CHECKING GUIDELINE CONFORMANCE OF MEDICAL PROTOCOLS USING MODULAR MODEL CHECKING

Arjen Hommersom Perry Groot Peter Lucas

Institute for Computing and Information Sciences, Radboud University Nijmegen
P.O.Box 9010 6500GL Nijmegen, The Netherlands
{arjenh, perry, peterl}@cs.ru.nl

Abstract

Medical guidelines and protocols are documents aimed at improving the quality of medical care by offering support in medical decision making in the form of management recommendations based on scientific evidence. Whereas medical guidelines are intended for nation-wide use, and thus omit medical management details that may differ among hospitals, medical protocols are aimed at local use within hospitals and, therefore, include detailed information. Although protocols are often constructed on the basis of medical guidelines, the question is to which extent a protocol conforms to the guideline. Formal methods are applied to shed light on this issue. A Dutch medical guideline regarding the treatment of breast cancer, and a Dutch protocol based on it, are taken as an example.

1 Introduction

Medical management is increasingly based on recommendations from the medical scientific community, summarised in medical guidelines. A medical guideline is an extensive document, developed by a working group of professionals involved in the management of the disorder covered by the guideline. It has been shown that medical guidelines can improve healthcare outcomes [15] and may even reduce the cost of care up to 25% [5]. In practice, however, local hospitals often work with local adaptations of medical guidelines called protocols. The need for a protocol in conjunction with a guideline is twofold: firstly, a guideline is an extensive document (e.g., the Dutch breast-cancer guideline is 121 pages in A4 format), and, therefore, it is not easy to locate relevant information; secondly, details about duration, dose, or actual procedure have been omitted from the guideline, and, thus, are added in a protocol to complement the information that is in the guideline. Hence, basically, a medical protocol is a summary of the most important sections that are in the guideline, mostly recommendations, supplemented with hospital-specific details concerning the treatment. This implies that many sections in a protocol may be very similar to related sections in a guideline. However, there may also be differences, partly due to differences in opinion between the guideline designers and protocol designers, and partly due to the difference in purpose of a guideline and protocol.

Although much research has already focused on designing computer-oriented languages and developing tools for their employment, not many researchers seem to note that guidelines and protocols are different levels of abstraction of medical management. This view raises a number of issues currently not addressed in literature. Firstly, guidelines are typically under-constrained thereby omitting many details about treatment order. This contrasts for example with [13] which considers a guideline to be some form of program, but in which no execution paths are excluded that are illogical for medical management in practice. Secondly, guidelines are usually considered to be solitary objects and their verification still takes a lot of effort. Clearly, reuse of verification efforts would be useful for the verification of protocols which are adapted from a guideline. No reference has been made to verifying adaptations of guidelines.

Sofar, protocol conformance to guidelines has only been looked at from an informal angle [9]. In this paper we address the problem of protocol conformance to guidelines using formal methods. This is done by interpretating guidelines as defining (logical) constraints on the medical management of patients performed in practice, whereas protocols are interpreted as more or less executable models. This approach was inspired by a statement by Wiersma and Burgers that “recommendations in guidelines should not only be based on

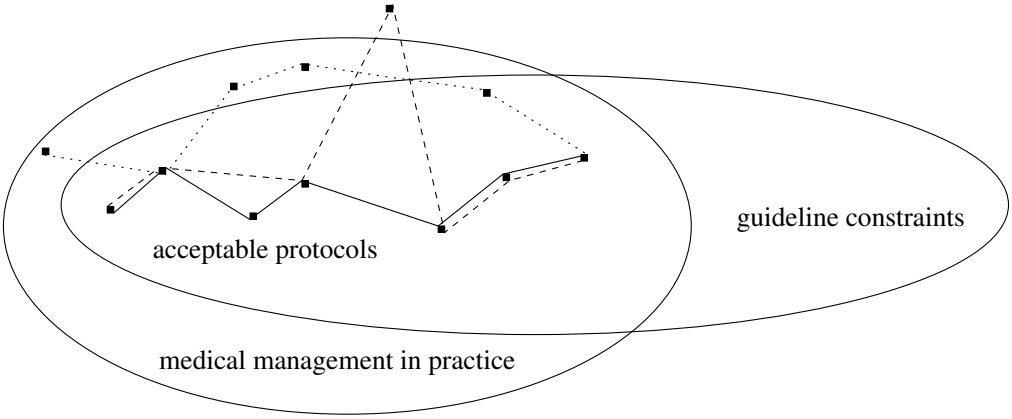

Figure 1: Sketch of medical management paths occurring in a protocol.

