Modeling and Executing a Guideline-based Clinical Decision Support System Using the PROforma Methodology

Shaimaa Saud AL-Fleit[†], Abdullah Saad AL-Malaise AL-Ghamdi[††], Mohammed Osama Nassif[†††]

[†, ††]Faculty of computing and information technology, King AbdulAziz University, Saudi Arabia
[†††]Faculty of medicine, King AbdulAziz University, Saudi Arabia

Summary
Clinical decision-support systems (CDSSs) are systems designed to influence clinician decision making regarding specific patients. They provide information related to a specific clinical situation and produce recommendations. Ambiguities in the cancer screening process and uncertainties in the identification of cancer symptoms, as well as concerns about providing referrals in primary care, lead to a complex cancer detection process for the general practitioner (GP). In addition, there is fragmentation in cancer care between primary and secondary care, which increases patients' feelings of uncertainty. Clinical interoperability guidelines are being used to help general primary care practitioners make appropriate decisions for the given clinical circumstances and improve adherence to paper-based guidelines. In this work, a CDSS is developed based on clinical practice guidelines (CPGs) and using the PROforma methodology, which employs the task network model (TNM). The TNM outperforms other models because it supports the guideline's steps, which are revealed over time, and it can explicitly model sequences of tasks or alternative pathways. CDSSs will assist GPs in challenges associated with detecting cancer in individual patients in a specific clinical situation.

Key words:
Clinical decision-support systems(CDSSs), Clinical practice guidelines(CPGs), PROforma, Task network model(TNM), Knowledge-based systems.

1. Introduction

Clinical decision support systems (CDSSs) provides knowledge and person-specific information to clinicians, patients, staff, or others in the clinical context, filtered intelligently or provided in appropriate times, to support health care [1]. The quality of the knowledge, among other factors, is the key determinant of the quality and utility of a knowledge-based CDSS; guideline-based healthcare practice has well-defined medical knowledge [2]. Clinical practice guidelines (CPGs) are documents that contain recommendations to ensure the best patient care as confirmed by an assessment of the benefits and harms of alternative care options based on the systematic review of evidence [3].

Cancer has recently become a big public health issue in many countries of the world, specifically breast cancer which considered as one of the common cancer types affecting women [4]. It is important to diagnose cancer at an early stage, where an early diagnosis is the first point to support survival in many cancers [5], and giving the general practitioners (GPs) the ability to identify those who need a rapid referral more easily can play an important role in delivering this earlier diagnosis [6]. Unfortunately, GPs are not always adequately prepared to meet the earlier cancer detection demands; in addition, the cancer diagnosis process is unclear for them because of the uncertainties in cancer symptom identification and referral decision determination in primary care which mostly results in misdiagnosis. This research is presenting a part of a master thesis aims to develop a guideline based CDSS using breast cancer screening and diagnosis CPG of the National Comprehensive Cancer Network (NCCN). To represent and execute the guidelines in this study, a formal knowledge representation language called PROforma [7], which have employed for extracting and representing the structure and content of a clinical guideline so that it can be interpreted by a computer using the Tallis software suite.

2. Related Work

Different knowledge-based and non-knowledge-based CDSSs for cancer are available, they can be classified based on the objectives of the systems. A review of the related work on these systems is provided in this section.

2.1 CDSSs for Risk assessment and diagnosis process

Risk Assessments in Genetics (RAGs) [8] is a system provided to the GPs that used family history data to evaluate a patient's genetic risk of cancer. it assesses the genetic risk by implementing a chosen guideline to the pedigree being examined using PROforma. It results in accurate, fast data entry and produce a suitable management decisions compared with systems using other methods. CADMIUM [9] is a system based on proforma, it was developed to investigate whether it could optimize the decision making on breast cancer screening. It was designed to classify the nature of the—whether it is benign or malignant—after automatically identifying micro-calcifications in breast tissue. The system resulted in raising the percentage of correct classifications of benign and malignant tumors and

Manuscript received July 5, 2019
Manuscript revised July 20, 2019

decreasing wrong diagnoses (false positives and false negatives). In [10], the researchers utilized knowledge from domain experts to develop a computer-aided diagnosis tool to categorize Breast Imaging-Reporting and Data System (BI-RADS) automatically using a fuzzy logic algorithm. A set of images, which had been previously analyzed by a group of experts, was analyzed using the Fuzzy Omega algorithm. Comparing the results, the accuracy for nodules was 76.67%, while the accuracy for calcifications was 83.34%. Another system with the purpose of mammogram classification was proposed in [11]; here, a computer-based diagnosis system for aiding radiologists in mammogram analysis was developed using wavelet co-occurrence features (WCFs) to extract features from four decomposition levels, and a genetic fuzzy system (GFS) was employed to maximize the classification accuracy. The GFS system performance was tested by mammograms extracted from the Mammographic Image Analysis Society (MIAS) database. The resulting system classification accuracy was 89.47% from only 16 features. A prototype of a CDSS was created for early diagnosis of breast cancer using case-based reasoning (CBR) to recommend diagnoses based on previous cases. Moreover, it aimed to incorporate fuzzy logic artificial intelligence technologies into the CBR system to reduce uncertainty in clinicians' reasoning and imprecision in case indexes. To evaluate the system, 50 historical records from St. Francis Mission Hospital for patients who had been diagnosed with breast cancer were used. The results showed that 48 out of 50 cases were diagnosed correctly using the system [12].

2.2 CDSSs for Treatment Selection Process

In [13], a linked data–based DSS for cancer treatment selection was proposed to overcome the lack of medical background knowledge in Health Information Systems (HIS), which will support doctors in specifying cancer treatment decisions. The proposed system used a similarity calculation algorithm with a semantic approach to provide the doctor with the most similar treatment cases from the historical database by calculating the similarity of the disease history and medication history of current and previous patients. A retrospective validation study conducted in a tertiary care cancer center in India [14] aimed to test the output of the Navya Expert System against oncologists' decisions in a tumor board decision. Navya Expert System is a CDSS that assists oncologist in providing a treatment decision for a specific patient by investigating medical guidelines and literature based on patient clinical data. Complex cases of women who had non-metastatic breast cancer with more than one treatment option were used to conduct the trial. The study concluded that the Navya Expert System, in combination with available clinical data, would provide expert treatment decisions; the Navya decisions were consistent with experts' clinical practice, and thus, the Navya Expert System can enhance the usage of evidence-based expert treatment decisions globally. In the same field, to assist oncologists with the treatment selection, a web CDSS for breast cancer patients was created [15]. The main system objective was to improve the treatment process and the assessment of the effect of prognostic factors on patient survival. In this system, three modeling methodologies are used to make a prognostic assessment based on patient characteristics, such as age at diagnosis and tumor size; these are Cox regression modeling, the Nottingham prognostic index (NPI), and a partial logistic ANN with automatic relevance determination (PLANN-ARD)). Each model provides an independent survival outcome for patients. All the prognostic models are combined with a stratification methodology based on regression trees to classify patients into a number of risk groups. The system will simply function after the patient's characteristics are inserted; the risk scores and the risk groups are calculated using the three models and the stratification method. The system was trained using a dataset of 743 breast cancer cases [15].

2.3 CDSSs for the Follow-up Process

For the breast cancer relapse risk estimation, a predictive model based on a naive Bayes (NB) classifier was developed in [16]. Two different NB models were tested, one using uniform (equally likely) priors and the other using empirical priors. Clinical, pathological, and therapeutic data were taken as a system input from a sample of 84 breast cancer patients with 82 months of follow up. The model performance was evaluated based on accuracy, specificity, sensitivity, the Area under the curve (AUC), and the balanced accuracy rate (BAR). The NB model shows good performance in predicting the risk of breast cancer relapse when using uniform priors in terms of sensitivity, specificity, and BAR.

