

Improving medical protocols through formalisation: a case study

Mar Marcos[†], Hugo Roomans[†], Annette ten Teije[‡] and Frank van Harmelen[†]

[†] Vrije Universiteit Amsterdam, Dept. of Artificial Intelligence,
De Boelelaan 1081a, 1081HV Amsterdam, Netherlands

[‡] Universiteit Utrecht, Institute of Information and Computing Sciences,
P.O. Box 80.089, 3508TB Utrecht, Netherlands

ABSTRACT: Medical practice protocols or guidelines contain more or less precise recommendations to assist practitioners and patient decisions about appropriate health care for specific circumstances. In order to reach their potential benefits, protocols must fulfill strong quality requirements. Medical bodies worldwide have made efforts in this direction, but mostly using informal methods such as peer review of protocols. In this paper we present a different approach, namely the quality improvement of medical protocols through formalisation. The research question that we try to answer in this paper is: *can formalisation contribute to improve the quality of medical protocols?* In order to answer this question, we have carried out two case studies on protocol formalisation, using two separate protocols. As a result of our formalisation effort, we have uncovered a significant number of anomalies in these two protocols. This is a surprising result, since the two selected protocols are of the highest quality produced by the medical profession.

I. INTRODUCTION

Medical practice protocols or guidelines are “systematically developed statements to assist practitioners and patient decisions about appropriate health care for specific circumstances” [4]. They contain more or less precise recommendations about the diagnosis tests or the interventions to perform, or about other aspects of clinical practice. These recommendations are based on the best empirical evidence available at the moment. Among the potential benefits of protocols, we can highlight the improvement of health-care outcomes [13]. More precisely, they can help to promote high quality practice, recommending interventions of proved benefit and discouraging those that are not supported by good evidence. They can also be used to reduce variations in care. Finally, protocols can be useful to improve cost efficiency, thanks to the standardisation of health-care. Indeed, it has been shown that adherence to protocols may reduce the costs of care upto 25% [2].

In order to enable their potential benefits, protocols must fulfill strong quality requirements. This is true not only for the final *product*, the protocol, but also for the development *process*. Medical bodies worldwide have made efforts in this direction, e.g. elaborating appraisal documents that take into account a variety of protocol aspects, of both protocols and their development process (see [6] for a compar-

ison of appraisal instruments). However, these initiatives are not sufficient since they rely on informal methods and notations.

In this paper we present a different approach, namely the quality improvement of medical protocols through formalisation. Currently, protocols are described using a combination of different formats, e.g. text, flow diagrams and tables. The underlying idea of our work is that making these descriptions more precise, with the help of a more formal language, will expose parts where the protocols are ambiguous, incomplete or even inconsistent. By pointing out these anomalous parts, and the reasons why they could be problematic, we expect to obtain useful indications for the improvement of the protocols. This idea is widely acknowledged in fields like software engineering, where formal methods are used as a tool for early detection of specification and design errors, but has been largely unexplored for medical protocols.

Thus, the research question that we try to answer in this paper is: *can formalisation contribute to improve the quality of medical protocols?* In order to answer this question, we have carried out a case study on protocol formalisation. For this purpose, a choice had to be made on the specific protocol representation language as well as on the medical protocols to be used.

Several languages exist for representing medical protocols (see [3] for a description of the most prominent ones). For our purposes we need a sufficiently formal and detailed enough language since only precise descriptions will allow us to uncover anomalies in the protocols. Although the trend is changing lately, many of the protocol languages in the literature (e.g. GLIF [8]) are not formal enough. Exceptions to this are PROforma [5] and Asbru [11]. In this work we have chosen Asbru, firstly because it is more precise in the description of various medical aspects, and secondly because Asbru protocols are more declarative, and thus they are more amenable to formal analysis.

Concerning the protocols object of our study, we have tried to select two examples covering different features. The first one comes from the set of protocols developed by the Dutch Association of General Practitioners. The second example is a specialist protocol developed by the American Academy of Pediatrics. The justification for the choice of these protocols is further discussed in section III.