evidence extracted from scientific literature, but take into account the context of daily medical practice as well" [14]. In principle this approach would allow one to discover flaws or suboptimal management actions in the medical management in practice, assuming that a given protocol and guideline are correct, or to find incorrect or suboptimal medical management decisions in a protocol or guideline, assuming that the medical management in practice is correct and optimal. In this paper, we assume that the guideline augmented with information from general medical practice is correct and use model checking techniques to find flaws in the protocol.

2 Approach

In this section, we first introduce a method for checking the conformance of protocols to guidelines. Subsequently, preliminaries to formal methods are presented.

2.1 Verification of Protocol Conformance to Guidelines

The premise in this paper is that clinical practice guidelines and protocols provide *necessary*, but not *sufficient* conditions for making medical decisions. In other words, the guideline requires interpretation of medical doctors to apply the advises in practice. Thus, treatments which are acceptable given a strict interpretation of the protocol, might not be acceptable in practice. For example, the Dutch breast cancer guideline does not exclude the possibility of informing the patient *after* treatment about the different treatment possibilities. Clearly, such incompleteness of the guideline or protocol will not cause problems in practice as the advises are not mechanically followed by the physician. Hence, differences related to such a (lack of) advise can be considered irrelevant in practice. This is abstractly represented in Figure 1, which shows several paths of a protocol compared with all paths allowed by guideline and medical manangement in practice. In this case, violations occuring in the dotted line are most interesting, as the violation in the dashed line will not occur in practice.

Thus, the question is whether or not a protocol, restricted to medical management in practice, conforms to the constraints that the guideline imposes. The problem of this approach is that this medical management in practice cannot easily be assessed, because it is based on general and abstract medical knowledge, which is difficult to articulate and express directly, which makes it difficult to elicitate and formalise. Therefore, the approach we have taken in this paper is that we have formulated a concrete description of treatment paths that are taken in practice, extracted from a medical textbook. This does not formalise the *general* medical practice, but provides a good estimation of what is usually done in practice for a specific domain. Then, the approach consists of the following steps. Firstly, concrete treatment paths are formulated, which are known to be part of medical management in practice. Secondly, this information is weakened up to a point it is consistent with the protocol, i.e., we investigate which paths of medical practice are consistent with the protocol. Finally, the protocol restricted by medical practice is compared with the constraints imposed by the guideline.

In other words, the guideline restricted to medical practice, i.e., the intersection of Figure 1, is defined

as the gold standard. Finding a deviation from this gold standard in a protocol might indicate a problem in the protocol, although it is possible that it only indicates an error in the medical textbook or guideline. Nonetheless, we believe that such information is more valuable to guideline and protocol developers than deviations which are not even part of common sense medical practice.

2.2 Formal Preliminaries

It has been shown in [8] that the step-wise, possibly iterative, execution of a guideline can be described precisely by means of temporal logic. The logic that we use here for specifying properties of medical guidelines is a combination of Computation Tree Logic (CTL) [3, 4] and Linear Temporal Logic (LTL) [10]. 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., all paths that are possible from a certain state. 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 four temporal operators used are **X**, **G**, **F**, and **U**. With $\mathbf{X}\varphi$ being true if φ holds in the next state, $\mathbf{G}\varphi$ if φ holds in the current state and all future states, $\mathbf{F}\varphi$ if φ holds in some state in the future (or is true in the current state), $\varphi\mathbf{U}\psi$ 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. LTL provides operators for describing events along a *single computation path*. Each formula is of the form $\mathbf{A}f$, with f being a path formula, which is either an atomic proposition or inductively defined as $\neg f$, $f \vee g$, $f \wedge g$, $\mathbf{X}f$, $\mathbf{F}f$, $\mathbf{G}f$, or $f\mathbf{R}g$ with f, g path formulas.

