommendation and also respond by performing a measurement and entering the value into the system for manual monitoring (BP and weight for AF patients and ketonuria for GDM patients).

2. Patients will have positive impressions of the system and will find it useful, especially for increasing their sense of safety and communications with clinicians.

3. Clinicians will be examining the PHR data (patient model) of patients during visits and in between visits.

The norm is that clinicians examine the patients' medical records only when the patients are coming in for scheduled visits. So, if the clinicians check the patients' data

more often while using the system, then this is a positive outcome, which shows that clinicians react to the changing patient model.

Note that the changes in patient model happen “under the hood”, within the MobiGuide architecture. Implicitly, the fact that clinicians are inspecting the patient model demonstrates the effect of the changes in the patient model on clinicians’ behavior.

4.2. Methods for assessing the hypotheses

Table 3 maps the hypothesis to the methods for their evaluation and the data sources used to evaluate each hypothesis.

Table 3. Methods and data sources for assessing the hypotheses

Hypothesis	Method	Data sources
1. sustainable usage (comparison with expected weekly usage)	-AF: total #days that patients used the system: mean, SD -GDM: ratio of days spent using the system out of the total number of days that each patient could potentially use the system (till delivery); -#recommendations and reminders sent to patients and clinicians; ratio of recommendations/ reminders viewed out of total generated -Graphed #weekly actions done by AF and GDM patients as a function of time	PHR: logs of monitoring (manual or download from sensors); reporting of symptoms; recommendations delivered. Patient GUI log: viewing the logbook; viewing recommendations and reminders; modification of user settings.
2. patients perceptions	End of study questionnaire. Six questions for assessing the general impression of the system and 11 questions regarding usefulness, as shown in Table 7. Questions regarding usefulness were designed to assess the main objectives of the MobiGuide system and the intention to use it beyond the life time of the project. Apart from 3 yes/no questions, all other questions were on a standard Likert scale of 1 through 5.	
3. clinicians views of patient data during and outside controlled visits	AF: one-tailed Wilcoxon signed rank test to compare the number of days that clinicians viewed a patient’s model to the expected number of visits. Similar analysis was done for the cohorts of patients. -Assess for how many patients clinicians changed diagnosis or therapy due to system advice - Find most common physician action type GDM patients had weekly visits on Tuesdays.	Clinicians’ actions were recorded in caregiver interface usage log. Patients’ actions were recorded in the Patient GUI log.

	For each patient, and each day of the week, we counted the total #times that clinicians viewed this patient’s model on that day. We graphed the weekly patterns	
--	---	--

5. Results

The MobiGuide system delivered correct recommendations and reminders according to the personal model. The results of the patients’ usage of the system are reported below, arranged according to the hypotheses, which are related to the effect of the personal patient model in users’ behavior. Clinical outcomes, including compliance to clinical recommendations, will be reported in a separate publication and are briefly mentioned in the Discussion section. We first present the characteristics of patients who participated in the study, followed by the results related to the three hypotheses.

5.1. Patient characteristics (demographics)

Table 4 compares the two patient populations in terms of the number of participating patients, their age and gender, level of education and technological skills, the average number of days spent using the system, dropout rate, and number of times the patients had switched context.

Table 4. Characteristics of the two patient populations (AF and GDM)

	AF	GDM
Number of patients	10	20
Age	66.3±9.2 years	35.2±3.9
Gender	7 males; 3 females	20 females
Level of education	Elementary(1); secondary school(3); high school(3); university(2)	Elementary(2); secondary school(2); high school(1); university(12)
Technological skill*	High*(2); Low (8);	Low (1); High (16)
Average and SD of number of days using MobiGuide	127.2±68.6 = 4.2 ± 2.3 months (and up to 8.6 months)	57.1 ± 21.5 = 8.2 ± 2.99 weeks (and up to 3 months)
Number of patients who dropped out	0	1
Number of patients who switched context	5	8
Number of patients eligible for pill-in-the-pocket scenario	0	Not applicable

Number of patients eligible for cardioversion advice delivered to physicians	1	Not applicable
--	---	----------------

* used Smartphone and laptop/tablet

5.2. Hypothesis 1: Sustainable usage by patients

Tables 5 and 6 report in the column “Days on MG” the total number of days that patients had used the system. These tables also report results concerning the correct delivery of decision-support, which is specific to the dynamic and reactive patient model. This includes delivery of recommendations and reminders according to CIG logic that addresses the personal patient model.

Tables 5 and 6 show the number of recommendations that were sent to clinicians and patients (including clinical and technical recommendations for patients); recommendations are sent in response to declarative knowledge expressions that the system monitors, and when they become true, the system reacts by sending relevant recommendations (see (MobiGuide Consortium 2016) for a list of the available recommendations). We did not have estimates for the expected frequency of recommendations as they depend on the specific conditions that trigger these recommendations, which relate to the patients’ dynamic clinical state. We could observe that AF patients reported AF symptoms for which they received recommendations to add extra ECG monitoring sessions. The mean number of recommendations received was once a month (with the patient receiving ECG recommendations most frequently at 4.1 times a month). Other recommendations, related to events that are associated with a risk for bleeding, were quite rare.

GDM patients received more clinical recommendations than AF patients: on average 1.9 clinical recommendations a week (8.1 a month), with the patient receiving clinical recommendations most frequently at 6.4 times a week. But the numbers are not comparable as the clinical conditions to trigger these recommendations are different. AF patients received recommendations when they reported risky events or AF symptoms whereas GDM patients received recommendations when patterns relating to BG, ketonuria, or BP values occurred; some of these patterns could occur weekly or even every two days (for ketonuria). The frequency of technical recommendations delivered to AF patients was about five times higher than the rate of clinical recommendations. In the case of GDM patients, the frequency of technical recommendations was much lower than clinical recommendations (60%). Technical recommendations are related to quality of data entered, need to charge the battery, and problems in wearing the ECG monitor.

Recommendations to AF clinicians were quite sparse, because only one patient met the criteria for cardioversion advice. Recommendations for GDM clinicians were generated for all patients, but were still quite low as compared to the recommendations to

patients (around 1.5 recommendations a month per patient to clinicians vs. 8.1 recommendations to patients).

In addition, tables 5 and 6 show the number of reminders that the system delivered to patients and the number of times that recommendations and reminders were viewed by patients, for which we had estimates. Nine of ten AF patients, and all of the GDM patients, chose to receive measurement reminders and 6 of 8 AF with drug prescriptions chose to receive medication reminders. Observing the results shown in Table 5, we can see that the ECG reminders were issued by the MobiGuide system at roughly the expected frequency of two daily reminders¹. The number of drug reminders delivered depended on the therapeutic plan of each patient (some might have only one medication to be taken daily, while others may have multiple doses of several drugs each day) which can also change over time, thus we did not expect a fixed frequency for these kind of reminders. GDM patients received BG and ketonuria reminders at a lower rate than expected. Technical problems were reported for some patients such as 2, 13, 14, and 17, who received a scarce number of reminders.

All of the columns of tables 5 and 6, except for the last one, focus on correct system behavior. But an important component of Hypothesis 1 concerns expected patient behavior. There was much similarity between AF and GDM patients in terms of their usage of the system. All patients used the system mainly to monitor, record and view the most important parameters of their health conditions, i.e. ECG and BP for AF patients and BG, ketonuria, and BP for GDM patients. Viewing of the logbook data was correlated with the amount of data entered for each parameter. The logbook data included monitored parameters, therapies, patient-reported symptoms, risky events, medication non-compliance for AF patients and diet non-compliance by GDM patients – reported in the last column of Table 5 and Table 6. As can be seen, AF patients viewed at least 95% of the total number of recommendations and reminders that were delivered to them and GDM patients viewed 92%. However, this is the lower limit, because, when clicking on any logbook view, all data collected until that time was displayed, so patients could have seen several recommendations or several data values within a single click. In addition, the calculations regarding the number of data visualization and access to different features of the MobiGuide application, such as downloading sensor data, were performed based on the actions registered in a mobile log while patients were using the application. According to the design of the mobile application, patients need to visualize data before entering new measurements either manually or automatically. For this reason, the patients' data visualization patterns were computed based on the total number of visualizations performed minus the total number of monitoring data updated to the backend, per day. This calculation has limitations because some patients entered several measurements at the same time (e.g. glycemia values were measured on average 4.1 times a day but they were downloaded 2.63 times per week).

Fig. 6 shows the amount of actions done by AF and GDM patients over time. As shown, the total counts of weekly actions performed by both groups of patients was high

¹ Reminders were not delivered When the phone was OFF

(average and standard deviation of actions per week for AF is 200.2 ± 56.2 ; and for GDM 150.4 ± 66.6). Although most patients experienced a decline in usage over time, their usage remained for the whole period (the average number of actions per week in the last two weeks was 157.4 for AF patients and 110.9 for GDM patients). This number is by far higher than the envisioned number indicating successful usage by patients, which was estimated to be at least 60 for AF patients and 53 for GDM patients. A few patients for some weeks (observe AF patient 7 at week 24-25 and AF patient 10 at weeks 9-11) were using the system less. AF patient 7 brought the smartphone to the doctor's office for checking his report about a potential problem of drug reminders; AF patient 10 was measuring her ECG and using the app only when her daughter was around). One GDM patient used the system less when she started insulin treatment (see GDM patient 4 at week 4) while two patients had a lower number of actions on the first week due to technical problems (see GDM patient 1 and patient 2).

Most patients performed a higher number of actions during the first week of usage (average and standard deviation of actions the first week for AF patients was 365.7 ± 112.1 and for GDM patients was 327.3 ± 121.7), most probably they spent additional time learning about different scenarios of the application.

Table 5. Delivery of DSS for AF patients: response to changes in dynamic patient model.

