



**IST-FP6-508794
PROTOCURE II**

*Integrating formal methods in the development process of
medical guidelines and protocols*

Specific Targeted Research Project
Information Society Technologies

**D5.2 How to use model-checking for critiquing
using medical guidelines**

Due date of deliverable: 31 May 2006
Actual submission date: 31 July 2006

Start date of project: 1 January 2004 **Duration:** 30 months

Organisation name of lead contractor: Universitat Jaume I

Revision 1

Project co-funded by the European Commissions within the Sixth Framework Programme(2002-2006)		
Dissemination Level		
PU	Public	√
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Protocure II, Workpackage 5

Deliverable 5.2

**How to use model-checking for critiquing using
medical guidelines**

Radu Serban, Annette ten Teije, Frank van Harmelen
Vrije Universiteit, (VUA), Amsterdam

Arjen Hommersom, Peter Lucas, Perry Groot
Radboud Universiteit (RU), Nijmegen

Albert Jovell, Sergi Blancafert
Fundacio Biblioteca Josep Laporte (FBJL), Barcelona

Jolanda Wittenberg, Kitty Rosenbrand, Joyce van Croonenborg
Dutch Institute for Healthcare Improvement

July 22, 2006

Contributions:

VUA (Lead participant):

- Responsible for the general structure of this deliverable, sections 1, 4, 5.1, 5.3, 6 and the appendix section B and C.
- Responsible for analyzing the patient data and providing mappings to ASBRU.

RU:

- Responsible for sections 2 and 3 and 5.2.
- Responsible for analysis of additional breast cancer case studies.

FBJL:

- Responsible for supporting the selection and medical interpretation/validation in the appendix section A, B and C.

CBO:

- Responsible for providing patient data and medical background knowledge.

Abstract

This deliverable describes the methodology for critiquing the reference guideline using model checking based on an SMV model generated from the ASBRU guideline model.

From a technical point of view, our work has shown that it is indeed possible in principle to use model-checking on formalized models in order to critique medical guidelines against patient-records. The strong aspect of this technology is the high degree of automation as compared to theorem proving, making it more accessible and easier to use.

The main difficulty encountered in critiquing the patient data using the guideline model is the different level of abstraction of the vocabularies which describe the patient in the clinical records and in guidelines. For that reason, it is difficult to have a valid critique of a guideline based on data interpretation.

Contents

1	Introduction and main findings	3
1.1	Introduction	3
1.1.1	Data	4
1.1.2	Experiments	5
1.2	Main outcome of the experiments	6
1.3	Overall Conclusions	6
1.4	Structure of this deliverable	7
2	Relevant Literature	8
2.1	Critiquing	9
2.2	Related Work	10
2.2.1	The ATTENDING Systems	10
2.2.2	ONCOCIN	10
2.2.3	HyperCritic	11
2.2.4	Critiquing using Medical Guidelines	11
2.3	Concluding Remarks	12
3	Model checking for Critiquing	13
3.1	Introduction	13
3.2	Comparison with other techniques	13
3.3	Specification and Methods	13
3.3.1	Temporal Logic	13
3.3.2	LTL and Modular Model Checking	14
3.3.3	Tools	15
3.4	Applying Model Checking for Critiquing	15
3.4.1	Method	15
3.4.2	Specification of model and traces	15
4	Approach to critiquing the guideline	18
4.1	Critiquing of patient data using Model Checking	18
4.1.1	Critiquing objectives	18
4.1.2	Mapping patient data to ASBRU	19
4.1.3	Validate the mappings by medical and technical interpretation	19
4.1.4	Reconstructing possible treatment paths	19
4.1.5	Refine the critiquing traces	20
4.1.6	Verify and interpret the critiquing properties	20

4.2	Critiquing breast cancer case studies from literature	21
5	Implementation and Results	22
5.1	Critiquing Patient Data	22
5.1.1	Patient Data	22
5.1.2	Profiles of prototypical patients	23
5.1.3	Case study 1: Patient 4	24
5.1.4	Case study 2: Patient 21	28
5.1.5	Case study 3: Patient 22	29
5.1.6	Case study 4: Patient 34	31
5.1.7	Concluding remarks	32
5.2	Critiquing Realworld Case Studies	32
5.2.1	Informal description	32
5.2.2	Attachment to the model	33
5.2.3	Translation to the SMV model	34
5.2.4	Results	34
5.2.5	Concluding Remarks	38
5.3	Analysis of the results	38
5.3.1	Analyzing non-compliant treatment traces	39
5.3.2	Suggesting treatment improvements based on Model Checking	39
5.3.3	Refinement of the critiquing property	39
6	Concluding Remarks	40
A	Appendix A - Sample Patient Traces	41
B	Mapping of Patient Data to the Asbru Model	45
C	Appendix C: Detailed Analysis of the Selected Patients	48

Chapter 1

Introduction and main findings

1.1 Introduction

The general goal of critiquing is to spot differences between a set of “ideal” actions and the actual actions performed on a specific patient. Among others, guidelines are one source to obtain such a set of ideal actions, to be used as the basis for critiquing. Another source of such a set of ideal actions would be a planning system that generates a treatment plan against which the actual treatment of the patient can be compared.

One way of realizing critiquing is to rely on formal methods. In particular model-checking is suitable for this purpose. The approach is then to describe guideline as a system description in a model-checking system, to describe the actions that have been performed on a specific patient as a temporal formula, and then check the consistency between the system description (= guideline) and the formula (= patient record). Model-checking is a well investigated approach for other kinds of systems, such as computer hardware, transportation systems, software etc. Until now, model-checking has been mainly applied to technical systems. The attempt to apply model-checking to medical knowledge (guidelines and patient data) is the innovative part of this work. This attempt is inspired by the general metaphor that drives the Protocure project, namely the analogy between a guideline and a software program (both being sets of instructions on how to act in certain circumstances). A more detailed description of this approach of using model-checking for guideline critiquing can be found in Chapter 3, in particular Section 3.4.

The goal of this deliverable is therefore:

Goal: to investigate the feasibility of off-the-shelf model-checking technology as a method for critiquing the breast cancer guideline against a realistic data-set.

The main requirement that must be fulfilled for this work is of course a sufficiently high degree of correspondence between the parameters and actions occurring in the guideline (or more precisely: in the formal model that we have constructed of the guideline), and the parameters and actions that have been recorded for particular patients.

We will therefore now turn to a description of the data-sets that we had available in our experiments, before describing the steps that we took in our experiments, leading to a brief summary of the main findings of this report.

1.1.1 Data

As in deliverable D3.2b [28], we used data of the Dutch Comprehensive Cancer Centers. In order to make the current report self-contained, we briefly repeat the details of this dataset from deliverable D3.2:

The Comprehensive Cancer Centers (CCC's) in The Netherlands maintain a population based cancer registry. This registry is an important source of information for cancer research, the planning of services, and the evaluation and implementation of guidelines. In the Netherlands there are 9 such CCC's. The data for the Protocure II project originates from the Comprehensive Cancer Center South. This centre was chosen because they collect more specific data of breast cancer patients than the other centres (in particular follow-up data).

All data concerning breast cancer patients for the period January 2003 - June 2004 was obtained for the Protocure II project. This resulted in 5507 patient records with 269 variables per patients. The variables include information about patient characteristics (e.g., age, date of diagnosis, etc), cancer characteristics (e.g., type of tumor, location of tumor, etc) and treatment information (e.g., type of surgery, radiotherapy, chemotherapy, etc).

As already noted in Deliverable D3.2, the real-life clinical data from the CCC data-set fails to match with the guideline. This mismatch already happens with the original text of the guidelines, and consequently also with the other models that we constructed based on the original guideline text: the MHB model, the Asbru model, and the formal model.

This observation points to the fact that the mismatch between patient data and guideline vocabulary is actually a problem caused by the clinical situations in the hospitals: data-collection is simply not in line with guideline terminology. Hence, the mismatch is not an artifact introduced by the modelling procedures adapted in the Protocure project. The mismatch is already present between data-set and the text of the original guideline; it was not introduced in the transformations from text to semi-formal and formal models.

Obviously, this mismatch formed a significant obstacle in our experiments to apply model-checking for critiquing, since a close correspondence is required between the terms in the "system description" (i.e., the guideline model) and the terms in the temporal formula to be checked (i.e., the patient record).

We now describe the approach we took to still enable our experiments in the face of this problem. In one approach we constructed a derived data-set from the CCC data with a sufficiently close match to the guideline terminology. In the second approach, we located an entirely different data-set.

Producing a derived data-set. The domain experts of FB JL constructed 7 prototypical patient-cases. These were inspired by the data from the Dutch CCC set, but were altered to represent "idealized" typical cases. Together, these 7 are most frequently occurring cases, and cover all the chapters of the guideline.

These 7 prototypical cases were expressed in the "TNM staging vocabulary" in use at FB JL to describe the extension of the tumour: T (local level of extension), N (lymph nodes), and M (distant metastasis).

The use of this TNM vocabulary meant that the 7 prototypical cases still used a different vocabulary from the terms deployed in the various guideline models. Hence, a manual mapping step was still required in order to map the TNM-vocabulary to the vocabulary used in the guideline and its models

This resulted in 7 prototypical cases that could be interpreted in terms of the formal and

semi-formal guideline models, and that could hence be used as the basis for critiquing. The results of this are reported in Section 5.1.

Using another data-set. In a second approach, we abandoned the Dutch case-data altogether, and obtained new case-studies from the publicly available Breast Cancer Case Index¹, which contains a small collection of relevant patient cases. In particular we used the following two cases:

1. A ductal carcinoma in situ in a male patient²
2. An infiltrating duct carcinoma with use of lymphoscintigraphy³

Our justification for selecting this alternative source of case-data is that the lack of background knowledge is the main reason why mapping of the CCC data to the guideline is so hard to do, while the BCI cases come with extensive background information, enabling us to make the required mapping of the BCI vocabulary to the terms used in the guideline. This manual mapping from the BCI-data to the vocabulary of the CBO guideline could hence be done with sufficient confidence. The results of the model-checking effort based on this data-set is described in Section 5.2.

1.1.2 Experiments

Our experiment on trying to apply off-the-shelf model-checking technology for guideline critiquing consisted of the following steps:

1. Take the KIV model of the CBO breast cancer guideline. This model was extensively described in Deliverable 2.2bcd ([20]).
2. Automatically compile this KIV model into the system-description required as input for the model-checker we deploy. In our case, this is the SMV model checker [21]. This compilation procedure is entirely automated, and is described in Deliverable 4.3 [11].
3. Manually translate the patient-data into a temporal formula. This translation is described in Section 3.4.2.
4. Check consistency between data and guideline using the Cadence SMV model-checker (again, a fully automated procedure).
5. If the model-checker detected an incompatibility between guideline and patient-data (formally: by constructing a counterexample), we explored two avenues:
 - (a) We tried to find the longest initial trace of the guideline that was still consistent with data. This might lead to further insight into the cause of the inconsistency between patient-data and guideline model. Also, we checked whether each of the actions occurring in the patient-record after the failure point could be individually explained in the light of the available data about this patient.
 - (b) In an alternative approach, we relaxed the ordering constraints that lead to the detected inconsistency and then reduced the set of actions performed until it became consistent with the model. Again, this would lead to further insight into the nature of the detected incompatibility between guideline and patient data.

¹ <http://www.bci.org.au/medical/caseindex.htm>

² <http://www.bci.org.au/cases/96-04/96-04p1.htm> and further pages

³ <http://www.bci.org.au/cases/99-03/p1.htm> and further pages

6. Ideally, in the final step, the detected inconsistency should be presented to a domain expert for comment and evaluation. This step was not performed, because of the artificial nature of the data (see the discussion above), and because we already spotted that some of the inconsistencies were artifacts caused by errors where the formal model deviated from the original guideline text.

1.2 Main outcome of the experiments

We give a brief summary of our results. A more detailed description is given in chapter 5.

In the experiment on the 7 prototypical patient-cases derived from the Dutch CCC dataset, some deviation was found between the guideline and each of the 7 prototypical cases. Interestingly, for 3 of these, some differences could indeed be explained by looking at the new 2004 revision of the guideline (see Deliverable 2.2bcd additional report [29] for a detailed description of this revision). For example, some patient-trace (Case-study 2, patient 21) contained a chemotherapy action not mentioned in the appropriate chapter of the guideline, but in the 2004 revision, Chapter 1 of the guideline was extended to include exactly this action. However, but due to the extensive re-interpretation of data we had to do (as described above), it is impossible to draw any hard conclusions from such an observation.

In the experiment on the cases from the BCI-data, the first case could be fully explained in terms of the guideline model. The second case was found to be inconsistent with the formal model of the guideline. This remained so even after removing all the ordering constraints (as described in step 5b above. The inconsistency could be traced to the occurrence of a combination of 2 specific actions in the patient-case. Manual inspection revealed that this combination is actually allowed by the guideline, but was not allowed by our guideline model. In other words, in this case model-checking revealed a modelling anomaly.

1.3 Overall Conclusions

The overall conclusion from this work is twofold:

Conclusion on technology. This work has shown that it is indeed possible in principle to use model-checking on formalized models in order to critique medical guidelines against patient-records. The strong aspect of this technology is the high degree of automation as compared to theorem proving, making it correspondingly easier to use.

Model-checking also provides additional value to critiquing via an operational version of the guideline (by running it through an interpreter, as described in deliverable D3.2 [28]). Such interpreter-based critiquing only checks the consistency of a patient record against a single trace through the guideline (namely the one chosen by the interpreter), while model checking compares the patient record against all possible traces through the guideline. This difference is of course crucial when the guideline contains non-deterministic choices (as is the case in all guidelines we have in our work to date).

The fully automated nature of model-checking also brings a corresponding weakness: model-checking only *detects* inconsistencies, but does not contribute to the interpretation of the inconsistency. In general, model-checking can construct a counter-example illustrating the inconsistency, which is often a very good guide towards tracing its source. However, this only works when model-checking universally quantified formulas, while in Chapter 3 we argue

that the required formula is existentially quantified, thereby making it impossible for the model-checker to generate a counter-example. In this deliverable we propose some general strategies to deal with this (repeated experiments with weaker models, changing assumptions, relax ordering constraints, find longest initial traces etc).

Conclusion on data-collection As pointed out above, the mismatch between patient-data and guideline is already apparent in the original textual version of the guideline, and is not an artifact of our formalizing transformations of this guideline. This mismatch is all the more surprising since the patients in the CCC data-set were treated under the CBO guideline. It was therefore reasonable to expect a high level of correspondence, particularly since “evaluation and implementation of guidelines” is listed as one of the reasons for which the CCC data is collected. But in fact the terminology of the data-set did not align in an obvious way with terminology used in the guideline. At the very least non-trivial medical knowledge is needed to interpret the data in the light of the guideline.

A general conclusion that can be drawn is that a closer correspondence is needed between the processes of guideline construction and data-collection. In fact, this is currently already being partially implemented by CBO: newly constructed guidelines are currently being equipped with a data-collection dictionary, which will ensure the correspondence between collected data and guideline terminology.

1.4 Structure of this deliverable

This document is structured as follows: in Chapter 2 we discuss the definition of critiquing of medical guidelines, and give a brief survey of the most important approaches to this problem in the literature. In Chapter 3 we introduce the basics of model checking, and how to apply this general method to our problem of critiquing medical guidelines. Chapter 4 outlines the approach we have taken in our experiments, and Chapter 5 discusses the results obtained using these methods and the limitations of our approach. Chapter 6 summarizes our conclusions.

Chapter 2

Relevant Literature

The trend of the last decades has been to base clinical decision making more and more on sound scientific evidence, which has been called *evidence-based* medicine [32, 35]. In practice this has led medical specialists in a particular area to develop medical guidelines, i.e., structured documents providing detailed steps to be taken by health-care professionals in managing the disease in a patient, for promoting standards of medical care. It has been shown that medical guidelines can improve health-care outcomes [34] and may even reduce the costs of care up to 25% [9]. However, medical guidelines can be extensive documents (e.g., the Dutch CBO breast cancer guideline is 121 pages in A4 format), and, therefore, it is not easy for physicians to locate relevant information in a short amount of time.

Researchers in Artificial Intelligence have picked up on these developments and are working on providing computer-based support for the development and deployment of guidelines by means of computer-oriented languages and tools. Computer-based decision support in health-care already is a field with a long standing tradition, e.g., [10, 13, 33, 18, 5, 7], dealing with complex problems in medicine such as diagnosing disease and assisting in the prescription of appropriate treatment. In [27], Rymon notes several ways in which computerized decision support could help physicians with their *patient management*:

- Supplementing physicians' expertise where it is lacking
- Supporting standardization of care
- Providing on-line Quality Assurance, by monitoring deviations from standards of care.

In this document we focus on the latter in the context of medical guidelines. One could build a decision support system that, according to a medical guideline, in response to any new information regarding the patient or management procedures, outputs a listing of recommended management plans. These recommended management plans can be taken as orders or more as guides for subsequent actions to be taken by the physicians. However, taking the recommended plans as orders does not take into account that physicians are autonomous and well capable of reasoning on their own on how to manage their patients. Such a system most likely will not be accepted within a clinical setting. Furthermore, besides being unacceptable, such a system would be too inefficient if the physician would need to refer every time to the guideline recommendations.

More promising would be to use the recommended treatment plans of the system to be used as a guide for the physician. The physician formulates his own treatment plans,

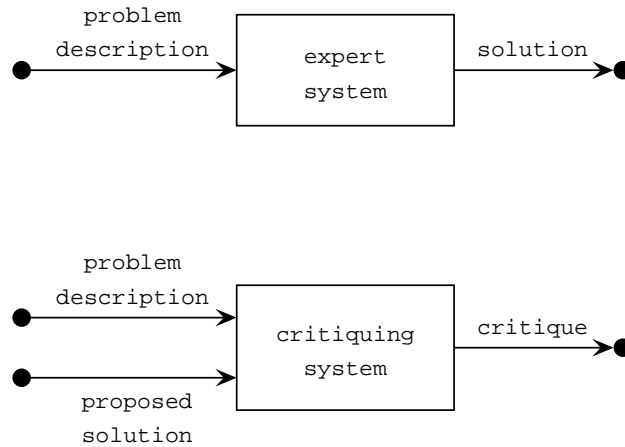


Figure 2.1: Differences between expert systems and critiquing systems

which are compared to the recommended treatment plans of the system. Only when the system detects problems within the treatment plan of the physician it will present its own recommended treatment plans and starts a dialog with the physician giving explanations and possible alternative courses of action in the context of the physician’s intended actions.

