Ontology-based Modeling of Clinical Practice Guidelines: A Clinical Decision Support System for Breast Cancer Follow-up Interventions at Primary Care Settings

Samina R. Abidi, Syed SR. Abidi, Sajjad Hussain, Mike Shepherd

[a] *NICHE Research Group, Faculty of Computer Science, Dalhousie University, Halifax, Canada*

Abstract

Breast cancer follow-up care can be provided by family physicians after specialists complete the primary treatment. Cancer Care Nova Scotia has developed a breast cancer follow-up Clinical Practice Guideline (CPG) targeting family physicians. In this paper we present a project to computerize and deploy the said CPG in a Breast Cancer Follow-up Decision Support System (BCF-DSS) for use by family physicians in a primary care setting. We present a semantic web approach to model the CPG knowledge and employ a logic-based proof engine to execute the CPG in order to infer patient-specific recommendations. We present the three stages of the development of BCF-DSS—i.e. (a) Computerization of the paper-based CPG for Breast Cancer follow-up; (b) Development of three ontologies—i.e. the Breast Cancer Ontology, the CPG ontology based on the Guideline Element Model (GEM) and a Patient Ontology; and (c) Execution of the Breast Cancer follow-up CPG through a logic-based CPG execution engine.

Keywords: Clinical Practice Guidelines, Breast Cancer, Decision Support System, Medical Ontology, Semantic Web

Introduction

Clinical Practice Guidelines (CPG) entail medical knowledge intended for clinical decision-making and standardization of clinical practice [1, 2]. Despite the potential benefits of CPG, reviews show that CPG are underutilized in clinical practice [3, 4], largely due to problems associated with their dissemination to physicians [2, 5]. CPG computerization involves the modeling and conversion of a paper-based CPG into an electronic and executable format that can both be accessed by physicians and be embedded within clinical decision-support systems at the point of care. CPG guided decision support systems are particularly useful in clinical settings where non-specialist health practitioners, such as family physicians or nurses, are required to deal with complex or unusual cases. In such situations, CPG based decision support systems can guide the healthcare practitioner's actions and suggest appropriate recommendations. One such situation is the discharge of Breast Cancer (BC) follow-up care by family physicians. Note that in Nova Scotia follow-up care sis currently being provided by cancer care specialists at tertiary care centers.

Recent advancements in BC treatment have significantly improved the rate of BC survivors in Nova Scotia. The follow-up care for patients in remission entails periodic visits for history, physical exams and mammogram surveillance [6]. Although specialized cancer clinics provide long-term follow-up care, there is a case for formally involving family physicians in breast cancer follow-up care. In fact, trials conducted in Canada and Britain show that family physician offer a viable alternative to specialized care clinics for offering follow-up care to women who are in remission from breast cancer [6]. However, for most family physicians BC follow-up care is a new and added responsibility, therefore they need clear clinical guidelines to effectively perform the follow-up activities, make correct decisions and provide the right recommendations. The Canadian Steering Committee on CPG for the Care and Treatment of Breast Cancer has developed and recently updated the guideline on follow-up care after treatment for BC [7] with special emphasis on the needs of primary care physicians. The challenge was to disseminate the CPG to the family physician and to integrate it within his/her clinical workflow so that the CPG is seamlessly executed whenever a patient undergoes BC follow-up in a primary care setting.

In this project, we collaborated with Cancer Care Nova Scotia to address the abovementioned challenges by promoting the (knowledge) translation of the CPG for the Care and Treatment of BC to family physicians, to support them in the delivery of BC follow-up care and patient education at their clinics. This will reduce the workload of specialist cancer centers within Nova Scotia. Our approach was to develop an interactive Decision Support System (DSS) that enables family physicians to (a) access and utilize the said CPG at the point of care to provide standardized follow-up care; and (b) offer customized patient educational information targeting disease management, lifestyle behaviours and psychosocial support.