Patient#	Days on MG	Recommendations to clinicians (total #)	Clinical recommendations to patients (total #)	Technical recommendations to patients (total #)	Reminders sent to patients (total #)	Ratio Recommendations & reminders viewed of total generated
1	96	0	13 (ECG-sym)	43	ECG: 192 BP: 11 Drugs: 236	0.92
2^	78	0	1 (risky event) 1 (ECG-sym)	4	ECG: 156 BP: 9 Weight: 1	1
3*	98	0	4 (ECG-sym)	6	ECG: 196 BP: 5	0.97
4	91	0	6 (ECG-sym)	- #	ECG: 182 BP: 58 Drugs: 40	- #
5	90	0	0	33	ECG: 180 BP: 13 Weight: 12 Drugs: 359	0.96
6	136	2 (cardioversion)	1 (anticoagulant pre cardioversion)	7	ECG: 272 Drugs: 305	0.97
7	259	0	10 (ECG-sym)	41	ECG: 518 BP: 41 INR: 11 Drugs: 776	0.97
8	89	0	5 (ECG-sym)	4	ECG: 178 BP: 1 Drugs: 155	0.99
9^^	249	0	7 (ECG-sym)	39	ECG: 498 Drugs: 155	0.95
10	86	0	1 (ECG-sym)	28	ECG: 172 Drugs: 14	0.81
Average & stdev (mode & %)	127.2±68.6 days ≅ 18±10week	Monthly average: 0.044 ± 0.13	Monthly average: 1.05 ± 1.11	Monthly average: 5.55 ± 4.4	Daily average: 3.6 ± 1.4	Average ratio of recommendations&reminders viewed out of total: 0.95 ± 0.60

* - Following ECG monitoring results, this patient was diagnosed to supraventricular tachycardia rather than AF, so no recommendations could be delivered apart from measurement reminders. ECG-sym – recommendation to start an ECG monitoring session due to symptom reporting. # - mobile log was not available for this patient; ^ - patient chose to not receive medication reminders; ^^ - patient chose not to receive measurement or medication reminders

Table 6. Delivery of DSS for GDM patients: response to changes in dynamic patient model

Patient#	Potential #days that patient could be using the System	Days on MG	Ratio #days on MobiGuide / potential days	Recommendations to clinicians (total #)	Clinical recommendations to patients (total #)	Technical recommendations to patients (total #)	Reminders sent to patients (total #)	Ratio Total #Recom. & reminders visualized
1	51	36	0.71	1	13	12	BG: 98 Keto: 32	0.96
2	42	34	0.81	0	11	1	BG: 0 Keto: 3	0.87
3	38	28	0.74	2	7	12	BG: 49 Keto: 12	0.85
4	64	64	1	1	7	21	BG: 42 Keto: 10	0.73
5	57	51	0.89	3	12	6	BG: 78 Keto: 21	0.94
6	54	47	0.87	5	4	0	BG: 38 Keto:6	1
7	90	90	1	2	36	48	BG: 183 Keto: 45	0.84
8	97	92	0.95	2	18	14	BG: 133 Keto: 32	0.92
9	116	77	0.66	2	13	9	BG: 148 Keto: 58	0.94
10	37	36	0.97	4	7	4	BG: 41 Keto: 14	0.94
11*	**-	7		-#	-#	-#	-#	0.92
12	50	49	0.98	2	20	10	BG: 87 Keto: 12	0.91
13	57	57	1	1	2	0	BG: 28 Keto: 0	0.88
14	74	72	0.97	8	54	25	BG: 132 Keto: 0	1
15	20	20	1	2	3	0	BG: 45 Keto: 15	0.95
16	92	91	0.99	2	14	16	BG: 236 Keto: 59	1
17	53	53	1	6	5	1	BG: 20 Keto:6	0.98
18	52	51	0.98	3	7	4	BG: 180	0.97

							Keto: 34	
19	74	73	0.99	4	8	9	BG: 208 Keto: 51	0.94
20	170	65	0.38	2	59	13	BG: 126 Keto: 10	0.95
Average &stdev	67.7 ±34.2d ≅ 9.7±4.9 week	57.16 ±21.52d ≅ 8.1±3.0 week	0.89 ± 0.16	Monthly average: 1.5± 1.1	Monthly average:8.1±6.7	Monthly average:4.9±4.4	Daily average: 2.1 ± 1.2	Average ratio of recommendations&reminders viewed: 0.92 ± 0.88

- mobile log was not available for this patient; * - this patient dropped out; ** - delivery date not available

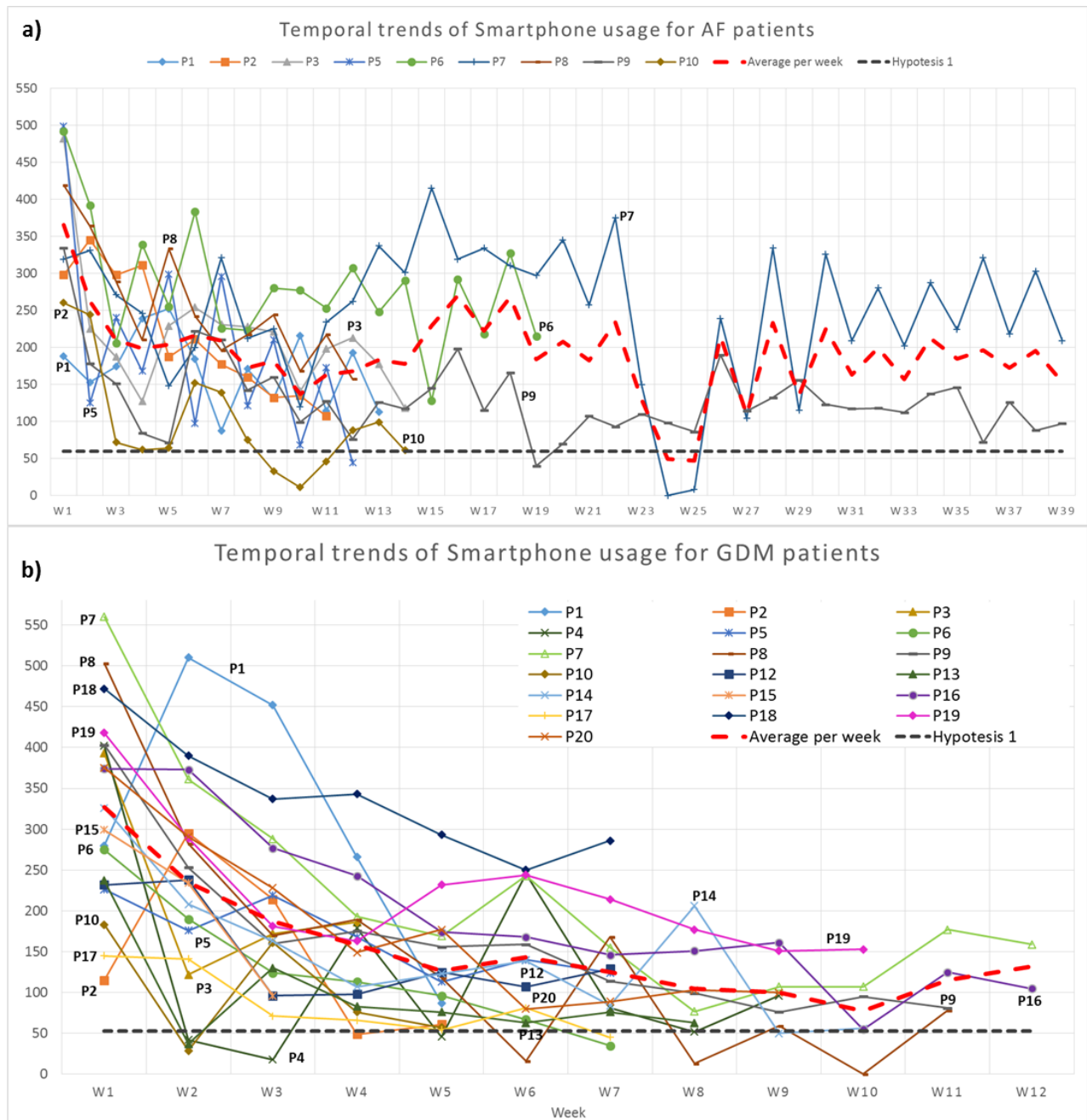


Fig. 6 Temporal trends of patients using the MobiGuide system. (a) AF patients; (b) GDM patients. Numbers on the y-axes represent total actions per week. Average number of actions shown in a dashed line. *Data for AF patient #4 was not available.

5.3. Hypothesis 2: Patients will have positive impression of the MobiGuide system

Table 7 reports the results of the patients' end-of-study questionnaires. The table provides the translation of the questions' exact phrasing for the AF and GDM domains. Patients

were asked to answer on a Likert scale of 1 (don't agree) through 5 (perfect agreement). The number of patients who gave each rating is provided. Nine of ten AF patients filled out the questionnaire. Two GDM patients failed to complete the questionnaire because they lived far away from the hospital and could not deliver the questionnaire after the pregnancy; the patient who dropped out after one week was not asked to fill out the questionnaire.

Table 7. Results of patients' end-of-study questionnaires

Question topic	Exact phrasing	AF N=9	GDM N= 17
General Impression			
App is interesting	The application is: 1=boring; 5=very interesting	1:0 2:0 3:0 4:4 5:5	1:0 2:0 3:1 4:8 5:8
App is easy to use	The application is: 1=very difficult to use; 5=very easy to use	1:0 2:0 3:3 4:4 5:2	1:0 2:0 3:1 4:8 5:8
Sequence of activities is clear	The sequence of activities that can be accomplished with the app is: 1=very confusing, 5=very clear	1:0 2:1 3:0 4:6 5:2	1:0 2:0 3:1 4: 11 5:5
Application response time	The response time of the app for most actions is: 1=very slow; 5=very fast	1:0 2:0 3:4 4:5 5:0	1:0 2:1 3:11 4:3 5:2
Errors were experienced	I have experienced errors with the system: 1=very frequently; 5=never	1:0 2:1 3:4 4:4 5:0	1:1 2:4 3:4 4:7 5:1
Ease of learning curve	Learning to use the app has been: 1=very difficult; 5=very easy	1:0 2:1 3:2 4:5 5:1	1:0 2:0 3:0 4:7 5:10
Usefulness			
System increased patients' confidence	AF: I think that they system helped me to feel safer about my decisions when taking care of AF: 1= totally disagree; 5= totally agree GDM: Do you think that MG has improved your perception regarding diabetes management? With MG I feel myself: 1=less safe; 5= more safe	1:1 2:1 3:2 4:4 5:0 N/A:1	1:0 2:1 3:4 4:8 5:4
Effective visualization	I think that the application has helped me to visualize and interpret my monitoring data in a faster and more effective way: 1=totally disagree; 5= totally agree	1:0 2:2 3:0 4:5 5:1 N/A:1	1:0 2:2 3:4 4:7 5:4