In the clinical domain, approaches that model knowledge in an explicit way are preferred because they provide justification of the recommendation [17]. The problem in this work was the need for a CDSS for breast cancer diagnoses. As illustrated in the literature above, most of CDSS applications that have been developed in this area until now include family history and genetic risk assessment tools, and interpreting medical images in general healthcare practice; moreover, in specialized healthcare practice, most of the work was concentrated on treatment selection, with evident interest in developing systems for follow-up procedures. Systems related to breast cancer diagnosis focused on mammogram classification and interpretation of medical images. However, there is still a gap in the area of developing CDSSs to support complex healthcare pathways, managing CWFs, and decision making, especially in the breast cancer diagnosis area. This work will contribute to the previous research focusing on a key part of the breast

cancer care pathway, which is a breast cancer diagnosis process in general practice, making GPs work as the gatekeepers to specialist services, in which they provide patients with evaluation service in primary care. The proposed system will establish a prototype of a CDSS that can assist GPs and medical students in the breast cancer diagnosis process by developing an executable guideline technology to support complex healthcare processes, workflows of patient management, and clinical decision making. This system will aid decision makers—GPs in this case—throughout the diagnostic process, starting from screening plan, moving to the required tests recommendation (imaging and biopsy) based on the specific patient situation, and reaching the final diagnosis, which will either be referral to the breast surgeon or another suitable decision based on the patient case.

3. Guideline Modeling Languages

3.1 Arden Syntax

Arden Syntax is a rule-based methodology that creates MLMs, which are self-contained, independent units, each containing the logic necessary for a single medical decision [18]. An individual MLM comprises the rule, logical condition to activate the rule, and action that should be done when the rule is being activated [19]. The syntax of an MLM is provided as a text stream stored in an ASCII file in statements called slots, each of which contains a name and body [20]. An XML version was also recently proposed, which reduces the need for custom-built compilers [21],[22]. Arden Syntax provides some basic data types, such as Boolean, number, string, time, and duration, which are essential to medicine. Arden Syntax has an MLM query language as an expression language. An implementation of the Arden Syntax prototype was developed in Prolog, and an EE implementation was carried out using C++ [23] and Java [24]. A tool called the MLM Builder was also provided [25].

3.2 PROforma

The PROforma language composes the basic method and technology for developing and publishing executable clinical guidelines [26]. It integrates object-oriented modeling and logic programming, and it is based on R2L Language [27]. In PROforma, the guideline is represented as a group of data items and tasks, shown in a hierarchical way. The PROforma task model has four classes of tasks, which are Actions, Enquiries, Decisions, and Plans. Each of these four subclasses has a designated icon to be represented graphically using a diagram-based convention. Tasks in PROforma also have attributes, where their values determine how they are to be interpreted. PROforma has an expression language to define its conditions and arguments, which is a special type of attribute; PROforma expression language contains arithmetic, logical, and comparison operators, and it includes functions that assist the task execution states. For execution, the PROforma process description is loaded into the PROforma engine, which keeps a process dynamic state record. Currently, there are two implementations of the PROforma engine. One is the Arezzo implementation [28], which is commercially available from InferMed Ltd. The second is the Tallis implementation [29] from the Advanced Computation Laboratory of Cancer Research. The Arezzo and Tallis implementations are similar in their components, but the basic difference between them is that Arezzo was intended to be run on a Microsoft Windows platform, while Tallis was designed for delivering web-based services and has a Java-based implementation [26].

3.3 GLIF

GLIF [30],[31] was developed for modeling the guidelines as a flowchart of structured steps. GLIF contains different classes with the attributes to address the knowledge in the CPG and its complex relationships. The guideline step class has five types of subclasses or tasks that can be represented and linked with each other in a flowchart to designate their scheduling through the guideline implementation [32]. These five main process-modeling entities are the decision, action, synchronization, branch, and patient state steps. These concepts are used to make an algorithm of the guideline and provide a general view of the guideline decision-making process [33]. Guideline Expression Language, Object-oriented (GELLO) [34],[35], which was recently approved as by Health Level 7 (HL7) and American Nation Standard Institute (ANSI) standard, is a vendor- and platform-independent, extensible, side-effect-free, object-oriented executable expression language for GLIF [36]. GLIF uses a specific EE called the Guideline Execution Engine (GLEE) [37] for guideline enacting purposes. GLEE is currently implemented in Java [32].

3.4 Asbru

Asbru is an intention-based and time-oriented skeletal-plan representation language that is used for the specification of clinical protocols. There are different concepts defined in a plan, which are as follows: (1) concepts for characterizing plan attributes, for example, conditions, intentions, and effects; (2) concepts for ordering plans; and (3) concepts for setting temporal dimensions of states and plans [2]. Asbru plans are written in XML, the schema specification is given in [38]. There are several tools that support the authoring and visualizing of Asbru guidelines. These include AsbruView [39]–[41], Delt/A [42], [43], CareVis [44], and the DeGeL framework [45]–[47]. Asbru's original execution environment is called the Asbru Run Time Modules (AsbruRTM) [48]–[50], and it is written in Java.

To translate the XML-based Asbru plans to Java classes, the EE facilitates Castor, an open source data-binding framework for Java, which produces a model of Java objects out of Asbru's XML schema. Another EE engine, called Spock, is included in the DeGeL framework [51], [52].

Arden Syntax has been supported because it represents the guidelines using interacting MLMs, although MLMs do not provide full support representing a guideline with a multistep that continuously unfolds [53]. Task network models (TNM) approach can overcome this issue. TNM languages basically provide specifically designed modeling primitives for the representation of complicated, multistep guidelines and the explanation of temporal and different relationships of component tasks. Distinct from rule-based systems, TNMs can represent sequences of tasks (i.e., control flow) or alternative pathways explicitly; they also have tools for plans visual representation and task organization within them [53]. The rest of the clinical guidelines modeling languages use TNMs for modeling the guidelines; their approaches are only slightly different when it comes to addressing specific modeling challenges. Peleg et al. [53] conducted a study for comparing the different CIG modeling approaches, including PROforma, GLIF, and Asbru. They concluded that all the approaches are similar in their basic aspects and had the same components, including the expression language, plan organization, data abstractions, medical concept model, and conceptual medical record model. In addition, there is no standard framework to permit judgment between them.

4. The PROforma Methodology

For representing and executing the guideline the PROforma language was chosen to extract and represent the content and structure of a clinical guideline to enable its interpretation by a computer. PROforma is an example of a task-based guideline modeling format that employs TNMs. It is an executable process modeling language that describes the protocols and care pathways of the guidelines and automates clinical processes. It is composed of knowledge representation language features, such as the methods using artificial intelligence, and formal specification language features, such as methods that use software engineering [54]. It also merges object-oriented modeling and logic programming [55].

The structure of the PROforma language is based on a model called the domino model [56]. The nodes stand for concepts, and the arrows stand for inference processes. The domino model facilitates the automation of the care process using the plan enactment (right-hand side) and decision-making parts (left-hand side), (see Fig. 1). PROforma uses data structures to model the decision making and automated inference mechanisms to illustrate the enactment part. The model describes the process of decision making and task management accomplished through the execution of clinical procedure, as shown in Fig. 1; starting from the appearance of the clinical problem that needs to be diagnosed, this is represented by arrow 1. Then, with the assistance of general medical knowledge, the computer will provide an alternative solution to the diagnosis problem, indicated by arrow 2. Then, each alternative's arguments is constructed, as shown by arrow 3. The system then recommends one of the diagnosis candidates, as shown by arrow 4. After commitment to the diagnosis results, another cycle will start, which is a treatment cycle; this is the same as the diagnosis cycle, and the system will provide the specific therapeutic procedure, arrow 5, and scheduling process for clinical tasks in terms of time and order, as represented by arrow 6.

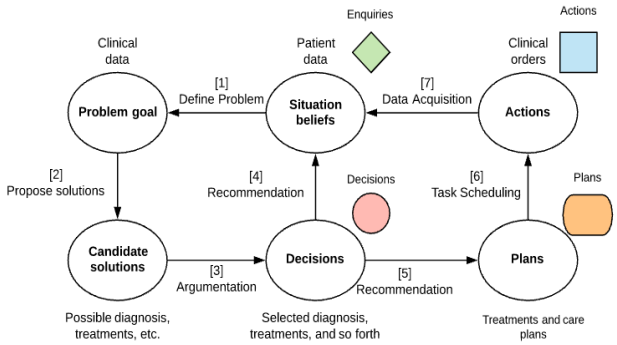

Fig. 1 Domino model, the basis model of the PROforma language[56].