In model checking literature, the approach of verifying a restricted part of the system (e.g., the protocol consistent with medical practice) is called *modular verification* (cf. [7]). 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 behaviour. In this paper, the assumed behaviour is written down in LTL and the guaranteed behaviour in CTL. The assume-guarantee assertions are written down as $[\varphi]M\langle\psi\rangle$, meaning that the CTL formula ψ holds in the computation tree that consists of all computations of the program, described by M, that satisfy the LTL formula φ. The tool used for performing the modular verification of medical guidelines is Cadence SMV.[1]

3 Medical Management in Breast Cancer

First, we give an informal description on the medical management as stated in the CBO guideline, the IKO protocol, and the textbook of Roses [11] that deals with locoregional treatment of operable breast cancer, i.e., T1-2 N0-1 M0 breast cancer according to the TNM classification system [6]. Thereafter, we give formalisations according to the approach described in the previous section.

3.1 Informal Description of Medical Management

According to the CBO guideline there are only two options for local treatment of operable invasive breast cancer: breast-conserving therapy (BCT) or modified radical mastectomy (MRM). BCT implies ample local excision of the tumour, an axillary staging procedure, and radiotherapy of the breast. MRM involves a total resection of the breast (mastectomy) and dissection of the axillary nodes (AND). The aim of BCT is to achieve a survival rate comparable to that following MRM with an optimal cosmetic result in terms of the treated breast. BCT is usually the preferred treatment unless the patient has a clear preference for MRM or there are contra indications for BCT, i.e., there is either (1) multicentricity (two or more tumour foci in different quadrants), (2) diffuse malignant microcalcifications, or (3) previous radiotherapy of the breast. Whereas these three contra indications are obtained *before* surgery, one other contra indication for BCT is obtained *during* surgery, i.e., (4) the margins of the local excision remain tumour-positive after repeated local excision attempts. In this case, local excision attempts are unsuccessful in removing the primary tumour and treatment therefore switches to MRM.

Treatment of the axillary nodes is also part of the treatment of breast cancer as the pathologic assessment of axillary lymph nodes remains the most important prognostic variable for the invasive breast cancer patient. An optimal assessment would be achievable by means of a complete axillary node dissection. However,

[1] `http://www.cis.ksu.edu/santos/smv-doc/`

AND may lead to morbidity, e.g., pain, limited shoulder movement. An alternative for axillary staging is the sentinel node procedure (SNP), which only dissects the sentinel nodes, i.e., those nodes that drain the area of the breast where the primary tumour is located and thus are most likely to contain metastasis. The SNP is currently the standard procedure for axillary staging in breast cancer provided that the contra indications do not hold, where contra indications of SNP are defined as (1) suspected or proven malignancy in the axillary nodes, (2) tumour $>$ T2, (3) multiple tumour foci, or (4) potentially disrupted lymph drainage due to recent axillary surgery or a large biopsy cavity following tumour excision. When the SNP is not possible, complete axillary node dissection should be carried out. Furthermore, treatment of the axilla is indicated (i.e., dissection, radiotherapy) for all forms of lymph node metastasis.[2]

Whereas the CBO guideline and IKO protocol lack many details about treatment order, [11] provides a more detailed description. In addition, according to [11], the sentinel node procedure (SNP) is started before segmental excision (i.e., used in BCT) or mastectomy. The sentinel nodes (SNs) are then immediately sent to the pathology lab, where they are examined during surgery. If the SNs are found to be positive, axillary dissection can be completed during the primary breast surgery in one setting. Furthermore, [11] differs with the CBO guideline and IKO protocol in the case of recurrent tumour positive resection margins in the BCT treatment. Whereas CBO and IKO recommend to switch the treatment to MRM, which includes axillary dissection, [11] only recommends a mastectomy with axillary dissection dependent on sentinel node histopathology.

3.2 Formalisation of Medical Management

Here, we introduce the constraint-based representation of the guideline, an executable model of the protocol, and the model of the medical management performed in practice.