App did/did not complicate patients' lives*	AF: I think that using the app has complicated my daily life: 1= totally disagree; 5= totally agree GDM: I think that using the app has NOT complicated my daily life: 1=totally disagree; 5= totally agree	1:2 2:1 3:4 4:2 5:0	1:0 2:0 3:2 4:10 5:5
I like the system's ability to adapt to context	I like the fact that the system can adapt to my daily life and context changes: 1=totally disagree; 5=totally agree	1:0 2:0 3:1 4:3 5:3 N/A:2	1:0 2:2 3:3 4:8 5:4
System reduces care providers' response time	System reduces the activation time of the healthcare personnel whenever an intervention is required (e.g., waiting time for a visit is shortened)	1:0 2:1 3:1 4:0 5:0 N/A:7	Not asked^
System improved interaction with clinicians	The system improved the interaction with my doctor	1:0 2:0 3:3 4:5 5:0 N/A:1	Not asked
System improve peace of mind on travel	The system improved my peace of mind when travelling away from home	1:0 2:0 3:1 4:6 5:1	Not asked
Recommending the system to other patients	I would recommend the system to other (AF/GDM) patients 1=totally disagree; 5=totally agree	1:0 2:0 3:0 4:0 5:9	1:0 2:1 3:0 4:9 5:7
Recommending the system to a friend	Would you recommend the system to a friend of yours? Y/N	Y:9 N:0	Y:15 N:1
Continue to use the system	Would you use the system in your daily routine (after this study)? Y/N	Y:8 N:1	Y:12 N:2
Paying for the system	Would you pay something for using it? Y/N	Y:7 N:2	Y:3 N:12

* Note the opposite phrasing of this question for AF vs. GDM

^GDM patients were not asked three questions that they could not assess because they did not have previous experience dealing with the disease and were not recommended to travel during pregnancy.

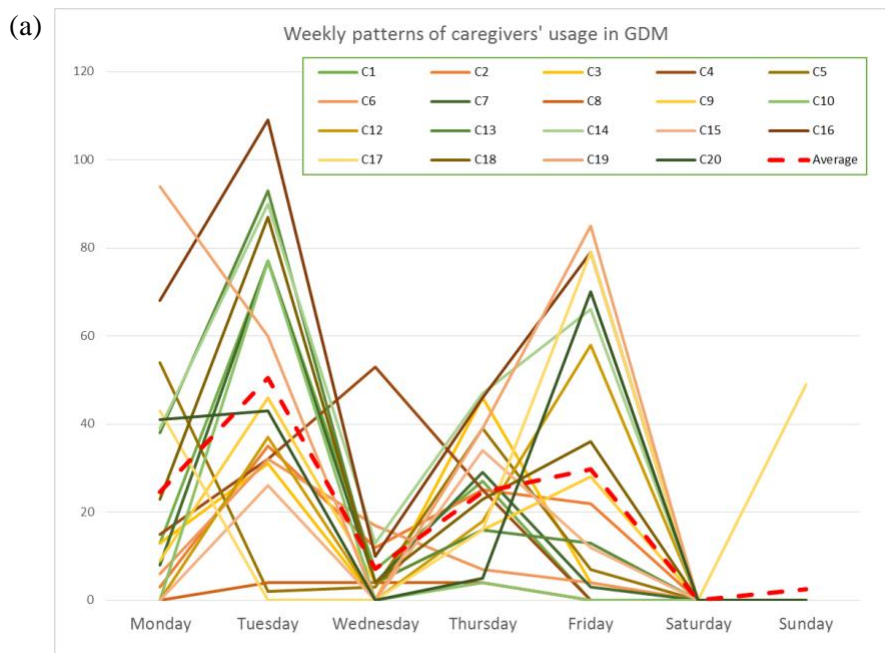
As shown in the table, AF and GDM patients agreed in general that the application was interesting, easy to use and with an easy learning curve, the sequences of activities were clear, the application's response time was appropriate, and patients had rarely experienced errors. We were very pleased by the patients' perceptions of usefulness; most patients agreed that the system has increased their confidence, that it increased their peace of mind while travelling (as decision support was available to them anytime and everywhere), that it improved their interaction with the clinical staff, that it did not

complicate their lives, and that the system has helped them visualize and interpret their data. Most patients liked the fact that the system could adapt to their personal context. Most patients said that they would recommend the system to other patients with their condition, to friends of theirs, and that they would continue to use the system themselves. In addition, AF patients were willing to pay for using the system while GDM patient were not. Patients did not feel that the system has improved the response time of the clinicians, perhaps because it was good also without MobiGuide.

5.4. Hypothesis 3: The clinicians will be examining the data (patient model) of patients during visits and in between visits

In the AF clinic, patients were expected to visit their doctor once a month. However, the AF nurse told us that visits of the MobiGuide AF patients were not formally scheduled and documented; patients were dropping in to see her more often than the planned visits of once every 1-2 months. Therefore, we compared the number of days in which clinicians had viewed the patient’s data (via their caregiver user interface) not to the expected monthly visit, but to a larger number of expected visits (once every 10 days), which accounts for patients unscheduled drop-in visits.

GDM patients had weekly visits each Tuesday. Fig 7a shows the clinicians total number of actions (mostly views) per day for each GDM patient. We were expecting to see a peak on Tuesday and were hoping to also see activity on other days (between visits). The GDM clinicians told us that they had sometimes used the patients’ smart phone GUIs in order to view some of the patients’ data. Fig 7b shows the logged patients daily actions. The actions on that graph represented the total actions by patients and by the clinicians who used the patients’ GUI on visits. Hence, we were expecting to see a peak on Tuesday on that graph as well. Peak activity for clinicians was observed on Tuesdays but also on Thursdays and Fridays. Peak activity for patients was observed on Tuesdays.



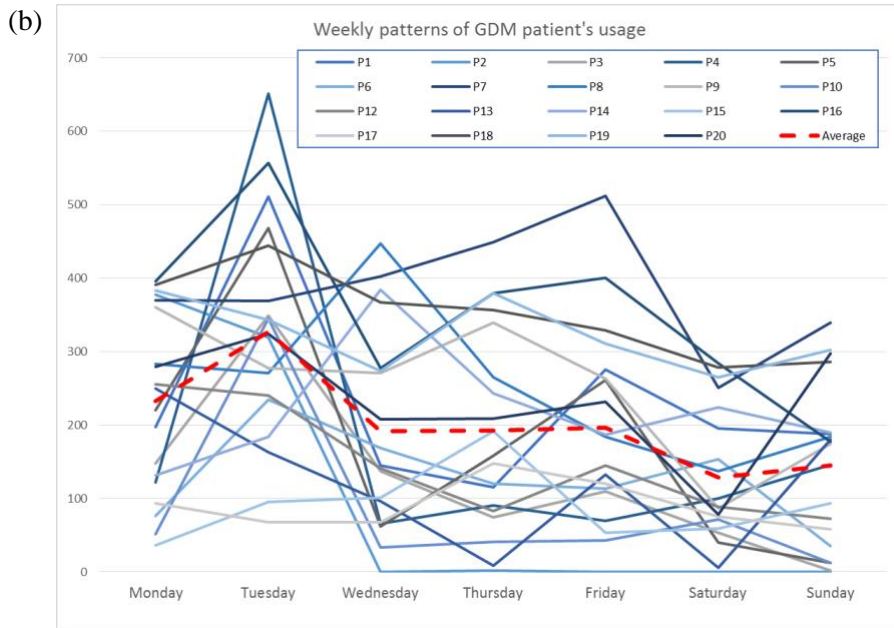


Fig. 7 (a) Total number of GDM clinicians actions performed for each week day for each patient; (b) total number of GDM patient actions performed for each week day by each patient

Tables 8 and 9 present the Wilcoxon signed rank test results for the AF and GDM patients, comparing the #days of caregiver views to the planned number of days of views (Total #days on MobiGuide / 10 for AF and Total #days on MobiGuide / 7 for GDM). As reported below each of these tables, the number of views was in both cases significantly greater than the planned number of visits, indicating that clinicians have viewed patient data between visits.

In addition, we performed similar tests for the cohort of patients, testing the difference in number of actual vs. planned day views for the entire patient population. Fig. 8 summarizes the results.

The personalized decision support of the MobiGuide system impacted clinicians. The AF doctors changed their diagnosis for two of ten AF patients (and as a consequence changed their therapy). One of these patients had complained of arrhythmias for years, but these had never been observed during monitoring sessions at the hospital. She was diagnosed with suspected AF. With the aid of MobiGuide, the patients was diagnosed as having an arrhythmia that is not AF. The other AF patient had been diagnosed with paroxysmal AF but self-monitoring using MobiGuide confirmed that in fact this patient had permanent AF. MobiGuide had also impacted the behavior of the GDM clinicians.

For eleven of the twenty GDM patients, the system had anticipated therapy change that was accepted by GDM clinicians; ten patients started insulin treatment after receiving MobiGuide recommendations and one had a diet therapy change (not included in the group of patients with insulin treatment).

Analyzing what the most common physician actions were, we found that for AF patients, the most frequent action was visualizing the synthesis of the ECG monitoring sessions performed by the patients. Other frequent actions related to the visualization and prescription of medications and the visualization of the recommendations. For GDM, the most frequent action was to visualize important recommendations regarding the patient status. The second most frequent action was the visualization of the synthesis of the BG monitoring sessions.

Table 8. Number of views by AF clinicians

Patient	#days of caregivers views	total #days on MobiGuide	total #days on MobiGuide / 10
1	24	95	10
2	34	77	8
3	28	97	10
4	30	90	9
5	19	89	9
6	30	135	14
7	98	274	27
8	41	89	9
9	63	273	27
10	23	86	9

The W-value is 0. The critical value of W for $N = 10$ at $p \leq 0.05$ is 10. Therefore, the result is significant at $p \leq 0.05$.