PROforma is distinguished from other schemes by its abstract nature, Thus, rather than describing all the guideline concepts, which contain a range of various concepts—such as diagnostics, imaging, clinical examination, gathering of data via laboratory tests, and treatment decisions—the semantics of PROforma deals with components, representing groups of objects [26]. The classes of objects are coordinated in an inheritance hierarchy that forms a PROforma component set (see Fig. 2). Data items contain values provided by an enquiry that is accessible to the tasks. A task can be an Enquiry, Action, Plan, or Decision. The candidates are objects representing decision options. Arguments are objects represent arguments for or against a specific decision option (candidate). Warning conditions are conditions to be checked when a value is assigned to a data item, while parameters are used to assign data to a specific task instance. Finally, sources represent information that is needed by an enquiry task.

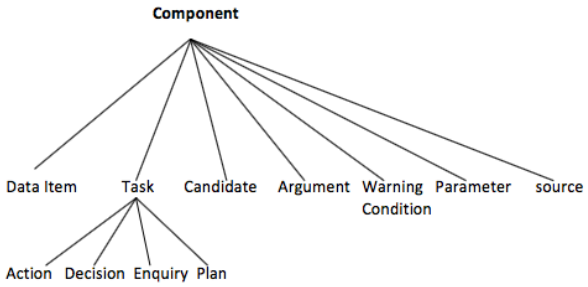

Fig. 2 PROforma component set [26].

CIG language must at least have the following two tools: an authoring tool and EE [57]. For acquiring the guideline knowledge, PROforma is composed of a two-step process, which makes up the PROforma composition method. first step is representing a task network graphically, second step is populating task templates. representing a task network graphically, by describing the high level of the guideline with the usage of a graphical design package. The graphical network results in providing the view of the basic clinical tasks, including temporal and logical interrelationships, in an understood form. The guideline is represented as a set of data items and tasks. The tasks are organized for a plan in a hierarchical way. The PROforma task ontology has four classes [7] (see Fig. 3). Actions represent procedures that have to be done in the external clinical environment, like the task of administering an injection. Enquiries represent information required from the user or external system to make a decision or complete a procedure. Decisions represent the choice that has to be made, either about what to believe or what to do, such as a diagnosis or therapy decision. Finally, plans are groups of tasks that are collected together to accomplish a clinical objective. The PROforma process description is represented as directed graphs, where nodes stand for tasks and arcs stand for scheduling constraints [58].

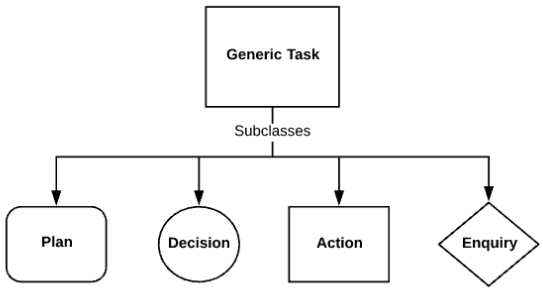

Fig. 3 PROforma task ontology[7].

Step 2 is populating task templates, which involves defining the specifics of the clinical tasks included in the guideline. Properties for the tasks are separated into the generic task properties, which are common for all tasks, and specific properties for each class of tasks [7]. Generic task properties that all classes may have values for are as follows:

i. Preconditions which are the logical conditions that should be true for the task to start;
ii. Goals, or the logical condition explaining the situation the task has to achieve; when the goal is achieved, the task will terminate;
iii. Descriptions, representing an explanation of the task;
iv. Cycles, which define under what constraints and conditions the task will be repeated.

With the previous generic properties, the decision, plan, enquiry, and action each has specific properties.
The decision properties are:

i. candidates—represent the decision options. Each candidate has multiple arguments;
ii. argument—is a logical condition with an incorporated weighting, and the support value for each candidate can be specified by calculating the weightings of the true arguments;
iii. A recommendation rule may also be associated with each candidate; this determines whether it is advisable to choose the candidate.

Plan properties are divided into the following:

i. Components which are the tasks included in the plan;
ii. Scheduling Constraints, which determine the task execution order;
iii. Abort Conditions, which are the logical conditions under which a plan will fail if they are found to be true;
iv. Termination Conditions, where the logical conditions under the plan will be terminated if they are true.

Enquiry Properties include:

i. Sources—representing the names of the required data items; the sources may be associated with an SQL query that will yield the requested values or a query that defines an external database.

Action Properties include:

i. action procedures—texts describing operations to be finished by a human or database update described through an SQL statement.

5. Tallis Software

A number of software tools have been developed to create, represent, and execute PROforma guidelines. One of them is Tallis, which will be used in this study. This is a software suite containing a group of components written in Java. It has two basic applications, which are as follows: the composer, a tool for graphically authoring knowledge for the editing, creation, and graphical representation of CIGs, and the tester, which tests and enacts the guideline. The composer and tester access and connect to other components of Tallis through their Application programming interfaces (API)s. Other components include an engine, which executes guidelines and can be manipulated by other applications. The Tallis tester calls on the engine to execute the guideline [26].

6. Model Development

To model the workflow of the cancer diagnosis procedures, we used the guidelines followed by the oncologists at King Abdulaziz University hospital, which is the breast cancer screening and diagnosis CPG of NCCN [59]. The PROforma modeling language was used to encapsulate the clinical care protocols into CIGs, Because of the lack of conventional graphical representation in PROforma [19], and for the models to be clear for domain experts to understand. The UML activity diagrams were used to make the initial design, which was translated into PROforma during system implementation. Before translating them to PROforma, for the CDSS model to be appropriate for the use of GPs, the UML models were revised and refined in an iterative process with assistance from Dr. Mohammed Nassif, a surgical oncologist and assistant professor in the Department of Surgery, King Abdulaziz University.

6.1 Knowledge Modeling

The models representing clinical knowledge after being revised by Dr. Nassif are presented in UML activity diagrams [60], [61]. The first model represents the first phase in the breast cancer screening and diagnosing guideline, which is the clinical encounter; this includes the assessment of breast cancer risk and clinical breast examination (CBE); Fig. 4. This model Start from the collect patient data activity, comprising age and other risk factor data, after that, the CBE activity is placed for the physical evaluation. Then, a decision contains two options to decide whether a patient has symptoms. Different activities follow each option. If it is a symptomatic case, the pathway will proceed to the next model. If it is an asymptomatic case, then the next activity will be based on the data entered earlier. If a patient has any of the risk factors or her age is 40 years or older, she will be categorized under the increased risk category and recommended to attend an annual clinical visit, receive an annual screening mammogram, and be aware of her breast characteristics; in cases where patients have genetic positivity for the disease, physicians should immediately send patients to the breast surgeon. Asymptomatic cases with no risk factors will be categorized under average risks.

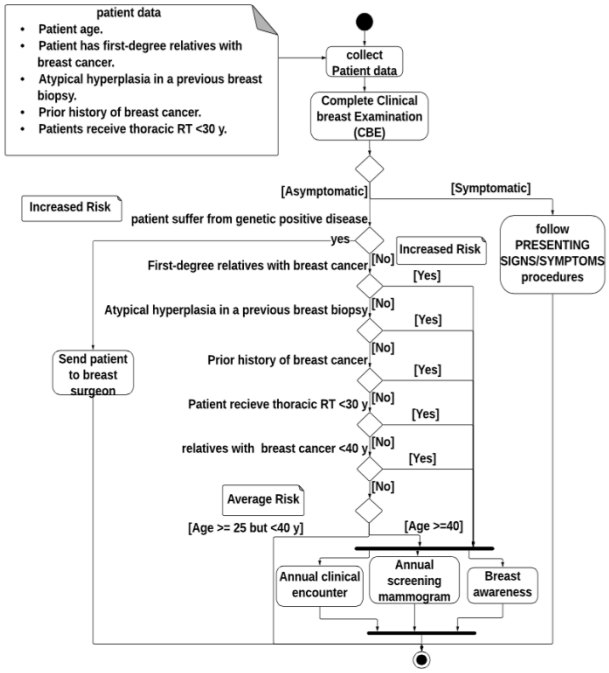

Fig. 4 Activity diagram for the clinical encounter procedures workflow.