3.2.1 Constraint-Based Representation of the CBO Guideline

The language we use for atomic propositions consists of medical actions $Actions$, medical plans $Plans$, and data structures $Data$:

$$\begin{aligned} Actions &: \{\text{tumour-excision}, \text{mastectomy}, \text{AND}, \text{SNP}\} \\ Plans &: \{\text{TREATMENT}, \text{BCT}, \text{MRM}, \text{AXILLA-STAGING}\} \\ Data &: \{\text{CI-BCT}, \text{CI-SNP}, \text{TF}, \text{SN}, \text{ITC}\} \end{aligned}$$

with CI-BCT, CI-SNP $\in \{\top, \bot\}$ denoting the contra indications for BCT and SNP respectively, SN $\in \{unknown, neg, pos\}$ denotes whether there is a metastasis found in the lymph nodes after performing the SN procedure, TF $\in \{unknown, \top, \bot\}$ denotes whether the re-section margins are tumour free, and ITC $\in \{unknown, \top, \bot\}$ denotes whether the tumour cells are isolated whten the sentinel node is found positive. In formulas, we write the variable name if it is meant that the variable is equal to $\top$, and the negated variable name is used to denote the respective variable is equal to $\bot$. The final representation in temporal logic of the medical management in the CBO guideline is shown in Figure 2.

Some constraints given by the guideline are not easily expressible in temporal logic as other modalities than treatment order are involved. For example, the preference for BCT over MRM and the preference for the SNP over axillary dissection for staging the axilla. Furthermore, certain assumptions regarding the patient data are implicit in the guideline. For example, the status of the resection margins (tumour free (TF) or not ($\neg$TF)) becomes known after tumour excision and the existence of metastasis (SN=pos or SN=neg) becomes known after the SNP. Here we have chosen not to consider these more implicit constraints.

3.2.2 Asbru Representation of the IKO Protocol

To transform the IKO protocol into a more or less executable model, which can be verified with respect to the constraints set by the CBO guideline, we have chosen to use the guideline representation language Asbru [12] as intermediate representation. We use Asbru, because its semantics has been defined precisely in previous research [1] and can be translated automatically into SMV for model checking purposes [2].

The Asbru model constructed (Figure 3) consists of nine plans ordered in a hierarchy. Arrows indicate sequentially executed sub-plans, dashed lines unordered executed sub-plans. The latter is in particular used when the protocol lacks any information about treatment order, e.g., `MRM` with sub-plans `AND` and

[2]The CBO guideline differs at this point with the IKO protocol as it makes an exception for isolated tumour cells.

Constraints related to control structure
(1) $\mathbf{AG}(\text{TREATMENT} \rightarrow \mathbf{AF}(\text{BCT} \vee \text{MRM}))$
(2) $\mathbf{AG}(\text{CI-BCT} \rightarrow \neg\text{BCT})$
(3) $\mathbf{AG}(\text{BCT} \rightarrow \mathbf{AF}(\text{AXILLA-STAGING} \vee \text{MRM}) \wedge \mathbf{AF}$ tumour-excision$)$
(4) $\mathbf{AG}(\text{MRM} \rightarrow \mathbf{AF}$ AND $\wedge\ \mathbf{AF}$ mastectomy$)$
(5) $\mathbf{AG}(\text{AXILLA-STAGING} \rightarrow \mathbf{AF}$ (AND $\vee$ SNP))
(6) $\mathbf{AG}(\text{CI-SNP} \rightarrow \neg\text{SNP})$
(7) $\mathbf{AG}(\text{tumour-excision} \rightarrow ((\neg\text{TF} \rightarrow \mathbf{AF}\ \text{MRM}) \wedge (\text{TF} \rightarrow \mathbf{AG}\ \neg\text{MRM}))$
(8) $\mathbf{AG}(\text{SNP} \rightarrow (\text{SN} = pos \wedge \neg\text{ITC} \rightarrow \mathbf{AF}$ AND$))$
(9) $[(\mathbf{G}\neg\text{MRM})]M\langle\mathbf{AG}(\text{SNP} \rightarrow \mathbf{AG}(\text{ITC} \rightarrow \mathbf{AG}\neg\text{AND}))\rangle$
(10) $\mathbf{AG}(\text{TREATMENT} \rightarrow (\text{CI-SNP} \rightarrow \mathbf{AF}$ AND$))$
Constraints related to data
(11) $(\text{CI-BCT} \rightarrow \mathbf{AG}\ \text{CI-BCT}) \wedge (\neg\text{CI-BCT} \rightarrow \mathbf{AG}\ \neg\text{CI-BCT})$
(12) $(\text{CI-SNP} \rightarrow \mathbf{AG}\ \text{CI-SNP}) \wedge (\neg\text{CI-SNP} \rightarrow \mathbf{AG}\ \neg\text{CI-SNP})$