The main contribution of this paper is an analysis of the

different types of problems in protocols that we have uncovered during the formalisation process, together with an assessment of the utility of the approach. We expect that our experiences will be valuable to those organisations and individuals interested in protocol quality improvement. The structure of the paper is as follows. First we briefly describe the Asbru protocol representation language, and then we give some background information about the protocols of our study and sketch their formalisation in Asbru. Next we go into the details of the formalisation and improvement of these protocols, and report on the findings of our case-study. Finally, we end with some conclusions and future work ideas.

II. THE ASBRU PROTOCOL REPRESENTATION LANGUAGE

Asbru is a semi-formal language intended to support the tasks necessary for protocol-based care [11]. It is a semi-formal language because its semantics, although more precise than in other protocol representation languages, is not defined in a formal way¹. This semi-formal feature makes Asbru suitable for an initial analysis but not for a systematic verification of protocols. Finally, Asbru can be defined as a time-oriented, intention-based, skeletal plan specification language.

In Asbru, protocols are expressed as plan schemata defined at various levels of detail, precise enough to capture the essence of clinical procedures but leaving space for flexibility during their application. Major features of Asbru are:

- explicit intentions and preferences can be stated for plans.
- intentions, conditions, and states are temporal patterns.
- uncertainty in temporal patterns and parameters can be expressed.
- plans can be executed in different compositions, e.g. in parallel (PARALLEL), in sequence (SEQUENTIALLY), in any possible order, sequential or not (UNORDERED), or every certain time (CYCLICAL PLAN); in addition, it can be defined whether all the steps should be executed (ALL) or not (e.g. ONE or NONE).
- conditions can be defined to control plan execution, e.g. to set applicability conditions (FILTER) or to determine when execution should be interrupted (ABORT).

Some of the above elements need additional explanation, namely intentions and temporal patterns. Intentions are high-level goals that can be associated with plans and subplans. They are patterns of states or actions to be achieved, maintained, or avoided (ACHIEVE, MAINTAIN or AVOID), during or after the execution of the plan (INTERMEDIATE or OVERALL).

Temporal patterns are crucial in Asbru. They are more or less complex logical propositions that are qualified with time annotations. The time annotations used in Asbru al-

¹Current work is extending Asbru with a formal semantics based on state-charts.

low the representation of uncertainty in starting time, ending time, and duration with intervals, as well as the use of multiple reference points. A time annotation is written in the form ([EarliestStarting, LatestStarting] [EarliestFinishing, LatestFinishing] [MinDuration, MaxDuration] REFERENCE). Thus, the temporal pattern (TSB-decrease=yes any [4 hours,-] [-,6 hours] [-,-] self) is a proposition becoming true if there is a decrease of TSB, in any context, between 4 and 6 hours after the activation of the current plan, self.

III. THE PROTOCOLS OF OUR CASE STUDY

Our study is based on two protocols: a general practitioner (GP) protocol for the management of diabetes mellitus type 2 [10] and a pediatrics protocol for the management of jaundice in healthy term newborns [1]. They have been developed, respectively, by the Dutch Association of General Practitioners (NHG) and the American Academy of Pediatrics (AAP). In this section we will give some more information on the protocols and motivate their choice.

There exists a wide variety of protocols. They differ along several dimensions, which can be related to the contents of the protocol or to its form, amongst others. Important dimensions are the **target users** of the protocol (e.g. GPs or nurses) and the **aspects of clinical practice** it covers (e.g. diagnosis and/or treatment). The **period of time** to which the protocol applies is also characteristic, varying between e.g. short time-span (hours to days) and long time-span protocols (weeks to years). Concerning the **form**, protocols are usually more or less structured textual descriptions, and sometimes combine alternative formats, such as tables, lists or flowcharts.