If the patient has symptoms, the pathway will proceed to the next model to let the user determine which symptoms to choose from the group of symptoms, namely, palpable mass or asymmetric thickening/nodularity, nipple discharge, skin changes, breast pain, or axillary mass (Fig. 5). Each presenting symptom follows a different pathway, with recommended procedures (see Fig. 6–9) except for the nipple discharge symptoms option, which only has one action—sending the patient to a breast surgeon.

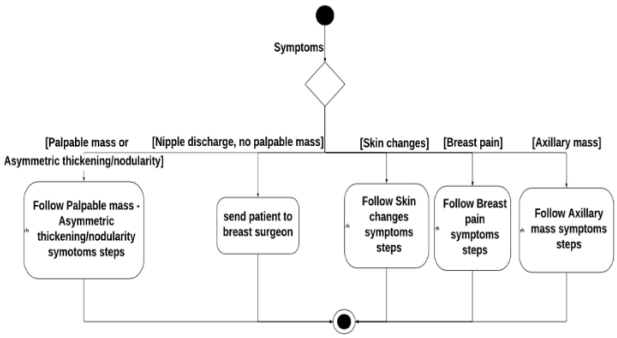

Fig. 5 Activity diagram for symptomatic patient procedures workflow.

Based on the selected symptoms type option from the previous model, the workflow will proceed to the model present in Fig. 6 If the option that has been selected is palpable mass or asymmetric thickening/nodularity, the next action will be based on the patient age, which was entered previously. If the patient age is less than 30 years, the recommended action is bilateral breast and axilla ultrasound. If the patient is older than 30, the recommended action is mammogram and bilateral breast and axilla ultrasound. The breast imaging result should be entered by the user. The results are split into five categories, as follows: BI-RADS1 (negative), BI-RADS2 (benign), BI-RADS3 (probably benign), BI-RADS4, and BI-RADS5, (suspicious) and (highly suggestive of malignancy) respectively. If the result is BI-RADS1, the resulting action will be symptomatic management, and subsequently, follow-up screening. If the result is BI-RADS2, lesion size is required from the user; if the lesion size is more than 3 cm, the next action will be to send the patient to the breast surgeon. If the lesion size is less than 3 cm, the lesion type is required from the user. If the type is infection (mastitis/abscess), the following action will be treating with antibiotics + drainage, followed by sending the patient to the breast surgeon. If the type is simple cyst, the recommended action will be considering drainage for symptom relief, followed by follow-up screening. If the type is neither infection nor simple cyst, the recommended action is patient reassurance and subsequent by follow-up screening. If the result is BI-RADS3, the lesion size is required from the user; if the lesion size is more than 3 cm, the followed action will be the same as that for a lesion greater than 3 cm in BI-RADS2, but if the lesion size is less than 3 cm, the recommended action will be follow up with a diagnostic mammogram and/or ultrasound for 6 months and subsequent follow-up screening. If the result is BI-RADS4/5, the following action will be core needle biopsy, followed by sending the patient to the breast surgeon.

Fig. 7 shows the actions following the skin changes symptom selection. The first action will be determined based on the age of the patient; bilateral breast and axilla ultrasound will be recommended for a patient under 30 years old. Patients older than 30 years should proceed with a diagnostic mammogram and bilateral breast and axilla ultrasound. If the breast imaging result is BI-RADS1–3, the recommended action will be treatment with antibiotics and follow up for 2 weeks concurrently. After completing the two actions, if the skin changes are resolved, the next action should be follow-up screening; if it persists, the following action should be sending the patient to the breast surgeon. If the BI-RADS result is 4 or 5, the recommended action is core needle biopsy, followed by sending the patient to the breast surgeon.

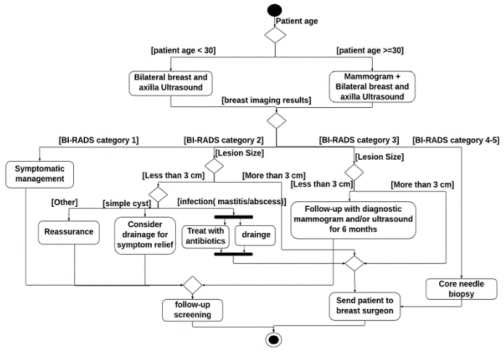

Fig. 6 Activity diagram for palpable mass or asymmetric thickening/nodularity symptoms procedures workflow.

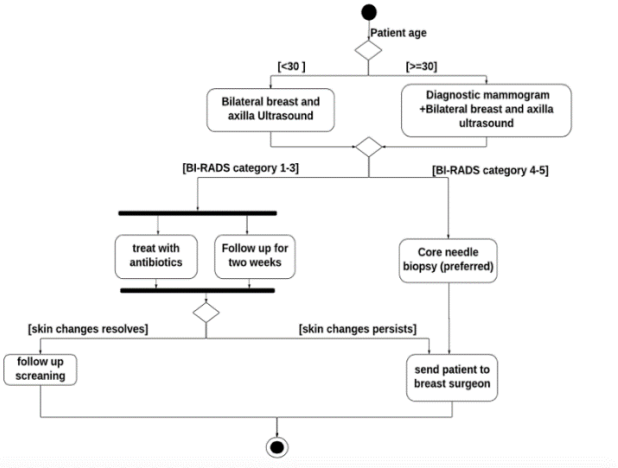

Fig. 7 Activity diagram for the skin change symptoms procedures workflow.

As shown in Fig. 8, the breast pain symptoms option will require the user to enter the result of complete history and CBE to determine the cause of pain; this will be breast mass or asymmetric thickening/nodularity, nipple discharge, skin changes, or no physical findings. If the result is breast mass or asymmetric thickening/nodularity, the flow will follow the palpable mass or asymmetric thickening/nodularity symptoms steps model (see Fig. 6). If the result is nipple discharge, the next action will be sending the patient to the breast surgeon. If the result is skin changes, the flow will follow the skin changes symptoms steps model (see Fig. 7). If there is no physical finding, the user should determine the pain characteristics—whether the pain is focal or non-focal, cyclic, diffuse pain with a size larger than a quadrant. If it is focal pain, the flow will follow the same steps for a palpable mass or asymmetric thickening/nodularity symptoms (see Fig. 6), except for patients with BI-RADS2 and masses less than 3 cm with neither infection nor simple cyst lesion. In this situation, the recommended action is symptomatic management of pain and reassurance. If the pain is non-focal, cyclic, and diffuse, with a size larger than a quadrant, the next action is to reassure the patient and provide treatment if needed.

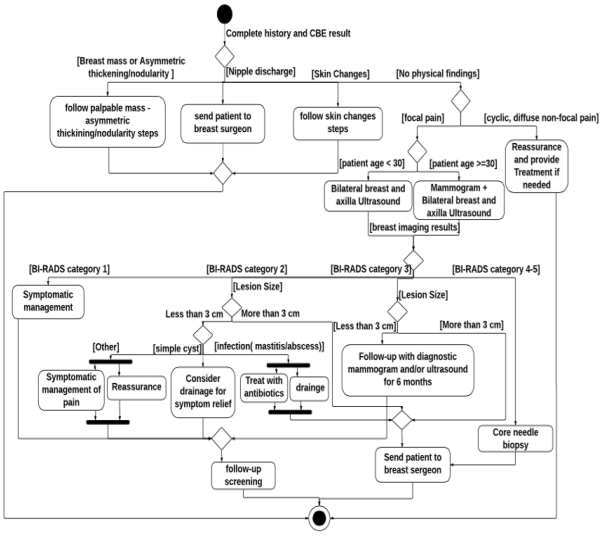

Fig. 8 Activity diagram for breast pain symptoms procedures workflow.