2.1 Critiquing

Within this work we will use the term *critiquing* as it was first coined by Miller in his pioneering work [23, 24] to refer to a computer program that critiques human-generated solutions. The common feature of a critiquing system is that the user of the system provides as input (1) a problem description (e.g., patient symptoms), and (2) a proposed solution (e.g., a treatment plan). This second input is what distinguishes critiquing systems from the more traditional expert systems, which only take a problem description as input [31, 12]. The second input to a critiquing system, i.e., a proposed solution, is typically the output of an expert system. In critiquing systems the output of the system is criticism on the solution provided by the user. The difference between critiquing systems and expert systems is illustrated in Figure 2.1.

Critiquing fits well with providing decision support to physicians because of the inherent variability in solutions to medical problems and the subjectivity of problem solvers involved. In medicine, often there is no one correct way for the management of patients and procedures taken may well depend on personal preferences of the physician. Traditional expert systems are therefore not well equipped to handle such situations. Firstly, solutions of expert systems, although possibly good ones, may not be the best solutions because of unquantifiable data about the patient, treatment procedures, or physician preferences. Secondly, telling a physician what to do may not be an acceptable way for physicians to communicate information. Also, it may lead to physicians taking the advice too literally without a careful evaluation of their own.

2.2 Related Work

The use of the term *critiquing* to describe a system that criticizes the solution provided by a human can be attributed to Perry Miller [23, 24] who developed his ATTENDING system for critiquing anesthesia management. Although critiquing has first been used for evaluating medical treatment plans, since then it has been applied to a wide variety of problems such as engineering design, decision making, word processing, knowledge base acquisition, and software engineering [31]. Here we give an overview of only those critiquing systems that have been used in the medical domain.

2.2.1 The ATTENDING Systems

The original ATTENDING system [23] was designed to critique the management of anesthesia for patients undergoing surgery. The system used a system of rules implemented as Augmented Decision Networks (ADNs), which were based on the model of Augmented Transition Networks (ATNs). The ADNs represent choices about the patients anesthesia, which are structured hierarchically to capture relations between decisions and sub-decisions. The arcs in the ADNs are connected to Problem Management Frames that indicate the implications of anesthetics of certain medical problems. Background knowledge of anesthesia and the risks involved allows the ATTENDING system to evaluate and to propose alternatives to the proposed treatment plan of the physician. A prose generator finally produces a more user friendly output of the critique generated by the system.

The ATTENDING system has led Miller to develop a whole family of critiquing systems to explore the critiquing approach in different domains to identify features of domains that would lend themselves to critiquing. The whole family of ATTENDING based systems, which can be found in Miller's book [24], include: HT-ATTENDING, a critic for the pharmacologic management of essential hypertension; VQ-ATTENDING, a critic for ventilator management for patients on mechanical respiratory support; PHEO-ATTENDING, a critic for the laboratory and radiologic work-up of patients for suspected pheochromocytoma (rare tumour of the adrenal gland); ICON, a critic for differential diagnosis in radiology; E-ATTENDING, a shell for the development of critiquing systems including a differential analyzer and prose generator.

Miller never took his critics beyond the prototype stage to focus mainly on research issues. His principal finding was that simpler differential analyzers apply to domains with a few correct answers to any task. More specifically, Miller learned when to use critiquing by reacting, by local risk analysis, and when by global plan analysis. Critiquing by reacting occurs where (1) specific rules can be written for each type of answer, (2) rules for assessing user solution optimality are objective and few, (3) only one or two possible correct outcomes exist for that task, and (4) each subtask can be critiqued independently of the others. In case conditions (1) through (3) deteriorate, critiquing requires evaluation of the risks and benefits of a number of alternative solutions. Finally, if (4) does not hold, critiquing must adopt a global view to evaluate what the user is proposing.

2.2.2 ONCOCIN

Whereas the ATTENDING system was primarily designed for the purpose of critiquing, the ONCOCIN system [16] started as an expert system for diagnosing cancer problems in a clinical

setting and was later converted into a critiquing system. This conversion was made as the original system was unacceptable to its users as they had great difficulty in justifying their own reasoning whenever they wanted to override the system’s decisions. Although ONCOCINs knowledge base was based on existing protocols for cancer treatment, physicians invariably wanted to vary from the protocol.

The conversion of ONCOCIN into a critic primarily involved the addition of a differential analyzer which would suppress the solutions made by the ONCOCIN system until the user had input his own solution. The comparison between system solution and user solution is then performed by formatting them into a hierarchical plan to which a hierarchical plan analyzer is applied.

The differences that are found by ONCOCIN between system solution and user solution are then presented to the user on an ‘agenda’ from which the user can select items for which he wants further explanation. In this way, the user can focus on only those items that he is interested in.

2.2.3 HyperCritic

The HyperCritic system [17] looks at patient data stored in automated medical records of patients with hypertension and critiques the therapy reported in those medical records. Critiquing is based on critiquing statements that are assigned to significant events through the use of four domain-independent critiquing tasks: preparation, selection, monitoring, and responding. The critiquing tasks are, in contrast with previous systems like ATTENDING, not stated in specific medical knowledge, but in more general properties (e.g., side effects, contra indications, etc.). In this way, domain knowledge is separated from critiquing knowledge. This is one of the main contributions of van der Lei. The second contribution is the focus on medical patient records, which gives additional support for critiquing as physicians find it unacceptable to input their treatment plans each time they want to use the system.

2.2.4 Critiquing using Medical Guidelines

A problem with the critiquing approaches proposed by Miller (ATTENDING [23]) and van der Lei (HyperCritic [17]) is that these systems are unable to cope with deviations from the underlying model that is used to generate system solutions. In particular, this problem may be serious when medical guidelines are used as the underlying model for critiquing the physicians management plan. Medical guidelines are typically under-constrained [14], specifying only *some* constraints for medical practice, and can therefore be seen as skeletal plans [22] that need to be further refined in order to be used. Hence, Physicians are required to interpret the guideline to determine the *intentions* that the guideline developers had in mind in order to determine how to act.

Shahar et al. [30] and Advani et al. [2, 1] therefore propose to perform critiquing by assessing the compliance of the physicians intentions with the intentions of the guideline. The meaning of intentions has already been looked at in depth in other domains like philosophy [4] and AI [26]. In the work of Shahar et al. intentions are viewed to be temporal patterns of provider actions or patient states, to be achieved, maintained, or avoided, i.e., temporally extended goals at various abstraction levels. Also of interest in the approach of Shahar et al., which contrasts with previous approaches, is that, besides provider actions, also patient states are considered for critiquing. More precisely, five comparison dimensions are identified:

(1) the guideline's prescribed actions vs. provider's actual actions; (2) guideline's intended plan vs. prover's (abstracted) plan; (3) guideline's intended patient state vs. provider's state intention; (4) guideline's intended state vs. the patient's (abstracted) actual state; (5) provider's intended state vs. the patient's (abstracted) actual state. Although critiquing based on intentions is an intriguing way to look at critiquing, at this moment still no support exists using guidelines currently developed.

2.3 Concluding Remarks

Both the ATTENDING systems and ONCOCIN favored an approach that uses a hierarchical plan decomposition for the critiquing task. This fits well with the Asbru language that was chosen within Protocure II to represent medical guidelines, as Asbru decomposes medical guidelines into a hierarchically organized number of treatment plans.

In ONCOCIN the main effort to implement critiquing has been the addition of a differential analyzer, which is also an integral part of the E-ATTENDING shell developed by Miller. Previous work [19, 14] within Protocure II, which looked into the problem of locating differences between medical guidelines and protocols, therefore seems also applicable within the context of critiquing. The model checking techniques developed in [19, 14] are therefore taken as a starting point and extended for the critiquing task.

When physicians have to input their treatment plans every time they want to receive a critique on their input, critiquing is often found unacceptable. Van der lei therefore built his HyperCritic system that used medical records instead as input. We will follow a similar approach using patient data collected in hospitals and typical patient cases obtainable through public internet resources. This also fits well with the current trend in medicine of moving to an infrastructure that uses electronic patient records instead of paper records.

Chapter 3

Model checking for Critiquing

3.1 Introduction

Model checking is an effective technique for verifying properties of a formal system. Taking as input a model, usually in the form of a transition system, and a property, a model checker searches all reachable states of the system to establish whether the property is a theorem of the model. In Section 3.2, we briefly discuss the advantage of using model checking compared to other verification and testing techniques. The language for representing the property is usually a temporal logic, which is described in Section 3.3. Finally, in Section 3.4, we discuss a method for applying model checking to critiquing.

3.2 Comparison with other techniques

Model checking represents a more thorough validation technique than testing executable specifications. The latter can only generate a single trace from the model, while model checking allows one to search through *all* traces in the non-deterministic model. This is particularly important in models where detail about ordering between actions is omitted, which was often found in the Dutch breast cancer guideline.

Another advantage of model checking over other verification techniques is the degree of to which the verification is automated. The main problem here is the specification of the model and the properties. After this, the verification is fully automated.

3.3 Specification and Methods

3.3.1 Temporal Logic

Temporal logic is a modal logic, where relationships between worlds in the usual possible-world semantics of modal logic is understood as time order. The logic that we use here for specifying properties of medical guidelines is a combination of Computation Tree Logic (CTL) [3, 8] and Linear Temporal Logic (LTL) [25].

In this paper we model a guideline as a Kripke structure M over a set of atomic propositions AP , which formally is defined as a four tuple $M = (S, S_0, R, L)$ where

- S is a finite set of states.

- $S_0 \subseteq S$ is the set of initial states.
- $R \subseteq S \times S$ is a transition relation that must be total, i.e., for every $s \in S$ there is a state $s' \in S$ such that $R(s, s')$.
- $L : S \rightarrow 2^{AP}$ is a function that labels each state with the set of atomic propositions true in that state.

A *path* in the model M from a state s is an infinite sequence $\pi = s_0s_1s_2\dots$ such that $s_0 = s$ and $R(s_i, s_{i+1})$ holds for all $i \geq 0$. With π^i we denote the suffix of π starting at s_i .

CTL uses atomic propositions and Boolean connectives (e.g., \neg, \vee, \wedge) to build up more complicated expressions for describing properties of states. Furthermore, CTL formulas can be composed of *path quantifiers* and *temporal operators* for describing properties of *computation trees*, i.e., the tree that is formed by designating a state in the Kripke structure as the initial state and then unwinding the structure into an infinite tree according to the transition relation R with the initial state as root. The path quantifiers are **A** and **E** to specify that all of the paths or some of the paths starting at a specific state have some property. The temporal operators describe properties of a path through the tree. The five temporal operators used are **X**, **G**, **F**, **U**, and **R**. With **X** φ being true if φ holds in the next state, **G** φ if φ holds in the current state and all future states, **F** φ if φ holds in some state in the future (or is true in the current state), φ **U** ψ if φ holds until ψ holds, i.e., there is a state on the path where ψ holds and in every preceding state φ holds, and φ **R** ψ if ψ holds along the path up to and including the first state where φ holds, however φ is not required to hold eventually.

In CTL there are two types of formulas: *state formulas*, which are true in a specific state, and *path formulas*, which are true along a specific path. The syntax of state and path formulas is defined as follows:

- Each atomic proposition is a state formula.
- If f and g are state formulas, then $\neg f$, $f \vee g$, and $f \wedge g$ are state formulas.
- If f is a path formula, then **E** f and **A** f are state formulas.
- If f and g are state formulas, then **X** f , **G** f , **F** f , f **U** g , and f **R** g are path formulas.

The semantics of CTL is defined with respect to a Kripke structure M . Given a state formula f , the notation $M, s \models f$ denotes that f holds in state s of the Kripke structure M . Assuming that f_1 and f_2 are state formulas and g_1 and g_2 are path formulas, the relation \models can be defined inductively as shown in Figure 3.1.

In contrast to CTL, LTL provides operators for describing events along a single computation path. Each formula is of the form **A** f , with f being a path formula, which is either an atomic proposition or inductively defined as $\neg f$, $f \vee g$, $f \wedge g$, **X** f , **F** f , **G** f , or f **R** g with f, g path formulas.

3.3.2 LTL and Modular Model Checking

In model checking literature, the approach of verifying a restricted part of the system (i.e., the protocol consistent with medical practice) is called *modular verification* (cf. [15]). In the assumed-guarantee paradigm, the specification of a module consists of a specification of guaranteed behaviour assuming that the system behaves in a certain way, i.e., the assumed

$M, s \models p$	\Leftrightarrow	$p \in L(s)$
$M, s \models \neg f_1$	\Leftrightarrow	$M, s \not\models f_1$
$M, s \models f_1 \vee f_2$	\Leftrightarrow	$M, s \models f_1$ or $M, s \models f_2$
$M, s \models f_1 \wedge f_2$	\Leftrightarrow	$M, s \models f_1$ and $M, s \models f_2$
$M, s \models \mathbf{E}g_1$	\Leftrightarrow	there is a path π from s such that $M, s \models g_1$
$M, s \models \mathbf{A}g_1$	\Leftrightarrow	for every path π starting from s such that $M, s \models g_1$
$M, \pi \models f_1$	\Leftrightarrow	s is the first state of π and $M, s \models f_1$
$M, \pi \models \neg g_1$	\Leftrightarrow	$M, \pi \not\models g_1$
$M, \pi \models g_1 \vee g_2$	\Leftrightarrow	$M, \pi \models g_1$ or $M, \pi \models g_2$
$M, \pi \models g_1 \wedge g_2$	\Leftrightarrow	$M, \pi \models g_1$ and $M, \pi \models g_2$
$M, \pi \models \mathbf{X}g_1$	\Leftrightarrow	$M, \pi^1 \models g_1$
$M, \pi \models \mathbf{F}g_1$	\Leftrightarrow	there exists a $k \geq 0$ such that $M, \pi^k \models g_1$
$M, \pi \models \mathbf{G}g_1$	\Leftrightarrow	for all $k \geq 0$, $M, \pi^k \models g_1$
$M, \pi \models g_1 \mathbf{U} g_2$	\Leftrightarrow	there exists a $k \geq 0$ such that $M, \pi^k \models g_2$ and for all $0 \leq j < k$, $M, \pi^j \models g_1$
$M, \pi \models g_1 \mathbf{R} g_2$	\Leftrightarrow	there exists a $k \geq 0$ such that $M, \pi^k \models g_1$ and for all $0 \leq j < k$, $M, \pi^j \models g_2$

Figure 3.1: Semantics of CTL with f_1 and f_2 representing state formulas and g_1 and g_2 representing path formulas

behaviour. In this paper, the assumed behaviour is written down in a linear temporal logic and the guaranteed behaviour in branching temporal logic. The assume-guarantee assertions are written down as $[\varphi]M\langle\psi\rangle$, meaning that the branching temporal formula ψ holds in the computation tree that consists of all computations of the program, described by M , that satisfy the linear temporal formula φ .

3.3.3 Tools

The choice of model checking tool for Protocure II is Cadence SMV. In deliverable 4.3, the reason for this choice is explained. Also, for complexity issues of more expressive logics, we refer to this deliverable.

3.4 Applying Model Checking for Critiquing

3.4.1 Method

Figure 3.2 depicts the steps performed to establish the doctor's compliance with the guideline recommendations. The parts drawn in light grey represent steps done semi-automatically as part of generating the specification, while the parts coloured in dark grey are the steps performed for the critiquing task.

3.4.2 Specification of model and traces

The model that is being used for critiquing is the same model as developed for the model checking task. Hence, it is based on abstracted Asbru models, translated using the operational semantics of Asbru.

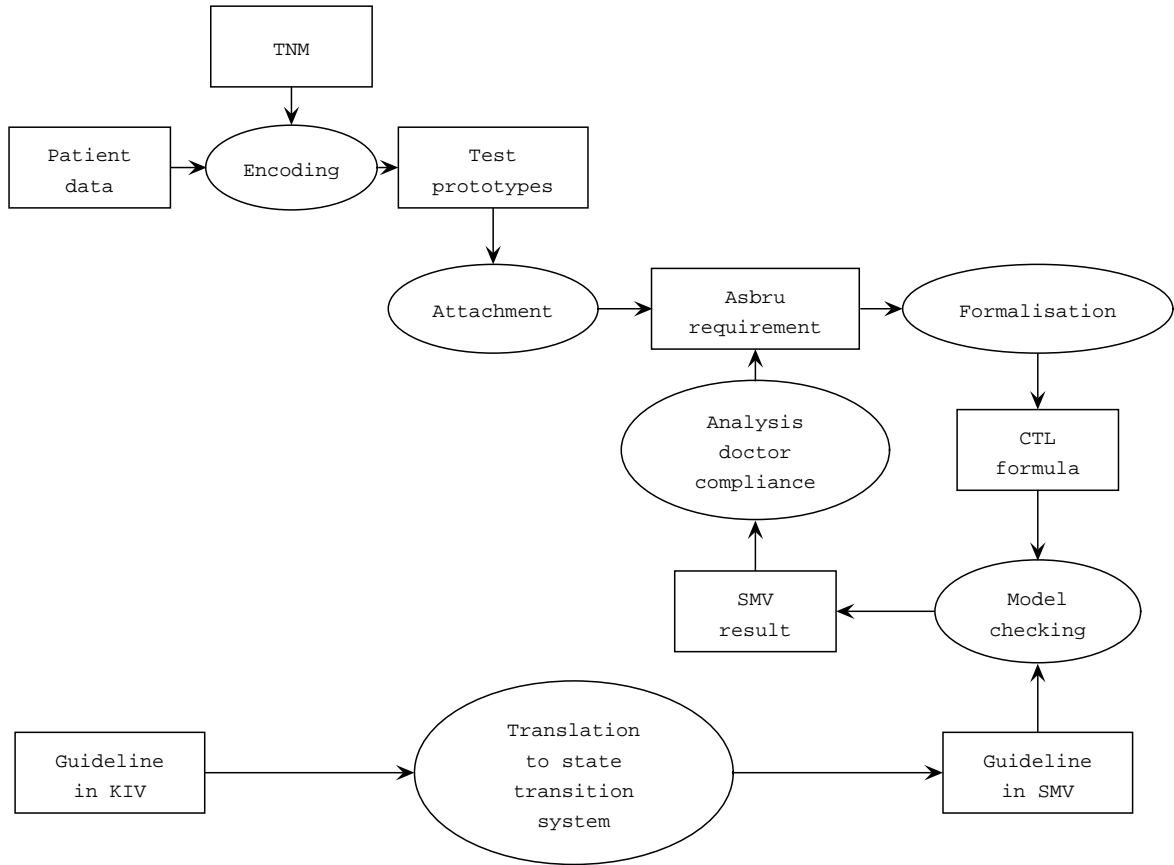


Figure 3.2: Steps for critiquing using model checking.