Figure 2: Constraint-based representation of the CBO guideline. BCT = breast conserving treatment, MRM = modified radical mastectomy, CI-BCT = contra indications for BCT, SN = result of sentinel node procedure, CI-SNP = contra indications for SNP, TF = tumour free resection margins. Constraint (9) is restricted to a part of the model, denoted by the assumption $\mathbf{G}\neg$MRM. All other constraints ψ are short for $[\top]M\langle\psi\rangle$.

`mastectomy`. The top level plan `treatment` first executes `BCT` unless there are contra indications (filter condition). If `BCT` successfully completes, so will `treatment` (wait-for condition), else `MRM` will be executed. The execution of `BCT surgery` and `MRM` by `BCT` is analogous. `BCT surgery` may execute its sub-plans in any order. To allow for a particular order we use a manual activation which we assume to occur eventually. The `SNP` and `tumour excision` may abort (abort condition) in case of positive SNs or not tumour free re-section margins respectively, which is then propagated up the hierarchy, resulting in `BCT` executing `MRM`. (cf. [1] for details about the Asbru semantics.)

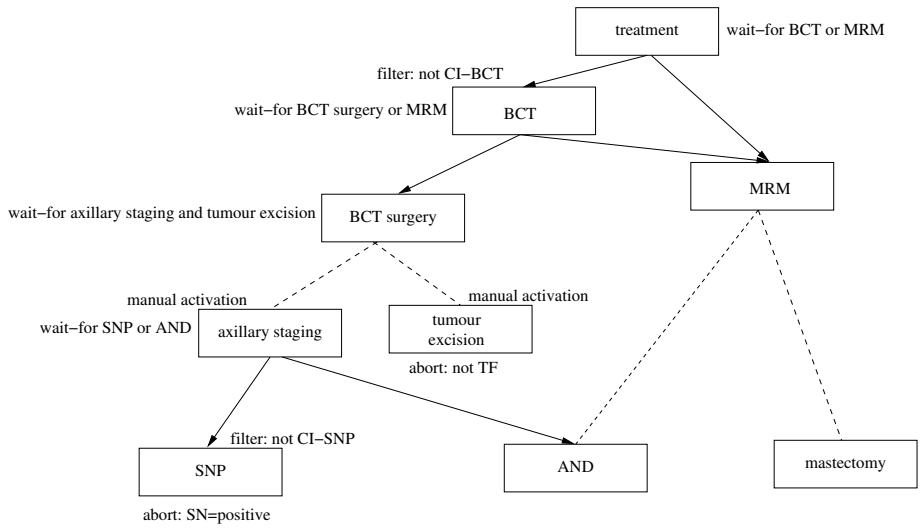

Figure 3: Asbru interpretation of the IKO protocol. Arrows represent sequential plans, dashed lines represent unordered sub-plans

In the SMV model, which is automatically constructed from the Asbru model, most variables dealing with patient data are initialised as *unknown* and receive an indeterministic value in the second step to make sure there is only one root of the model. Furthermore, we assume that they do not change during the treatment. The only variables that are initialised at a later stage are the status of the sentinel node, which becomes known during the SNP and whether or not the re-section margins are tumour free, which becomes known after excising the tumour. Furthermore, fairness constraints have been added to ensure that the manual activation of both the axillary staging and the tumour excision eventually occurs. In other words,

the patient will not wait indefinitely for the treatments to start.