In this paper, we present an ontology-based Breast Cancer Follow-up Decision Support System (BCF-DSS) based on the CPG for the Care and Treatment of Breast Cancer. We take a semantic web approach to model the CPG knowledge and to reason over the ontology to provide 'trusted' CPG-driven recommendations. We have developed three ontologies: (a) *CPG ontology* that models the structure of the CPG based on the Guideline Element Model (GEM); (b) *Breast Cancer Ontology* that represents the medical knowledge encapsulated within the CPG and general BC related concepts; and (c*) Patient Ontology* that models the patient's parameters. The ontologies are developed using Protégé and are in OWL format. We have developed a logic-based reasoning engine that reasons over the knowledge from these three ontologies. Our BCF-DSS allows family physicians to collect patient data and assists them to make CPG mediated decisions, recommendations and referrals for BC survivors. We present the three stages of the development of BCF-DSS—i.e. (a) Computerization of the paper-based CPG for the Care and Treatment of Breast Cancer; (b) Development of the ontologies, in particular the Breast Cancer Ontology; and (c) Execution of the BC follow-up CPG through a logic-based CPG execution engine.

Computerization of BC Follow-up CPG

Computerization of the CPG involved (a) selection of a CPG modeling formalism; and (b) capturing and representing the CPG knowledge based on the modeling formalism. We selected the Guideline Representation Model (GEM) to model the BC follow-up CPG. GEM is based on XML that renders it operable in a semantic web environment by allowing semantically salient indexing and searching of CPG knowledge. We used the GEM Cutter tool to annotate the BC follow-up CPG with GEM tags (or elements). The conversion task involved determining the function of a specific CPG text and annotating it using the relevant GEM tag. It may be noted that GEM constitutes 100 tags covering a wide variety of concepts. For our purposes, the most salient concepts were the 'Knowledge Components' that store and categorize the knowledge present in a CPG. The knowledge components have sub-components called recommendations which are categorized as either *imperative*—i.e. directed towards entire target population, or *conditional*—i.e. act on the decision variables and results in appropriate actions.

From a knowledge modeling perspective the main challenge was to resolve the (medical and semantic) ambiguities inherent within the BC follow-up CPG. To resolve the ambiguities we (a) consulted with BC oncologists, in particular the author of the BC follow-up CPG; (b) reviewed available literature; and (c) applied our personal clinical experience. Examples of ambiguities included phrases such as, 'vaginal bleeding is present in the absence of obvious cause', 'physiological causes of fatigue' and 'other risk factors of osteoporosis'. The phrase 'vaginal bleeding in the absence of obvious cause' was resolved to the term 'Menstruation', and other ambiguous phrases were similarly resolved by mapping them to explicit concepts. Finally, through the use of GEM we managed to create an executable representation for the BC follow-up CPG.

Development of Breast Cancer Ontology

The BC ontology models the knowledge encapsulated within the BC follow-up CPG. We used Protégé ontology editing environment to build our BC ontology in OWL (Web Ontology Language) using Protégé OWL.

The BC ontology is largely derived from the contents of the knowledge components—i.e. the 'Imperative' or Conditional recommendations—in the GEM representation of the BC follow-up CPG. More specifically, the conditional recommendation element, which comprises sub-elements such as 'decision.variable', 'action' and 'logic' elements, was used to develop the BC ontology. Given conditional recommendations, the challenge was to identify the decision variables, the actions to be taken and the Boolean logical operations in the recommendations, so that the resultant ontology was compatible with our logical reasoning engine. In this regard, two design constraints were addressed: (a) Our CPG execution engine does not processes statements containing 'OR' and 'NOT' logical operators. Therefore a rule such as:

"**IF** *age* >65 **OR** *family history of* osteoporosis **OR** *menstrual status of* premature menopause due to treatment,
THEN screen with bone mineral density and treat accordingly with bisphosphonates",

was required to be decomposed into three smaller rules, each with a single decision variable so that the OR operator was eliminated; and (b) The BC ontology classes that have multiple domains or ranges could not be executed safely. Therefore, we ensured that all properties have a single domain and range.

Specifying BC Ontology Classes

Considering the above constraints we defined eight main classes, namely; Patient_Type, Physician_Type, Illness, Menstrual_Status, Recommendation, Symptom, Diagnostic_Tests, Treatment, Age, Risk_Factor, Weight_Status and Patient_Wish Next, we specified the disjoint classes, where classes are disjoint when an individual cannot be an instance of more than one of these classes. In the BC ontology the only classes which are not disjoint are the Recommendation and Diagnostic Tests, and Recommendation and Treatment since they share some instances, for example "Screening with bone mineral analysis" is an instance of two classes i.e. Diagnostic_Test and Recommendation; and "Bisphosphonates" belong to class Treatment and Recommendation.