Table 9. Number of views by GDM clinicians

Patient	#days of caregivers views	total #days on MobiGuide	total #days on MobiGuide / 7
1	4	36	5
2	5	34	5
3	4	28	4
4	9	64	9
5	9	51	7
6	6	47	6
7	7	90	13
8	2	92	13
9	6	77	11
10	1	36	5
12	4	49	7

13	7	57	8
14	16	72	10
15	4	20	3
16	17	91	13
17	10	53	8
18	10	51	7
19	16	73	10
20	9	65	9

The W-value is 29. The critical value of W for $N = 19$ at $p \leq 0.05$ is 53. Therefore, the result is significant at $p \leq 0.05$.

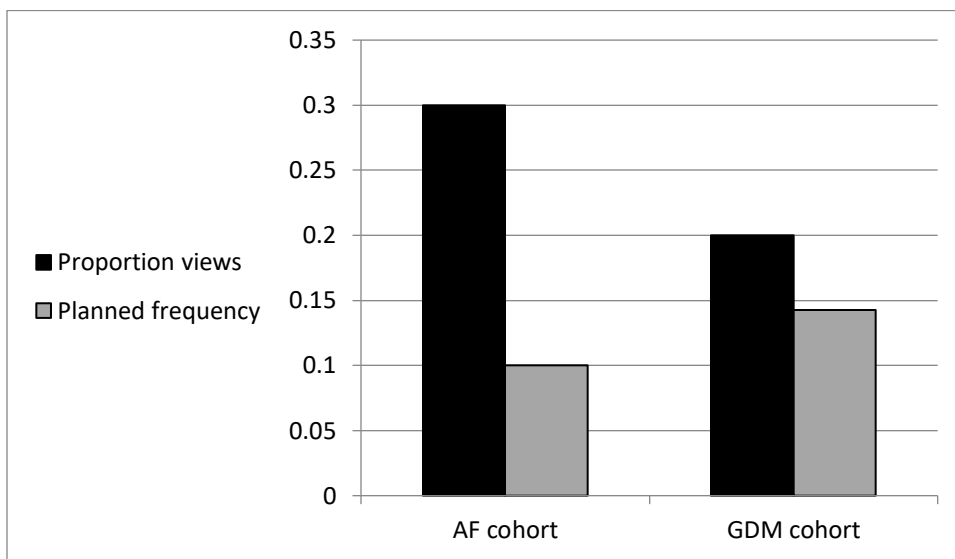


Fig. 8. Bar chart comparing the planned frequency of visits to the actual views by clinicians for the cohorts of AF and GDM patients

6. Discussion

MobiGuide is an interactive decision-support system that can be customized and adapt itself dynamically to its users' incrementally changing personal model, which includes longitudinal clinical data and personal preferences and context. In this paper we discussed the role of patient models in the adaptation process. Most of the current CIG-based DSSs are used only by clinicians, who are guided by the patient-specific recommendations provided by the CIG's declarative knowledge, which constitutes the patient model. The declarative knowledge in other CIG-based DSSs is not customizable and it refers only to clinical concepts rather than to patients' personal context and preferences. MobiGuide, on the other hand, is innovative both in having patients as users and in having a patient model that is adaptive to the patients' personal local and global preferences and contexts. These different aspects of personalization are addressed using

innovative means, including projections and DIRCs. Moreover, we have developed methods for eliciting parallel workflows, psycho-social and demographic context and personal utilities that once specified, can support knowledge projection, context customization and personalization, and global decisions based on decision trees with personal preferences (utilities) (Rubrichi et al. 2015)(Parimbelli et al. 2015).

Our evaluation study used the same underlying MobiGuide system to create DSSs for two very different clinical domains with different instances of the knowledge base (see Table 1), differences in some of the sensors, and difference hospital EMRs from which data has been imported into the MobiGuide PHR. This architecture can potentially be used in other clinical domains after appropriate CIGs are encoded and EMR exporters are written to export existing medical records into the MobiGuide PHR. The feasibility of the HL7 virtual medical record – the data model used by MobiGuide’s PHR – to cover a large range of medical data items has already been shown (Gonzalez-Ferrer 2016). In addition, the Asbru language used in MobiGuide and its supporting tools have been successfully used in various other clinical domains (Shalom et al. 2015).

The study reported in this paper focuses on three hypotheses that stem from the personal patient model and how it impacts the behavior of patients and clinicians.

Hypothesis 1 - sustainable usage by patients

Our first hypothesis was confirmed. It is a well known phenomenon that patients get tired of using (medical) applications after several weeks, especially for monitoring applications (Consumer Health Information Corporation 2012). Yet, our results show that on average, AF patients used the MobiGuide application consistently for a mean of 4.2 months (and up to 8.6 months) and GDM patients used it for 8.2 weeks (and up to 3.2 months). Moreover, AF patients were performing an average of 200.2 weekly actions per patient and GDM patients were performing an average of 150.5 weekly actions per patient (Fig. 6). While AF is a chronic condition, GDM does not start before the fifth month of pregnancy (and patients may have joined MobiGuide even later than that) and lasts till delivery. Hence the amount of time that GDM patients had used the system was naturally shorter than the AF patients. The fraction of days on which GDM patients had used the system out of the potential number of days on which they could have used the system (from enrollment till delivery) was 89%, which means that patients had been using the system most of the time (and continuously, as shown in Figure 8). Patient 20 stayed a shorter period than expected due to the end of the project and six patients stopped using the system one week before delivery, decreasing the expected usage time per patient.

Patients’ actions were in accordance with the CIG-based reminders and recommendations, which were adaptive to the patient model. AF patients performed ECG monitoring at an overall rate of 0.65 ECGs per day, which is encouraging, especially considering that these were older adults not experienced in using technology. Compliance of AF patients to BP monitoring in routine context was 0.75. However, compliance with BP monitoring in semi-routine context was very low (0.03 ± 0.03). This may be due to over-burdening these patients with extra measurements (see Hypothesis 3).

Compliance rates of GDM patients were even higher, with observed compliance of 0.82 and 0.96 to recommended BP and ketonuria plans. Compliance with BG monitoring plans (either daily or twice weekly) was 0.99 and compliance with four measurements a day was 0.87. The latter matches the BG monitoring compliance rate of 0.87 ± 0.28 that was observed for a historical cohort of 247 patients diagnosed with GDM that was followed at the same hospital department two years before our experiment, during 2010-2013 (Villaplana 2015). However, the patients in that cohort were requested to measure at least (and not exactly) 4 BG measurements a day and their compliance was calculated such that measuring 4 BG measurements a day yields full compliance (of 1.0) but more measurements could yield a compliance greater than 1. Using the same calculation for the MobiGuide cohort yielded a compliance rate of 1.01 ± 0.10 , which was significantly higher than that of the historical cohort (p-value of 0.0312). A similar baseline could not be obtained for the AF domain, as per current best practice, AF patients are not provided with ECG monitors at home nor with the ability to report symptoms in real time; hence the MobiGuide system empowers patients beyond current practices.

While the compliance of GDM patients to monitoring plans (reminder-based) was higher than that of AF patients, their compliance with the more sparse clinical recommendations (regarding diet due to abnormal BG or ketonuria) was lower (0.3-0.5) than the perfect compliance of AF patients. However, this may be due to the fact that only two AF patients received one recommendation each (related to risky events and to anticoagulants) whereas all GDM patients received several recommendations each, related to BG or ketonuria control.

While monitoring of the most important health parameters (ECG and BP for AF and glycemia, ketonuria, and BP in GDM) was well followed in both domains, monitoring of other measurements was not followed as well. This may reflect the perceived importance that patients attribute to different recommendations and their connection to their primary health concern. Other reasons could relate to the fact that the other measurements were prescribed to be done less frequently; INR was to be recorded every two weeks by one patient, and weekly weight measurements by two patients, and the measured values were almost always within the normal range.

Part of patients' empowerment is their ability to report things to the system, such as their symptoms and measurements that are not automatically uploaded from mobile biosensors and also view their data recommendations. AF patients often reported symptoms related to their condition, with a mean of one symptom per month. The percentage of recommendations and reminders viewed by patients was also very high (95% for AF patients and 92% for GDM patients of total recommendations and reminders generated).

The patients' sustainable usage of the system for monitoring their health and following personalized recommendations is in line with the fact that only one of the 20 GDM patients and none of the ten AF patients had dropped out and that three of the AF patients did not want to return the equipment at the end of the study and kept using the system while budget permitted for several additional weeks.

Hypothesis 2 – positive impressions by patients

Despite the large difference between the two evaluation studies of the MobiGuide system, the system was effective and appreciated by the two patient populations (Table 7). Patients found the application useful. Overall, the sense of safety that the system has provided to the patients was its greatest asset. When interviewing the patients in person, many of them chose to tell us about this quality of the system. One patient said “With the system I feel the doctor by my side, as if he is hugging me”.

Hypothesis 3 – clinicians will examine system-collected patient data also outside visits

The clinicians examined the patient-personal data that justified the personalized recommendations output by the system. For the AF domain, the most important data was ECG sessions and medication prescriptions. For the GDM domain, the clinicians viewed the MobiGuide recommendations sent to patients and correlated them with patient data, including BG, BP and ketonuria. Clinicians also examined patients’ global and local preferences (e.g., meal times) to make their decisions.

For both AF and GDM clinicians, the frequency of actions done by clinicians was significantly larger than the planned number of visits; the clinicians viewed the patients’ models also between patient visits. The AF nurse looked at the entire ECG signal every day, and if there was something worth noting, she alerted the cardiologist. Due to the pilot nature of the evaluation study, patients were invited to schedule visits with their clinicians often (every 1-2 months for AF patients and every week for GDM patients), yet the safety of patients who chose to only be monitored from home was maintained. In fact, some AF patients only had the enrollment and the end-of-study visit, and used the system from home for the whole duration of the study. This is exactly the intention of the MobiGuide system: that using the system the patients would be monitored and safely stay in their normal environment and only come in to the clinics when required by their clinical state.

The GDM clinicians told us that a few days after each Tuesday visit, and especially just before the weekend, they checked again the patients’ model to see how the patient was doing. In line with the statement of GDM clinicians that during the weekly patient visits they used the patient’s smart phone to check the patient’s data, which was more efficient for them, we observed a peak activity on the patients’ GUI on Tuesdays. An additional smaller peak in patient’s actions was observed on Fridays, in line with the request that patients should download their data every 3 days or so.