The last model is the axillary mass symptoms steps (see Fig. 9); this requires the user to specify the location of the mass, namely, whether it is bilateral or unilateral. If it is bilateral, the followed action should be local and systemic evaluation and images. If the results show that the patient has systemic disease, the resultant action will be sending the patient to the medical oncologist. However, if the results show that the patient does not have systemic disease, the procedures will be those of the unilateral axillary mass steps, which first recommend bilateral breast and axilla ultrasound for patients less than 30 years old and a bilateral breast and axilla ultrasound with a bilateral breast and axilla diagnostic mammogram for patients aged 30 years old or older. The next actions will be based on the imaging results. If it is negative or benign axilla -BI-RADS1–3-, the recommended action is close follow up for 3 months; if it is suspected -BI-RADS4–5- , the recommended action will be sending the patient to the breast surgeon.

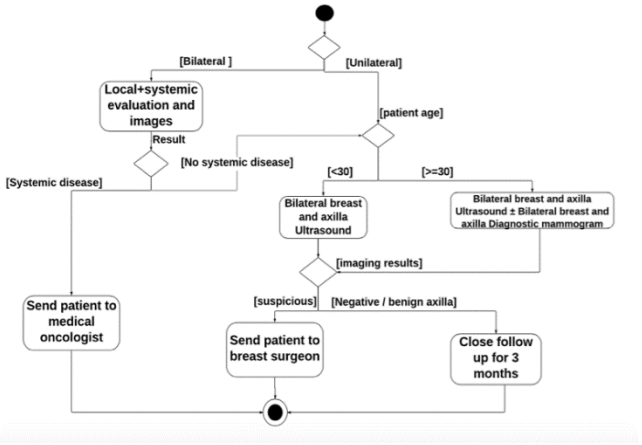

Fig. 9: Activity diagram for axillary mass symptoms procedures workflow.

6.2 Models Transformation Using the PROforma Composition Method

In this subsection, the UML activity diagrams illustrated in the previous section will be translated into a PROforma methodology.

6.2.1 Graphical Representation of the Task Network

First, a network of objects developed in PROforma, called a tasks network was created, which represents the actions, decisions, plans and enquiries required for implementing the guideline. Fig. 10 gives a screenshot of the PROforma protocol editor presenting the top-level plan for the breast cancer screening and diagnosing guideline task network, also called a process description. As can be seen in Fig. 10, the task network of the top-level plan of the breast cancer screening and diagnosing guideline consists of one decision, three actions, five enquiries, and four plans. The decision is asymptomatic patient recommendation. The actions are complete CBE, send patient to the breast surgeon, reassurance, and treatment. The enquiries are collect patient data, symptom presence, symptoms type, complete history and CBE results, and pain characteristics. Finally, the four plans are increased risk recommendations, axillary mass symptoms procedures, skin changes symptoms procedures, and palpable mass or asymmetric thickening/nodularity symptoms. PROforma plans decompose into other tasks. In the main plan, there are four other plans with tasks that compose them. Each contains a task network representing the workflow for that plan.
attribute templates for entering the attribute values. Templates lead the application designer in creating the task specification for the application.

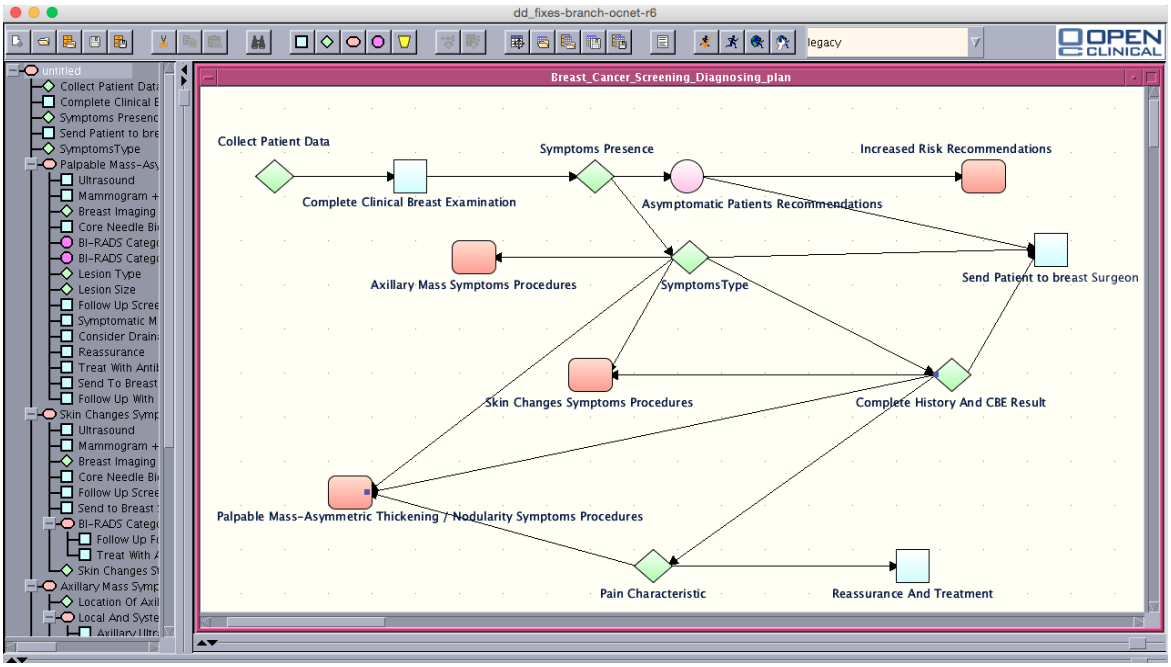

Fig. 10 View of breast cancer screening and diagnosing procedures as a task network (as displayed in the Tallis authoring tool). This shows the main plan, which represents the workflow of the breast cancer screening and diagnosing guideline in both the network (right side) and tree view(left side).

6.2.2 Populating Task Templates

The task networks presented in the previous subsection are just a "design sketch" of the process description, and they are not sufficient for developing the complete application. Detailed knowledge must be provided to enact each component task. For each task, there is a set Each task has two templates to allow values to be entered in addition to the sketch templates presented in the previous subsection. The left side represents the generic attributes fields, which are common for all types of tasks, while the right side represents the specific attributes for each task type; Figures 11 shows the general and specific

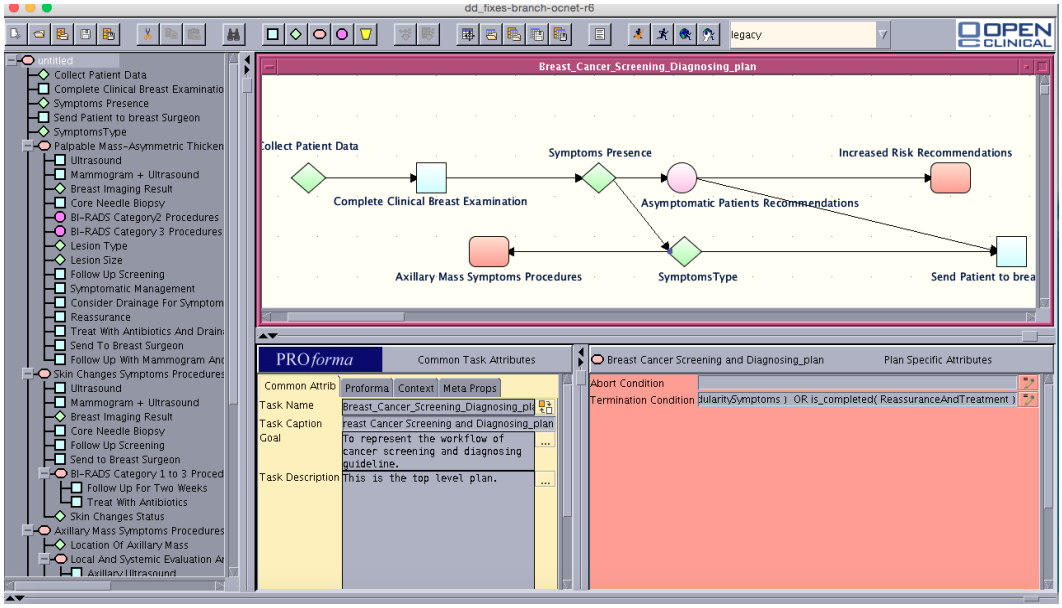

Fig. 11 Details of the breast cancer screening and diagnosing plan are added using a specialized editor: specific attributes (right), generic attributes (left).

attributes for the breast cancer screening and diagnosing top level plan.
The rest of the PROforma tasks was populated with the generic and specific attributes for the breast cancer screening and diagnosing guideline.