The two selected protocols not only come from different organisations but also differ in their characteristics. The diabetes protocol is addressed to GPs and covers the diagnosis and treatment of the illness over a long period of time. The jaundice protocol is also devoted to the diagnosis and treatment but, unlike the diabetes one, targets non-specialist clinicians and is intended to be applied in a hospital setting, over a short time-span. Regarding the form, the text layout of diabetes protocol² contrasts with the rather unstructured text of jaundice. The latter, however, includes complementary descriptions such as a flow diagram describing the protocol algorithm. Finally, the two protocols can be considered of high quality –one is in daily use by Dutch GPs and the other is included in the repository of the National Guideline Clearinghouse³. All the above means that the chosen protocols cover different aspects and, at the same time, fulfill certain quality standards, which makes them good candidates in our study.

²The diabetes protocol comes in two versions: a booklet and two quick reference cards.

³See <http://www.guideline.gov/>.

IV. AN OVERVIEW OF THE PROTOCOLS IN ASBRU

In this section we will sketch the high level structure of the diabetes and jaundice protocols in Asbru.

Figure 1 presents the overall plan structure of the diabetes protocol in Asbru. Notice that the notation used in this figure and the next one does not correspond to the usual XML syntax of the Asbru language, but it is a more readable representation⁴. Moreover, for space reasons, it is mainly focused on the hierarchical decomposition of plans into subplans. This is described in the **plan-body** and **wait-for** parts, with an indication of the type of composition (parallel, sequentially, and so on) and of the subplans that should be executed, i.e. those for which the plan should wait.

Like the original protocol, the Asbru version has as main parts the Diagnostics and the Policy plans. The diagnosis of diabetes is performed by measuring the glucose levels in the blood. If the Diagnostics plan concludes that there is diabetes mellitus type 2, the next step is applying the Policy plan. The figure focuses on the details of the latter, which roughly consists in instructing the patients in the most important aspects of their illness and then searching for an effective treatment (Treatments-and-Controls plan). This search for a treatment is performed along with the treatment of risks of cardiovascular diseases, and with the periodical (quarterly and yearly) controls. Besides, during the application of this treatment and control part, it could be the case that a more specific policy needs to be applied under certain conditions (e.g. hypoglycemic coma or insulin adjustments). This is implemented in the specific policy plans and in the Consultation-and-referring plan.

The high-level structure of the jaundice (or hyperbilirubinemia) protocol is shown in figure 2. Here the focus is on the diagnosis part, which consists of different anamnesis questions and blood tests to rule out the possibility of any serious underlying disease, followed by the plan Jaundice-determination. If there are no signs of a severe disease and the diagnostics plan concludes that the baby has jaundice, the treatment plan has to be activated. This roughly tries to reduce the high bilirubin levels in the blood responsible for the jaundice, usually applying some kind of phototherapy treatment.

V. PROTOCOL IMPROVEMENT BY FORMALISATION

During the Asbru formalisation of the protocols, numerous anomalies became apparent. Formalisation is a labour intensive process by which the protocol contents are transformed into their equivalent Asbru version. In some cases, the Asbru translation demands making explicit elements which are not explicitly part of the protocol. This is typically the case when describing the decomposition of plans into subplans. In other situations, the Asbru syntax requires

⁴The full XML version of the protocols described in this paper can be found in <http://www.protocolure.org/>.

details that otherwise could go unnoticed. An example of this is the minimal and maximal delay necessary for the specification of the retry time in a cyclical plan. In any case, any problem in this transformation process has been considered as a possible anomaly in the protocol. In addition to this, the resulting Asbru transcription is more amenable to analysis than the original protocol and hence it has helped in the identification of problematic parts.

In a general sense, we have used the term anomaly to refer to any issue preventing a satisfactory interpretation of the original protocol. The fact that these anomalies were detected during formalisation is already a surprising result, since the two selected protocols can be considered of the highest quality produced by the medical profession.

In the rest of the section we give examples of the different types of anomalies we found, together with an explanation of the potential problems they might cause. For presentation purposes we have grouped them into general categories such as ambiguity and incompleteness. However, as we will discuss later, these anomalies can sometimes be interpreted in different ways according to the alternative causes that might be at their origin.