Perhaps the most impressive outcome was that as a result of observing the changes in the patient model and the patient-specific recommendations delivered to them, the AF clinicians changed the diagnosis for two AF patients and the GDM clinicians started insulin treatment for two of the twenty GDM patients (ten patients started insulin treatment after MobiGuide recommendations, but for two patients, the system detected the need to start insulin earlier than clinicians did).

Limitations of our approach

One of the largest limitation of our approach is that knowledge-acquisition requires a lot of effort. Even the CIG Specification stage, which is done in non-personalized clinical DSSs is very laborious, despite the fact that it begins with a narrative clinical practice guideline as a starting point. However, as in the work of Camerini et al. (Camerini et al. 2011) “the design of a personalized health system, must consider the balance between evidence based medical guidelines, the feasibility of their implementation, and the modeling of the system”. Here, we need to elicit from clinicians information that is not contained in the original guideline regarding opportunities for involving and empowering the patient using mobile technology at the patient’s normal environment. These opportunities are usually not open to clinicians and they are not experienced in thinking about them. The elicitation methods that we have developed have been successful, but they require much time and effort from busy clinicians. Another problem, is that reuse of encoded plans by other CIGs is in most cases not suitable.

Another disadvantage is that our personalization addresses personal preferences and context but only partially addresses comorbidities that are particular to different patients by delivering recommendations to drugs for comorbidities, unlike the more complete approach for handling comorbidities addressed in (Riaño et al. 2012) (Grandi et al., 2012)(Grandi 2016). It is not realistic to think that older chronic patients will suffer from a single disease, yet clinical guidelines focus on a single disease. The AF and GDM CIGs addressed the hypertension comorbidity explicitly. In addition, we have added to the declarative knowledge, concepts representing each medication that any of the patients were taking so that they could receive personalized reminders. Nevertheless, systematic methods for integrating CIGs, each focusing on a different disease, should be developed as future research to support detection of conflicts and their resolution (Peleg 2013).

Except for the availability of BG monitoring compliance rates of a historical GDM cohort, we have not evaluated user behavior against a control; some of the impacts might have been achieved simply by providing the sensors and reminders, without the DSS.

An important limitation is that there was not much opportunity to demonstrate the impact of CCCs, but we think that CCCs would prove important for other clinical domains. Here, the different CCCs that were added during the customization of the two CIGs did not affect the care plans to a large extent; in the case of AF, the semi-routine context impacted only the frequency of measurement of BP, as compared to the frequency in routine context. For GDM, the different clinical recommendations did not change for different context, only the ability to adjust the reminder times per context.

The fact that certain system bugs were not noticed during the pre-pilot stage also limited our potential analysis. Specifically, in the AF system, reminders were not always working as planned; ECG reminders were the only ones that were always delivered, but other reminders were not working consistently. Hence we could not fairly evaluate the impact of the routine vs. semi routine CCCs. In addition, due to the small size of cohorts of patients, we did not have a chance to evaluate the full scope of the CIG knowledge, including the two other AF CCCs: intensive physical activity and 24 hour ECG

monitoring. Nor could we check the impact of variability of patient's local preferences regarding reminders on patients' behavior, because most patients did not change their reminder settings of ON/OFF between contexts. Finally, detailed analysis concerning clinical impact and compliance are beyond the scope of this paper and will be reported elsewhere.

Conclusion

MobiGuide is a personalized and patient-centric clinical DSS based on clinical practice guidelines that involves patients in the management of their disease in their normal environments. The results of our multi-national feasibility study demonstrated clearly the feasibility of the MobiGuide architecture and its adaptive behavior that adjusts itself to the patient's personal model. The evaluation has confirmed the three hypotheses regarding the effect of the personal patient model on users' behavior; hence the personalization has met its goal. First, the results demonstrate that the users of the system, both patients and physicians, have used the system consistently for substantial periods of time and found that it provided them with multiple benefits, which stem from the personalized patient model. Second, the most substantial benefit to patients was the increase in their sense of safety as well as in their involvement (demonstrated through a high compliance rate), which was enhanced by the personalization and patient-centrality features of the system. Third, clinicians have used the system to follow up on their patients' models outside the planned visit, which fits the intention of the system to save visits for well-controlled patients. In addition, the system has affected clinicians' behavior resulting in change of diagnosis for two of the ten AF patients and anticipated change in therapy for eleven of the twenty GDM patients.

References

- Boaz D and Y Shahar. 2005. "A Framework for Distributed Mediation of Temporal-Abstraction Queries to Clinical Databases." *Artificial Intelligence in Medicine* 34 (1): 3–24.
- Bodenreider O. 2004. "The Unified Medical Language System (UMLS): Integrating Biomedical Terminology." *Nucleic Acids Research* 32 (suppl 1): D267-D270.
- Camerini L, M Giacobazzi, M Boneschi, PJ Schulz, and S Rubinell. 2011. "Design and Implementation of a Web-Based Tailored Gymnasium to Enhance Self-Management of Fibromyalgia." *User Modeling and User-Adapted Interaction* 21: 485–511.
- Chittaro L, E Carchietti, L De Marco, and A Zampa. 2011. "Personalized Emergency Medical Assistance for Disabled People." *User Modeling and User-Adapted Interaction* 21(4): 407-440.
- Consumer Health Information Corporation. 2012. "Motivating Patients to Use Smartphone Health Apps." <http://www.consumer-health.com/motivating-patients-to->

use-smartphone-health-apps/.

- Fux, A, M Peleg, and P Soffer. 2012. "How Does Personal Information Affect Clinical Decision Making? Eliciting Categories of Personal Context and Effects." *AMIA Symp.*, 1741.
- García-Sáez G, M Rigla, I Martínez-Sarriegui, E Shalom, M Peleg, T Broens, B Pons, E Caballero-Ruíz, E J Gómez, and M Elena Hernando. 2014. "Patient-Oriented Computerized Clinical Guidelines for Mobile Decision Support in Gestational Diabetes." *J Diabetes Sci Technol* 8 (2): 238–46.
- García-Sáez G, M Rigla, E Shalom, M Peleg, E Caballero, E J Gómez, and ME Hernando. 2013. "Parallel Workflows to Personalize Clinical Guidelines Recommendations: Application to Gestational Diabetes Mellitus." *13th Mediterranean Conf on Medical & Biological Engineering & Computing*, 1409-12.
- González-Ferrer A, M Peleg, M Marcos, J A Maldonado. "Analysis of the process of representing clinical statements for decision-support applications: a comparison of openEHR archetypes and HL7 virtual medical record". *J Med Syst* 2016;40(7):163-72.
- Grandi, F. 2016. "Dynamic Multi-Version Ontology-Based Personalization." *Journal of Computer and System Sciences* 82 (1): 69–90.
- Grandi F, F Mandreoli, and R Martoglia. 2012. "Efficient Management of Multi-Version Clinical Guidelines." *Journal of Biomedical Informatics* 45 (6): 1120–36.
- Kahneman D, and A Tversky. 1982. "The Simulation Heuristic." In *Judgment Under Uncertainty: Heuristics and Biases*, edited by A. J, Kahneman, D., Slovic, P. & Tversky, 201–8. Cambridge University Press.
- Lanzola G, Parimbelli E, Micieli G, Cavallini A, Quaglini S. Data quality and completeness in a web stroke registry as the basis for data and process mining. *J Healthc Eng.* 2014;5(2):163-84.
- Lasierra N, A Alesanco, S Guillén, and J García. 2013. "A Three Stage Ontology-Driven Solution to Provide Personalized Care to Chronic Patients at Home." *Journal of Biomedical Informatics* 46 (3). Elsevier Inc.: 516–29.
- Lindgren, H. 2011. "Towards Personalized Decision Support in the Dementia Domain Based on Clinical Practice Guidelines." *User Modeling and User-Adapted Interaction* 21 (4): 377–406.
- Marcos C, A González-Ferrer, M Peleg, and C Cavero. 2015. "Solving the Interoperability Challenge of a Distributed Complex Patient Guidance System: A Data Integrator Based on HL7's Virtual Medical Record Standard." *Journal of American Medical Informatics Association* 22 (3): 587-99.
- Martins S, Y Shahar, D Goren-Bar, M Galperin, H Kaizer, et al. 2008. "Evaluation of an Architecture for Intelligent Query and Exploration of Time-Oriented Clinical Data." *Artif Intell Med* 43 (1): 17-34.
- Miksch S, Y Shahar, and P Johnson. 1997. "Asbru: A Task-Specific, Intention-Based, and Time-Oriented Language for Representing Skeletal Plans." In *7th Workshop on Knowledge Engineering: Methods & Languages*, 1–25.
- MobiGuide Consortium. 2016. "Monitored Patterns, Notifications and Recommendations