Specifying Properties for the BC Ontology Classes

Properties for Patient Class

The '*Patient_Type'* is the most important class because most conditional recommendations are targeted towards the patient. To specify different patient types we defined a range of properties. The patient properties represent the patient profile—i.e. an instance of the class *Patient_Type*. We defined object properties to establish link between the classes so that recommendations can be associated with a patient profile. The class *Patient_Type* has the most object properties with their domain being *Patient_Type* but their range includes instances from other classes in the BC ontology, for example, the properties 'has*_history_of*' and *'has_illness'* have individuals of the class Illness as their range.

In our CPG execution engine most of the properties are treated as *decision variables* that serve as the premise of a logical rule. The conclusion of the rule is an *action variable* that corresponds to a recommendation, treatment, or statement directed towards a patient. To account for the patient-centric action variables, we specified two *Patient_Type* properties—i.e. *is_Recommended* and *possible_cause_can_be* with an unspecified range as their range can be any individual from any class. The action variable *is_Recommended* refers to the any recommendation, diagnostic test or treatment suggested to a patient. The action variable *possible_cause_can_be* provides the physician information regarding the cause of certain sign or symptom, for example the CPG statement "emotional distress, may be the underlying cause of subjective complaints of impaired cognitive functioning."

Properties For Other Classes

The properties for the other classes were derived from conditional statements in the recommendation element of the GEM representation of the BC follow-up CPG. These conditional statements specify variables that non-patient specific data, for example consider the CPG statement "If the purpose is to detect distant metastasis, then routine lab and radiographic exam should not be carried out". Such as statement was modeled by the class 'Diagnostic_Test' through two properties; *has_purpose_to_detect* whose range is Illness and *test_apply_to* whose no specified range. In total, we specified 40 properties for all other classes excluding the patient class.

Properties For Statements Having the Not Logical Operator

Modeling of some statements in the BC follow-up CPG required the use of the 'NOT' operator. For instance consider the statement, "When such bleeding (vaginal bleeding) is present in the absence of obvious cause, endometrial biopsy should be carried out". Here the phrase 'in the absence of obvious cause', really means *not* obvious cause. We handled such situations by specifying a new property, for instance *is_not_caused_by*, for the class Symptom. Note that the rationale for creating such as property is because our execution engine is unable to handle the 'NOT' logical operator.

Specifying Property Characteristics

Certain properties such as *has_age*, *has_weight_status* and *has_menstural_status* are functional properties since a patient can have only one age, weight status (i.e. can either be overweight, under-weight or have correct weight) and menstrual status (i.e. can either be premenopausal, postmenopausal or premature menopause due to treatment). Most of the properties are not functional and allow multiple values. We also specified some inverse object properties such as 'is*_recommended_for_illness*' which is the property of class 'Treatment' is the inverse property of '*is_treated_by*' which is the property of class 'Illness'. This means that 'Bisphosphonates' *is_recommended_for_illness* 'Osteoporosis' and 'Osteoporosis' *is_treated_by* Bisphosphonate. Note that Osteoporosis is the individual of the class Illness and Bisphosphonate is the individual of the class Treatment.

Specifying Individuals (Instances)

In the next step of BC ontology development we specified the individuals (or instances) from the conditional recommendations of the BC follow-up CPG. For example, individuals for class Symptom include Anxiety, Back_Pain, Cognitive_Impairment, Fatigue, Impaired_sexual_function, Menopausal_Symptoms, Vaginal_Bleeding, and Vaginal_Dryness. The class Patient_Type has the most individuals since each recommendation is valid for a patient with a particular set of clinical characteristics, thus each patient type refer to a particular patient profile in accordance to the said CPG.

Specifying Relationships Between the Classes

The relationships among diferent classes were modeled using the class properties, and can be best understood by the following example. In order to model the recommendation statement "When such bleeding (vaginal bleeding) is present in the absence of obvious cause endometrial biopsy should be carried out". Here, "obvious cause of bleeding" means Menstruation. We established a relationship between the classes *Patient Type*, *Symptom*, *Menstural_Status* and *Diagnostic_Test* as follows: Patient_Type_12 is an individual of class Patient Type who has Vaginal_Bleeding which is the value for its object type property *has symptom*. Vaginal_Bleeding is also an individual of class Symptom, which has an object type property called *is_not_caused_by*, whose value is Mensturation_or_Obvious_cause, which in turn is an individual of another class called Mentrual_Status. The class Menstrual_Status has an object property called *ms_apply_to_diagnostic_test* whose values in this case will be Endometrial_Biopsy which is an instance to the class Diagnostic_Test (See Figure 1). By adding the property *is-not-caused_by* we ensured that the recommendation is logically valid i.e. endometrial biopsy is not recommended whenever the patient has vaginal bleeding. In this way we are able to establish an inter-class relationship that can be used to infer that if a patient has vaginal bleeding and bleeding is not caused by menstruation or any other obvious cause, then endometrial biopsy is the recommended test.