A. Ambiguity

Ambiguous protocol parts are inherently hard to interpret. Two examples follow. In the jaundice protocol, there are sentences formulated in such a way that it is not clear whether they should be read in connection to each other or not (*ambiguous text*). Again in the jaundice protocol, very similar terms (e.g. “jaundiced” and “clinically jaundiced”) appear in different parts with no clarification about whether they are used indistinctively or not (*ambiguous terms*). The problem with ambiguity is that it leaves space for wrong interpretations by the medical practitioner, and therefore may have a negative effect on the decisions taken.

B. Incompleteness

Incompleteness can be related either to insufficient information or to completely missing pieces of information. In either case, incompleteness hinders a correct interpretation of the protocol.

Missing information items are by far the most common anomaly in the studied protocols. A simple case is the lack of definition for the terms “jaundiced” or “clinically jaundiced” in the jaundice protocol (*missing definition of terms*). Another example is the use of the abstract notion “rapidly rising TSB⁵ levels”, which should obviously be based on the quantitative TSB measurement, without an indication of when the TSB rise can be considered rapid (*missing definition of abstract terms*). At various places in the protocols a test is requested or a question is asked to the patient but the results are not referred to anywhere in the protocol, i.e. there is no information about the utilisation of the results in

⁵Total serum bilirubin.

```

(1) plan Diabetes-Mellitus-Type-2
(2)   plan-body type=unordered, wait-for-optional-subplans=yes
(3)   wait-for one
(4)     plan Diagnostics
(5)     ...
(6)     plan Policy
(7)       plan-body type=sequentially
(8)       wait-for all
(9)         plan Education-DMT2
(10)        ...
(11)        plan-body type=unordered, retry-aborted-subplans=yes
(12)        wait-for Treatments-and-Controls
(13)          plan Treatments-and-Controls
(14)            plan-body type=parallel
(15)            wait-for none
(16)            plan DMT2-treatments
(17)              plan-body type=sequentially, wait-for-optional-subplans=yes
(18)              wait-for Non-insulin-DMT2-treatments
(19)                plan Non-insulin-DMT2-treatments
(20)                ...
(21)                plan Insulin-DMT2-treatments
(22)                ...
(23)                plan Treatment-of-CV-disease-risk-factors
(24)                ...
(25)                cyclical-plan
(26)                  plan Quarterly-control
(27)                  ...
(28)                  retry-delay minimum=3mon, maximum=3mon
(29)                cyclical-plan
(30)                  plan Annual-control
(31)                  ...
(32)                  retry-delay minimum=1year, maximum=1year
(33)                plan Policy-for-simultaneous-diseases
(34)                ...
(35)                plan Policy-for-hypoglycemic-coma
(36)                ...
(37)                plan Consultation-and-referring
(38)                ...

```

Fig. 1. Essential steps from the treatment part of the diabetes protocol

the reasoning process (*missing information on data utilisation*). As examples of tests results not used we can cite the child blood typing in the jaundice protocol and the creatinin test in the diabetes one. Regarding data requested but not used, in diabetes we can find questions about alcohol consumption and physical exercise habits. A related problem is the lack of specification of data storage mode, e.g. it is not clear if the tests/data requested in diabetes periodical controls, such as weight and blood pressure, should be kept in a historical record or not (*missing information on data man-*

agement). Lastly, a completely different kind of information is missing in the diabetes protocol, in which the alternative insulin treatments are described without specifying how to decide on the most adequate one (*missing information on decision making*).