- Used in the AF and GDM CIGs of MobiGuide.”
http://mis.hevra.haifa.ac.il/~morpeleg/MobiGuide_Patterns.pdf.
- Parimbelli E, L Sacchi, S Rubrichi, A Mazzanti, and S Quaglini. 2015. “UceWeb: A Web-Based Collaborative Tool for Collecting and Sharing Quality of Life Data.” *Methods Inf Med* 54 (2): 156–63.
- Peleg M. 2013. “Computer-Interpretable Clinical Guidelines: A Methodological Review.” *J Biomed Inform* 46 (4): 744–63.
- Peleg M and A Gonzalez-Ferrer. 2014. Chapter 16: Guidelines and Workflow Models. In: *Clinical Decision Support: The Road to Broad Adoption*, 2nd Edition, RA Greenes (editor). 435-464 Academic Press.
- Peleg M, Y Shahar, and S Quaglini. 2013. “Making Healthcare More Accessible, Better, Faster, and Cheaper: The MobiGuide Project.” *European Journal of ePractice* 20: 5–20.
- Peleg M, SW Tu, J Bury, P Ciccarese, J Fox, et al. 2003. “Comparing Computer-Interpretable Guideline Models: A Case-Study Approach.” *J Am Med Inform Assoc* 10 (1): 52–68.
- Pitts, MG, and GJ Browne. 2007. “Improving Requirements Elicitation: An Empirical Investigation of Procedural Prompts.” *Information Systems Journal* 17 (1): 89–110.
- Quaglini S, S Miksch, Y Shahar, M Peleg, M. Peleg, M. Rigla, C. Napolitano, A. Pallàs, E. Parimbelli, and L. Sacchi, “Supporting Shared Decision Making within the MobiGuide Project,” in *AMIA Symp.*, 2013, pp. 1175–1184.
- Riaño D, F Real, JA López-Vallverdú, F Campana, S Ercolani, et al. 2012. “An Ontology-Based Personalization of Health-Care Knowledge to Support Clinical Decisions for Chronically Ill Patients.” *Journal of Biomedical Informatics* 45 (3): 429–46.
- Rubrichi S, C Rognoni, L Sacchi, E Parimbelli, C Napolitano, A Mazzanti, and S Quaglini. 2015. “Graphical Representation of Life Paths to Better Convey Results of Decision Models to Patients.” *Med Decis Making* 35 (3): 398–402.
- Sacchi L, A Fux, C Napolitano, S Panzarasa, M Peleg, et al. 2013. “Patient-Tailored Workflow Patterns from Clinical Practice Guidelines Recommendations.” *Stud Health Technol Inform.* 2013, 192:392–96.
- Shahar, Y. 1997. “A Framework for Knowledge-Based Temporal Abstraction.” *Artificial Intelligence* 90 (1-2): 79–133.
- Shahar Y, S Miksch, and P Johnson. 1998. “The Asgaard Project: A Task-Specific Framework for the Application and Critiquing of Time-Oriented Clinical Guidelines.” *Artificial Intelligence in Medicine* 14 (1-2): 29–51.
- Shahar Y, and M A Musen. 1996. “Knowledge-Based Temporal Abstraction in Clinical Domains.” *Artificial Intelligence in Medicine* 8 (3): 267-298.
- Shahar, Y. 1998. “Dynamic Temporal Interpretation Contexts for Temporal Abstraction.” *Annals of Mathematics and Artificial Intelligence* 22 (1-2): 159–92.
- Shalom, E, Y Shahar, Y Parmet, and E Lunenfeld. 2015. “A Multiple-Scenario Assessment of the Effect of a Continuous-Care, Guideline-Based Decision Support System on Clinicians’ Compliance to Clinical Guidelines.” *International Journal of*

Medical Informatics 84 (4): 248–62.

Shalom E, Y Shahar, and E Lunenfeld. 2016. “An Architecture for a Continuous, User-Driven, and Data-Driven Application of Clinical Guidelines and Its Evaluation.” *J Biomed Inform.* pii: S1532-0464(15)00243-9. doi: 10.1016/j.jbi.2015.11.006..

Shalom E, Y Shahar, M Taieb, D Goren-Bar, A Yarkoni, et al. 2008. “A Quantitative Evaluation of a Methodology for Collaborative Specification of Clinical Guidelines at Multiple Representation Levels.” *J Biomed Inform* 41 (6): 889-903.

Villaplana M, Pons B, Morillo M, Aguilar A, Mendez A, Tirado R, et al. Early introduction of insulin in gestational diabetes seems to prevent from birth weight abnormalities. *Diabetes, Hypertension, Metabolic Syndrome & Pregnancy Symposium*. 2015.

1) Mor Peleg. Ph.D.

Department of Information Systems, University of Haifa, Haifa, Israel, 3498838
Mor Peleg is Assoc. Prof at the Dept. of Information Systems, University of Haifa, and has been Department Head in 2009-2012. Her BSc and MSc in Biology and PhD in Information Systems are from the Technion, and her post-doc was at Stanford BioMedical Research. She was awarded the New Investigator Award by the American Medical Informatics Association, is International Fellow of the American College of Medical Informatics and is Associate Editor of Journal of BioMedical Informatics. Her research concerns knowledge representation, decision support systems, and process-aware information systems in healthcare. <http://mis.hevra.haifa.ac.il/~morpeleg/>

2) Yuval Shahar, M.D., Ph.D., FACMI

Department of Information Systems Engineering, Ben Gurion University of the Negev, Beer-Sheva, Israel, 84105

Yuval Shahar, M.D. (Hebrew University), M.Sc. (Computer science, Yale University), Ph.D. (Medical Informatics, Stanford University), heads the Medical Informatics Research Center at Ben Gurion University, in which he is a professor of software and information systems engineering, and the Joseph Erteschik Chair of Information Systems Engineering. His research focuses on temporal reasoning, therapy planning, knowledge representation, and decision analysis in biomedical domains, and he has more than 200 publications in these areas. He is a recipient of an NIH 5-year FIRST career award, an IBM Faculty Award, and an HP Worldwide Innovation Program award. He is an International Fellow of the American College of Medical Informatics (ACMI).

3) Silvana Quaglini, PhD

Dipartimento di Ingegneria Industriale e dell'Informazione, University of Pavia, Italy, 27100

Silvana Quaglini, PhD, is full professor in the bioengineering track at the University of Pavia. Her research interests are in the decision support systems in medicine and telemedicine, decision theory, probabilistic systems, biomedical statistics. The application areas include computerized guidelines, homecare systems, economic evaluations based on decision-analytic models, and health systems management workflow. She has participated and participates in several projects funded by the European Union, the Ministry of Health and the Ministry of Education, including GAMES (General Architectures for Medical Expert Systems), GAMES II, HC - ReMa (Health Care Resource Management), Patman (Patient Management), Homey (Homecare) , MyHeart (prevention and treatment of cardio - cerebrovascular diseases) and MobiGuide (Guiding patient anytime everywhere). She is the author of about 250 papers on scientific journals and international conference proceedings (H- index = 35, source Scopus).

4) Adi Fux, MA

University of Haifa, Department of Information Systems, Haifa, Israel 3498838

Adi Fux is a PhD student in the Department of Information Systems, University of Haifa. His MA research, in logistic information systems, presenting generalized algorithm to define the optimal policy of the repairing lines. His PhD research focused on developing layered decision support models to improve systems develop and maintenance. He has been a researcher in the MobiGuide project, in which he had led the socio-demographic knowledge acquisition and representation in the decision support system. His research interests include business process modeling and management, process mining, requirements engineering and conceptual modeling.

5) Gema García-Sáez, PhD

Universidad Politécnica de Madrid, Biomedical Technology Center (CTB), ETSI Telecomunicación, Av Complutense 30, Madrid, Spain 28040.

Dr. García-Sáez is Assistant Professor and Senior Researcher in Biomedical Engineering. She received her Master degree in Telecommunication Engineering and her Ph.D. degree in Biomedical Engineering from the Polytechnical University of Madrid. Her research areas are: chronic patients support through eHealth and mHealth services, decision support systems in telemedicine services and development of an artificial pancreas.

6) Dr. Ayelet Goldstein, PhD

Department of Information System Engineering, Ben Gurion University of the Negev, Beer-Sheva, Israel, 84105

Ayelet goldstein received her B.Sc. in computer Science and MBA from the Hebrew University of Jerusalem. She completed her Ph.D. in 2015 at Ben Gurion University of the Negev. She has been a senior researcher in the Mobiguide project, in which she had led the medical knowledge acquisition and representation process. Her research interests include automatic summarization, knowledge acquisition and representation, data abstraction, natural language generation and temporal data mining.

7) Maria Elena Hernando, PhD

Universidad Politécnica de Madrid, Biomedical Technology Center (CTB), ETSI Telecomunicación, Av Complutense 30, Madrid, Spain 28040.

Dr. M.E. Hernando is Professor of Biomedical Engineering, director of the ICT for Personalized Medicine lab at the CTB. Dr Hernando receives her M.S. degree in

Telecommunication Engineering and her PhD degree in Biomedical Engineering. Dr. Hernando has worked in several areas of telemedicine and artificial intelligence, including ambulatory patient monitoring, communication with medical devices and mobile applications.

8) Denis Klimov, PhD

Department of Information System Engineering, Ben Gurion University of the Negev, Beer-Sheva, Israel, 84105

Denis Klimov recently finished PhD at Ben Gurion University of the Negev, where he also obtained his M.Sc. degree. He is a senior researcher at the Medical Informatics Research Center at Ben Gurion University. His main research interests include the visual analytics, information visualization, machine learning, data mining and decision support. In the MobiGuide project he led the intelligent data analysis work package.

9) Iñaki Martínez-Sarriegui

Centre for Biomedical Technology (CTB), ETSIT, Universidad Politécnica de Madrid, and Networking Research Centre for Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Spain, 28040

BSc and MSc in Telecommunication Engineering (2005) and MSc in Telemedicine and Bioengineering (2008) from the Technical University of Madrid (Spain). I have more than 10 years of experience in the field of Telemedicine and Telecommunications focused on creating mobile solutions to support diabetic patients' self-empowerment and the development of trust and reputation mechanisms applied to medical systems. I received a PhD (cum laude) in Biomedical Engineering (2016) also from the Technical University of Madrid in the latter area. I have been engaged in a wide range of European, National and Regional projects proposing innovative solutions and the use of disruptive technologies. In the last ten years, I have produced more than 20 peer-reviewed publications: <https://scholar.google.es/citations?hl=es&user=OjBjfBcAAAAJ>

10) Carlo Napolitano MD, PhD

IRCCS Foundation "Salvatore Maugeri", Molecular Cardiology, Pavia, Italy 27100

Carlo Napolitano (ORCID: <http://orcid.org/0000-0002-7643-4628>) is currently vice director of Cardiology and Molecular Cardiology. After MD degree (State Univ. of Milan) he achieved the Cardiology diploma and the PhD in Cardiovascular Pathophysiology. He trained in clinical cardiology, cellular electrophysiology (Univ. of Florida and Washington Univ., St Louis) and molecular biology (DIBIT, Milan). He served as associate professor of cardiovascular genetics at New York University. His primary research focus is on cardiovascular epidemiology and genetics of hereditary cardiac diseases. He published more than 250 full papers on peer-reviewed journals and has a h-index of 68 (Scopus).

11) Enea Parimbelli, PhD

Dipartimento di Ingegneria Industriale e dell'Informazione, University of Pavia, Italy, 27100

Dr. Enea Parimbelli is currently a post-doc fellow at the laboratory of biomedical informatics "Mario Stefanelli" at the University of Pavia. He completed his MS in computer engineering at his current institution and worked on the MobiGuide project during his Ph.D. and fellowship. Dr Parimbelli's research interests include patient-centered

medicine, decision support, telemedicine and artificial intelligence methodologies applied to the medical field.