7. Testing and Evaluation

To test the model, the process description created using the Tallis composer is tested using the Tallis tester, which provides a test environment to check the correct execution of the application. After investigating the execution paths of the algorithm, different variables are identified. Using the identified variables, patient data are extracted from 30 real patient cases that were previously diagnosed and used as test cases. These data have been collected using observation and structured interviews with the surgical oncologist, Dr. Nassif, in his clinics, and 10 simulated cases have added to satisfy the branch testing criteria.

7.1 Enacting process description in Tallis Tester

A verification tool provided by Tallis was used to search for syntactic and semantic errors in the process description. After verifying the process description, zero errors were found, and the process description was executed using the Tallis tester. Fig. 12 displays the main screen of the Tallis tester enacting the breast cancer screening and diagnosing guideline.
As can be seen in Fig. 12, the Tallis tester's main screen contains five panels. The requested data panel displays the sources when the enquiry state is in progress, actions are displayed in the confirmable actions/keystones panel, and the procedure panel displays the procedure of the action in progress. Moreover, decisions are displayed in the confirmable decisions panel, and finally, decision candidates are displayed in the candidates panel.

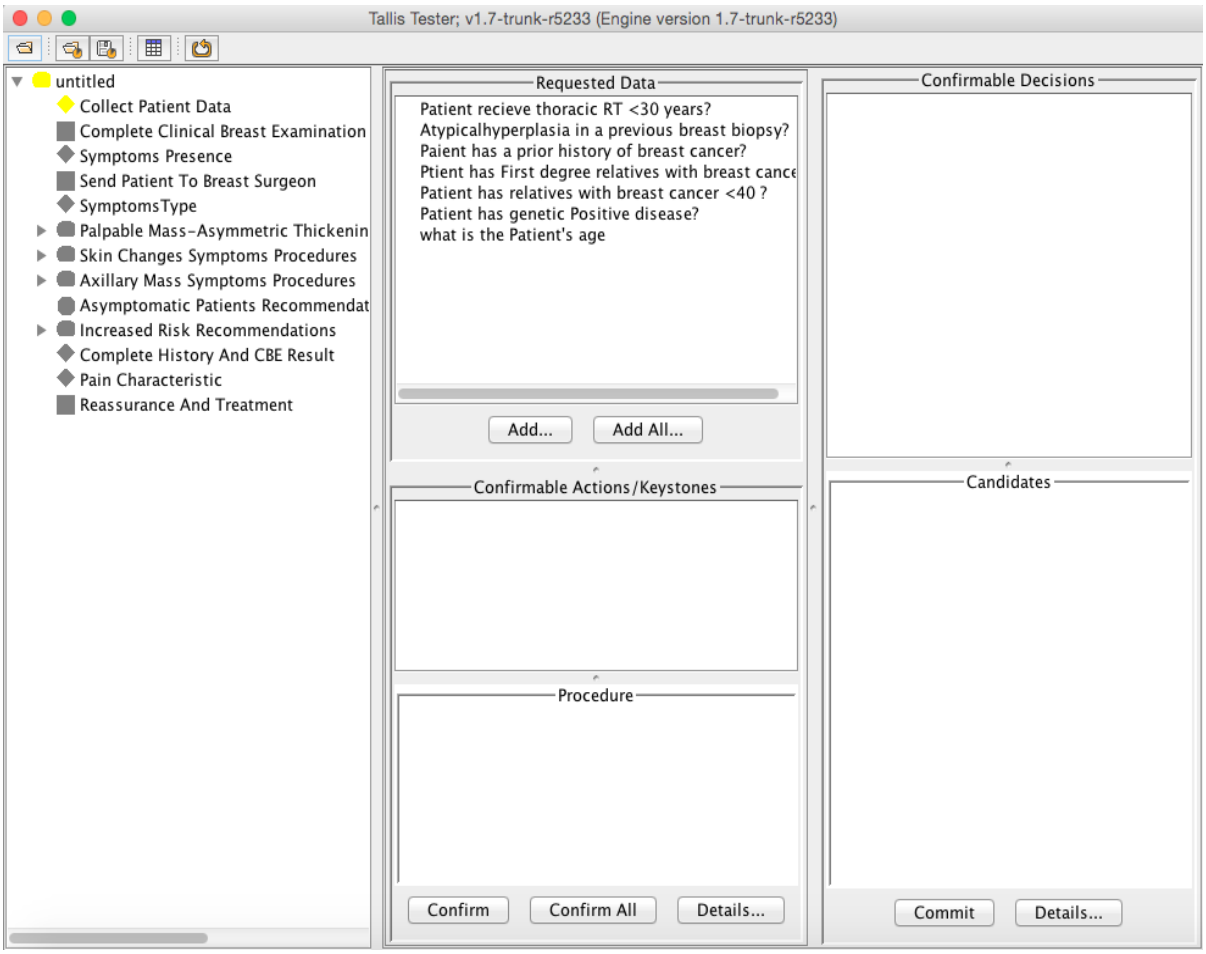

Fig. 12 Main screen of the Tallis tester while enacting the breast cancer screening and diagnosing guideline

7.2 Performance Evaluation

After finishing testing the 40 test cases, the simulated case recommendations results were reviewed by the specialist, who approved their correctness. The real case recommendation results were compared to the specialist's diagnosis. The mapping in Table 1 was used for comparing the specialist's final recommendation to the CDSS for evaluation; as can be seen, each final action recommended was mapped to a class, either likely cancer or unlikely cancer; this mapping was done with the assistance of the specialist.

Table 1: Mapping Each Final Recommendation to Its Classification

Final recommendation	Class
Send patient to breast surgeon	Likely cancer
Annual clinical visit with annual screening mammogram and breast awareness	Unlikely cancer
Follow-up screening	
Treat with antibiotics and drainage followed by sending patient to breast surgeon	
Send patient with BI-RADS2 and lesion more than 3 cm to breast surgeon	
Send patient to medical oncologist	

The accuracy, sensitivity, and specificity of the CDSS were calculated for cases classified as likely cancer and unlikely cancer. There were 8 recommendations produced by the CDSS as likely cancer and 22 as unlikely cancer recommendations. Twenty-seven cases correctly matched the specialist's recommendations, resulting in 90% accuracy. The sensitivity result was 86% and specificity was 91%, and the system was better in identifying the unlikely cancer, or in other words, the negative cases. Table 2 presents the results in a confusion matrix. The three unmatched recommendations were two cases diagnosed incorrectly as likely cancer cases and one case diagnosed incorrectly as an unlikely cancer case. In these cases, the system correctly followed the guideline instructions, but due to some exceptional reasons related to special patient cases, in the real-world processes, the specialist did not prefer to follow the guideline recommendations. Instead, he chose to provide different recommendations based on the patient's situation.

Table 2: Confusion Matrix for Cancer Diagnosis

		Specialist diagnosis		
		Likely cancer	Unlikely cancer	
CDSS diagnosis	Likely cancer	6	2	
	Unlikely cancer	1	21	
		Sensitivity: 86%	Specificity: 91%	Accuracy: 90%

The disease prevalence in the real test cases was 23%, where only 7 patients out of 30 real cases were diagnosed with having a higher suspicion of breast cancer. Consequently, they were referred to a breast surgeon. The Positive Predictive Value (PPV), or precision, of the system was 75%, indicating that the system had a relatively stable performance in confirming likely cancer cases (positive cases). The Negative Predictive Value (NPV) was 95%, indicating the powerful performance of the system in confirming the unlikely cancer cases (negative cases).