Insufficient information issues occur less often. As an example we can cite a sentence of jaundice protocol suggesting a procedure for the choice of monitoring times and treatment but without further details on it (*insufficient information on decision making*):

```

(1) plan Hyperbilirubinemia
(2)   plan-body   type=unordered
(3)   wait-for   Diagnostics-and-Treatment-hyperbilirubinemia
(4)     plan Check-for-rapid-TSB-increase
(5)     ...
(6)     plan Check-for-jaundice-after-2-weeks
(7)     ...
(8)     plan Check-for-jaundice-after-3-weeks
(9)     ...
(10)  plan Diagnostics-and-Treatment-hyperbilirubinemia
(11)    plan-body type=sequentially
(12)    wait-for  none
(13)    ask term-child
(14)    ask age-child
(15)    plan Diagnostics-hyperbilirubinemia
(16)      plan-body type=sequentially
(17)      wait-for  all
(18)        pathologic-reason←no
(19)      plan Anamnesis-abnormal-signs
(20)      ...
(21)      plan Blood-tests
(22)        plan-body type=sequentially, wait-for-optional-subplans=yes
(23)        wait-for  one
(24)          plan Test-blood-test-mother
(25)          ...
(26)          plan Perform-blood-test-child
(27)          ...
(28)        plan Anamnesis-hemolytic-disease
(29)        ...
(30)        plan Jaundice-determination
(31)        ...
(32)    plan Treatment-hyperbilirubinemia
(33)    ...

```

Fig. 2. Essential steps from the diagnostic part of the jaundice protocol

“Determination of the rate of rise of TSB and the infant’s age may help determine how often to monitor bilirubin levels and whether to begin phototherapy”

Like ambiguity, incompleteness can lead to wrong interpretations and possibly have negative consequences in the final decision. Besides, it may result in an efficiency loss in case the physician needs to resort to additional sources to fill in the information gaps.

C. Inconsistency

Inconsistencies are elements that may somehow result in different (and even conflicting) decisions given the same patient data. The jaundice protocol presents several inconsistencies. In general they involve differences between al-

ternative formulations of the same aspects. For example, the list of signs that should be investigated to rule out any underlying disease as it appears in the tables is slightly different from the one in the flowcharts (and both differ from the text). Another example, more directly related to the recommendations issued, is a sentence stating that continued observation may be an alternative to repeated TSB measurement and phototherapy. This is not reflected elsewhere in the protocol, specially not in the table specifying the appropriate therapy based on the TSB value. As a final example we can mention the existence in the text of recommendations for children jaundiced in their first day of life. These children are not considered healthy and hence are out of the scope of the protocol in flowcharts. This is particu-

larly problematic since the recommendations directly conflict with the applicability conditions of the protocol.

The problems derived from inconsistent elements are very serious and as such must be avoided.

D. Redundancy

Whenever some part of the protocol can be removed without any (noticeable) effect in the resulting prescriptions, it can be considered redundant. The only redundancy we have detected is in the diabetes protocol and involves a repeated question about weight in the annual control (*redundant data requests*). Although redundant questions are in principle harmless, the main problem with redundancy is that it may cause unnecessary inconveniences for the patient and efficiency loss in the care process.

So far we have described the different types of anomalies found in the protocols. To give a better idea of the extent of uncovered anomalies, some concrete numbers follow (see also table I). In the case of jaundice protocol we found 1 ambiguity, 10 incompleteness anomalies, 6 inconsistencies and no redundancy. Regarding the diabetes protocol, we identified 4 ambiguities, 38 incompletenesses and 2 redundancies, but no inconsistency. Here, it is important to note that these high numbers are due to the complexity of the protocol (e.g. it comprises 5 therapies while in the case of jaundice there is only 1) and thus should not be interpreted as an indicator of a lower quality. Besides, we are aware that this task of interpretation of anomalies and absolute numbers must be carried out by medical practitioners with the pertinent knowledge.

VI. DISCUSSION

In the previous section we have presented a collection of protocol anomalies, highlighting the reasons why they could be problematic. Below we discuss some anomaly aspects, both in connection with the protocols we have studied and general ones.