The main work of this task consisted of the specification of sequences of paths. Using model checking, we can verify whether a sequence of actions performed by the clinician is a valid path in the model of the guideline. In practice, we do not have a complete description of all the actions done, i.e., the description of the actual treatment is under-specified. We will assume there are types of information available:

- Information that holds during the *whole* trace, e.g., the primary diagnosis is usually assumed to be constant during the execution of the guideline model
- Existence of actions in the treatment
- Order between actions

In CTL, the existence of an action α can be represented as follows:

$$EF \alpha$$

i.e., in some trace, somewhere in the future, α is executed. Note that properties used for critiquing always start with an E, i.e., it is established that *some* path exists. Things become more complicated in case there are two actions. In a more expressive logic, called CTL*, the specification is:

$$E((F \alpha_1) \wedge (F \alpha_2))$$

However, this combination (i.e., the conjunction of two state formulas containing modalities) is not a CTL formula and thus not supported by Cadence SMV, nor by other model checkers such as NuSMV. Hence, if, for example, we know that α_1 was executed before α_2 , then the correct CTL formula is:

$$EF(\alpha_1 \wedge EX EF \alpha_2)$$

In case we do not have this information, the situation may be represented as:

$$EF(\alpha_1 \wedge EF \alpha_2) \vee EF(\alpha_2 \wedge EF \alpha_1)$$

i.e., all the combinations between the execution of α_1 and α_2 are acceptable. It is obvious how to extend this for more than two actions, though the complexity of the formula exponentially grows in case the order between actions is unknown.

The same problem occurs in case we have information which is relevant through the whole trace. For example, suppose we know that action α occurs and β does not. Then, this can be represented as:

$$E(\neg\beta U (\alpha \wedge EG \neg\beta))$$

i.e., β does not occur in this path, until α occurs after which β never occurs either. Clearly, if more actions have been performed by the clinician, the situation becomes even more difficult to express. As the complexity of model checking depends on the size of the formula, this will slow down the model checking performance.

An alternative to this approach is modular model checking, where the model is restricted to those traces where the assertions are valid. To prove the *existence* of a path in this approach, it is required to verify that the model is not empty. Hence, if

$$[\varphi]M\langle\perp\rangle$$

is true, then this means that no counter model (i.e., trace from the model) could be provided and therefore, we know that the model is empty and thus φ does not specify a possible trace. In contrast, if it false, then the counter model is provided consistent with φ proving that such a trace exists. For example, suppose we would like to find out if a trace exists consisting of α_1, α_2 and never β . Then, we check that:

$$[F \alpha_1, F \alpha_2, G \neg\beta]M\langle\perp\rangle$$

is false. An equivalent CTL formula is:

$$E(\neg\beta U (\alpha_1 \wedge E(\neg\beta U (\alpha_2 \wedge EG \neg\beta)))) \vee E(\neg\beta, U (\alpha_2 \wedge E(\neg\beta U (\alpha_1 \wedge EG \neg\beta))))$$

which is quite unintuitive for even this small example.

In general, it is difficult to say which of the logics is better for this task. If intuitive formulas are important for the analysis of the critiquing, then LTL seems to be the best choice. With respect to complexity, our hypothesis is that it greatly depends on the type of information (e.g., whether or not the order is given) resulting in a blow-up for the CTL case and the size of the model.

Chapter 4

Approach to critiquing the guideline

We discuss here different approaches to critiquing:

1. critiquing of patient data using Model Checking, and
2. critiquing of additional case studies from literature.

The latter is meant as an exercise to verify whether in absence of concrete patient data it is possible to check compliance of treatments with respect to the guideline.

Based on a few selected examples, we compare the critiquing by Model Checking, which can be performed semi-automatically, with critiquing by manual inspection of the ASBRU plan hierarchy.

4.1 Critiquing of patient data using Model Checking

The main type of critiquing has been by using Model Checking, in which we were aided by the availability of a large set of breast cancer patient treatment data that were treated according to the initial version of the selected guideline.

We analyzed this data in order to find mappings between the patient data fields and the parameters of our ASBRU models. This analysis has led to a small number of mappings. The main reason is that the patient data was part of clinical records, i.e. descriptive data, not meant for an operational setting, and the guideline was focused on a more operational aspect of breast cancer treatment, with little reference in the recommendations to general parameters used in the treatment of breast cancer.

4.1.1 Critiquing objectives

Critiquing in our approach has several objectives, trying to answer the following questions:

1. Can the treatment be linked to one of the treatment paths prescribed by the guideline recommendations?
2. Does the treatment follow exactly the treatment sequence mentioned in the recommendations? If not, is the treatment correct?

3. Are there several patients from the same category with the same non-compliant behaviour?

The main goal of critiquing is to establish whether the actual treatment trace of a real patient complies with the guideline recommendations.

Another goal is to establish how frequent non-compliant behaviour can be discovered, and whether the problem lies with the guideline.

The third goal is to determine whether the knowledge gaps between the guideline recommendations and medical practice can be covered or assessed accurately.

4.1.2 Mapping patient data to ASBRU

The mappings between patient data parameters and ASBRU parameters are given in appendix (section B).

Due to the difficulty of generating one-to-many mappings between patient data actions and ASBRU plans, to aid us in reconstructing action traces used in critiquing, we split the critiquing process into several smaller steps: validation of data by medical interpretation, reconstructing possible medical treatment paths through manual mappings, refine the property used as input for critiquing, verify it, and analyze the potential alternative solutions.

4.1.3 Validate the mappings by medical and technical interpretation

Using manual mappings and supported by a program that computes semantic similarities between action names (based on the overlaps between composing words), we produced a large list of candidate mappings.

These mappings were validated by medical experts and have undergone a further technical refinement: when the mappings referred to plans situated in different chapters of the guideline, first an evaluation on whether the plans (and their local plan hierarchies) could be merged. If that was not possible, then we selected the mappings that refer to a self-standing component of the guideline model (this could be the SMV model of one chapter or of a subpart of that chapter).

In fact, in this step, a review of the mappings was done, which resulted in selecting only those mappings which targeted a self-standing component of the formal guideline model.

4.1.4 Reconstructing possible treatment paths

In this step, we generated a **critiquing trace**, which means that for each of the patient treatment actions we selected exactly one of the mappings obtained in the previous step.

For instance, if patient P had medical condition $C = C_1, C_2, \dots, C_n$ and has undergone treatment $A1 + A2 + \dots + A5$ (a maximum of 5 actions was recorded), and given a set of mappings established and validated in the previous step:

$C_i = ONE - OF(Par_{i1}, Par_{i2}, \dots, Par_{ik}), i = \overline{1, nr_patient_parameters}$ and $A_j = ONE - OF(P_{j1}, P_{j2}, \dots, P_{jl}), j = \overline{1, 5}$ (max 5 action treatment codes were present in patient data)

then we generate a critiquing trace and a critiquing property for all combinations of these actions:

This is the generic form of most properties used in critiquing. For each patient, $Par_{i,j}$ represents a parameter of the patient, $val(Par_{i,j}, C_k)$ represents one possible value of that

GENERAL FORM OF A PROPERTY CRITIQUING SEQUENCES OF ACTIONS

$$\begin{aligned}
 & EF ((Par_{11} = val(Par_{11}, C_1) \vee Par_{12} = val(Par_{12}, C_1) \vee \dots \vee \\
 & Par_{1k} = val(Par_{1k}, C_1)) \wedge \dots \wedge \\
 & (Par_{n1} = val(Par_{n1}, C_n) \vee Par_{n2} = val(Par_{n2}, C_n) \vee \dots \vee Par_{nk} = val(Par_{nk}, C_n)) \wedge \\
 & (\text{assuming there exists one patient with the parameter values corresponding to} \\
 & \text{medical condition } C_i, \text{ denoted } val(Par_{ij}, C_i)) \\
 & P_{1i}.state = inactive \wedge P_{2i}.state = inactive \wedge \dots \wedge P_{5i}.state = inactive \wedge \\
 & EF ((P_{1i}.state = activated \wedge P_{2i}.state = inactive \wedge \dots \wedge P_{5i}.inactive) \wedge \\
 & EF ((P_{1i}.state = completed \wedge P_{2i}.state = inactive \wedge \dots \wedge P_{5i}.inactive) \wedge \\
 & EF ((P_{1i}.state = inactive \wedge P_{2i}.state = activated \wedge \dots \wedge P_{5i}.inactive) \wedge \dots) \\
 & (\text{each plan gets activated in sequence})))
 \end{aligned}$$

Figure 4.1: General form of a property used for critiquing a medical treatment sequence by model checking

parameter, which could be mapped to the condition C_k , and each P_{mn} represents one possible mapping of treatment actions to ASBRU plans.

4.1.5 Refine the critiquing traces

After the possible treatment paths have been reconstructed, some of them might include inconsistent or illogical combinations (such as, removal of the breast, followed by a diagnosis, or followed by asking patient's permission), and therefore have to be eliminated from the list of possibilities.

For this reason, an additional validation of the reconstructed treatment traces takes place, which further reduces the set of reconstructed possible traces that explain what the doctor might have performed. The role of this step is to select the most likely treatment paths and to reduce the complexity of the critiquing property that has to be verified, by using correlations between parameters and plans (for instance, if we know that a parameter *local_excision_of_DCIS* = *false* and has the same value along the entire treatment, then a trace in which the plan *local_excision_of_DCIS* is included is not valid, and the action can be dropped.

4.1.6 Verify and interpret the critiquing properties

Once the critiquing traces have been refined, we have to verify whether a property such as the one in figure 4.1 is verified.

If the property is true, the medical sequence of treatment actions performed by the doctor is considered correct.

If the critiquing property is not verified, one has to split it into simpler properties and investigate whether they hold. We are interested to answer propositions of the following form:

1. check whether plan P_{ij} can be activated, provided condition C is true;
2. verify that condition C remains unmodified when plan P_{ij} is completed;

3. verify that there exists a path when plan P_{ij} gets activated before $P_{(i+1)j}$;
4. verify that on all paths where condition C is true, plan P_{ij} always gets activated before $P_{(i+1)j}$

4.2 Critiquing breast cancer case studies from literature

The same steps as before are applied, only in this case the medical interpretation and reconstruction of the medical process are supported by more background information and additional medical knowledge of the team that performed the critiquing.

Also the steps performed for dealing with the results of the critiquing were different: in case of failure, the property used in critiquing was refined by relaxing the ordering constraints, instead of reducing the sequence of actions, adding additional patient assumptions and studying specific combinations between medical conditions and actions in isolation, as done based on real patient data.

Chapter 5

Implementation and Results

5.1 Critiquing Patient Data

This section describes the steps undertaken by us, according to the methodology described in section 4, for critiquing the patient data.

5.1.1 Patient Data

For critiquing we used patient data obtained from the Dutch Comprehensive Cancer Centre South, which maintains a population based cancer registry in the Netherlands used for cancer research, planning of services, and evaluation and implementation of guidelines. The data for the Protocure II project originates from the Comprehensive Cancer Centre South, because they collect more specific data of breast cancer patients than the other centres (i.e. follow-up data).

All data concerning breast cancer patients for the period January 2003 - June 2004 was obtained for the Protocure II project. This resulted in 5507 patient records with 269 variables per patients. The variables include information about patient characteristics (i.e. age, date of diagnosis, etc), cancer characteristics (i.e. type of tumor, location of tumor, etc.) and treatment information (i.e. type of surgery, radiotherapy, chemotherapy, etc).

This information has been analyzed by medical experts and by guideline modellers, attempting to establish mappings with the ASBRU parameters and plans. Based on their annotations and a codebook for this data, a Java application has been built, which outputs a textual description of the essential parameters of the patient model.

Since the main goal of data collection has not been to check the compliance of the patient treatment with the recommendations of the selected CBO breast cancer guideline, a large knowledge gap exists between the patient data and the ASBRU model. An example in which this knowledge gap has been closed, under specific assumptions about the meaning of patient action names, is given below (patient 4 from appendix section A):

```
APPLICATION OUTPUT:
patient_trace(4,
% PATIENT CONDITION
patient_data([sex(female),age(79)]),
diagnosis([name('Invasive ductal carcinoma (mixed invasive and in-situ)'),
tumour_info(
[localization('overlapping lesion of breast'), laterality('right'),
tumour_extension(3), tumour_grade(2),
tumour_differentiation_grade('moderately differentiated'),
```

```

topology('overlapping lesion of breast'),
behaviour('malignant/invasive, primary site'),
morphology('ductal carcinoma'),
nr_lymph_nodes([total(2),positive(0)]),
metastases([],nr_metastases_in_lymph_nodes([total(0),investigated(0)]),
additional_tumour_codes([typeV1('rectum'),typeV2(''),typeV3('')]),
clinical_staging_results ([
stage_grouping_clinical('4'),
clinical_tumour_staging('not applicable'),
ct('tumour having between 1cm and 2cm in greatest dimension'),
cn('regional lymph nodes cannot be assessed'),
cm('no distant metastasis'),
ctnmt('primary tumour cannot be assessed'),
ctnmn('no regional lymph node metastasis'),
ctnmm(' staging code =1C')]),
pathological_staging_results([
stage_grouping_pathological('(1C,X,0)'),
pathological_tumour_staging('not applicable'),
pt('tumour having between 1cm and 2cm in greatest dimension'),
pn('regional lymph nodes metastasis cannot be assessed
(regional lymph nodes not removed for study or previously
removed)'),
pm('no distant metastasis'),
ptnmt('tumour having between 1cm and 2cm in greatest dimension'),
ptnmn('regional lymph nodes metastasis cannot be assessed
(regional lymph nodes not removed for study or previously
removed)'),
ptnmm('no distant metastasis')]) ])
% PATIENT TREATMENT SEQUENCE
treatment([
i1('sentinel node biopsy'),
i2('organ specific surgery (organ localization code=14)='lumpectomy)])

```

An exact translation of the treatment sequence could not be found in the ASBRU model, due to the different terminologies used for patient data and in the guideline, but the following approximate translation of the treatment trace could be found and connected with the ASBRU model of chapter 1 of the CBO guideline.

APPROXIMATE TRANSLATION OF PATIENT 4 IN ASBRU TERMINOLOGY

INPUT PARAMETERS:
diagnosis=DCIS AND (tumour-range > 1cm and tumour-range < 2cm)
TREATMENT SEQUENCE:
Axillary-staging-by-SN+ bct

As shown in this example, provided the mappings between patient data parameters and parameters of the ASBRU model are correct, a number of typical input ASBRU parameters can be established. In fact, a model of the prototypical patient can be built, provided comprehensive information about treatment is available. Also, under the assumption that the treatment strictly followed the guideline recommendations, test objectives can be generated based on this patient data model.

5.1.2 Profiles of prototypical patients

Starting from the profiles of a few prototypical patients, which represent a large category of patients, a medical analysis was made, to establish whether the treatments commonly associated with these classes of patients, according to the guideline, correspond with the actual treatments performed.

The analysis was made based on a natural language description of the patient data, annotated by medical experts with a set of conditions and assumptions. It was difficult to establish the annotated (and interpreted) treatment path, as, more prognostic factors had to be considered, which were not immediately present in the patient data.

Then these "actual" treatment traces were mapped, based on the medical condition, to parts of the guideline where their diagnosis is handled.

An example of such a prototypical treatment trace, obtained based on more than one patient:

MEDICAL PATIENT PROTOTYPE
MEDICAL CONDITION: Invasive ductal carcinoma grade 2 mixed with carcinoma in-situ. Size between 1 and 2 cm. Lymph nodes: 2. No positive lymph nodes. TREATMENT: organ specific surgery and sentinel node biopsy and (optional: if age < 50 (premenopausal status) then do chemotherapy after biopsy).

This prototypical treatment trace represents the actual treatment for a class of patients, namely those diagnosed with Ductal Carcinoma in Situ, average tumours, and no positive lymph nodes. As a large percentage of the patient data have this diagnosis, critiquing this treatment give an indication of the most typical treatment used in practice for the patients falling in this category.

5.1.3 Case study 1: Patient 4

Let us illustrate the steps of critiquing for patient 4 in patient data records analyzed.

SELECTED PATIENT #4
Medical condition: 79 years-old woman. Overlapping lesion of right breast. Invasive ductal carcinoma grade 2 mixed with carcinoma in-situ. Size between 1 and 2 cm. Lymph nodes: 0/2. Treatment: sentinel node biopsy + lumpectomy (= breast-conserving surgery without axillary clearance).

Medical expert comments: "When having a breast conserving surgery, patients should be treated with complementary radiotherapy. The size of the carcinoma in-situ component would determine the kind of surgical procedure."

The patient has undergone a sentinel node biopsy (denoted i1), followed by a lumpectomy (denoted i2), which is a special case of breast-conserving surgery without axillary clearance.

By mapping these actions to ASBRU, we arrive at the following approximate translation of the treatment:

```
i1 = ONE-OF(  
  1.1. histological-biopsy(hibi/P79/CH5)  
  1.2. biopsy(bio/P116/CH6) |  
  1.3. ultrasound-guided-core-biopsy(ugb/P50/CH1) |  
  1.4. axillary-staging-by-SN(asbSN/P2/CH1) |  
  1.5. sentinel-node-procedure(snp/P105/CH5)
```

```
i2 = ONE-OF(
  2.1. limited-surgery-manual(lsm/P59/CH3) |
  2.2. breast-conserving-therapy(bct/P65/CH5) |
  2.3. bct (bct/P5/CH1))
```

That is, for each action in the medical sequence of treatments, we propose a list of possible mappings, and select for each action i_1 , i_2 , and so on, the ASBRU equivalent action. In the list above, *histological – biopsy* is an ASBRU plan name, *hibi* is a shortcut used for its KIV name, *P79* indicates that this plan has index 79 in a reference list of plans, and *CH5* indicates that this plan has been encountered in the guideline model of chapter 1.

As the diagnosis of the patient is DCIS (Ductal Carcinoma In Situ), we decide to map the actions to chapter 1 (which deals with Ductal Carcinoma In Situ), therefore obtaining the following treatment sequence, which is the most likely to be the actual treatment trace to be critiqued:

MOST LIKELY TREATMENT SEQUENCE FOR PATIENT 4:

1. axillary-staging-by-SN(asbSN/Plan2/Chapter 1) + 2. bct (bct/P5/CH1)

When we specify this property in SMV, for the model checker, we use the generic form described in figure 4.1 in section 4.1. To specify and then verify whether breast cancer therapy (denoted *bct*) can take place after axillary staging by sentinel node procedure (denoted *asbSN*), the following property describes the treatment sequence which patient 4 has undergone.