3.2.3 Decision Tree of Medical Management in Practice

Information from [11] can be represented in a decision tree as shown in Figure 4, which deals with the ordering of medical actions treating the primary tumour (BCT and MRM) and the axilla (SNP and/or AND).[3] Nodes represent medical actions or plans, arcs represent constraints. A path from the root node to a leaf node represents a treatment path, which defines the order of medical actions when the constraints on the path are satisfied. Leaf nodes labelled with BCT or MRM mean the wrap-up of the medical strategy followed.

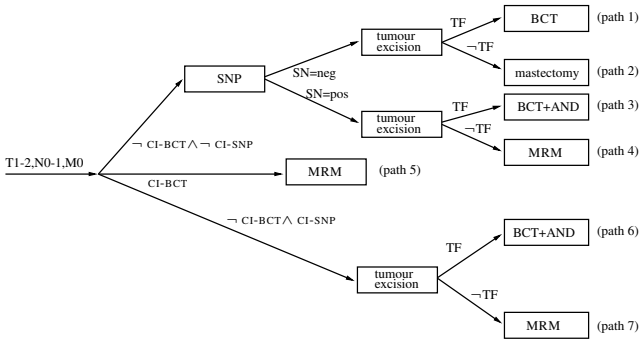

Figure 4: Background knowledge: possible treatment paths for surgery of operable invasive breast cancer. CI-BCT = contra indications BCT, CI-SNP = contra indications SNP, TF = tumour free resection margins, AND = axillary node dissection.

The formal interpretation of this decision tree is two-fold. First, we can give a CTL characterisation of these paths in terms of an existential path operator, denoted by Δ, which allows us to verify whether each of these paths exists in the protocol. Each action that occurs in the decision tree has been interpreted as a plan activation in the execution of the Asbru model, although the technical details of this have been omitted here for clarity. The first two paths are described as follows:

(1) $\mathbf{EX}(\neg\text{CI-BCT} \wedge \neg\text{CI-SNP} \wedge \mathbf{EF}(\text{SNP} \wedge \text{SN} = neg \wedge \mathbf{EF}(\text{tumour-excision} \wedge \text{TF} \wedge \mathbf{AG}(\neg\text{mastectomy} \wedge \neg\text{AND}))))$

(2) $\mathbf{EX}(\neg\text{CI-BCT} \wedge \neg\text{CI-SNP} \wedge \mathbf{EF}(\text{SNP} \wedge \text{SN} = neg \wedge \mathbf{EF}(\text{tumour-excision} \wedge \neg\text{TF} \wedge \mathbf{EF}(\text{mastectomy}) \wedge \mathbf{AG}(\neg\text{AND}))))$

On the other hand, we may view the decision tree as a number of constraints on medical management in practice. To this end, we extract a number of assumptions from the decision tree about the normal implementation of the advises given in the protocol. It is quite obvious to guarantee that such statements are sound with respect to the decision tree. In this case study, we consider the following LTL assumptions Γ.

(1) $(\neg\text{CI-BCT} \wedge \neg\text{CI-SNP}) \leftrightarrow \mathbf{F}\ \text{SNP}$
(2) $(\mathbf{F}\ \text{SNP}) \rightarrow ((\neg\text{tumour-excision}\ \mathbf{U}\ \text{SNP}) \wedge \mathbf{F}\ \text{tumour-excision})$
(3) $((\mathbf{F}\ \text{SN} = neg) \wedge (\mathbf{F}\ \text{TF})) \rightarrow (\neg(\mathbf{F}\ \text{AND}) \wedge \neg(\mathbf{F}\ \text{MRM}))$
(4) $((\mathbf{F}\ \text{SN} = neg) \wedge (\mathbf{F}\ \neg\text{TF})) \rightarrow ((\mathbf{F}\ \text{mastectomy}) \wedge \neg\mathbf{F}\ \text{AND})$
(5) $((\mathbf{F}\ \text{SN} = pos) \wedge (\mathbf{F}\ \text{TF})) \rightarrow ((\mathbf{F}\ \text{AND}) \wedge \neg(\mathbf{F}\ \text{MRM}))$
(6) $((\mathbf{F}\ \text{SN} = pos) \wedge (\mathbf{F}\ \neg\text{TF})) \rightarrow \mathbf{F}\ \text{MRM}$
(7) $(\text{CI-BCT} \rightarrow (\neg(\mathbf{F}\ \text{tumour-excision}) \wedge \mathbf{F}\ \text{MRM}))$
(8) $(\neg\text{CI-BCT} \wedge \text{CI-SNP}) \rightarrow \mathbf{F}\ \text{tumour-excision}$
(9) $(\neg\text{CI-BCT} \wedge \text{CI-SNP} \wedge (\mathbf{F}\ \text{TF})) \rightarrow ((\mathbf{F}\ \text{AND}) \wedge \neg(\mathbf{F}\ \text{MRM}))$
(10) $(\neg\text{CI-BCT} \wedge \text{CI-SNP} \wedge (\mathbf{F}\ \neg\text{TF})) \rightarrow \mathbf{F}\ \text{MRM}$