The False Positive Rate (FPR) or fallout was 9%, representing the proportion of the negative cases that were falsely diagnosed as positive breast cancer cases. The False Negative Rate (FNR) or miss rate was 14%, representing the proportion of positive breast cancer cases that were incorrectly diagnosed as negative cases. Table 3 shows the percentages described above.

Table 3: Other Measurements Derived from the Confusion Matrix

Measurement	Prevalence	Precision	NPV	FPR	FNR
Percentage	23%	75%	95%	9%	14%

8. Conclusion

Cancer is a serious problem around the world. Deaths due to cancer have been increasing rapidly, and breast cancer is one of the most common cancers affecting women.

Detection of cancer as early as possible is the first point toward patient survival in many cancers. Cancer cases are often detected via primary health care, so giving GPs the ability to more easily identify those who need a rapid referral can play an important role in delivering this earlier diagnosis. However, GPs are not always adequately prepared to meet the earlier cancer detection demands which will sometimes lead to misdiagnosis. Thus, we identified a need for CDSSs to help practitioners in primary care address the problem of cancer detection. We also found that the available studies in this area concentrated on the cancer classification problem, and there was a lack of CDSSs developed to support GPs in the complex healthcare pathways and decision making, especially for breast cancer diagnosis.

To represent and execute the guideline in this study, the PROforma language—which can be defined as an executable process modeling language to describe guidelines, care pathways, and protocols and automate clinical processes—was employed to represent the guideline using TNMs. PROforma was used to extract and represent the structure and content of the breast cancer screening and diagnosis CPG from the NCCN, interpreting it via a computer using the Tallis software suite. Moreover, because of the lack of conventional graphical representation in PROforma and to ensure that the models would be clear for a domain expert to understand, UML activity diagrams were used to make the initial design. The developed model was executed in the Tallis tester using 30 real test cases; then, 10 simulated test cases were added to satisfy the branch testing criteria. After testing the 40 test cases, the simulated case recommendations results were reviewed by the specialist, who approved their correctness. The real case

recommendations results were compared with the diagnoses made by the specialist, showing that the CDSS accurately diagnosed most of the cases, with 27 cases out of 30 concordant with the specialist's recommendations, representing 90% accuracy. The incorrect diagnosing occurred because of the specialist's preferences in dealing with exceptional cases. In addition, the sensitivity was 86% and specificity was 91%, indicating the better performance of the system in identifying the negative cases than diagnosing positive ones. The PROforma guideline–based CDSS was better in its performance than the CDSS models for cancer diagnosis in [10], [11].

References

[1] J. A. Osheroff, J. M. Teich, B. Middleton, E. B. Steen, A. Wright, and D. E. Detmer, "A roadmap for national action on clinical decision support.," J. Am. Med. Inform. Assoc., vol. 14, no. 2, pp. 141–5, 2007.

[2] J. L. Mathe, "The precise construction of patient protocols: modeling, simulation and analysis of computer interpretable guidelines," 2012.

[3] Institute of Medicine (U.S.). Committee on Standards for Developing Trustworthy Clinical Practice Guidelines. and R. Graham, Clinical practice guidelines we can trust. National Academies Press, 2011.

[4] "WHO | Cancer," WHO, 2018. [Online]. Available: http://www.who.int/cancer/en/. [Accessed: 27-Jun-2018].

[5] T. Green, K. Atkin, and U. Macleod, "Cancer detection in primary care: insights from general practitioners.," Br. J. Cancer, vol. 112 Suppl, no. Suppl 1, pp. S41-9, Mar. 2015.

[6] Z. Gross et al., "Using Cancer Decision Support Tools to support the early diagnosis of cancer Accelerate,."

[7] J. Bury, J. Fox, and D. Sutton, "The PROforma guideline specification language: progress and prospects," Stud. Heal. Technol. Informatics., 2001.

[8] A. S. Coulson, D. W. Glasspool, J. Fox, and J. Emery, "RAGs: A Novel Approach to Computerized Genetic Risk Assessment and Decision Support from Pedigrees."

[9] P. Taylor, J. Fox, and A. T. Pokropek, "The development and evaluation of CADMIUM: a prototype system to assist in the interpretation of mammograms," Med. Image Anal., vol. 3, no. 4, pp. 321–337, Dec. 1999.

[10] G. H. B. Miranda and J. C. Felipe, "Computer-aided diagnosis system based on fuzzy logic for breast cancer categorization," Comput. Biol. Med., vol. 64, pp. 334–346, Sep. 2015.

[11] M. M. Pawar and S. N. Talbar, "Genetic Fuzzy System (GFS) based wavelet co-occurrence feature selection in mammogram classification for breast cancer diagnosis," Perspect. Sci., vol. 8, pp. 247–250, Sep. 2016.

[12] P. R. J. Simeon, "A DECISION SUPPORT SYSTEM FOR THE DIAGNOSIS OF BREAST CANCER USING FUZZY LOGIC AND CASE BASED REASONING WANYONYI." 2016.

[13] J. Hu, H. Cai, B. Xu, and C. Xie, "A Linked Data Based Decision Support System for Cancer Treatment," in 2014 Enterprise Systems Conference, 2014, pp. 39–44.

[14] N. Nair et al., "Abstract P4-16-01: Validation of a software based clinical decision support system for breast cancer treatment in a tertiary care cancer center in India," Cancer Res., vol. 75, no. 9 Supplement, pp. P4-16-01-P4-16-01, May 2015.

[15] A. S. Fernandes, P. Alves, I. H. Jarman, T. A. Etchells, J. M. Fonseca, and P. J. G. Lisboa, "A Clinical Decision Support System for Breast Cancer Patients," Springer, Berlin, Heidelberg, 2010, pp. 122–129.

[16] S. Tortajada et al., "EP-1210 Naïve bayes models for predicting the risk of loco-regional relapse in breast cancer patients," Radiother. Oncol., vol. 111, p. S59, 2014.

[17] F. Velickovski et al., "Clinical Decision Support Systems (CDSS) for preventive management of COPD patients.," J. Transl. Med., vol. 12 Suppl 2, no. Suppl 2, p. S9, Nov. 2014.

[18] J. L. Mathe et al., "A Model-Integrated, Guideline-Driven, Clinical Decision-Support System," Software, IEEE, vol. 26, no. 4, pp. 54–61, 2009.

[19] D. R. Sutton, P. Taylor, and K. Earle, "Evaluation of PROforma as a language for implementing medical guidelines in a practical context," BMC Med. Inform. Decis. Mak., vol. 6, pp. 1–11, 2006.

[20] G. Hripcsak, P. Ludemann, T. A. Pryor, O. B. Wigertz, and P. D. Clayton, "Rationale for the Arden Syntax," Comput. Biomed. Res., vol. 27, no. 4, pp. 291–324, Aug. 1994.

[21] R. M. Sailors, "ArdenML: The Arden Syntax Markup Language (or Arden Syntax: It's Not Just Text Any More!).," AMIA, 2001.

[22] S. Kim, P. J. Haug, R. A. Rocha, and I. Choi, "Modeling the Arden Syntax for medical decisions in XML," Int. J. Med. Inform., vol. 77, no. 10, pp. 650–656, Oct. 2008.

[23] R. A. Kuhn and R. S. Reider, "A C++ framework for developing medical logic modules and an Arden Syntax compiler," Comput. Biol. Med., vol. 24, no. 5, pp. 365–370, Sep. 1994.

[24] M. Gietzelt et al., "Arden2ByteCode: A one-pass Arden Syntax compiler for service-oriented decision support systems based on the OSGi platform," Comput. Methods Programs Biomed., vol. 106, no. 2, pp. 114–125, May 2012.

[25] R. M. Sailors, "MLM Builder: An Integrated Suite for Development and Maintenance of Arden Syntax Medical Logic Modules."

[26] D. R. Button and J. Fox, "The syntax and semantics of the PRO forma guideline modeling language," J. Am. Med. Informatics Assoc., vol. 10, no. 5, pp. 433–443, 2003.

[27] J. (John) Fox, S. K. Das, and American Association for Artificial Intelligence., Safe and sound : artificial intelligence in hazardous applications. AAAI Press/MIT Press, 2000.