12) Mercedes Rigla Cros, PhD

Endocrinology and Nutrition Department, Hospital de Sabadell, Sabadell, Spain, 08208
Mercedes Rigla Cros is Associate professor at the School of medicine, Autonomous University of Barcelona. She is the head of the Endocrinology and Nutrition Department, Parc Taulí University Hospital, Sabadell, Spain. Her current research interests include diabetes mellitus and particularly technology applied to diabetes.

13) Lucia Sacchi, PhD

Dipartimento di Ingegneria Industriale e dell'Informazione, University of Pavia, Italy, 27100

Dr. Lucia Sacchi is assistant professor at the Department of Electrical, Computer and Biomedical Engineering at the University of Pavia, Italy. Her research interests are related to data mining, with particular focus on temporal data, clinical decision support systems, and technologies for biomedical data analysis. She has co-authored more than 60 scientific peer-reviewed publications on international journals and international conferences. She is the Vice-chair of the IMIA working group on Data Mining and Big Data Analytics, she is member of the board of the Artificial Intelligence in Medicine (AIME) Society, and of the Italian Society of Biomedical Informatics (SIBIM). She is part of the Editorial Board of BMC Medical Informatics and Decision Making and she is a member of the Scientific Program Committee of several international conferences and workshops.

14) Erez Shalom, PhD

Department of Information System Engineering, Ben Gurion University of the Negev, Beer-Sheva, Israel, 84105

Erez Shalom earned a BSc in Industrial Engineering and an MSc and PhD in Information System Engineering from Ben-Gurion University. He was the Research Manager of the Ben-Gurion University team in the MobigGuide project, focusing on distribution of decision-support and guideline formalization. Erez interested in medical decision-support applications, general methods for knowledge representation and its uses within clinical domains in particular. He is the CTO at Medilogos, a HealthCare IT spin-off of Ben-Gurion University.

15) Pnina Soffer, PhD

University of Haifa, Department of Information Systems, Haifa, Israel 3498838
Pnina Soffer is an Associate Professor and the head of the Information Systems Department at the University of Haifa. She received her Ph.D. in from the Technion - Israel Institute of Technology (2002). Her research deals with business process modeling and management, process mining, requirements engineering and conceptual modeling. Pnina has published over 100 papers in journals and conference proceedings. She has served as a guest editor of a number of journal special issues, and in organizing and editorial positions at numerous conferences, including CAiSE, BPM, ER, ICIS, and others.

Appendix A. Care provider first interview: psycho-social and demographic concepts elicitation

The questionnaire below refers to the personal information that is gathered during the patient interview (history taking), work up, treatment and follow-up treatments. The medical information of a patient, created during physicians-patient encounters, is usually stored in an electronic medical record (EMR) of a specific healthcare facility. The medical record is designed per the requirements of each medical facility, but nevertheless there are generic parts that are common to all medical records, such as family history or personal history (social history) that contains the personal data that affects the patient and therefore impacts treatment decisions.

This questionnaire is designed to identify psycho-social and demographic concepts. These concepts induce context that may have effects on patient management. As well, the questionnaire will help to identify these effects.

For our first part – please think of the main psycho-social and demographic concepts that have the greatest influence on treatment recommendations:

- a. What is the psycho-social or demographic concept?
- b. What is the main effect of the contexts induced by the concepts on clinical goals or recommendations?

For the second part – we will refer to a situation where a patient has entered your clinic. The patient complains that he/she suffers from chest pain; last lab exam revealed that he has high cholesterol and he/she seems over weight.

The patient-doctor encounter will usually include the following steps of data collection:

- | | |
|---|------------------------|
| a. Personal details | [coded options] |
| b. Chief complaint [in our case – chest pain] | [coded options] |
| c. Current disease [elaboration of chief complaint] | [free text] |
| d. Current problem list [emphasis on relevant diseases] | [coded options] |
| e. Past medical problems | [coded options] |
| f. Family history. | [coded options + text] |

Please describe the patient-doctor encounter, with emphasis on psycho-social and demographic concepts:

1. Basic psycho-social and demographic concepts, please refer to specify:
 - 1.1. What are the main psycho-social and demographic concepts you collect during the medical interview?
 - 1.2. How do you take this information into consideration during the patient workup?
 - 1.3. How does each concept affect the treatment goals and diagnosis/work up process?
[sending the patient to other exams, next meetings and other considerations]

[Basic psycho-social and demographic concepts include: degree of routineness of the daily schedule and diet, degree of family support, etc.]

For the next part we will refer to a patient with chronic medical condition who requires treatment, monitoring, and follow-up.

2. Please think of a patient of yours who already has an established diagnosis (e.g., AF, hypertension). Can you describe the process of patient management for such a patient?

- 2.1. Specify the major decision points and treatment goals during current line of treatment for this patient.
- 2.2. What was the role of psycho-social and demographic context in each treatment goal decision?
- 2.3. Based only on clinical information (ignoring the psycho-social and demographic contexts), what are the other treatment options available?
- 2.4. Per treatment recommendation – define different psycho-social and demographic contexts that will change the treatment recommendation.

For the next part we will assume that we have a new patient guidance system that integrates hospital and monitoring data into a Personal Health Record (PHR) accessible by patients and care providers.

The system includes Decision Support System (DSS) at the back end (health center), and on the front end (patient), by utilizing monitoring technologies. The distributed DSS provides information to the patient and care providers. Customization will be achieved by considering psycho-social/demographic data.

The following description refers to the domain of gestational diabetes. Please refer to this example and create an analogous scenario from your clinical domain.

Pazit is 39 years old; she has 3 children and now expecting the 4th one. She is a manager in a small projects company. Her duty requires a lot of car driving, meetings and out of the office scheduling.

Pazit was diagnosed with Gestational Diabetes Mellitus (GDM).

Pazit has been enrolled in the new system (the MobiGuide system). She has received a Smartphone, and devices to monitor her condition. Her state requires insulin therapy, so the patient monitoring needs to be strict.

She received recommendations to monitor blood glucose levels 4-6 times per day (fasting and 3 postprandial -1 hour- measurements (3 measurements taken 1 hour after each meal), and possible addition of pre-lunch and pre-dinner). She is instructed to monitor insulin (doses and times), intake deviations from the established diet recommendations (daily), ketonuria (daily), and body weight (weekly). The Smartphone will monitor her physical activity according to the defined guidelines and the patient health state (in some cases the physical exercise could have medical or obstetrical contraindications, so the goal will be to monitor the physical activity to guarantee that the patient is resting).

She has access to MobiGuide via a mobile application and also to a web application. The next visit is scheduled 2 weeks later, if there are no relevant events.

Pazit is now in her 28th week and in the next two days Pazit has a big event and presentation in a conference.

3. In the process of patient follow-up, treatment and work up, what is the information that you expect to gain from the system?
 - 3.1. Specify the decision making points in the process, indicating the reasons for selecting each decision option.
 - 3.2. What are the mobile monitoring devices that the patient can use to supply the physicians with reliable information that is indicative of her state and to what extent do you think that they are reliable (for example glucose meter, a pedometer, an accelerometer and a pulse-meter)?
 - 3.3. What is the personal information required in this monitoring process?
 - 3.4. What is the effect of each item of psycho-social or demographic data on the treatment recommendations?
 - 3.5. Per treatment recommendation – define the different psycho-social and demographic data that will change treatment recommendations.

- 3.6. Specify data values that will generate notification or will cause you to consider rescheduling the next meeting with the patient?
4. For each psycho-social and demographic data item:
 - 4.1. What are the main values that the data item may take (scale)?
 - 4.2. Who should enter the information into the system (physician, nurse, other care provider or patient)?
 - 4.3. Who should be able to change the information?
 - 4.4. Would you consider the option of subscribing mechanism (e.g. RSS-like) for checking the events (clinical or non-clinical) that had happened in the last period (hour/day/week) which would trigger change in the value of a psycho-social or demographic data item?
 - 4.5. What are the specific values of psycho-social or demographic data items that require immediate notification?
 - 4.5.1. Who should get the notification (Patient/Physician/Nurse/other care provider)
5. In cases a monitored patient doesn't have web connection, namely in case the patient has only the smartphone and the monitoring system (no connection option to main server):
 - 5.1. What do you expect that the system do to notify the patient (only continue monitoring? notify changes and suggest actions? Etc.)
 - 5.2. What are the decisions the system should take?
 - 5.3. What situations must be addressed immediately?

Thank you for your cooperation.

Appendix B. Patients' first interview: psycho-social and demographic concept elicitation

The questionnaire below refers to the personal information that is gathered during interview with the care providers (history taking), work up, treatment and follow-up treatments. The medical information that gather during those meetings is usually stored in an electronic medical record (EMR) of a specific healthcare facility.

This questionnaire is designed to identify psycho-social and demographic concepts. These concepts induce context that may have effects on patient management. As well, the questionnaire will help to identify these effects.

For our first part – we will refer to a series of treatments that you had with your care provider.

Please describe the course of the meetings, with emphasis on treatments and decisions in a time period:

1. What were the main psycho-social and demographic concepts to which you had referred during the meetings/s?
2. How did the treatments influence your daily routine, in terms of:
 - 2.1. Difficulties you encountered.
 - 2.2. Personal decision points you took, during daily treatments?
3. Family support:
 - 3.1. Did you have family support during the treatments?
 - 3.2. What was the level of the family member involvement in treatments, and the decisions impacted?
4. How did it reflect in the process (daily activities, meeting, etc.)? What habits and daily functions were changed during the treatment, and how it was reflected in the process?
5. Can you specify significant event\|s that made a (good\|bad) change in the treatment process or daily routine?
6. What is the most burdening element, why and how did you handle it?
7. Did the treatment influence your daily functions:
 - 7.1. How did the treatment process affect your work?
 - 7.2. How did the treatment process affect your hobbies?
 - 7.3. How did the treatment process affect your habits?
8. Did you change your diet, based on treatment recommendation?
 - 8.1. Did you keep the nutrition habits and for how long?

For the next part we will assume that we have a new patient guidance system that integrates hospital and monitoring data into a Personal Health Record (PHR) accessible by patients and care providers.

The system includes Decision Support System (DSS) at the back end (health center), and on the front end (patient), by utilizing monitoring technologies. The distributed DSS provides information to the patient and care providers. Personalization will be achieved by considering patient preferences and psycho-social and demographic context.