SMV SPECIFICATION OF CRITIQUING PROPERTY FOR PATIENT 4:

=====

```
SPEC EF (
-- medical condition
(diagnosis=DCIS AND
-- initial plan states
chap1p_state=activated & asbSNp_state=inactive & bctp_state=inactive)
AND EF (
-- plan states after event 1: asbSN is activated
(chap1p_state=activated & asbSNp_state=activated &
bctp_state=inactive) & EF (
-- plan states after event 2: asbSN is completed
(chap1p_state=activated & asbSNp_state=completed &
bctp_state=inactive) & EF (
-- plan states after event 3: bct is activated
(chap1p_state=activated & asbSNp_state=inactive &
bctp_state=activated) & EF (
-- plan states after event 4: bct is completed
(chap1p_state=activated & asbSNp_state=inactive &
bctp_state=completed) & EF (
-- plan states after event 5: top level plan is completed
chap1p_state=completed
))))))
```

As it can be seen, this property is not specific to agent 4, but in fact says: "for any patient having diagnosis Ductal Carcinoma in Situ, check whether, according to the model, there exists a situation in which initially no plan is active, then axillary staging by sentinel node becomes activated, then it is completed, after which breast conserving therapy becomes activated, and then completes successfully".

We reduce the property, to say only that activation order is correct, without saying anything about the completion order:

```
SPEC EF (
(diagnosis=DCIS &
chap1p_state=activated & asbSNp_state=inactive & bctp_state=inactive) &
EF ((chap1p_state=activated & asbSNp_state=activated &
bctp_state=inactive) & EF (
(chap1p_state=activated & asbSNp_state=inactive & bctp_state=activated)
)))
```

This property is false, with a counterexample trace of length 34.

Model checking results

=====

```
(EF (((((diagnosis=DCIS)&(chap1p_state=activated))&(asbSNp_state=inact.....false
```

Resources used

=====

```
user time.....304.55 s
system time.....0.42 s
BDD nodes allocated.....3010610
```

This counterexample indicates a situation (in which the tumour size is large) when bct becomes activated without asbSN to become active. In other words, according to the ASBRU model of the guideline chapter describing the treatment of Ductal Carcinoma in Situ, the sequence of action performed by the doctor is incorrect for the patients with the same diagnosis as patient 4. It could also mean that, according to the model, at least one of the actions cannot be started, or that the two actions can activate, but not in the expected sequence.

We have to identify now the explanation of this anomaly.

If we reduce the sequence to either of the two actions, the property becomes true (this was actually expected). The new conclusion is that either under some specific circumstances the two actions cannot be activated in this sequence, or the ordering should be reversed.

If the ordering of these two actions in the sequence are reversed, the property still remain false:

Model checking results

=====

```
-- Reverse sequence: bct + asbSN
-- PROPERTY IS FALSE, counterexample of length 1
SPEC EF ((diagnosis=DCIS & chap1p_state=activated &
asbSNp_state=inactive & bctp_state=inactive) &
EF ((chap1p_state=activated & asbSNp_state=inactive &
bctp_state=activated) &
```

```

        EF ( (chap1p_state=activated & asbSNp_state=activated &
              bctp_state=inactive) )))

-- Reduced sequence (one action: bct):
-- PROPERTY IS TRUE
SPEC EF ((diagnosis=DCIS & chap1p_state=activated &
          bctp_state=inactive) &
          EF ((chap1p_state=activated & bctp_state=activated )))

-- Reduced sequence (one action: asbSN):
-- PROPERTY IS TRUE
SPEC EF ((diagnosis=DCIS & chap1p_state=activated &
          asbSNp_state=inactive) &
          EF ((chap1p_state=activated & asbSNp_state=activated )))

```

The intriguing part is the counterexample of the reverse ordering. It has length 1 and it contains the initial situation in which several variables are unknown.

We investigate what would happen if more precise information is known about the patient. What if all variables present among the activation conditions of either of the two actions (bct and asbSN) are known? This would clarify that with additional patient information, we can clarify the circumstance in which either of bct or asbSN can still activate.

If we add an additional fairness assumption, namely that it is possible for asbSN to leave the inactive state (and become at least considered), we could rule out the fact that asbSN is an illegitimate action for patient 4.

We can then verify again whether the property is true, given the following fairness assumption:

```
FAIRNESS !asbSNp_state = inactive
```

In this case, the original critiquing property (sequence asbSN + bct is possible) is still empty, with a counterexample of length 104. In this counterexample, plan bct (breast-conserving therapy) gets completed, but asbSN does not get activated. On a closer investigation, a sibling of asbSN (complete-axillary-node-dissection, shortly cand) can get activated, so apparently the activation of asbSN's parent plan (dealing-with-axilla) takes place.

It can be derived that, according to the guideline model, complete-axillary-node-dissection would have been a valid alternative to axilla-staging-by-SN. In this case the answer of the critiquing module would be:

```
CRITIQUE(PATIENT 4): FOR PATIENTS WITH DIAGNOSIS=DCIS, IT IS
EXPECTED TO HAVE COMPLETE-AXILLARY-NODE-DISSECTION, AND
NOT AXILLA-STAGING-BY-SN.
```

It can also be the case that the treatment sequence of patient is incorrect due to incompatible activation conditions between the two plans. We decide to become more specific about the patient condition, by iteratively adding fairness assumptions about the variables that make up the activation conditions of each of the two plans.

For instance, the two plans asbSN and bct have the following activation condition:

```
asbSNp_filter_condition :=
```

```
(a_sensitivity_rate_of_95_is_achievable=1)&
(((suspected_or_proven_malignancy_in_the_axillary_nodes=1)|
(tumour_range=grmedium)|
(multible_tumour_foci=1)|
(!(potentially_disrupted_lymph_drainage=1))))|
((pardcis=1)&(tumour_range=gelarge)));
```

```
bctp_filter_condition := patient-prefers-bct=1;
```

In other words, the axilla-staging-by-SN does (asbSN) should only be activated depending on the values of the several parameters. This information about the patient can in principle located with the patient record. If that is the case, we can add the following assumptions:

```
a_sensitivity_rate_of_95_is_achievable!=unkown
multible_tumour_foci != unkown
tumour_range != unkown
```

This would ideally give a more concrete counterexample (if at all), which could be matched easier with the patient record. It would only help us to identify an explanation on why an alternative action has been selected by the doctor, and would not change our critique with respect to the ordering of medical actions.

For instance, apart from signaling that asbSN + bct is not the right sequence, the critiquing would also indicate the right alternative (with the help of medical experts):

CRITIQUE(PATIENT 4): AN ALTERNATIVE ACTION TO asbSN (NAMELY, cand) SHOULD HAVE BEEN EXECUTED. THE CURRENT CHOICE MIGHT BE CORRECT UNDER THE FOLLOWING ASSUMPTIONS:

1. multidisciplinary decision on BCT is taken: *mdBCTp_is_terminated*
2. *a_sensitivity_rate_of_95_is_achievable* = 1; this parameter is not present in the patient record, its value has to be clarified with the help of the medical expert;
3. tumour has multiple foci: *multiple_tumour_foci* = 1
4. tumour is large: *tumour_range* > *medium*; this is less likely to be the case, as the patient's tumour is between 1 and 2 cm is not large
5. informing the patient about SN procedure is terminated: *SNipp_is_terminated*; this cannot be retrieved from clinical record.

These alternatives can be reused for other patients with the same characterization (diagnosis=DCIS) and with the same treatment sequence (or a longer treatment sequence that respects the same ordering of actions as in the treatment of patient 4).

5.1.4 Case study 2: Patient 21

Patient 21 is described by the following information.

SELECTED PATIENT #21

Medical condition: 44 years-old woman. Lower-outer quadrant lesion of right breast. Invasive ductal carcinoma grade 2 mixed with carcinoma in-situ. Size between 1 and 2 cm. Lymph nodes: 0/2. Treatment: organ specific surgery, sentinel node biopsy and chemotherapy.

This patient's diagnosis "Invasive ductal carcinoma (mixed invasive and in-situ)", is dealt with in chapter 1. The more exact mapping of the patient record to the ASBRU model terminology looks like this:

```
patient_data[age_range(middle-aged), menopausal_status(pre), age('44')],
diagnosis('Invasive ductal carcinoma (mixed invasive and in-situ)'),
tumour-range(small).
```

Treatment: (attached to chapter 1)

```
1.1:axillary-staging-by-SN(andp);
2.1:bct(bct);
3.1:microscopically-complete-tumour-excision(mcte) OR bct(bct);
4.1.chemotherapy (??)
```

In the model of chapter 1, no action could be associated with chemotherapy. Based on the information present in the medical record which was revealed by this application, the medical expert could not define a specific correct treatment. Chemotherapy is a complementary option to surgical treatment if the tumour is ER or PR negative. This decision is not that clear when having ER or PR positive tumours. In any case, if the tumour is ER or PR positive, the patient should receive hormonal treatment, unless medical contraindications apply, after receiving chemotherapy. On the other hand, when having a breast conserving surgery, patients should be treated with complementary radiotherapy. The size of the carcinoma in-situ component would determine the kind of surgical procedure.

Apart from not including any reference to chemotherapy, the guideline model of chapter 1, to which this treatment trace has been attached, the guideline model was build with the assumption that treatment of DCIS is not a cyclical plan, therefore it does not allow to do breast-conserving therapy twice.

So, a treatment trace consisting of all four actions is not possible. If we reduce it to 3 steps, the remaining property is still false, as the sequence asbSN + bct has been shown to be incorrect (an informal analysis shows that this is still an acceptable ordering of events, but not captured in the model). The only subtrace acceptable according to the SMV model is the sequence bct-1+microscopically-complete-tumour-excision.

The trace is incorrect, first of all because no chemotherapy is mentioned in the model. The 2004 version of the guideline introduces a reference to chemotherapy. In 2002, when this patient was treated, it was not mentioned in the recommendations, but the doctor administered it. This shows that in some cases, critiquing can view a correct doctor trace as incorrect according to the current model

5.1.5 Case study 3: Patient 22

Patient 22 has a different diagnosis:

SELECTED PATIENT #22

Medical condition: 80 years-old patient. Central portion lesion of right breast. Lobular and ductal carcinoma (in-situ or invasive) grade 1. Size greater than 5 cm. Lymph nodes: 0/10. Treatment: mastectomy, external radiotherapy and hormone-therapy.

The representation of this patient in the terminology of ASBRU model for chapter 1 (which covers the treatment for the diagnosis "Lobular Carcinoma in Situ") is given below:

```
patient_data([age_range(old),menopausal_status(post), age('80')]),
diagnosis('Lobular and ductal carcinoma (in-situ or invasive)'),
tumour-range(large).
tumour-size>5cm, tumour-staging=2B.
```

Treatment:

i1(mastectomy)

1.1:mastectomy(mast/P26/CH1);

1.2:mastectomy-proper(mapro/P27/CH1);

i2(hormone therapy)

2.1:hormone-therapy(hothe/P141/CH6),

2.2:secondary-adjutant-hormone-therapy(sahth/P102/CH5),

2.3:adjutant-hormonal-therapy(aht/P51/CH3),

2.4:secondary-adjutant-hormone-treatment(saht/P101/CH5),

2.7:anti-estrogens-therapy(aethe/P113/CH6),

2.10:combination-of-trastuzumab-and-paclitaxel(cotap/P124/CH6),

2.12:herceptin-treatment(hertr/P140/CH6),

i3(external radiotherapy)

3.3:initial-radiotherapy(ir/P21/CH1),

3.4:subsequent-radiotherapy(sr/P47/CH1),

3.5:dcis-radiotherapy(dcis/P10/CH1),

3.15:locoregional-postoperative-radiotherapy(lrpr/P23/CH1),

3.16:primary-radiotherapy-of-axilla(praoa/P40/CH1)

No correct mapping could be established with a guideline chapter. A logical attachment to chapter 1 cannot be completed, due to the fact that this chapter's recommendations do not mention hormone therapy.

The subtraces: mastectomy + primary-radiotherapy of axilla AND mastectomy + locoregional-postoperative-therapy AND mastectomy + radiotherapy-to-chest-wall are all correct/possible in the 2002 version, but according to the 2004 model, only subtrace mastectomy + radiotherapy-to-chest-wall is allowed.

The only valid sequence accepted by the model is axilla=staging-by-SN+mastectomy-proper+radiotherapy-chest-wall. If the model is assumed to be correct and complete, the sequence of doctor actions is incorrect. However, a more detailed informal analysis of this patient suggests the doctor actions are possible.

5.1.6 Case study 4: Patient 34

Another interesting case, which based on a preliminary analysis seems correct is that of patient 34:

SELECTED PATIENT #34

Medical condition: 82 years-old woman. Overlapping lesion of the left breast. Adenocarcinoma. Size greater than 5 cm. Lesion with oedema or ulceration of the skin or satellite skin nodes. Unknown grade. Axillary status unknown. Metastasis in supraclavicular lymph nodes. Treatment: chemotherapy and hormone therapy.

The relevant parameters, expressed in the terminology of the ASBRU models, are:

```
patient_data([age_range(old), menopausal_status(post), age('82')]),
diagnosis('Adenocarcinoma, NOS')
clinical_staging_results_before_therapy([stage_grouping_clinical('2B'),
clinical_tumour_staging('not applicable'),ct('oedema (including peau
d-orange) or ulceration of the skin of the breast, or satellite skin
nodules confined to the same breast'),cn('metastasis in
supraclavicular lymph node(s)'),cm('no distant
metastasis'),ctnmt('tumour of more than 5cm in greatest
dimension'),ctnmn('no regional lymph node metastasis')
```

Treatment:

1. hormone-therapy + 2. chemotherapy

i1(hormone-therapy)

```
1.1:hormone-therapy(hothe/P141/CH6),
1.2:secondary-adjvant-hormone-therapy(sahth/P102/CH5),
1.3:adjvant-hormonal-therapy(aht/P51/CH3),
1.4:secondary-adjvant-hormone-treatment(saht/P101/CH5),
1.5:combination-of-trastuzumab-and-paclitaxel(cotap/P124/CH6),
1.6:herceptin-treatment(hertr/P140/CH6),
```

i2(chemotherapy)

```
2.1:cmf-chemotherapy(cmf/P123/CH6),
2.2:palliative-chemotherapy(pache/P150/CH6),
2.3:neoadjuvant-chemotherapy(neo/P60/CH3),
2.4:six-courses-anthracycline-chemotherapy(scac/P62/CH3),
2.5:single-course-anthracycline-chemotherapy(sgcac/P63/CH3),
2.6:secondary-adjvant-chemotherapy(sac/P100/CH5),
2.7:anthracycline-containing-regimen(acr/P112/CH6),
2.9:lumbally-citostatic(lucit/P147/CH6),
```

The treatment of this patient can be associated with chapter 6, which deals with treatment of metastatic breast cancer. As can be seen, there are several critiquing properties that can be associated with this treatment trace. Some of them, for instance hormone-therapy+palliative-chemotherapy, are TRUE.

5.1.7 Concluding remarks

It turns out that, due to the difficulty of mapping real patient data to the ASBRU plans, and particularly representing the patient condition, a lot of the actual treatment traces in fact are non-compliant with the guideline recommendations.

For instance, from the 7 patient prototypes, only 4 can be considered to follow the guideline recommendations. One of the causes of this non-compliance, apart from incomplete mappings of patient data to the model, is the fact that other sections of the guideline might well mention the situations encountered in real practice, but they were not included in the recommendations, and therefore not present in the formal guideline model.

5.2 Critiquing Realworld Case Studies

Here, we present two case studies, extracted from the New South Wales Breast Cancer Institute, Australia ¹. In the next subsection, we present these informally, then we formalise them by attaching (mapping) the patient state and treatments given to the Asbru model. Finally, we apply model checking to find out whether or not the treatment given is consistent with the model.

5.2.1 Informal description

We will first briefly list the description of two patients and the treatment that was given to them.

Patient 1: Ductal carcinoma in situ in a male patient

The first patient is a male diagnosed with a ductal carcinoma in situ by means of an ultrasound and a bilateral mammography. The findings were as follows.

- Lesion was found in the left breast using an ultrasound
- Bilateral mammography: "fatty tissue is seen in the right breast only. On the left side there is coarse reticular stranding in the tissues deep to the nipple with indrawing of the nipple consistent with previous treatment. No focal dominant or suspicious mass was seen. No clusters micro calcification was identified."

The treatment for this patients consisted of a total mastectomy (with removal of the nipple). No treatment of the axilla or systemic therapy was conducted.

Patient 2: Infiltrating ductal carcinoma / use of lymphoscintigraphy

The second patient we consider is a female with an infiltrating ductal carcinoma in the right breast. Two lesions in the right breast were found in a histopathology. The finding for this patient were as follows:

- Lump in the 3 o'clock position of right breast (upper inner quadrant) and a second lump just above this. No palpable axillary nodes or other abnormalities.

¹<http://www.bci.org.au/medical/caseindex.htm>

- Mammography: no focal mass, grouped microcalcifications or anatomic distortion seen.
- Histopathology : two lesions: both infiltrating duct carcinoma, 20mm in size, similar morphology

The sentinel nodes were mapped using lymphoscintigraphy and a biopsy was taken of a right axillary lymph node and an internal mammary node. In the right axillary lymph node, no malignancy was found. However, in the internal mammary node, metastatic carcinoma (5mm x 3mm) was identified.

The treatment consisted of a total mastectomy of the right breast with immediate reconstruction. The axilla was treated by means of an axillary clearance and resection of two further internal mammary nodes at higher levels (these were sampled partly because of the original pathology finding and partly because of ready access to the IMC.)

5.2.2 Attachment to the model

For critiquing, we will focus on the treatments and some specific interventions that are extensively discussed in the guideline (e.g., sentinel node biopsy). that the patient received. The guideline we use for critiquing has paragraphs dealing with diagnosis, but its main focus is on treatments, which makes it harder to reliably attach some diagnostic tests to the model. We will also abstract from the breast reconstruction, as the guideline only mentions this shortly and does not provide conclusive recommendations, i.e., it only provides *slight preferences*.

Patient 1

The patient was diagnosed with ductal carcinoma in situ as no indications that the tumour invaded was found. Hence, we assume the diagnosis is DCIS. The only treatment that the patient is mastectomy, which is called `mastectomy` in the Asbru model.

Thus, the idea is to select those traces in the model that does not contain DCIS and some time in the future mastectomy is activated.

In CTL, this can be formalised as:

$$\begin{aligned} & E(DCIS \wedge \neg \text{dealing-with-axilla} \\ & \quad U (\text{mastectomy} \wedge EG(DCIS \wedge \neg \text{dealing-with-axilla}))) \end{aligned}$$

The modular proof obligation is:

$$[G(DCIS \wedge \neg \text{dealing-with-axilla}), F \text{mastectomy}] M \langle \perp \rangle$$

In case this generates a counter model, then we know that this is a valid trace in the model.

Patient 2

The second patient was diagnosed with in infiltrating breast cancer. For a proper attachment to the model, the question is whether or not the carcinoma is operable. According to the guideline, operable invasive BC is defined as T1-2 N0-1 M0, i.e., a tumour smaller than 5cm, with maximally 1 lymph node positive and no distant metastasis. Clearly, on basis of the diagnostic tests, this classification may be assumed for this patient, and consequently, the diagnosis is assumed to be `operable-invasive-BC`.