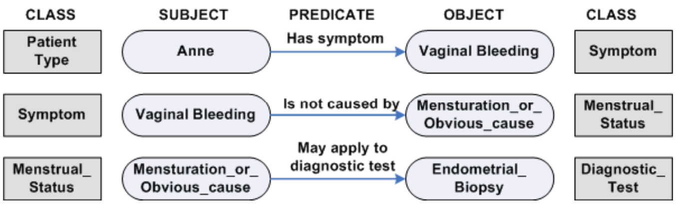

Figure 1 - RDF Triples depicting the relationships between classes Patient_Type, Symptom, Menstrual_Status, and Diagnostic_Test to model a recommendation.

CPG Execution Engine

In keeping with our Semantic Web approach we developed a CPG Execution Engine (CPG-EE) that constitutes (a) multiple ontologies to model the domain and CPG knowledge; (b) a logic-based proof engine that leverages the ontologies and CPG specific rules to infer CPG mediated recommendations; and (c) a justification trace to describe the rationale for the inferred recommendations; this is to establish 'trust' in the proposed recommendations (Figure 2). The CPG-EE provides the functionality to define CPG-specific decision logic rules based on the decision variables in the CPG and to execute the rules based on patient clinical data to provide CPG based recommendations. The CPG-EE comprises two main modules: (i) Rule Authoring Module and (ii) Rule Execution Module.

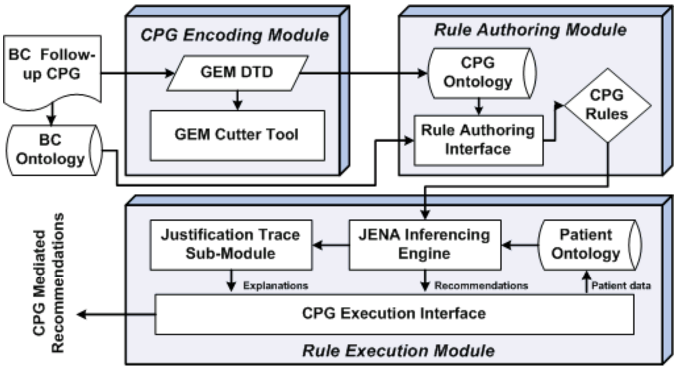

Figure 2 – System diagram of the CPG-EE

Rule Authoring Module

The Rule Authoring Module provides users an interface to specify decision logic rules, using a native CPG rule syntax, based on the decision logic inherent within the CPG. The rule authoring process is guided by the knowledge, relationships and constraints represented within the CPG ontology and the domain ontology (in our case the BC ontology). In this way, rule authoring is constrained by pre-defined knowledge and hence ensures the sanity of the decision rules.

The CPG ontology is designed to semantically model the structure of a CPG in order to annotate the decision variables and logic structures inherent within the CPG. Our CPG ontology is based on the GEM DTD which provides a characterization of different elements of a CPG. In particular we utilized the knowledge component of the GEM DTD and mapped it to a *Recommendation* class that entails the procedural, conditional or imperative knowledge of the CPG. The decision and action variables that constitute the premises and conclusions of a CPG rule, respectively, are explicitly stated in the CPG ontology, and these variables are utilized in authoring CPG rules. In the CPG ontology, the decision variables are represented as a sub-class. For execution purposes we added a new property variable.name to the decision.variable, such that its value is derived from all properties in the Domain Ontology.