9. What are your expectations from the system, regarding the following aspects:
 - 9.1. Messages and type of information that you want to receive from the system?

9.2. What are the main features, you think, the system should present?

9.3. What new options the system will help you achieve?

10. How do you think the system will influence your treatment process?

11. How the system should help you follow treatment recommendations (reminders, medications, etc.)?

12. What don't you want that the system to do?

13. In cases you don't have web connection, namely in case you have only the smartphone and the monitoring system (no connection option to main server): What do you expect that the system will do? (notification, reminders, decisions support, suggest actions, etc.)?

Thank you for your cooperation.

Appendix C. Care providers' second interview: Refining and generalizing psychosocial and demographic concepts

We conducted preliminary qualitative studies consisting of interviews and questionnaires with different stakeholders. Based on the questionnaire analysis we defined the emerging psychosocial and demographic concepts that induce contexts which influence patient treatment recommendation and treatment goals.

Part 1 – defining scale for psycho-social and demographic concepts and the potential effects of the contexts that they induce on decision-making

The following list presents general psycho-social and demographic concepts defined by physicians, nurses and patients:

Ability to comply with treatment, communication level, cooperation level, desire to know truth about prognosis, education level, language level, trust level, need for accompanying person for visits, degree of family support, degree of routine of daily schedule and diet, distance from medical center, financial capability, living area accessibility, living area pollution, and family support level.

1. Per each concept, please define the following information:
 - 1.1. Scale (e.g. distance from medical center = close, medium distance, far, very far away)
 - 1.2. If relevant, what are the minimum and maximum units (e.g., close = 15 Km)
 - 1.3. What are the typical use cases scenarios of each concept-induced context, and can you think of personal events that induce these contexts?
2. Can we gather concepts into groups:
 - 2.1. Please specify the group and the concepts under it.
 - 2.2. What are the main characteristics of each group?
3. Per each context - What are the potential effects on treatment goals and recommendation:
 - 3.1. During physician-patient encounter (treatment goals, clinical measures, medication doses, limitations, etc.)
 - 3.2. During daily routine (clinical measures, medication doses, limitations, treatment options etc.)

Part 2 –effect of each context

4. For each concept – please consider all possible values. For each context corresponding to a concept value, what would be the effect on treatment recommendations?
5. Is it possible to define independent measures (or combination of metrics) that can establish that the context is present (e.g. high BP and rapid pulse can indicate patient being in a state of high stress)
6. For the main potential effects listed below – can you define alerts that we may add to the guideline?

Main effects on: treatment goals, referrals, measurement type and schedule, drug type and schedule, exercise, diet routine, appointment schedule, daily routine (work, hobbies).

Thank you for your cooperation.

Appendix D. Elicitation instrument for customizing clinical guidelines

The motivation of this questionnaire is to detect which psycho-social and demographic contexts affect treatment goals and plans. The questionnaire should be used with a specific clinical guideline considering the clinical decision points established in the guideline.

The process of elicitation of CIG-customized contexts (CCCs) is shown in Fig. 4 of the paper and is described here. In the first part of the elicitation process (top right of figure), knowledge engineers use general guiding principles for improving adherence and for reducing risks and complications, along with the psycho-social and demographic context-inducing concepts and general insights gleaned from the interviews that were previously conducted for eliciting these concepts and their effect types (as was described in part A of Section 3.2.2). Using these principles and insights, the knowledge engineers identify in the clinical guideline, and in its corresponding flowchart, sections in which risk, complications, and issues of non-adherence may arise. For example, adherence to short-acting insulin is difficult to achieve because the patient should have a routine daily schedule and diet; a non-routine schedule or diet increases the overall risk and the likelihood of complications. After the knowledge engineers focused on the sections that should be examined in the free-text clinical guidelines, in the second step (top left of the figure), expert clinicians define domain-specific scenarios that will focus on points of non-adherence to treatment, process risks, and process complications.

The next step is based on the domain-specific scenarios and on the psycho-social and demographic contexts identified by that point. In this step, the clinical experts define the domain-specific CCCs (e.g., semi-routine schedule) and their effects on treatment goals and recommendations (e.g., increase the number of daily measurements of blood glucose or blood pressure), which will enable the customization of the clinical guideline.

Psycho-social and demographic concepts describe different aspects of a patients' condition and their environment, including family support, daily routine, stress, etc.

Each psycho-social and demographic concept is represented as a data item that has a range of values on a categorical scale. The range may depend on the clinical domain. A CCC may be defined as an expression over psycho-social and demographic concepts that limit the allowed range of the concept (e.g., relating to medium stress level). A patient may be in a context for a time interval.

Having a flowchart representing the structure of the guideline, we would like to go over it and check which psycho-social and demographic contexts could change a specific treatment step. Therefore, a first step would be to prepare this flowchart, focusing on decision points related to clinical state of the patient. Please number the decision points in this flowchart, so that we can refer to them during this interview.

We have identified, after more than 50 interviews with physicians and nurses, the following psycho-social and demographic concepts and their ordinal scales (a-e):

psycho-social / demographic concept	a	b	c	d	e
Ability to comply with treatment	No ability	Minimal ability	Normal ability	Excellent ability	
Communication level*	Low	Medium	High		
Need for accompanying person for visits	Very high	High	Medium	Low	
Support level	No support	Some support	Night support	Frequent support	Full support
Education level	Low	Medium	High		
Daily/diet routine	No routine	Semi routine	Routine		
Fitness level	Rest	Minimal	Normal	Moderate	Extreme
Language level	No knowledge	Basic	Fluent		
Living area's accessibility	Difficult	Moderate	Easy		
Living area's pollution level	Very high	High	Medium	Low	
Distance from medical center	Isolated	Distant	Near by	Immediate	
Stress status	High	Medium	Low		
Financial capability	Low	Medium	High		

* Communication level is an abstraction over cooperation level, desire to know truth about prognosis, education level, language level, and trust level

From our previous interviews, the treatment steps affected by the psycho-social and demographic contexts are usually of the following types:

- Medication change – changes in dose, schedule or type;
- Measurement modification – changes in schedule, type and amounts;
- Allowed concept threshold – changes in maximum or minimum values;
- Physical activity routine – changes in physical activity routine;
- Diet modification– changes to recommended diet;
- Appointment schedule – changes to timing of future appointments.

The questionnaire's goal is to establish the treatments steps in the guideline that are affected based on the psycho-social and demographic context.

1. Could you specify for the psycho-social and demographic concepts, which ones should be filled out by physicians (because the patient's input might not be realistic)?
2. For the first part – looking at the guideline's flowchart, please provide information regarding the following items:

- Main clinical reasons that lead to non-compliance during treatments (For example, change in daily routine that leads to changes in meal composition or timing or even skipping a meal that will change the time of taking insulin in diabetes patients).
 - Think of clinical goals that are defined in the guideline, and identify which are the reasons that patient may not achieve these goals (For example, if asthma is not controlled using the current treatment regimen, treatment should be stepped up until control is achieved– which of the psycho-social and demographic contexts could be the reasons for not achieving controlled asthma state?).
 - Looking at the decision points, try to identify psycho-social and demographic contexts that would change the treatment step (For example, if there are several options to treat edema in a chronic patient and the best option requires family member support (concept), in case family members are not available (context), the physician may change the option to another option that the patient can handle in the new situation).
3. Please specify patients’ scenarios that you managed where his/her psycho-social or demographic context affected the treatment
 4. Please use the following table to define for each psycho-social and demographic concept, the scenario that it affected and what the effect was.

psycho-social and demographic concepts	Affected scenario (considering context)	Effect on treatment or clinical goal
Ability to comply with treatment		
Communication level*		
Need for accompanying person for visits		
Support level	If someone lives alone and had cerebral vascular accidents then starting insulin treatment may have a very high risk of hypoglycemia not being noticed	then he doesn't give insulin
Education level		
Daily/diet routine		
Fitness level		
Language level		
Living area’s accessibility		
Living area’s pollution level		
Distance from medical center		
Stress status		
Financial capability		

Appendix E. Pre-pilot testing

Before the evaluation study, we performed extensive testing during a pre-pilot period, to see that the system behaved correctly. Correct behavior implied correct and timely response to user input, to context switches, and to patterns discovered in the data. It also meant that the BEE-DSS changed plans and projected new knowledge to the mDSS. We tested that the mDSS and BE-DSS delivered correct and timely recommendations to care provider and patients. We further tested that the mDSS delivered reminders to patients according to current patient model (that considers current patient clinical data and current therapy), while addressing timing of reminders and active context. The testing for the AF domain was done by nine AF and seven GDM patients who used the system for 2 months. They asked the volunteers to test all monitoring and reminders, and the “pill in the pocket” evidence-based protocol, that for eligible patients (depending on their medication therapy and clinical and social parameters) provided indication on when to take an emergency drug that the patient was instructed to keep in his pocket and take when an AF event was experienced. The protocol involved responding to patients reports of unacceptable symptoms, sending them a recommendation to conduct a 30-minute ECG monitoring. If an AF episode was detected in high certainty and the patient has not taken the emergency drug in the past 4 hours, s/he was sent a recommendation to take the emergency pill and then test the ECG after 1 hour.

The AF clinicians tested the care provider’s parallel workflow consistently uses (like the patient workflow) the patient model and provides appropriate recommendations. In addition, the clinicians tested the cardioversion advice protocol.

To test the GDM system during the pre-pilot stage, we developed a simulation program that generated patient data for long periods of time because the temporal patterns that were monitored spanned several weeks (e.g., 30 days of good blood glucose values). The aim of the simulation program was to test all the different workflows covered by the GDM guideline. Four clinicians and three knowledge engineers involved in the guideline elicitation tested the system for two months in a daily basis, in terms of EMR data imports with simulated data, entry of GDM relevant monitoring parameters, checking of the content of technical and clinical recommendations and reminders generated both at patient and caregiver GUIs. Monitoring data entered by users during this period included simulated glucose values, diet, ketonuria, insulin, measurement of blood pressure with a sphygmomanometer and real physical activity sessions performed with a Physical Activity Detector (PAD) integrated in the mobile application. Bugs in software components and in the knowledge base were fixed for both domains.