The first action that is taken here is a biopsy of the internal mammary node and an additional axillary lymph node. The latter biopsy is not mentioned in the guideline. Therefore, we will only attach it to the `axilla-staging-by-SN`. Furthermore, a mastectomy is performed, which can be attached to `mastectomy`, similarly to the previous patient. Finally, the axillary node dissection can be attached to `axillary-node-dissection`, even though the the additional resection of two further internal mammary node is not mentioned by the guideline.

The only explicit information about the order in which the actions are performed is that the sentinel node procedure is performed before axillary node dissection and mastectomy. Thus, we define a property which states that first sentinel node procedure is done after which the axillary node dissection and mastectomy may be done in any order.

In CTL, this property is represented as:

$$\begin{aligned}
& E (\text{operable-invasive-BC} \\
& \quad U (\text{axilla-staging-by-SN} \\
& \quad \quad \wedge (E (\text{operable-invasive-BC} \\
& \quad \quad \quad U (\text{axillary-node-dissection} \\
& \quad \quad \quad \quad \wedge E (\text{operable-invasive-BC} \cup (\text{mastectomy} \\
& \quad \quad \quad \quad \quad \quad \wedge EG \text{operable-invasive-BC})))) \\
& \quad \quad \vee E (\text{operable-invasive-BC} \\
& \quad \quad \quad U (\text{mastectomy} \\
& \quad \quad \quad \quad \wedge E (\text{operable-invasive-BC} \cup (\text{axillary-node-dissection} \\
& \quad \quad \quad \quad \quad \quad \wedge EG \text{operable-invasive-BC}))))))
\end{aligned}$$

Using modular model checking, we may represent the same as follows:

$$[G \text{operable-invasive-BC}, F(\text{axilla-staging-by-SN} \wedge F \text{axillary-node-dissection} \wedge F \text{mastectomy})] M \langle \perp \rangle$$

5.2.3 Translation to the SMV model

There are a number of differences between the specification of properties presented earlier and the implementation in SMV. First, the logical language used in SMV is slightly different than standard notation. However, as it is completely obvious what is meant, we will ignore this. Second, the Asbru model in SMV abbreviates some of the names, e.g., `dwa` = `dealing-with-axilla`, `mast` = `mastectomy`. In context of the properties, this will be obvious. Furthermore, the execution of actions is interpreted as usual in Asbru, i.e., action α is performed iff $\alpha = \text{activated}$. Finally, the SMV of breast cancer has a number of initialisations of variables in the first transition of the model. Therefore, variables may be in an `unknown` in the initial state. Consequently, all formulas starting with a global modality (G) are taken from the second step, e.g., $EG \varphi$ is translated to $EX EG \varphi$.

5.2.4 Results

Patient 1

The CTL property in SMV is presented as follows:

```

prop1_ctl:
SPEC EX E ( (diagnosis = DCIS & ~(dwap_state = activated))
            U ( mastp_state = activated
              & EG (diagnosis = DCIS & ~(dwap_state = activated)))
            );

```

and was proven to be true:

Model checking results

=====

prop1_ctl.....true

Resources used

=====

user time.....801.35 s

system time.....3.15 s

BDD nodes allocated.....16414408

Using modular model checking, we have:

```

patient1_a: assert X G (diagnosis = DCIS);
patient1_b: assert G ~(dwap_state = activated);
patient1_c: assert F (mastp_state = activated);
false: SPEC 0;

```

using patient1_a, patient1_b, patient1_c prove false;

resulting in

Model checking results

=====

false.....false

Resources used

=====

user time.....33.06 s

system time.....0.17 s

BDD nodes allocated.....1655198

Patient 2

The CTL representation of the SMV property:

```

prop2_ctl:
SPEC EX E ( (diagnosis = operable_invasive_BC)
            U ( asbSNp_state = activated
              & ( E ( (diagnosis = operable_invasive_BC)
                  U ( andp_state = activated
                    & E ( (diagnosis = operable_invasive_BC)
                      U ( (mastp_state = activated)

```

```

        & EG diagnosis = operable_invasive_BC
    )))
)
| E ( (diagnosis = operable_invasive_BC)
  U ( mastp_state = activated
    & E ( (diagnosis = operable_invasive_BC)
      U ( (andp_state = activated)
        & EG diagnosis = operable_invasive_BC
      )))
)
));

```

which results in

Model checking results
 =====

prop2_ctl.....false

Resources used
 =====

user time.....575.67 s
 system time.....2.39 s
 BDD nodes allocated.....17624319

The modular proof obligation is as follows:

```

patient2_a: assert X G (diagnosis = operable_invasive_BC);
patient2_bcd: assert F ( asbSNp_state = activated
  & F ( andp_state = activated
    & mastp_state = activated));

```

false: SPEC 0:

giving:

Model checking results
 =====

false.....true

Resources used
 =====

user time.....544.56 s
 system time.....2 s
 BDD nodes allocated.....15254313

possible error in the treatment of the patient. It is not clear how to continue exactly now, because we do not acquire a counter-model. Continuing with modular model checking, suppose we relax the constraints and ignore the order between the two sentinel node dissection and further treatment. We then prove:

```

patient2_a: assert X G (diagnosis = operable_invasive_BC);
patient2_b: assert F (asbSNp_state = activated);
patient2_c: assert F (andp_state = activated);
patient2_d: assert F (mastp_state = activated);
false: SPEC 0;

```

```

using patient2_a, patient2_b, patient2_c, patient2_d prove false;

```

which shows that this is also not possible:

Model checking results

=====

```

false.....true

```

Resources used

=====

```

user time.....227.34 s
system time.....0.98 s
BDD nodes allocated.....7264893

```

Now, the question is which of the actions of {axilla-staging-by-SN, axillary-node-dissection, mastectomy} is not allowed according to the model. This can, again, be investigated using model checking by removing one of the actions from the list of assertions. Here we summarise the result:

- {axilla-staging-by-SN, mastectomy}: trace exists

Resources used

=====

```

user time.....216.97 s
system time.....0.86 s
BDD nodes allocated.....7645471

```

- {axillary-node-dissection, mastectomy}: trace exists

Resources used

=====

```

user time.....265.95 s
system time.....1.38 s
BDD nodes allocated.....6374944

```

- {axilla-staging-by-SN, axillary-node-dissection}: traces does *not* exist

Resources used

=====

```

user time.....803.14 s
system time.....58.83 s
BDD nodes allocated.....20676737

```

Hence, in this case, the cause for the mismatch seems to be that the axillary-node-dissection cannot be combined with axilla-staging-by-SN according to the model.

The reason is that plans axilla-staging-by-SN and axillary-node-dissection are sub-plans of dealing-with-axilla. This plan is defined as follows:

```
<plan name="dealing-with-axilla">
  <plan-body>
    <subplans retry-aborted-subplans="no" type="any-order"
      wait-for-optional-subplans="no">
      <wait-for><one/></wait-for>
    <plan-activation>
      <plan-schema name="axilla-staging-by-SN"/>
    </plan-activation>
    <plan-activation>
      <plan-schema name="complete-axillary-node-dissection"/>
    </plan-activation>
    (...)
  </plan>
```

Hence, as soon as either `axilla-staging-by-SN` or `complete-axillary-node-dissection` has completed, then `dealing-with-axilla` completes. Whether or not this is corresponding to the guideline is not the issue here, as we are assuming this model is correct (although there are strong reasons to suspect the model is incorrect). In this case we could report back to the clinician that he should make a choice between the sentinel node procedure and axillary node dissection.

5.2.5 Concluding Remarks

A number of conclusions can be drawn from these examples. First, CTL formulas are generally unintuitive when the order of actions performed is omitted. The alternative, using a combination of LTL and CTL model checking seems to be feasible for the examples that were shown here. Most of time spent model checking is used for building the BDD model, while the actual verification of the property was relatively brief. Theoretically speaking, the complexity of LTL model checking is much higher, but this was not observed in practice only the few examples presented here.

With respect to model checking, it was not completely clear what to do when the treatment path provided by the clinicians does not comply with the guideline. Model checking results do not provide any clue why it was incorrect, as it never provides a usable counter-example. The only thing that can be done is relaxing the constraints that the trace imposes, but again, not obvious how to relax exactly. Further work is needed to relax the constraints imposed by the property under investigation more systematically.

5.3 Analysis of the results

One desirable property of critiquing, apart from proposing an alternate solution to the one executed, is to explain the cause of non-compliance between a treatment path and the guideline recommendations. This would enable guideline developers to propose changes of the medical practice or to suggest penalties for faulty behaviour.

By studying statistics of compliance with evidence-based practice, the quality of healthcare for a particular hospital can be evaluated.

5.3.1 Analyzing non-compliant treatment traces

Once a treatment has been found to be non-compliant with the guideline recommendations, we have to detect the root cause of this non-compliance: is it because the guideline did not mention the case at all, is it because of an exceptional situation (such as, the patient taking part in a clinical trial, not covered yet by the guideline), or is it a fault of the doctor?

There are several possibilities: the guideline is ambiguous and incomplete and requires update; or the guideline is correct and the doctor did not follow its recommendations. In the latter case, if no exceptional situation can explain the non-compliance of the treatment with the guideline, a penalty has to be devised.

It should also be investigated whether the deviation of the treatment from the guideline recommendations are small or serious.

5.3.2 Suggesting treatment improvements based on Model Checking

In general, given the form of the properties we verify using Model Checking ($EFa \wedge EFb$), the counterexamples generated in case of non-compliance are very unreliable.

If the counterexample can be traced back to the guideline text, and interpreted by the medical expert, then the following situations can be envisioned: 1. The treatment trace is correct and does not conflict with the guideline recommendations. In this case a request-for-evidence about that case, or even a request-for-update of the guideline text can be issued, to cover the specific case the doctor encountered. 2. The treatment trace is correct but conflicts with one or several recommendations; in this case a request-for-update of the guideline text can be issued, to resolve this conflict. 3. The treatment trace is faulty. In this case, a warning can be delivered to the doctor.

5.3.3 Refinement of the critiquing property

As part of explaining the root cause of non-compliance, the critiquing property can be change, to accomodate a less restrictive property. For instance, if patient P has undergone treatment T, we would like to answer questions such as: "Under which set of conditions CE which extend P's medical condition C, T is not taken". This would allow us to find the relevant part of the guideline which could explain why T is not executed.

One type of critiquing is to check whether T represents a treatment for a complementary condition of C. In that case T might have been wrongly applied: $C \rightarrow AF\neg T$.

Another type of extended critiquing is to check whether T does not represent a contra-indicated treatment for C: $AF(C \rightarrow \neg T)$.

If the failing critiquing property is: $EF(C \wedge T)$, then we would like to find out instances in which another alternative treatment could have been selected: $EF(C \wedge \neg T)$, as well as to guarantee that if T is executed, then always C is the case: $AF(T \rightarrow \neg C)$.

Chapter 6

Concluding Remarks

This deliverable describes the methodology for critiquing the reference guideline (CBO guideline for treatment of breast cancer, [6]) using model checking based on an SMV model generated from the ASBRU guideline model.

From a technical point of view, this work has shown that it is indeed possible in principle to use model-checking on formalized models in order to critique medical guidelines against patient-records. The strong aspect of this technology is the high degree of automation as compared to theorem proving, making it correspondingly easier to use.

The main difficulty encountered in critiquing the patient data using the guideline model is the different level of abstraction of the vocabularies which describe the patient in the clinical records and in guidelines. Different names are used for treatments, and even encoded, and without a clear specification of the encodings, it is very difficult to have a valid critique of a guideline based on data interpretation. This problem has been solved by manually building a set of mappings between the two terminologies, but in the future this can be done when the guideline is proposed. A possible improvement from which critiquing might benefit would be to include in the appendix of the guideline a patient data model and action model, along with the measurable parameters of the medical process, which allows clinical records to include information used for critiquing.

Appendix A

Appendix A - Sample Patient Traces

This section includes the most common prototypes for treatment of breast cancer, sampled from the patient data. They were selected to illustrate how two patients with the same condition receive different treatments and what medical conditions are most commonly treated with a standard treatment.

A.1 - Selected Patient Traces

SELECTED PATIENT #4

Medical condition: 79 years-old woman. Overlapping lesion of right breast. Invasive ductal carcinoma grade 2 mixed with carcinoma in-situ. Size between 1 and 2 cm. Lymph nodes: 0/2. Treatment: organ specific surgery and sentinel node biopsy.

SELECTED PATIENT #21

Medical condition: 44 years-old woman. Lower-outer quadrant lesion of right breast. Invasive ductal carcinoma grade 2 mixed with carcinoma in-situ. Size between 1 and 2 cm. Lymph nodes: 0/2. Treatment: organ specific surgery, sentinel node biopsy and chemotherapy.

SELECTED PATIENT #22

Medical condition: 80 years-old patient. Central portion lesion of right breast. Lobular and ductal carcinoma (in-situ or invasive) grade 1. Size greater than 5 cm. Lymph nodes: 0/10. Treatment: mastectomy, external radiotherapy and hormone-therapy.

SELECTED PATIENT #34

Medical condition: 82 years-old woman. Overlapping lesion of the left breast. Adenocarcinoma. Size greater than 5 cm. Lesion with oedema or ulceration of the skin or satellite skin nodes. Unknown grade. Axillary status unknown. Metastasis in supraclavicular lymph nodes. Treatment: chemotherapy and hormone therapy.

SELECTED PATIENT #36

Medical condition: 70 years-old patient. Upper-inner quadrant lesion of the left breast. Lobular carcinoma grade 1. Pre-treatment size between 2 and 5 cm. Lymph nodes: 0/1. Treatment: organ specific surgery, sentinel node biopsy and external radiotherapy. Post-treatment size between 1 and 2 cm.

SELECTED PATIENT #54

Medical condition: 56 years-old patient. Overlapping lesion of the left breast. Ductal carcinoma mixed with tubular carcinoma. Grade 1. Size between 0.5 and 1 cm. Lymph nodes: 1/5. Treatment: organ specific surgery, sentinel node biopsy and external radiotherapy.

SELECTED PATIENT #71

Medical condition: 50 years-old patient. Overlapping lesion of the left breast. Morbus Paget and ductal carcinoma in situ. Grade 3. Pre-treatment size between 2 and 5 cm. Lymph nodes: 1/5. Internal mammary lymph nodes with microscopic metastasis detected by sentinel node dissection but not clinically apparent. Pre-treatment metastasis in movable ipsillary axillary lymph node. Treatment: mastectomy, chemotherapy and radiotherapy. Post-treatment size between 1 and 2 cm.

A.2 - Medical observations on patient treatment traces

Among the patient data actions, the following actions are most frequently encountered:

1. axillary-node-dissection: done in all tumours when lymph node extension should be assessed
2. axillary-staging-by-SN (staging of the axilla by Sentinel Node procedure)
3. bct and breast-conserving therapy: this is the same as tumorectomy or lumpectomy; applicable to (T1 N0) tumours
4. biopsy: this means tumour biopsy; it is always performed, in any stage of the tumour
5. limited-surgery-manual, grm, grm-or-mst, locoregional-grmmst: these concepts correspond to some types of mastectomy, but should be confirmed by those who proposed the concepts, they do not look compliant with the BC treatment taxonomy
6. mst and mastectomy and mastectomy-proper, possible-mastectomy, salvage-mastectomy, radical-surgery-manual, re-excision, surgery-manual and surgery-proper, microscopically-complete-tumour-excision, tumour-excision: all these refer to surgery

Based on the test prototypes defined by the medical experts, the following paths of actions are most common in breast cancer treatment:

1. Surgery: ONE-OF(radical mastectomy, mastectomy, or lumpectomy/bct)
2. Surgery + adjuvant chemotherapy (if hormone positive)
3. Surgery plus postoperative radiotherapy and/or chemotherapy
4. Chemotherapy without surgery (metastatic breast cancer)

All other actions represent intermediate actions for disease staging.

Treatment options

Surgery can be one of the following actions: 1. Radical mastectomy (breast extirpation and lymph nodes extirpation) 2. Mastectomy (breast extirpation) 3. Tumorectomy or lumpectomy (breast conservative surgery)

In the typical case, always surgery is done to address the breast cancer, except in cases of advanced breast cancer in which only palliative care is provided

Breast cancer reconstruction: cosmetic surgery 1 year after the treatment

Radiotherapy:

- It is performed to reduce tumour size (presurgical approach) or to prevent recurrence (postsurgery).
- It can be 2D or 3D with different dosage and placed on the tumour area or in tumour area and local lymph nodes.

Chemotherapy:

- To treat advanced breast cancer (metastatic) and/or to prevent recurrence when there is a chance of it.
- It can be: Palliative: postsurgery or without surgery; or Adjuvant chemotherapy

Hormonotherapy:

- It is a chemotherapy based on antiestrogens for estrogen positive tumours: inhibitors aromatase, tamoxifen, raloxifen (in patient dataset you might find new treatments under randomized controlled trials)
- A new treatment is proposed for herceptin positive cases.

Medical experts classify breast cancer types according to the so-called TNM classification (T is tumour size, N is lymph node affectation and M is metastasis):

T Tumour size: • Tis Non invasive BC

- T1 less than 2 cms of diameter
- T2 More than 2 cms and less than 5 cms
- T3 More than 5 cms
- T4 Any size and affecting chest wall or skin

N Lymph node affectation: • N0 Not spread to lymph nodes

- N1 Spread to the lymph nodes under the arm on the same side as the BC.
- N2 Spread to 4 to 9 lymph nodes under the same side as the BC
- N3 Spread to 10 or more lymph nodes under the same arm side as the BC and top other areas around the breast.

M Metastases: • M0 No distant cancer spread

- M1 Cancer has spread to any other organ

The relations between the breast cancer types as characterized by their TNM attributes and the chapters of the CBO guideline are given below.

Chapter 1 covers TNM-stage T1-T2 No0-N1 M0, with tumour size less than 5 cm (in some part of the text less than 4 cm) independent of lymph node affectation (N0-N1) and no metastases. It focuses also on BCT. Chapter 2 refers to the indication of adjuvant chemo. The T value on the TNM staging system is indicated by tumour size. When the guideline refers to N positive nodes, it does not refer to lymph node affectation, but to being positive or negative to hormone response. Chapter 3: refers to T3 T4 N0 N1 M0 Chapter 4: refers to patient follow up with three specific cases: recurrence, contralateral BC and metastasis. Chapter 5: recurrent breast cancer, T3 T4. Chapter 6: refers to M1, existence of metastasis. In real patient data all affectation outside the breast is classified as M1 to simplify analysis.