The most common anomaly in both protocols is incompleteness. It can be argued that incompleteness is intentional in most cases, either because it refers to common background knowledge that can be assumed in the target user-group of the protocol or because it concerns points upon which there is no agreement. Other anomalies only occur in one of the protocols. Thus, no inconsistency was found in the diabetes protocol whereas several ones were found in the case of jaundice protocol, mainly between alternative formulations of the same recommendations (text, tables and/or flowcharts). The use of structured notations such as flowcharts can be a very useful complement to textual descriptions, making the protocol information more complete. However, parallel formulations must be handled with great care since inconsistency is more likely to occur in this case.

Second, it is important to note that anomalies are not errors but only signs of possible errors. As has been already mentioned, different errors may be responsible for the anomalies. For example, the repeated question about weight in diabetes control might be due to a duplication error (and hence a redundancy) if one of the questions is deemed unnecessary, or to an erroneous question if it is considered that different information should be asked in one of the cases, e.g. weight gain instead of weight. In addition to this, some anomalies are “two sides of the same coin”, in the sense that solving one of them will immediately solve the other. For example, adding the definitions of terms that are missing will help to remove the ambiguity related to their utilisation in the protocol. In this analysis of anomalies, the participation of medical experts with knowledge about the questions addressed in the protocol is crucial.

Lastly, emphasising that the analysis and interpretation of anomalies, as well as their removal, are crucial to avoid potential problems during the application of the protocol in the medical environment. We think that anomalies related to ambiguity, inconsistency and redundancy must be solved. Regarding incompleteness, as mentioned before, it can be seen as intentional from the point of view of protocol designers. If the reason for the incompleteness is the lack of agreement, it is important to put it explicitly as a point where the physician must make a choice, together with the pros and cons for the different alternatives, if known.

VII. CONCLUSIONS

Medical practice protocols are becoming increasingly important in health-care, because of their potential to promote high-quality practice and reduce variations in care while improving cost efficiency. However, in order to reach these goals protocols themselves need to be of high quality. As we have shown in our study, a significant number of anomalies can be detected, even in high quality protocols through Asbru formalisation. These anomalies are protocol parts in which potential problems might arise, and therefore indicate points where improvements are possible. It is important to note here that our aim is supporting the improvement of the original protocol rather than producing an enhanced version thereof in Asbru. For the time being, we believe that this goal is more realistic than trying to win the acceptance by the medical profession on the utilisation of a language such as Asbru.

Although it can be argued that the anomalies we found could have been detected by alternative means, the fact is that formalisation has proved to be useful for this purpose. We think that formalisation does provide a good foundation for detection of anomalies, since the use of precise notions enforces a critical examination of the protocol [9].

Solving the anomalies detected through formalisation not only can contribute to the enhancement of protocol quality but also would allow for a wider range of applications, including the use of protocols as a basis for the development

TABLE I
DIFFERENT TYPES OF ANOMALIES FOUND IN THE JAUNDICE AND DIABETES PROTOCOLS

	ambiguity		incompleteness				inconsistency		redundancy
	text	terms	missing definitions	missing info. on data	missing info. on decision	insuff. info. on decision	recom. out of scope	inconsistent recom.	repeated data requests
diabetes	1	3	6	24	1	7	0	0	2
jaundice	1	0	3	4	0	3	2	4	0

of computerised tools. It is important to note that the removal of anomalies only increases the internal consistency of the protocol. As for the external quality, it can only be enhanced by medical experts, with an analysis of the best empirical evidence, and/or the validation of the protocol against the practice of authorised experts. Following the latter idea, we have explored the possibility of using the decisions of an expert as a way to improve the external quality of existing protocols [7].

It is also worth to mention that our work is along the line of some medical approaches to protocol improvement. For instance, the appraisal instrument developed by the AGREE Collaboration [12] stresses protocol aspects that can be easily spotted through formalisation, such as the specificity and unambiguity of recommendations. The strength of formalisation with respect to other methods is, as has been already pointed out, in the use of precise notions.

The main drawback of our approach is the high cost of formalisation. It can be argued that this cost is too high with respect to the number and seriousness of the problems we have detected. In our view, it is a price that must be paid if we want to ensure the reliability of medical practice protocols. Besides, the costs of formalisation should be also weighed up against several other benefits.