Rule Authoring is performed by defining decision rules in the logic tag of CPG ontology as follows: *Step 1:* Select decision variables from the Decision Variable List, which represents the body (premises) of the rule; *Step 2:* Select the action variable from the Decision Variable List, which represents the head (conclusion) of the rule; *Step 3:* For each decision and action variable in the rule, an equality/inequality relation is defined with either a variable, a value, a binary algebraic formula, another decision variable or list of decision variables. We give a rule authoring example, where we assign the variable names i.e. properties to *decision variables* (coded as dv and each with a unique #) as well as action variables (coded as *av* and each with a unique #). In case of a rule

IF dv1 i.e. *Patient_is_on_medication* **=** Tamoxifen (property of class Patient_Type) **AND dv2** i.e. *Rx_apply_to_recommendation* **=** ? (property of class Rx_Recommended)
THEN a1 i.e. *Patient_is_recommended.* (property of class Patient_Type) **= dv5**

The derivation for this rule is as follows. The Patient_Type_1 which is an instance of the class Patient_Type is *on medication*, Tamoxifen. Patient_Type_1 is the resource for this rule. The treatment i.e. Tamoxifen is an instance of class Treatment, which has a property *apply to recommendation*, whose value is ‘query about vaginal bleeding’. Since we have specified in the rule that the value for *a1* (*Patient_is_recommended*) is same as the value for *dv2* (*Rx_apply_to_recommendation*), which according to the BC ontology is ‘query about vaginal bleeding’, the recommendation for this patient type is to query about vaginal bleeding.

Rule Execution Module

The Rule Execution Module executes the CPG rules based on a patient instance to infer CPG based recommendations. Rule execution is performed by a logic-based inference engine—i.e. JENA. The processing of this module is as follows: (i) The CPG rules are transformed from their native syntax to JENA rule syntax; (ii) The patient data is acquired through the CPG execution interface (see Figure 4) to form an instance of a patient, based on the Patient Ontology, that incorporates patient properties such as *age*, *gender*, *medical history*, etc. The values of the patient properties serve as input to the execution engine; (iii) The JENA inference engine uses the CPG rules and the patient instance to build an inference model using backward chain reasoning. The outcome is a set of inferred recommendations based on the patient data; (iv) A justification trace of the inferred recommendations is generated to explain the reasons for the proposed recommendations.

BCF-DSS in Action

We present an example to demonstrate BCF-DSS in action. The clinical case is: *“A BC patient who is overweight, complains of fatigue and is experiencing vaginal bleeding in the absence of obvious cause. She has a family history of osteoporosis and is on Aromatase inhibitors. She also wishes to get pregnant and wants to know whether this is a viable option”.*

The family physician records the patient’s properties using the BCC-DSS user interface (shown in Figure 3). The physician presses the *Recommendation* button and is provided five recommendations (shown in the bottom left box). The physician can seek an explanation for any recommendation by highlighting it and pressing the *Explanation* button. Figure 4 shows the explanation interface that includes the CPG description (upper left box) for the recommendation, the reasons for the proposed recommendation (upper right box) and the related references (lower middle box); all explanation material is derived from the annotated BC CPG. Finally, the justification trace (see Figure 5) for the inferred recommendations is as follows:

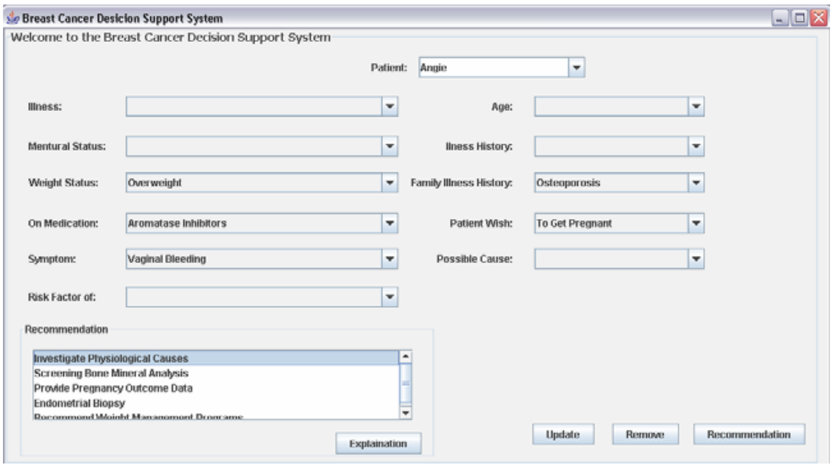

Figure 3 – The CPG execution interface for BCS-DSS, used to collect the patient data and to give recommendations