It should be noted that the selected patient treatment traces are more specific, but can in principle be mapped to one of the breast cancer prototypes build by medical experts based on TNM-staging:

1. Stage 0: Ductal carcinoma in situ: Tis N0 M0: surgery
2. Stage 0: Lobular carcinoma in situ: Tis N0 M0: surgery
3. Stage I: T1 N0 M0: surgery
4. Stage IIA: T0 or T1 N1 M0 or T2 N0 M0: surgery and possible other approaches
5. Stage IIB: T2 N1 M0 or T3 N0 M0: surgery plus other approaches
6. Stage IIIA: T0 N2 M0, T1 N2 M0, T2 N2 M0, T3 N1 M0, T3 N2 M0 (Important M0 and surgery plus different approaches)
7. Stage IIIC: Any T N3 M0: surgery plus other approaches
8. Stage IIV: Any T Any N M1 (palliative care that may include surgery).

Appendix B

Mapping of Patient Data to the Asbru Model

This section contains all mappings between parameters in patient data and ASBRU parameters, plus the mappings between patient treatment actions and ASBRU plans (tables [B.1](#) and [B.2](#), respectively).

Patient parameter	ASBRU equivalent
age	age-range,menopausal-status
tumour size	tumour-range
diagnosis	diagnosis
ctnmn,cn,ptnmn,pn (TNM parameters in clinical/pathological tumour staging)	metastasis
tumour grade	tumour-greater-T2, tumour-state-cT4
amount of positive nodes	amount-of-positive-nodes
tumour morphology	multiple-tumour-foci
tumour differentiation grade	differentiation-grade
hormonal sensitivity	hormon-receptor-levels,ER,PgR

Table B.1: Mapping patient parameters to ASBRU

Patient parameter	ASBRU equivalent
radiotherapy NNO	<i>radiotherapy</i> <i>radiotherapy – for – T1N0 – bc</i> <i>initial – radiotherapy</i> <i>subsequent – radiotherapy</i> <i>dcis – radiotherapy</i> <i>highdose – radiotherapy</i> <i>radiotherapy – chest – wall</i> <i>lowdose – reradiation</i> <i>radiation</i> <i>hyperthermia</i> <i>locoregional – radiotherapy – manual</i> <i>great – boost</i> <i>standard – boost</i> <i>optional – boost</i> <i>locoregional – postoperative – radiotherapy</i> <i>primary – radiotherapy – of – axilla</i>
chemotherapy NNO	<i>cmf – chemotherapy</i> <i>palliative – chemotherapy</i> <i>neoadjuvant – chemotherapy</i> <i>six – courses – anthracycline – chemotherapy</i> <i>single – course – anthracycline – chemotherapy</i> <i>secondary – adjuvant – chemotherapy</i> <i>anthracycline – containing – regimen</i> <i>lumbally – citostatic</i> <i>palliative – treatment</i>
hormone therapy NNO	<i>secondary – adjuvant – hormone – therapy</i> <i>adjuvant – hormonal – therapy</i> <i>secondary – adjuvant – hormone – treatment</i> <i>androgens – therapy</i> <i>estrogens – therapy</i> <i>anti – estrogens – therapy</i> <i>aromatase – inhibitors – therapy</i> <i>biophosphonates</i> <i>combination – of – trastuzumab – and – paclitaxel</i> <i>high – dose – dexamethasone</i> <i>herceptin – treatment</i> <i>palliative – treatment</i>

Table B.2: Mapping patient actions to ASBRU

Appendix C

Appendix C: Detailed Analysis of the Selected Patients

This section contains details of the steps of preparing the knowledge for the critiquing process.

The 7 patients selected as representative with respect to the diagnosis and treatment combination, are analyzed from a medical viewpoint, based on the output of an application from processing their (anonymized) patient record. Patient 4 means "patient in row 4" in the file containing patient records.

For critiquing of medical actions for treatment of breast cancer, using the recommendations of the CBO guideline for treatment of breast cancer, we need to establish the mappings between the actions that occur in the formal model of the guideline and the actions that are recorded in the patient data.

This section summarizes the critiquing based on mappings of the names of actions and other details present in patient data (denoted PatientData.Actions and PatientData.Fields, respectively) to names used in the guideline model (this is an executable model in ASBRU). We sketch the interpretation of a few examples of real patient treatment paths (each patient has different diagnosis). We used a program that reads the anonymized patient data recorded by the Cancer Institute and based on the data encoding rules from the codebook produces a RAW data analysis which is further used in the critiquing of the medical process.

C.1. Patient 4

```
=====  
Machine readable data on PATIENT 4:  
=====
```

```
patient_trace(4,patient_data([sex(female), age(79),  
birthdate('1923')]), diagnosis([name('Invasive ductal carcinoma (mixed  
invasive and in-situ)'), basisDiagnosisCode('7')]),  
tumour_info([nr_of_tumour(2),  
total_tumours(2),localization('overlapping lesion of breast'),  
laterality('right'), tumour_extension(3), tumour_grade(2),  
tumour_differentiation_grade('moderately differentiated'),  
topology('overlapping lesion of breast'),  
behaviour('malignant/invasive, primary site'), morphology('ductal
```

```

carcinoma'),
nr_lymph_nodes([total(2),positive(0)]),metastases([],
nr_metastases_in_lymph_nodes([total(0),investigated(0)]),
additional_tumour_codes([typeV1('rectum'),typeV2(''),typeV3('')]),
clinical_staging_results_before_therapy([stage_grouping_clinical('4'),
clinical_tumour_staging('not applicable'),
ct('tumour having between 1cm and 2cm in greatest
dimension'),cn('regional lymph nodes cannot be assessed'),cm('no
distant metastasis'),ctnmt('primary tumour cannot be
assessed'),ctnmm('no regional lymph node metastasis'),ctnmm('unknown
clinical classification code for staging<1C>'))]),
pathological_staging_results_after_therapy([
stage_grouping_pathological('unknown tumour stage group(1C,X,0)'),
pathological_tumour_staging('not applicable'),
pt('tumour having between 1cm and 2cm in greatest
dimension'),pn('regional lymph nodes metastasis cannot be assessed
(regional lymph nodes not removed for study or previously
removed)'),pm('no distant metastasis'),ptnmt('tumour having between
1cm and 2cm in greatest dimension'),ptnmm('regional lymph nodes
metastasis cannot be assessed (regional lymph nodes not removed for
study or previously removed)'),ptnmm('no distant metastasis'))]) ]

treatment([ i1('sentinel node biopsy'), i2('organ specific surgery
(organ localization code=14)')])

```

Medical interpretation of raw data for Patient 4

Information needed to decide the optimal treatment:

- Comorbidity.
- Hormonal status of the tumour.
- Size of carcinoma in-situ component.

Based on the information given we cannot define a specific treatment. It depends on other factors which are not specified, such as the already mentioned. So, if the tumour is ER1 or PR2 positive, the patient should receive hormonal treatment if having no medical contraindications. On the other hand, when having a breast conserving surgery, patients should be treated with complementary radiotherapy. The size of the carcinoma in-situ component would determine the kind of surgical procedure.

Technical interpretation of raw data for patient 4

```

=====
APPROXIMATE ASBRU TRANSLATION:
=====
PATIENT CONDITION:

```

Invasive ductal carcinoma grade 2 mixed with carcinoma in-situ.
Tumour Size between 1 and 2 cm. Lymph node negative.

TREATMENT SEQUENCE:

1. i1(sentinel-node biopsy) + 2. i2(breast conserving therapy)

ATTACHMENT (mappings between patient data treatment actions and guideline actions):

```
i1('sentinel node biopsy') ?  
['axilla-staging-by-SN']  
  ['ultrasound-guided-core-biopsy']  
  ['sentinel-node-procedure']
```

```
i2('organ specific surgery (organ localization code=14)') =  
  lumpectomy (=breast conserving therapy?) =  
  ['bct']  
  ['breast-conserving-therapy']
```

=====

ASBRU TRANSLATION:

=====

PATIENT CONDITION:

diagnosis = DCIS AND age = 79 AND age-range=old AND
(tumour-size-range >1cm and tumour-size-range < 2cm)

TREATMENT SEQUENCE:

```
i1 = ONE-OF(  
  histological-biopsy(hibi/P79/CH5)|biopsy(bio/P116/CH6)|  
  ultrasound-guided-core-biopsy(ugb/P50/CH1)|re-excision(reex/P42/CH1)|  
  axillary-staging-by-SN(andp/P2/CH1)|sentinel-node-procedure(snp/P105/CH5)|  
  tumour-excision(te/P108/CH5)|)
```

```
i2 = ONE-OF(  
  limited-surgery-manual(lsm/P59/CH3)|  
  salvage-mastectomy(salma/P103/CH5)|  
  breast-conserving-therapy(bct/P65/CH5)|  
  bct (bct/P5/CH1)|)
```

MOST LIKELY SEQUENCE:

1. axillary-staging-by-SN(andp/Plan2/Chapter 1) + 2. bct (bct/P5/CH1)

Formal translation of the real treatment trace as a critiquing property, to be verified in the model checker SMV against a specification of the guideline recommendations.

SMV TRANSLATION: Attached to chapter 1:

```

EF (
-- medical condition
(diagnosis=DCIS AND
-- initial plan states
chap1p_state=activated & asbSNp_state=inactive & bctp_state=inactive)
AND EF (
-- plan states after event 1: asbSN is activated
(chap1p_state=activated & asbSNp_state=activated &
bctp_state=inactive) & EF (
-- plan states after event 2: asbSN is completed
(chap1p_state=activated & asbSNp_state=completed &
bctp_state=inactive) & EF (
-- plan states after event 3: bct is activated
(chap1p_state=activated & asbSNp_state=inactive &
bctp_state=activated) & EF (
-- plan states after event 4: bct is completed
(chap1p_state=activated & asbSNp_state=inactive &
bctp_state=completed) & EF (
-- plan states after event 5: top level plan is completed
chap1p_state=completed
))))))

```

PROPERTY is FALSE.

Based on the output of the model checker and on the analysis detailed in section 5.1, we conclude that this patient has not been treated according to the guideline model. An informal analysis by the medical expert shows that, given the data available about patient, the medical expert cannot establish the correct sequence.

C.2. Patient 21

```

=====
Machine readable data on PATIENT 21:
=====
patient_trace(21,patient_data([sex(female), age(44),
birthdate('1957')]), diagnosis([name('Invasive ductal carcinoma (mixed
invasive and in-situ)') ]), tumour_info(localization('lower-outer
quadrant of breast'), laterality('right'), tumour_extension(3),
tumour_grade(2), tumour_differentiation_grade('moderately
differentiated'), topology('lower-outer quadrant of breast'),
behaviour('malignant/invasive, primary site'), morphology('ductal carcinoma'),
nr_lymph_nodes([total(2),positive(0)]),metastases([]),
nr_metastases_in_lymph_nodes([total(0),investigated(0)]),
additional_tumour_codes([typeV1('skin'),typeV2('skin'),typeV3('skin')])),
clinical_staging_results_before_therapy([
stage_grouping_clinical('4'),

```

```

clinical_tumour_staging('not applicable'),
ct('primary tumour cannot be assessed'),
cn('no regional lymph node metastasis'),cm('no distant metastasis'),
ctnmt('no evidence of primary tumour'),
ctnmn('no regional lymph node metastasis'),
ctnmm('unknown clinical classification code for staging<1C>')),
pathological_staging_results_after_therapy([
stage_grouping_pathological('unknown tumour stage group(1C,X,0)'),
pathological_tumour_staging('not applicable'),
pt('tumour having between 1cm and 2cm in greatest dimension'),
pn('regional lymph nodes metastasis cannot be assessed
(regional lymph nodes not removed for study or previously
removed)'),pm('no distant metastasis'),ptnmt('tumour having between
1cm and 2cm in greatest dimension'),ptnmn('regional lymph nodes
metastasis cannot be assessed (regional lymph nodes not removed for
study or previously removed)'),ptnmm('no distant metastasis')]
]),status_record([patient_status(unknown),tumour_residu('no residual
tumour after treatment'), deceased('unknown'),

treatment([ prescribed_therapies([i1('sentinel node biopsy'),
i2('organ specific surgery (organ localization code=14)'), i3('organ
specific surgery (organ localization code=12)'), i4('chemotherapy
NNO')])]),

actual_treatment(
[event1(surgery1,'11-06-2002',1023746400000),
event2(consult,'06-11-2002',1036537200000),
event3(surgery2,'26-11-2002',1038265200000),
event4(chemotherapy,'01-01-2003',1041375600000)]).

```

Medical interpretation of raw data for Patient 21

Information needed to decide the optimal treatment:

- Comorbidity.
- Hormonal status of the tumour.
- Size of carcinoma in-situ component.
- Type of surgical procedure: breast conserving surgery or mastectomy.

Based on the information given we cannot define a specific treatment. It depends on other factors which are not specified, such as the already mentioned. Chemotherapy is a complementary option to surgical treatment if the tumour is ER or PR negative. This decision is not that clear when having ER or PR positive tumours. In any case, if the tumour is ER or PR positive, the patient should receive hormonal treatment if having no medical contraindications, always after receiving chemotherapy if it is considered a good option. On the other hand, when having a breast conserving surgery, patients should be treated with

complementary radiotherapy. The size of the carcinoma in-situ component would determine the kind of surgical procedure.

Technical interpretation of raw data for patient 21

APPROXIMATE ASBRU TRANSLATION:

PATIENT CONDITION= Invasive ductal carcinoma grade 2 mixed with carcinoma in-situ. Size between 1 and 2 cm.

TREATMENT SEQUENCE:

i1('sentinel node biopsy') +
i2('organ specific surgery (code=14)')=lumpectomy +
i3('organ specific surgery (code=12)')=ablation+
i4('chemotherapy NNO')

ATTACHMENT (mappings between patient data treatment actions and guideline actions):

i1('sentinel node biopsy') ?
['axilla-staging-by-SN']
['ultrasound-guided-core-biopsy']
['sentinel-node-procedure']

i2('organ specific surgery (organ localization code=14)') =
lumpectomy (=breast conserving therapy?) =
['bct']
['breast-conserving-therapy']

i3('organ specific surgery (organ localization code=12)') =
ablation (=breast conserving therapy?) =
['bct']
['breast-conserving-therapy']

i4('chemotherapy NNO') ?
['cmf-chemotherapy']
[' palliative-chemotherapy ']
['neoadjuvant-chemotherapy?']
['six-courses-anthracycline-chemotherapy?']
['single-course-anthracycline-chemotherapy?']
['secondary-adjuvant-chemotherapy?']
['anthracycline-containing-regimen?']
['high-dose-dexamethasone?']
['lumbally-citostatic?']

FORMAL TRANSLATION:

PATIENT CONDITION

(diagnosis=DCIS AND tumour-size-range > 1cm AND tumour-size-range < 2cm)

MOST LIKELY TREATMENT SEQUENCE:

1. axilla-staging-by-SN +
2. bct OR surgery-proper (11-06-2002) +
3. breast-reconstruction OR microscopically-complete-tumour-excision
OR bct(26-11-2002)
4. cmf-chemotherapy (01-01-2003)

CRITIQUING:

CHECK WHETHER

WHEN diagnosis=DCIS THEN

IS_POSSIBLE_SEQUENCE (

axilla-staging-by-SN +

bct +

microscopically-complete-tumour-excision +

cmf-chemotherapy)

MOST GENERAL MAPPING:

i1 = ONE-OF(

axillary-staging-by-SN(andp/P2/CH1) |

sentinel-node-procedure(snp/P105/CH5) |

i2 = ONE-OF(

breast-conserving-therapy(bct/P65/CH5) |

surgery-proper (sp/CH1) |

bct (bct/P5/CH1) |))

i3 = ONE-OF(

bct(bct/P5/CH1) |

limited-surgery-manual(lsm/P59/CH3) |

mst (Mamma Sparende Therapie)(mst/P92/CH5) |)

i4 = ONE-OF(

cmf-chemotherapy(cmf/P123/CH6) |

palliative-chemotherapy(pache/P150/CH6) |

neoadjuvant-chemotherapy(neo/P60/CH3) |

six-courses-anthracycline-chemotherapy(scac/P62/CH3) |

single-course-anthracycline-chemotherapy(sgcac/P63/CH3) |

secondary-adjuvant-chemotherapy(sac/P100/CH5) |

anthracycline-containing-regimen(acr/P112/CH6) |

combination-of-trastuzumab-and-paclitaxel(cotap/P124/CH6) |

high-dose-dexamethasone(hdd/P138/CH6) |

herceptin-treatment(hertr/P140/CH6) |

```
lumbally-citostatic(lucit/P147/CH6)|
palliative-treatment(paltr/P152/CH6)|)
```

SMV TRANSLATION: Attached to chapter 1:

A: axillary-staging-by-SN(andp/P2/CH1) + surgery-proper (sp/CH1) + bct (bct/P5/CH1)

B: axillary-staging-by-SN(andp/P2/CH1) + bct (bct/P5/CH1) + bct (bct/P5/CH1)

A:

```
SPEC EF ((diagnosis=DCIS & chap1p_state=activated &
andpp_state=inactive & spp_state = inactive & bctp_state=inactive) &
EF ((chap1p_state=activated & andpp_state=activated & spp_state =
inactive & bctp_state=inactive)) & EF ((chap1p_state=activated &
andpp_state=completed & spp_state = inactive & bctp_state=inactive) &
EF ((chap1p_state=activated & andpp_state=inactive & spp_state =
activated & bctp_state=inactive) & EF ((chap1p_state=activated &
andpp_state=inactive & spp_state = completed & bctp_state=inactive) &
EF ((chap1p_state=activated & andpp_state=inactive & spp_state =
inactive & bctp_state=activated) & EF ((chap1p_state=activated &
andpp_state=inactive & spp_state = inactive & bctp_state=completed) &
EF (chap1p_state=activated)))))) )
```

PROPERTY A is FALSE.

B:

```
SPEC EF ((diagnosis=DCIS & chap1p_state=activated &
andpp_state=inactive & bctp_state=inactive) & EF
((chap1p_state=activated & andpp_state=activated &
bctp_state=inactive)) & EF ((chap1p_state=activated &
andpp_state=completed & bctp_state=inactive) & EF
((chap1p_state=activated & andpp_state=inactive &
bctp_state=activated) & EF ((chap1p_state=activated &
bctp_state=completed) & EF ((chap1p_state=activated &
bctp_state=activated) & EF((chap1p_state=activated &
bctp_state=completed) & EF (chap1p_state=activated))) )) )
```

PROPERTY B is FALSE.

Anomaly detected in the attachment phase: no reference to chemotherapy in recommendations of chapter 1, despite the medical interpretation of it being a valid action. It might be caused by insufficient information in the guideline text (including insufficient link between chapter 1 and other chapters).