Assumption (1) and (2) deals with the use of sentinel node procedure and the order between this and the excision of the tumour. Assumptions (3) to (6) are concerned with paths (1) to (4). Assumption (7) deals with path (5). Finally assumptions (8) and (9) deals with path (6) and (7).

[3] We abstract from radiotherapy and isolated tumour cells.

4 Model Checking Results

As explained in Subsection 2.1, the medical management stated by the IKO protocol is less precise than the medical management performed in practice. Typically, one would expect the medical management in the protocol to be under-constrained when compared to the medical management in practice. With the SMV model checker we were able to verify that all paths described by Δ, except (2), can occur in the IKO protocol. Path (2) does not hold in the IKO protocol because it recommends a MRM whereas [11] recommends a mastectomy, i.e., axillary dissection is included in the medical management according to the protocol, but not according to [11]. The guideline does not provide specific evidence related to this advice, which suggests that both possibilities are acceptable. Nonetheless, such differences should be discussed with medical experts to find out whether the protocol or textbook is incomplete or incorrect.

Because treatment path (2) from the medical textbook is not part of the protocol, it follows that sentence (4) of Γ can not be coherent with the model (i.e., from (4) it follows the antecedent of (4) is false), so in this form it is not usable. We could therefore either adapt the assumption so that it corresponds to the guideline or omit it. Here, we have omitted it. Let Γ' be Γ without (4), then we verify each guideline constraint $[\varphi]M\langle\psi\rangle$ by model checking $[\varphi, \Gamma']M\langle\psi\rangle$ using SMV on the Asbru model of the IKO protocol. This shows that constraint (9) does not hold in the Asbru model of the IKO protocol, indicating a difference between protocol and guideline with respect to medical management in practice. The reason for this difference can be tracked back to preliminary evidence stated in the guideline, which has a low certainty degree, i.e., a low *level of evidence*. Although, in this case the difference between protocol and guideline is clear and could also have more easily been found through an informal analysis, this is largely because the protocol and guideline have a very similar structure and their recommendations are almost identical. However, the approach taken is independent of the underlying structure of the protocol and guideline. Therefore, this case study shows that formal techniques can be used to compare guideline and protocol independent of their underlying document structure.

The resources to check that paths from the decision tree exists were minimal. The number of nodes in the binary decision diagram (BDD) was only 100.000, and terminated within 3 seconds on a modern pc. The number of BDD nodes allocated for checking that the protocol conforms with the guideline was 250.000, and time spent on the verification of all the constraints took 175 seconds. While it is to be expected that more resources will be needed for larger case-studies, we have not used advanced techniques such as bounded model checking, variable ordering, or proof decomposition to reduce the computational complexity.

5 Discussion

The aim of this work was to use formal methods for obtaining insight into the differences and similarities between guidelines and protocols, based on the assumption that protocols should be looked upon as local modifications of guidelines. In the work presented here the view was taken that medical guidelines and protocols usually only specify necessary but not sufficient constraints on medical management.