[28] "Clinical Decision Support and Pathway Guidance | Elsevier." [Online]. Available: https://www.elsevier.com/solutions/arezzo/arezzo-users. [Accessed: 07-Jul-2019].

[29] "Tallis | COSSAC." [Online]. Available: https://cossac.org/tallis.html. [Accessed: 20-Jun-2019].

[30] L. Ohno-Machado et al., "The guideline interchange format: a model for representing guidelines.," J. Am. Med. Inform. Assoc., vol. 5, no. 4, pp. 357–72.

[31] "OpenClinical: GLIF." [Online]. Available: http://www.openclinical.org/gmm_glif.html. [Accessed: 06-May-2019].

[32] D. Wang et al., "Design and implementation of the GLIF3 guideline execution engine," J. Biomed. Inform., vol. 37, no. 5, pp. 305–318, Oct. 2004.

[33] J. Choi, L. M. Currie, D. Wang, and S. Bakken, "Encoding a clinical practice guideline using guideline interchange format: A case study of a depression screening and management guideline," Int. J. Med. Inform., vol. 76, pp. S302–S307, Oct. 2007.
[34] M. Sordo, O. Ogunyemi, A. A. Boxwala, and R. A. Greenes, "GELLO: an object-oriented query and expression language for clinical decision support.," AMIA ... Annu. Symp. proceedings. AMIA Symp., p. 1012, 2003.
[35] "HL7 Standards - Master Grid | HL7 International." [Online]. Available: http://www.hl7.org/implement/standards/product_matrix.cfm?Search=Gello. [Accessed: 05-Jul-2019].
[36] N. Mulyar, W. M. P. van der Aalst, and M. Peleg, "A Pattern-based Analysis of Clinical Computer-interpretable Guideline Modeling Languages," J. Am. Med. Informatics Assoc., vol. 14, no. 6, pp. 781–787, 2007.
[37] D. Wang and E. H. Shortliffe, "GLEE--a model-driven execution system for computer-based implementation of clinical practice guidelines.," Proceedings. AMIA Symp., pp. 855–9, 2002.
[38] A. Seyfang, A. Seyfang, R. Kosara, and S. Miksch, "Asbru's Reference Manual, Asbru Version 7.3," 2002.
[39] K. Hammermm and S. Miksch, "A Workkow Model for the Asgaard Project: A Time-Oriented, Skeletal Planning Workbench in Medicine."
[40] R. Kosara, S. Miksch, Y. Shahar, and P. Johnson, "AsbruView: Capturing Complex, Time-Oriented Plans — Beyond Flow Charts," in Diagrammatic Representation and Reasoning, London: Springer London, 2002, pp. 535–549.
[41] "AsbruView: Integrated Visualization of Computerized Protocols and Temporal Patient Data." [Online]. Available: http://www.openclinical.org/dld_asbruview.html. [Accessed: 05-Jul-2019].
[42] P. Votruba, "Document Exploration and Linking Tool (with Addons) (DELT/A)," 2006.
[43] P. Votruba, S. Miksch, A. Seyfang, and R. Kosara, "Tracing the formalization steps of textual guidelines.," Stud. Health Technol. Inform., vol. 101, pp. 172–6, 2004.
[44] W. Aigner and S. Miksch, "CareVis: Integrated visualization of computerized protocols and temporal patient data," Artif. Intell. Med., vol. 37, no. 3, pp. 203–218, Jul. 2006.
[45] "OpenClinical: DeGeL." [Online]. Available: http://www.openclinical.org/gmm_degel.html. [Accessed: 05-Jul-2019].
[46] Y. Shahar et al., "A Distributed, Collaborative, Structuring Model for a Clinical-Guideline Digital-Library," AMIA Annu. Symp. Proc., vol. 2003, p. 589, 2003.
[47] Y. Shahar et al., "A framework for a distributed, hybrid, multiple-ontology clinical-guideline library, and automated guideline-support tools," J. Biomed. Inform., vol. 37, no. 5, pp. 325–344, Oct. 2004.
[48] "Plan Execution." [Online]. Available: http://www.asgaard.tuwien.ac.at/exec_unit/index.html. [Accessed: 05-Jul-2019].
[49] C. Fuchsberger and S. Miksch, "Asbru's Execution Engine: Utilizing Guidelines for Artificial Ventilation of Newborn Infants."
[50] "OpenClinical downloads: Asbru Interpreter." [Online]. Available: http://www.openclinical.org/dld_asbruInterpreter.html. [Accessed: 05-Jul-2019].
[51] O. Young and Y. Shahar, "The Spock System: Developing a Runtime Application Engine for Hybrid-Asbru Guidelines," Springer, Berlin, Heidelberg, 2005, pp. 166–170.
[52] O. Young et al., "Runtime application of Hybrid-Asbru clinical guidelines," J. Biomed. Inform., vol. 40, no. 5, pp. 507–526, Oct. 2007.
[53] M. Peleg et al., "Comparing Guideline Models : A Case-study Approach," pp. 52–69, 2003.
[54] J. Fox, N. Johns, and A. Rahmanzadeh, "Disseminating medical knowledge: the PROforma approach," Artif. Intell. Med., vol. 14, pp. 157–181, 1998.
[55] M. Marcos, "Safe and Sound: Artificial Intelligence in Hazardous Applications: John Fox, Subrata Das, AAAI Press, Menlo Park, CA, and MIT Press, Cambridge, MA/London, UK, 2000, 326 pp., References, Index, Illus., $ 40.00 hardcover, ISBN 0-262-06211-9," Artif. Intell. Med., vol. 27, no. 1, pp. 103–106, Jan. 2003.
[56] "https://cossac.org/technologies/proforma.html." [Online]. Available: https://cossac.org/technologies/proforma.html.
[57] M. Peleg, "Computer-interpretable clinical guidelines : A methodological review," J. Biomed. Inform., vol. 46, no. 4, pp. 744–763, 2013.
[58] J. Fox, N. Johns, A. Rahmanzadeh, and R. Thomson, "PROforma: A method and language for specifying clinical guidelines and protocols," Stud. Health Technol. Inform., vol. 34, pp. 516–520, 1996.
[59] NCCN, "Breast Cancer Screening and Diagnosis," NCCN, vol. Version 3., 2018.
[60] J. Valacich, J. George, and J. Hoffer, "Essentials of Systems Analysis and Design." [Online]. Available: https://dl.acm.org/citation.cfm?id=2721609. [Accessed: 15-Feb-2019].
[61] K. E. Kendall and J. E. Kendall, Systems analysis and design. Pearson Prentice Hall, 2011.

Shaimaa Saud AL-Fleit received her B.S. degree in Management Information Systems from faculty of economics and administration at King Abdulaziz University in 2014 , respectively. During 2016-2019, she resumed her study to get a master's degree in computer information systems from faculty of computing and Information Technology , King Abdulaziz University, whereas this research is a part of her thesis in partial fulfilment of the requirements for the degree.

Abdullah Saad AL-Malaise AL-Ghamdi received his PhD from George Washington University of USA. He is an associate professor of IS and CS, Information Systems Department, Faculty of Computing and Information Technology, King Abdulaziz University. He has worked as a chairman of IS Department. Currently, he is the vice dean for Graduate Studies and Scientific Research. He has overall 20 years of experience in education field, mainly in teaching and curriculum building of several courses.

Mohammed Osama Nassif received his MBcH from King Abdulaziz University, Jeddah, KSA in 2004 . Then he involved in General Surgery Residency from McGill University in the city of Montreal, Canada in 2011, Followed by getting his Master of Science Msc from McGill University in Montreal, Canada in

2013 along with getting his Complex Surgical Oncology Fellowship from the Temple University Fox Chase Cancer Center in Philadelphia, USA in 2015. he worked as an Assistant Professor in the department of Surgery at KAU Hospital, Jeddah, KSA, from 2016 to 2017 . After he was working as a Fellow in the department of Surgery at Foxchase Cancer Center in Philadelphia, Pennsylvania, USA from 2013 to 2015 . Earlier he worked as a Consultant in the department of surgery at Moncton instituation in the city of Moncton , New Brunswick, Canada from 2011 to 2013 .