C.3. Patient 22

```
=====
Machine readable data on PATIENT 22:
```

```

=====
patient_trace(22,patient_data([sex(female), age(80),
birthdate('1923')]), diagnosis([name('Lobular and ductal carcinoma
(in-situ or invasive, or a combination thereof)') ,
basisDiagnosisCode('7')]), tumour_info([nr_of_tumour(2),
total_tumours(2),localization('central portion of breast'),
laterality('right'), tumour_extension(3), tumour_grade(1),
tumour_differentiation_grade('well differentiated'), topology('central
portion of breast'), behaviour('malignant/invasive, primary site'),
morphology('Lobular or ductal carcinoma'),
nr_lymph_nodes([total(10),positive(0)]),
metastases([],nr_metastases_in_lymph_nodes([total(0),investigated(0)]),
additional_tumour_codes([typeV1('nipple'),typeV2(''),typeV3('')]),
clinical_staging_results_before_therapy([stage_grouping_clinical('unknown
tumour stage group(0,0,3)'),clinical_tumour_staging('not
applicable'),ct('primary tumour cannot be assessed'),cn('no regional
lymph node metastasis'),cm('no distant metastasis'),ctnmt('no evidence
of primary tumour'),ctnmn('no regional lymph node
metastasis'),ctnmm('unknown clinical classification code for
staging<3>'))]),pathological_staging_results_after_therapy([
stage_grouping_pathological('2B'),pathological_tumour_staging('not
applicable'),pt('tumour of more than 5cm in greatest
dimension'),pn('no regional lymph node metastasis'),pm('no distant
metastasis'),ptnmt('tumour of more than 5cm in greatest
dimension'),ptnmn('no regional lymph node metastasis'),ptnmm('no
distant metastasis')) ]]),
treatment([ [i1('organ specific surgery (organ localization
code=11)'), i2('drug-based hormone therapy'), i3('external
radiotherapy')])])

```

Medical interpretation of raw data for Patient 22

Information needed to decide the optimal treatment:

- Comorbidity.
- Precise pathologic information of the two components: in-situ and invasive carcinoma.
- Hormonal status of the tumour.

Based on the information given, another treatment sequence would be also correct, always considering an ER or RP positive tumour. Locally advanced tumours with positive ER or PR can be treated up-front with neoadjuvant hormone-therapy therefore giving the option of a breast conserving surgery.

The indication of hormonal treatment should be based on the hormonal status of the primary tumour, which is something that is not specified.

Technical interpretation of raw data for patient 22

APPROXIMATE ASBRU TRANSLATION:

PATIENT CONDITION:

Invasive ductal carcinoma grade 2 mixed with carcinoma in-situ.
Size between 1 and 2 cm.

TREATMENT SEQUENCE:

i1('organ specific surgery (organ localization code=11)')=radical mastectomy+
i2('drug-based hormone therapy')+
i3('external radiotherapy')

ATTACHMENT (mappings treatment actions to guideline actions):

i1('organ specific surgery (organ localization code=11)')=
=radical mastectomy = ONE-OF(
mastectomy(mast/P91/CH5)|grm-2 (Modified Radical
Mastectomy)(grm2/P76/CH5)|
radical-surgery-manual(rsm/P61/CH3)|
limited-surgery-manual(lsm/P59/CH3)|)

i2('drug-based hormone therapy') = ONE-OF(
hormone-therapy(hothe/P141/CH6)|
secondary-adjuvant-hormone-therapy(sahth/P102/CH5)|
adjuvant-hormonal-therapy(aht/P51/CH3)|
systemic-therapy-3(syth3/P162/CH6)|
secondary-adjuvant-hormone-treatment(saht/P101/CH5)|
androgens-therapy(anthe/P115/CH6)|
estrogens-therapy(esthe/P132/CH6)|
anti-estrogens-therapy(aethe/P113/CH6)|
aromatase-inhibitors-therapy(aithe/P114/CH6)|
biphosphonates(bioph/P117/CH6)|
combination-of-trastuzumab-and-paclitaxel(cotap/P124/CH6)|
high-dose-dexamethasone(hdd/P138/CH6)|
herceptin-treatment(hertr/P140/CH6)|
lumbally-citostatic(lucit/P147/CH6)|
palliative-treatment(paltr/P152/CH6)|)

i3('external radiotherapy?) = ONE-OF(
radiotherapy(radio/P158/CH6)|
radiotherapy-for-T1N0-bc(rT1N0/P99/CH5)|
initial-radiotherapy(ir/P21/CH1)|
subsequent-radiotherapy(sr/P47/CH1)|
dcis-radiotherapy(dcis/P10/CH1)|
highdose-radiotherapy(hira/P80/CH5)|
radiotherapy-chest-wall(rcw/P41/CH1)|
lowdose-reradiation(lore/P85/CH5)|
radiation(radio/P97/CH5)|

```
intensive-remedial-therapy(irt/P21/CH1)|
hyperthermia(hyp/P81/CH5)|
locoregional-radiotherapy-manual(lrm/P58/CH3)|
locoregional-postoperative-radiotherapy(lrpr/P23/CH1)|
primary-radiotherapy-of-axilla(praoa/P40/CH1)|
```

```
=====
ASBRU TRANSLATION:
=====
```

```
PATIENT CONDITION
(diagnosis=DCIS AND tumour-size-range > 5cm)
```

```
TREATMENT SEQUENCE:
1. hormone-therapy (01-01-2004) +
2. mastectomy (01-09-2004) +
3. external radiotherapy=ONE-OF(
    locoregional-postoperative-radiotherapy|
    subsequent-radiotherapy) (16-02-2004)
```

```
SMV: Attached to chapter 1:
hormone therapy = undefined,
mastectomy = mast,
external radiotherapy = ONE-OF(lrpr|sr)
```

```
SPEC EF ((chap1p_state=activated & mastp_state=inactive &
lrprp_state=inactive & srp_state=inactive) & EF
(chap1p_state=activated & mastp_state=activated & lrprp_state=inactive
& srp_state=inactive & EF (chap1p_state=activated &
mastp_state=completed & (lrprp_state=inactive & srp_state=inactive) &
EF (chap1p_state=activated & (lrprp_state=activated |
srp_state=activated) & EF (chap1p_state = activated &
(lrprp_state=completed | srp_state=completed))))))
```

PROPERTY is FALSE. The counterexample has 64 steps.

Anomaly: Aside from not being able to detect any meaningful mapping for hormone-therapy in chapter 1 model, the last two actions in sequence cannot take place in the given order. Apparently, after mastectomy is completed, none of the two radiotherapy plans can get activated.

C.4. Patient 34

```
=====
Machine readable data on PATIENT 34:
```

```

=====
patient_trace(34,patient_data([sex(female), age(82),
birthdate('1921')]), diagnosis([name('Adenocarcinoma, NOS') ,
basisDiagnosisCode('7')]), tumour_info([nr_of_tumour(2),
total_tumours(2),localization('overlapping lesion of breast'),
laterality('left'), tumour_extension(3), tumour_grade(9),
tumour_differentiation_grade('unknown (not established)'),
topology('overlapping lesion of breast'),
behaviour('malignant/invasive, primary site'),
morphology('adenocarcinoma'),
nr_lymph_nodes([total(0),positive(unknown)]),metastases([]),
nr_metastases_in_lymph_nodes([total(9),investigated(99)]),
additional_tumour_codes([typeV1('nipple'),typeV2(''),typeV3('')]),
clinical_staging_results_before_therapy([stage_grouping_clinical('2B'),
clinical_tumour_staging('not
applicable'),ct('oedema (including peau d-orange) or ulceration of the
skin of the breast, or satellite skin nodules confined to the same
breast'),cn('metastasis in supraclavicular lymph node(s)'),cm('no
distant metastasis'),ctnmt('tumour of more than 5cm in greatest
dimension'),ctnmn('no regional lymph node metastasis'),ctnmm('unknown
clinical classification code for staging<9>'))]),

treatment([ i1('drug-based hormone therapy'), i2('chemotherapy NNO')]),

```

Medical interpretation of raw data for Patient 34

Information needed to decide the optimal treatment:

- Comorbidity.
- Hormonal status of the tumour.
- Extension tumour examinations.

When considering the information given we can not assure the primary site of the tumour is the breast. It could be a metastasis of another primary tumour. So, in this situation, neither the hormone therapy nor the chemotherapy schemes considered would be of any benefit.

If the breast is considered to be the primary site of the tumour, we should determine the hormonal status of the tumour. If the tumour is ER or PG positive, we should consider the option of hormonal therapy. If the tumour hormonal receptors are negative, we should consider the option of chemotherapy, taking account that this treatment should be used with caution in this 82 years-old patient. Another option could be palliative radiotherapy.

Technical interpretation of raw data for patient 34

```

=====
APPROXIMATE ASBRU TRANSLATION:
=====

```

PATIENT CONDITION:

diagnosis=OPERABLE-INVASIVE-BC AND tumour-size-range= >5cm

TREATMENT SEQUENCE:

chemotherapy (1-02-2004) +
hormone therapy (01-02-2004)

ATTACHMENT:

i1('drug-based hormone therapy') = ONE-OF(
hormone-therapy(hothe/P141/CH6) |
secondary-adjuvant-hormone-therapy(sahth/P102/CH5) |
adjuvant-hormonal-therapy(aht/P51/CH3) |
systemic-therapy-3(syth3/P162/CH6) |
secondary-adjuvant-hormone-treatment(saht/P101/CH5) |
androgens-therapy(anthe/P115/CH6) | estrogens-therapy(esthe/P132/CH6) |
anti-estrogens-therapy(aethe/P113/CH6) |
aromatase-inhibitors-therapy(aithe/P114/CH6) |
biophosphonates(bioph/P117/CH6) |
combination-of-trastuzumab-and-paclitaxel(cotap/P124/CH6) |
high-dose-dexamethasone(hdd/P138/CH6) |
herceptin-treatment(hertr/P140/CH6) |
lumbally-citostatic(lucit/P147/CH6) |
palliative-treatment(paltr/P152/CH6) |

i2('CHEMOTHERAPY NNO') = ONE-OF(
cmf-chemotherapy(cmf/P123/CH6) |
palliative-chemotherapy(pache/P150/CH6) |
neoadjuvant-chemotherapy(neo/P60/CH3) |
six-courses-anthracycline-chemotherapy(scac/P62/CH3) |
single-course-anthracycline-chemotherapy(sgcac/P63/CH3) |
secondary-adjuvant-chemotherapy(sac/P100/CH5) |
anthracycline-containing-regimen(acr/P112/CH6) |
combination-of-trastuzumab-and-paclitaxel(cotap/P124/CH6) |
high-dose-dexamethasone(hdd/P138/CH6) |
herceptin-treatment(hertr/P140/CH6) |
lumbally-citostatic(lucit/P147/CH6) |
palliative-treatment(paltr/P152/CH6) |)

Most likely this patient can be attached to an operable invasive BC without metastasis, therefore the attachment to chapter 6 can be ignored.

=====
ASBRU TRANSLATION:
=====

WHEN

MEDICAL CONTEXT=(diagnosis=operable-invasive-BC AND tumour-range = large)

EXECUTE TREATMENT SEQUENCE:

Neoadjuvant-chemotherapy +
Adjuvant-hormonal-therapy

The SMV formal specification of a critiquing property expressing the possibility of a treatment sequence such as the one above, is:

:

SMV TRANSLATION: Attached to chapter 1:

A. Attached to chapter 3: chemotherapy = neo, hormone therapy = aht

```
SPEC EF ((loadp_state=activated & neop_state=inactive &
ahtp_state=inactive) & EF (loadp_state=activated &
neop_state=activated & ahtp_state=inactive & EF (loadp_state=activated
& neop_state=completed & ahtp_state=inactive & EF
(loadp_state=activated & ahtp_state=activated & EF
(loadp_state=activated & ahtp_state=completed & (EF (loadp_state =
completed))))))
```

PROPERTY A is FALSE. The counterexample has 35 steps.

Apparently, the top-level plan (load) can complete without neoadjuvant chemotherapy (neo) becoming active. Under the assumption that the critiquing property has been correctly attached and the guideline model is correct, the treatment sequence does not conform to the guideline.

C.5. Patient 36

```
=====
Machine readable data on PATIENT 4:
=====
```

```
patient_trace(36,patient_data([sex(female), age(70),
birthdate('1932')]), diagnosis([name('Lobular carcinoma, NOS') ,
basisDiagnosisCode('7')]), tumour_info([nr_of_tumour(2),
total_tumours(2),localization('upper-inner quadrant of breast'),
laterality('left'), tumour_extension(3), tumour_grade(1),
tumour_differentiation_grade('well differentiated'),
topology('upper-inner quadrant of breast'),
behaviour('malignant/invasive, primary site'), morphology('Lobular
carcinoma'),
nr_lymph_nodes([total(1),positive(0)]),metastases([]),
nr_metastases_in_lymph_nodes([total(0),investigated(0)]),
additional_tumour_codes([typeV1(''),typeV2(''),typeV3('')]),
clinical_staging_results_before_therapy([stage_grouping_clinical('4'),
clinical_tumour_staging('not
```

```

applicable'),ct('tumour of between 2cm and 5cm in greatest
dimension'),cn('no regional lymph node metastasis'),cm('distant
metastasis cannot be assessed'),ctnmt('no evidence of primary
tumour'),ctnmn('regional lymph nodes cannot be
assessed'),ctnmm('unknown clinical classification code for
staging<1C>'))],pathological_staging_results_after_therapy([
stage_grouping_pathological('unknown
tumour stage group(1C,X,X)'),pathological_tumour_staging('not
applicable'),pt('tumour having between 1cm and 2cm in greatest
dimension'),pn('regional lymph nodes metastasis cannot be assessed
(regional lymph nodes not removed for study or previously
removed)'),pm('distant metastasis cannot be assessed'),ptnmt('tumour
having between 1cm and 2cm in greatest dimension'),ptnmn('regional
lymph nodes metastasis cannot be assessed (regional lymph nodes not
removed for study or previously removed)'),ptnmm('distant metastasis
cannot be assessed')]) ]),

```

```

treatment([
i1('sentinel node biopsy'),
i2('organ specific surgery (organ localization code=14)'),
i3('external radiotherapy')]),

```

Medical interpretation of raw data for Patient 36

Information needed to decide the optimal treatment:

- Comorbidity.
- Hormonal status of the tumour.
- Type of surgical procedure: breast conserving surgery or mastectomy.

Based on the information given, another treatment sequence would be also correct. Locally advanced tumours with positive ER or PR can be treated up-front with neoadjuvant hormone-therapy. If the tumour is ER and PR negative, we could also consider the option of neoadjuvant chemotherapy. In both cases, the up-front treatment gives the patient the opportunity of a breast conserving surgery.

This patient should receive adjuvant hormone-therapy if the hormonal status of the tumour is positive.

Technical interpretation of raw data for patient 36

```

=====
APPROXIMATE ASBRU TRANSLATION:
=====

```

```

PATIENT CONDITION:
(diagnosis=LCIS AND tumour-size-range= >2cm <5cm)

```

```

TREATMENT SEQUENCE:

```

Sentinel node biopsy (ultrasound-guided-core-biopsy?) +
bct (13-12-2002) +
external radiotherapy= ONE-OF(
 locoregional-postoperative-radiotherapy,
 subsequent-radiotherapy) (23-01-2003)

ATTACHMENT:

i1(sentinel node biopsy) = ONE-OF(
histological-biopsy(hibi/P79/CH5)|
biopsy(bio/P116/CH6)|
ultrasound-guided-core-biopsy(ugb/P50/CH1)|
re-excision (reex/P42/CH1)|
axillary-staging-by-SN(andp/P2/CH1)|
sentinel-node-procedure(snp/P105/CH5)|
tumour-excision(te/P108/CH5)|)

i2(lumpectomy) = ONE-OF(
limited-surgery-manual(lsm/P59/CH3)|
breast-conserving-therapy(bct/P65/CH5)

i3(external radiotherapy) = ONE-OF(
radiotherapy(radio/P158/CH6)|radiotherapy-for-T1N0-bc(rT1N0/P99/CH5)|
initial-radiotherapy(ir/P21/CH1)|subsequent-radiotherapy(sr/P47/CH1)|
dcis-radiotherapy(dcis/P10/CH1)|highdose-radiotherapy(hira/P80/CH5)|
radiotherapy-chest-wall(rcw/P41/CH1)|lowdose-reradiation(lore/P85/CH5)|
radiation(radio/P97/CH5)|intensive-remedial-therapy(irt/P21/CH1)|
hyperthermia(hyp/P81/CH5)|previously-irradiated-treatment(pit/P96/CH5)|
locoregional-radiotherapy-manual(lrm/P58/CH3)|great-boost(gb/P18/CH1)|
standard-boost(sb/P44/CH1)|optional-boost(ob/P32/CH1)|
locoregional-postoperative-radiotherapy(lrpr/P23/CH1)|
primary-radiotherapy-of-axilla(praoa/P40/CH1)|)

=====
ASBRU TRANSLATION:
=====

PATIENT CONDITION

(diagnosis=DCIS AND tumour-size-range > 2cm AND tumour-size-range < 5cm)

TREATMENT SEQUENCE:

Sentinel node biopsy (axilla-staging-by-SN (asbSN)) +
bct (13-12-2002) +
external radiotherapy= ONE-OF(
 locoregional-postoperative-radiotherapy(lrpr),
 subsequent-radiotherapy(sr)) (23-01-2003)

SMV: Attached to chapter 1:

sentinel node biopsy = axilla-staging-by-SN (asbSN)

```
lumpectomy = bct
external radiotherapy = ONE-OF(lrpr|sr)
```

```
SPEC EF ((chap1p_state=activated & asbSNp_state=inactive &
bctp_state=inactive & lrprp_state=inactive & srp_state=inactive) & EF
(chap1p_state=activated & asbSNp_state=activated &
bctp_state=inactive & lrprp_state=inactive & srp_state=inactive & EF
(chap1p_state=activated & asbSNp_state=completed & bctp_state=inactive
& (lrprp_state=inactive & srp_state=inactive) & EF
(chap1p_state=activated & asbSNp_state = inactive &
bctp_state=activated & (lrprp_state=inactive & srp_state=inactive) &
EF (chap1p_state=activated & asbSNp_state = inactive &
bctp_state=completed & (lrprp_state=inactive & srp_state=inactive) &
EF (chap1p_state=activated & asbSNp_state = inactive &
bctp_state=inactive & (lrprp_state=activated | srp_state=activated) &
EF (chap1p_state = activated & asbSNp_state = inactive &
bctp_state=inactive & (lrprp_state=completed | srp_state=completed &
EF chap1p_state=completed))))))
```

PROPERTY is FALSE. Counterexample length: 61 steps.

Apparently, chapter1 top-level plan gets completed without the sequence axillary staging by Sentinel Node (asbSN), followed by breast conserving therapy (bct) to be activated. This problem has been noticed in case of patient 4 as well. These observations can be stored and used to reduce the critiquing effort and to build statistics of the most frequently encountered "violations" of the guideline model.