In our study we have setup a modular model checking approach for checking the conformance of a protocol to the guideline from which it has been adapted. In this approach, medical guidelines and protocols are considered to be constraints on medical management. Medical guidelines are representable as temporal logic formulas whereas protocols are interpretable as more or less executable models. Furthermore, medical management as used in practice was used as additional background knowledge for restricting the guideline and protocol thereby rejecting treatment paths which are illogical for medical management in practice. In this paper we have applied the modular model checking approach to the CBO breast cancer guideline, the IKO protocol, and used [11] for additonal background knowledge. We have shown that with this approach interesting differences between guidelines and protocols with respect to background knowledge can be obtained. These differences can then be communicated to health care professionals for further clarification. In our case, one of the medical oncologists responsible for the IKO protocol confirmed that some differences found could be traced back to low levels of evidence, suggesting that such differences may be ignored.

The research presented in this paper is a novel approach in locating differences between protocols and guidelines giving a promising starting point for further investigating the relations between guidelines, protocols, and medical management in practice. Some questions still remain for further research. Firstly, a limitation of this research is that it was only based on one reference protocol on breast cancer treatment, selected at the start of this research. A second protocol for this type of breast cancer by the NKI (the Netherland Cancer Institute) is available, however, this protocol is very different from the first IKO protocol and

the guideline, resulting in other challenges than those discussed in this paper. Secondly, a question that emerged during the course of our research was whether the level of evidence as indicated by the oncologist can be incorporated in our approach as differences based on low levels of evidence can be ignored. Thirdly, formally obtaining differences between protocol and guideline may be useful for protocol designers. In this research, we have only been able to find end point protocols; as a consequence, the transformation process could only be described as consisting of a single step, which in reality may be a more iterative process.

References

[1] M. Balser, C. Duelli, and W. Reif. Formal semantics of Asbru - an overview. In *Proceedings of the International Conference on Integrated Design and Process Technology*, Passadena, June 2002. Society for Design and Process Science.

[2] S. Bäumler, M. Balser, A. Dunets, W. Reif, and J. Schmitt. Verification of medical guidelines by model checking – a case study. In A. Valmari, editor, *Proceedings of 13th International SPIN Workshop on Model Checking of Software*, volume 3925 of *LNCS*, pages 219–233. Springer-Verlag, 2006.

[3] M. Ben-Ari, Z. Manna, and A. Pnueli. The temporal logic of branching time. *Acta Inf.*, 20, 1983.

[4] 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.

[5] P. Clayton and G. Hripsak. Decision support in healthcare. *International Journal of Biomedical Computing*, 39:59–66, 1995.

[6] F.L. Green, D.L. Page, I.D. Fleming, A. Fritz, M.C. Balch, D.G. Haller, and M. Morrow. *AJCC Cancer Staging Manual*. Springer-Verlag, New York, 2002.

[7] O. Kupferman and M.Y. Vardi. Modular model checking. *Lecture Notes in Computer Science*, 1536:381–401, 1998.

[8] M. Marcos, M. Balser, A. Teije, and F. Harmelen. From informal knowledge to formal logic: A realistic case study in medical protocols. In *Proceedings of EKAW*, pages 49–64. Springer, 2002.

[9] 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.

[10] A. Pnuelli. A temporal logic of concurrent programs. *Theoretical Computer Science*, 13:45–60, 1981.

[11] D.F. Roses. *Breast Cancer*. Elsevier, Philadelphia, PA, 2nd edition, 2005.

[12] Y. Shahar, S. Miksch, and P. Johnson. The Asgaard project: A task-specific framework for the application and critiquing of time-orientied clinical guidelines. *AIME*, 14:29–51, 1998.

[13] A. ten Teije, M. Marcos, M. Balser, J. van Croonenborg, C. Duelli, F. van Harmelen, P.J.F. Lucas, S. Miksch, W. Reif, K. Rosenbrand, and A. Seyfang. Improving medical protocols by formal methods. *Artificial Intelligence in Medicine*, 36(3):193–209, 2006.

[14] J.J.E. van Everdingen, J.S. Burgers, W.J.J. Assendelft, J.A. Swinkels, T.A. van Barneveld, and J.L.M. van de Klundert, editors. *Evidence-based Guideline Development*, chapter Formulating Recommendations, page 171. Bohn, Houten, 2004.

[15] 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.