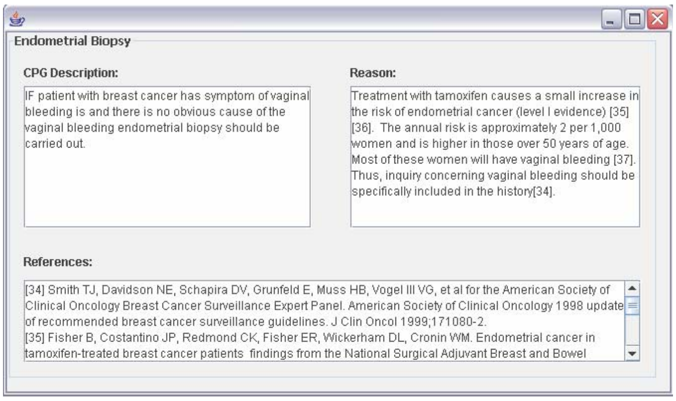

Figure 4 – The explanation interface of the BCF-DSS

Angie → has weight status = Overweight
Overweight → associated with recomm = Recommend Weight Management Programs
Angie → is Recommended = *Recommend Weight Management Programs*

Angie → has symptom = Vaginal Bleeding
Vaginal Bleeding → is not cause by = Mensturation or Obvious Cause
Mensturation or Obvious Cause → ms apply to diagnostic test = Endometrial Biopsy
Angie → is Recommended = *Endometrial Biopsy*

Angie → has family history of Illness = Osteoporosis
Osteoporosis → illness apply to recomm = Treatment Should Include Bisphosphonates
Angie → is Recommended = *Treatment Should Include Bisphosphonates*

Angie → has wish = To Get Pregnant
To Get Pregnant → associated with recommendation = Provide Pregnancy Outcome Data
Angie → is Recommended = *Provide Pregnancy Outcome Data*

Angie → is on medication = Aromatase Inhibitors
Aromatase Inhibitors → Rx apply to Diagnostic Test = Screening Bone Mineral Analysis
Angie → is Recommended = *Screening Bone Mineral Analysis*

Angie → has symptom = Fatigue
Fatigue → apply to recommendation = Investigate Physiological Causes
Angie → is Recommended = *Investigate Physiological Causes*

Figure 5 – Justification trace for the recommendations

Concluding Remarks

We have developed a CPG based interactive clinical decision support system for the BC follow-up to be used in the primary care setting. Our approach is innovative since we have linked the CPG ontology to the breast cancer domain ontology from which rules were derived. This aproach can also be applied to CPG in other medical specialties. The objective of this project is to promote knowledge translation to primary care settings in Nova Scotia so that family physicians can take on the responsibility for the BC follow-up care, thereby reducing the strain on specialist cancer centers within Nova Scotia. The project also aims to create an interactive environment for family physicians to facilitate customized patient management and educational information for an individual patient.

Acknowledgement

This project is funded by Nova Scotia Health Research Foundation (Canada), through an innovation research grant. Special thanks to Dr. Grunfeld for the BC-CPG and her support.

References:

[1] Pryor DB, Fortin DF. "Managing the delivery of health care: care-plans/managed care/practice guidelines". *Int J Biomed Comput*. 1995;39:105–9.

[2] Zielstorff RD, "Online Practice Guidelines, Issues, Obstacles and Future Practice", *JAMIA*. 1998;5:227–236.

[3] Lomas J, Anderson GM, Domnick-Pierre K, "Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians". *N Engl J Med*. 1989; 321:1306–1311.

[4] Greco PJ, Eisenberg JM. "Changing physicians' practices". *N Engl J Med*. 1993;329:1271–1273.

[5] Sanders GD, Nease RF, Jr., Owens DK, Design and Pilot Evaluation of a System to Develop Computer-based Site-specific Practice Guidelines from Decision Models; *Med Decis Making* 2000;20:145–159.

[6] Grunfeld E, Levine M, Julian J, Prtichard K, Coyle D, Mirsky D, Verma S, Dent S, Sawka C, Ginsburg D. "A randomized controlled trial of routine follow-up for early stage breast cancer: a comparison of primary care versus specialist care". *J Clin Oncol* 2004; 22:43

[7] Grunfeld E, Dhesy-Thind S, Levine M, "Guidelines for the Care and Treatment of Breast Cancer: Follow-up after treatment for breast cancer (2005 update)", *Can. Med. Assoc. J.*, May 2005; 172: 1319 - 1320

Address for correspondence

Syed Sibte Raza Abidi, Faculty of Computer Science, Dalhousie University, Halifax, Canada. Email: sraza@cs.dal.ca