C.6. Patient 54

```
=====
Machine readable data on PATIENT 54:
=====
```

```
patient_trace(54,patient_data([sex(female), age(56),
birthdate('1946')]), diagnosis([name('Duct carcinoma and tubular
carcinoma''Infiltrating duct carcinoma mixed with other types of
carcinoma') , basisDiagnosisCode('7')]), tumour_info([nr_of_tumour(2),
total_tumours(2),localization('overlapping lesion of breast'),
laterality('left'), tumour_extension(3), tumour_grade(1),
tumour_differentiation_grade('well differentiated'),
topology('overlapping lesion of breast'),
behaviour('malignant/invasive, primary site'), morphology('Duct
carcinoma mixed with other types, e.g. tubular carcinoma'),
nr_lymph_nodes([total(5),positive(1)]),
metastases([]),nr_metastases_in_lymph_nodes([total(0),investigated(1)]),
additional_tumour_codes([typeV1(''),typeV2(''),typeV3('')]),
clinical_staging_results_before_therapy([stage_grouping_clinical('4'),
clinical_tumour_staging('not applicable'),ct('primary tumour cannot be
```

```

assessed'),cn('no regional lymph node metastasis'),cm('no distant
metastasis'),
ctnmt('no evidence of primary tumour'),ctnmn('no regional lymph node
metastasis'),ctnmm('unknown clinical classification code for
staging<1B>'))],pathological_staging_results_after_therapy([
stage_grouping_pathological('2A'),pathological_tumour_staging('not
applicable'),pt('tumour having between 0.5cm and 1cm in greatest
dimension'),pn('unknown clinical classification code for
staging<1MS>'),pm('no distant metastasis'),ptnmt('tumour having
between 0.5cm and 1cm in greatest dimension'),ptnmn('unknown clinical
classification code for staging<1MS>'),ptnmm('no distant
metastasis'))]) ])
```

```

treatment([ [
i1('organ specific surgery (organ localization code=14)'),
i2('sentinel node biopsy'),
i3('external radiotherapy')]])
```

Medical interpretation of raw data for Patient 54

Information needed to decide the optimal treatment:

- Comorbidity.
- Hormonal status of the tumour.
- Precise pathologic information of the two components: ductal and tubular carcinoma.
- Type of surgical procedure: breast conserving surgery or mastectomy.

This is a lymph node-positive young patient. The risk of relapse is high when receiving no adjuvant treatment. If having no medical problems that contraindicate chemotherapy, this patient should receive adjuvant chemotherapy with anthracyclines and taxans. Afterwards, if the tumour is ER or PR positive, this patient should receive adjuvant hormonal therapy.

Technical interpretation of raw data for patient 54

```

=====
APPROXIMATE ASBRU TRANSLATION:
=====
PATIENT CONDITION:
diagnosis=DCIS AND
(tumour-size-range > 0.5cm AND tumour-size-range < 1cm)
TREATMENT SEQUENCE:
I1(lumpectomy )= bct (5-09-2003) +
I2(sentinel node biopsy ??) +
I3(external radiotherapy)
```

ATTACHMENT:

```
i1(bct) = ONE-OF(  
limited-surgery-manual(lsm/P59/CH3)|  
breast-conserving-therapy(bct/P65/CH5)|)
```

```
i2(sentinel node biopsy) = ONE-OF(  
histological-biopsy(hibi/P79/CH5)|biopsy(bio/P116/CH6)|  
ultrasound-guided-core-biopsy(ugb/P50/CH1)|re-excision (reex/P42/CH1)|  
axillary-staging-by-SN(ands/P2/CH1)|sentinel-node-procedure(snp/P105/CH5)|  
tumour-excision(te/P108/CH5)|)
```

```
i3(external radiotherapy) = ONE-OF(  
radiotherapy(radio/P158/CH6)|radiotherapy-for-T1N0-bc(rT1N0/P99/CH5)|  
initial-radiotherapy(ir/P21/CH1)|subsequent-radiotherapy(sr/P47/CH1)|  
dcis-radiotherapy(dcis/P10/CH1)|highdose-radiotherapy(hira/P80/CH5)|  
radiotherapy-chest-wall(rcw/P41/CH1)|lowdose-reradiation(lore/P85/CH5)|  
radiation(radio/P97/CH5)|intensive-remedial-therapy(irt/P21/CH1)|  
hyperthermia(hyp/P81/CH5)|previously-irradiated-treatment(pit/P96/CH5)|  
locoregional-radiotherapy-manual(lrm/P58/CH3)|great-boost(gb/P18/CH1)|  
standard-boost(sb/P44/CH1)|optional-boost(ob/P32/CH1)|  
locoregional-postoperative-radiotherapy(lrpr/P23/CH1)|  
primary-radiotherapy-of-axilla(praoa/P40/CH1)|)')
```

```
=====  
ASBRU TRANSLATION:  
=====
```

```
PATIENT CONDITION:  
(diagnosis=DCIS AND tumour-size-range >0.5cm AND tumour-size-range < 1cm)  
TREATMENT SEQUENCE:  
bct (bct) +  
Sentinel node biopsy (axilla-staging-by-SN (asbSN)) +  
ONE-OF(radiotherapy-chest-wall, primary-radiotherapy-of-axilla,  
dcis-radiotherapy, primary-radiotherapy-of-axilla(praoa/P40/CH1))
```

Formal translation to SMV of this treatment sequence as a critiquing property has the following form.

```
SMV: Attached to chapter 1:  
sentinel node biopsy = axilla-staging-by-SN (asbSN)  
lumpectomy = bct  
external radiotherapy = ONE-OF(lrpr|sr|dcis|praoa)
```

```
SPEC EF ((chap1p_state=activated & asbSNp_state=inactive &  
bctp_state=inactive & (lrprp_state=inactive & srp_state=inactive &  
dcisp_state=inactive & praoap_state=inactive)) &  
EF (chap1p_state=activated & asbSNp_state=activated &
```

```

bctp_state=inactive & (lrprp_state=inactive & srp_state=inactive &
dcisp_state=inactive & praoap_state=inactive) &
EF (chap1p_state=activated & asbSNp_state=completed &
bctp_state=inactive & (lrprp_state=inactive & srp_state=inactive &
dcisp_state=inactive & praoap_state=inactive) &
EF (chap1p_state=activated & asbSNp_state = inactive &
bctp_state=activated & (lrprp_state=inactive & srp_state=inactive &
dcisp_state=inactive & praoap_state=inactive) &
EF (chap1p_state=activated & asbSNp_state = inactive &
bctp_state=completed & (lrprp_state=inactive & srp_state=inactive &
dcisp_state=inactive & praoap_state=inactive) &
EF (chap1p_state=activated & asbSNp_state = inactive &
bctp_state=inactive & (lrprp_state=activated | srp_state=activated |
dcisp_state=activated | praoap_state=activated) & EF (chap1p_state =
activated & asbSNp_state = inactive & bctp_state=inactive &
(lrprp_state=completed | srp_state=completed | dcisp_state=completed |
praoap_state=completed) & EF chap1p_state=completed)))))))))

```

PROPERTY is FALSE, due to the same reasons as the property of patient 36. Counterexample has 61 steps.

C.7. Patient 71

```

=====
Machine readable data on PATIENT 4:
=====

```

```

patient_trace(71,patient_data([sex(female), age(50),
birthdate('1951')]), diagnosis([dutch_name('Morbus Paget en ductaal
CIS'), basisDiagnosisCode('7'), tumourMorphologyCode(8543),
tumour_behaviour(3)]), tumour_info([nr_of_tumour(2),
total_tumours(2),localization('overlapping lesion of breast'),
laterality('left'), tumour_extension(3), tumour_grade(3),
tumour_differentiation_grade('poorly (insufficiently)
differentiated'), topology('overlapping lesion of breast'),
behaviour('malignant/invasive, primary site'), morphology('unknown
morphology code=8543'),
nr_lymph_nodes([total(5),positive(1)]),
metastases([],nr_metastases_in_lymph_nodes([total(1),investigated(1)]),
additional_tumour_codes([typeV1('skin'),typeV2(''),typeV3('')]),
clinical_staging_results_before_therapy([
stage_grouping_clinical('4'),clinical_tumour_staging('not
applicable'),ct('tumour of between 2cm and 5cm in greatest
dimension'),cn('metastasis in movable ipsillary axillary lymph
node'),cm('distant metastasis'),ctnmt('primary tumour 2cm or less in
greatest dimension'),ctnmn('metastasis in movable ipsillary axillary
lymph node'),ctnmm('unknown clinical classification code for

```

```

staging<1C>'))],pathological_staging_results_after_therapy([
stage_grouping_pathological('4'),pathological_tumour_staging('not
applicable'),pt('tumour having between 1cm and 2cm in greatest
dimension'),pn('internal mammary lymph nodes with microscopic
metastasis detected by sentinel lymph node dissection but not
clinically apparent (not detected by clinical examination or by
imaging studies, excluding lymphoscintigraphy)'),pm('distant
metastasis'),ptnmt('tumour having between 1cm and 2cm in greatest
dimension'),ptnmn('internal mammary lymph nodes with microscopic
metastasis detected by sentinel lymph node dissection but not
clinically apparent (not detected by clinical examination or by
imaging studies, excluding lymphoscintigraphy)'),ptnmm('distant
metastasis')) ]),status_record([patient_status(2),tumour_residu('no
residual tumour after treatment'),
deceased('10-Oct-2004'),tumour_detection_date('03-Jul-2002'),
last_record_update('06-Apr-2005'), last_contact_date('10-Oct-2004'),
last_confirmed_contact_date('10-Oct-2004'),
last_tumour_data_update_date('25-Nov-2003'))]),

treatment([
i1('chemotherapy NNO'),
i2('organ specific surgery (organ localization code=11)'),
i3('external radiotherapy')])

```

Medical interpretation of raw data for Patient 71

Information needed to decide the optimal treatment:

- Comorbidity.
- Hormonal status of the tumour.

This patient should receive adjuvant hormone-therapy if the hormonal status of the tumour is positive after ending chemotherapy. Based on the information given, we consider this patient should receive chemotherapy in a neoadjuvant or in an adjuvant setting. This patient should undergo a total axillary dissection, not only a sentinel node procedure, because it had been detected a pre-treatment axillary lymph node.

Technical interpretation of raw data for patient 71

```

=====
APPROXIMATE ASBRU TRANSLATION:
=====

```

```

PATIENT CONDITION:
(diagnosis=locally-advanced-BC OR diagnosis=metastasized-BC AND
tumour-size-range >1cm AND tumour-size-range < 2cm)

```

```

TREATMENT SEQUENCE:

```

mastectomy (14-01-2003) +
chemotherapy (25-02-2003) +
radiotherapy (25-02-2003)

ATTACHMENT:

i1(chemotherapy NNO) = ONE-OF(
neoadjuvant-chemotherapy(neo/P60/CH3) |
six-courses-anthracycline-chemotherapy(scac/P62/CH3) |
single-course-anthracycline-chemotherapy(sgcac/P63/CH3) |

i2(organ specific surgery (organ localization code=11))=
radical mastectomy) = ONE-OF(
surgery(sm/CH3) |
radical-surgery-manual(rsm/P61/CH3) |
limited-surgery-manual(lsm/P59/CH3) |)

i3(external radiotherapy) = ONE-OF(
locoregional-radiotherapy-manual(lrm/P58/CH3) |

=====
ASBRU TRANSLATION:
=====

PATIENT CONDITION:
((diagnosis=locally-advanced-BC OR diagnosis=metastasized-BC) AND
tumour-range = medium)

TREATMENT SEQUENCE:
1.mastectomy (sm)+
2.neoadjuvant-chemotherapy(neo/P60/CH3) +
3.locoregional-radiotherapy-manual(lrm/P58/CH3)

The formalization of this sequence with respect to chapter 3 model:

SPEC EF ((diagnosis=locally-advanced-BC & loadp_state=activated &
smp_state=inactive & neop_state = inactive & lrtrp_state=inactive) &
EF ((loadp_state=activated & smp_state=activated & neop_state =
inactive & lrtrp_state=inactive) & EF ((loadp_state=activated &
smp_state=inactive & neop_state = activated & lrtrp_state=inactive) &
EF ((loadp_state=activated & smp_state=inactive & neop_state =
inactive & lrtrp_state=activated) & EF ((loadp_state=completed))))))

PROPERTY IS FALSE. The counterexample trace has 35 steps.

We conclude that, assuming the guideline model of chapter 3 represents an accurate description of this patient's condition, and the mapping of patient data terminology to ASBRU model terminology is correct, then the doctor's actions are not conforming to the guideline recommendations.

Bibliography

- [1] A. Advani, K. Lo, and Y. Shahar. Intention-based critiquing of guideline-oriented medical care. In *Proceedings American Medical Informatics Annual Symposium*, pages 483–487, 1998.
- [2] A. Advani, Y. Shahar, and M. A. Musen. Medical quality assessment by scoring adherence to guideline intentions. *Journal of the American Medical Informatics Society*, 9:S9–S97, 2002.
- [3] M. Ben-Ari, Z. Manna, and A. Pnueli. The temporal logic of branching time. *Acta Inf.*, 20, 1983.
- [4] M.E. Bratman. *Intention, Plans, and Practical Reason*. CSLI publications, 1987.
- [5] B. G. Buchanan and E. H. Shortliffe, editors. *Rule-based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Addison-Wesley, June 1984.
- [6] CBO. *Guideline for the Treatment of Breast Carcinoma*. van Zuiden, 2002. PMID: 12474555.
- [7] W. J. Clancey and E. H. Shortliff. *Readings in medical artificial intelligence: the first decade*. Addison-Wesley Series in Artificial Intelligence. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1984.
- [8] E. M. Clarke and E. A. Emerson. Design and synthesis of synchronization skeletons using branching time temporal logic. In *Logic of Programs: Workshop*, number 131 in LNCS. Springer, May 1981.
- [9] P. Clayton and G. Hripsak. Decision support in healthcare. *International Journal of Biomedical Computing*, 39:59–66, 1995.
- [10] F. T. de Dombal, D. J. Leaper, J. R. Staniland, A. P. McCann, and J. C. Horrocks. Computer-aided diagnosis of acute abdominal pain. *British Medical Journal*, 2:9–13, 1972.
- [11] A. et al. Dunets. Deliverable d4.3: Model-checking of selected guideline properties. Technical Report Protocure-D4.3-2006, Protocure2 Project, 2006.
- [12] A. S. Gertner. *Critiquing: effective decision support in time-critical domains*. PhD thesis, Department of Computer & Information Science, University of Pennsylvania, 1995.
- [13] G. A. Gorry and G. O. Barnett. Experience with a model of sequential diagnosis. *Computers and Biomedical Research*, 1:490–507, 1968.

- [14] A. Hommersom, P. Groot, P. Lucas, M. Marcos, and B. Martínez-Salvador. A constraint-based approach to medical guidelines and protocols. In *ECAI 2006 Workshop – AI techniques in healthcare: evidence based guidelines and protocols*, 2006.
- [15] O. Kupferman and M.Y. Vardi. Modular model checking. *Lecture Notes in Computer Science*, 1536:381–401, 1998.
- [16] C. P. Langlotz and E. H. Shortliffe. Adapting a consultation system to critique user plans. *International Journal of Man-Machine Studies*, 19:479–496, 1983.
- [17] J. van der Lei. *Critiquing based on Computer-Stored Medical Records*. PhD thesis, Erasmus University, 1991.
- [18] P.J.F. Lucas, L.C. van der Gaag, and A. Abu-Hanna. Bayesian networks in biomedicine and health-care. *Artificial Intelligence in Medicine*, 30:201–214, 2004.
- [19] M. Marcos, B. Martínez-Salvador, A.J. Hommersom, P. Groot, P.J.F. Lucas, A. Jovell, and S. Blancafort. Case-study in transformations for protocol development from guidelines. Technical Report D5.1, Protocure II, 2006.
- [20] M. et al. Marcos. Deliverable d2.2bcd: Models of selected guideline in intermediate, asbru and kiv representations. Technical Report Protocure-D2.2bcd-2005, Protocure2 Project, 2005.
- [21] Ken McMillan. The cadence smv model checker, 1993.
- [22] S. Miksch, Y. Shahar, and P. Johnson. Asbru: A task-specific, intention-based and time-oriented language for representing skeletal plans. In E. Motta, F. van Harmelen, C. Pierret-Golbreich, I. Filby, and N. Wijngaards, editors, *7th Workshop on Knowledge Engineering: Methods & Languages (KEMML-97)*. Milton Keynes, UK, 1997.
- [23] P. Miller. *A critiquing approach to Expert Computer Advice: ATTENDING*. Pittman Press, London, 1984.
- [24] P. Miller. *Expert Critiquing Systems: Practice-Based Medical Consultation by Computer*. Springer-Verlag, New York, 1986.
- [25] A. Pnuelli. A temporal logic of concurrent programs. *Theoretical Computer Science*, 13:45–60, 1981.
- [26] M. Pollack. The use of plans. *Artificial Intelligence*, 57(1):43–68, 1992.
- [27] R. Rymon. *Diagnostic Reasoning and Planning in Exploratory-Corrective Domains*. PhD thesis, Department of Computer & Information Science, University of Pennsylvania, 1993.
- [28] R. Serban, A. Ten Teije, and F. et al. Van Harmelen. Deliverable d3.2b: Testing the selected guideline. Technical Report Protocure-D3.2b-2006, Protocure2 Project, 2006.
- [29] A. et al Seyfang. Additional report on modelling changes in the guideline. Technical Report Protocure-D22bcd-extra, Protocure2 Project, 2006.

- [30] Y. Shahar, S. Miksch, and P. Johnson. A task-specific ontology for the application and critiquing of time oriented clinical guidelines. In *Proceedings of the sixth Conference on Artificial Intelligence in Medicine in Europe*, pages 51–61, 1997.
- [31] B. G. Silverman. Survey of expert critiquing systems: Practical and theoretical frontiers. *Communications of the ACM*, 35(4):106–127, 1992.
- [32] S. E. Strauss, W. S. Richardson, P. Glasziou, and R. B. Haynes. *Evidence-based Medicine - How to Practice and Teach EBM*. Churchill Livingstone, 2005.
- [33] H. R. Warner, A. F. Toronto, L. G. Veasy, and R. Stephenson. A mathematical approach to medical diagnosis: application to congenital heart disease. *Journal of the American Medical Association*, 177:177–183, 1961.
- [34] S. Woolf, R. Grol, A. Hutchinson, M. Eccles, and J. Grimshaw. Potential benefits, limitations, and harms of clinical guidelines. *British Medical Journal*, 318:527–530, 1999.
- [35] S. H. Woolf. Evidence-based medicine and practice guidelines: an overview. *Cancer Control*, 7(4):362–367, 2000.