In the future we plan to experiment on a further formalisation of medical protocols, using a variant of temporal logic suitable for expressing the different types of actions that can be found in protocols. In addition to the benefits of semi-formal languages like Asbru, the use of a logic formalism has the advantage of allowing for a systematic verification of protocols, using machine-based theorem provers. The idea is evaluating the feasibility and benefits of the approach. At the same time, in order to obtain a more objective assessment, we plan to include in the process the feedback from medical experts, which has been lacking until now.

ACKNOWLEDGEMENTS

This work has been partially supported by the European Commission's IST program, under contract number IST-2001-33049-Protocure. We also want to thank Geert Berger and Freya Kruizinga for their contribution to the formalisation of the protocols, and to Silvia Miksch, Andreas Seyfang and Robert Kosara for their support in using Asbru.

REFERENCES

- [1] American Academy of Pediatrics, Provisional Committee for Quality Improvement and Subcommittee on Hyperbilirubinemia. Practice parameter: management of hyperbilirubinemia in the healthy term newborn. *Pediatrics*, 94:558–565, 1994.
- [2] P.D. Clayton and G. Hripcsak. Decision Support in Healthcare. *International Journal of Bio-Medical Computing*, 39:59–66, 1995.
- [3] P.L. Elkin, M. Peleg, R. Lacson, E. Bernstam, S. Tu, A. Boxwala, R.A. Greenes, and E.H. Shortliffe. Toward Standardization of Electronic Guidelines. *MD Computing*, 17(6):39–44, 2000.
- [4] M.J. Field and K.N. Lohr, editors. *Clinical Practice Guidelines: Directions for a New Program*. National Academy Press, Washington D.C., USA, 1992.
- [5] J. Fox, N. Johns, C. Lyons, A. Rahmzadeh, R. Thomson, and P. Wilson. PROforma: a general technology for clinical decision support systems. *Computer Methods and Programs in Biomedicine*, 54:59–67, 1997.
- [6] I. Graham, L. Calder, P. Hébert, A. Carter, and J. Tetroe. A comparison of clinical practice guideline appraisal instruments. *International Journal of Technology Assessment in Health Care*, 16(4):1024–1038, 2000.
- [7] M. Marcos, G. Berger, F. van Harmelen, A. ten Teije, H. Roomans, and S. Miksch. Using critiquing for improving medical protocols: harder than it seems. In *Proc. of the 8th European Conference on Artificial Intelligence in Medicine (AIME-2001)*, pages 431–441, Cascais, Portugal, July 2001. Springer-Verlag.
- [8] L. Ohno-Machado, J.H. Gennari, S.N. Murphy, N.L. Jain, S.W. Tu, D.E. Oliver, E. Pattison-Gordon, R.A. Greenes, E.H. Shortliffe, and G. Octo Barnett. Guideline Interchange Format: a model for representing guidelines. *J. of the American Medical Informatics Association*, 5(4):357–372, 1998.
- [9] H. Roomans. Formalisation of a medical guideline. Usability investigation of the Asbru modelling language. Master's thesis, Universiteit van Amsterdam, 2001.
- [10] G.E.H.M. Rutten, S. Verhoeven, R.J. Heine, W.J.C. de Grauw, P.V.M. Cromme, K. Reenders, E. van Ballegooie, and T. Wiersma. NHG-Standaard Diabetes Mellitus Type 2 (eerste herziening). *Huisarts en Wetenschap*, 42(2):67–84, February 1999. First revision.
- [11] Y. Shahar, S. Miksch, and P. Johnson. The Asgaard project: a task-specific framework for the application and critiquing of time-oriented clinical guidelines. *Artificial Intelligence in Medicine*, 14:29–51, 1998.
- [12] Appraisal of Guidelines for Research & Evaluation (AGREE) Instrument, May 2001.
- [13] 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.