



Model-Based Safety Analysis of Human-Robot Interactions: the MIRAS Walking Assistance Robot

Jérémie Guiochet, Quynh Anh Do Hoang, Mohamed Kaâniche, David Powell

► To cite this version:

Jérémie Guiochet, Quynh Anh Do Hoang, Mohamed Kaâniche, David Powell. Model-Based Safety Analysis of Human-Robot Interactions: the MIRAS Walking Assistance Robot. International Conference on Rehabilitation Robotics (ICORR), Jun 2013, Seattle, United States. pp.1-7. hal-00839296

HAL Id: hal-00839296

<https://hal.science/hal-00839296>

Submitted on 27 Jun 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

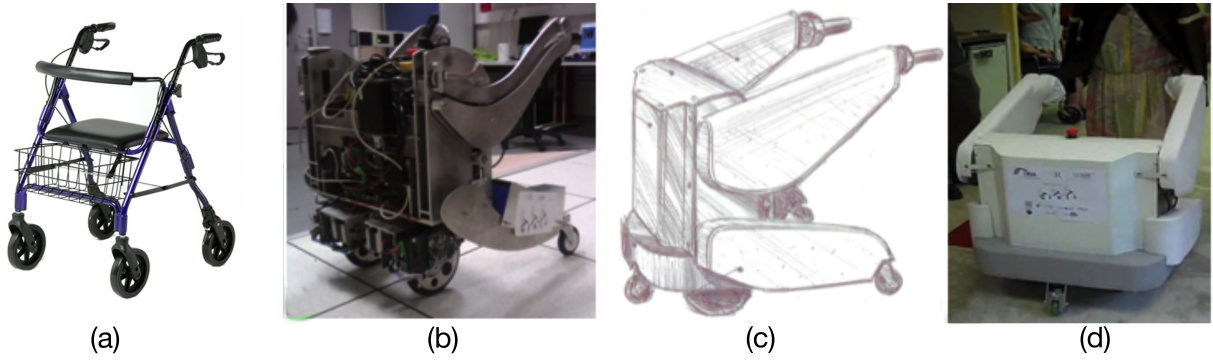


Fig. 1. (a) Classic "Rollator", (b) MIRAS experimental robot, (c) Design with packaging (d) Prototype during clinical investigation

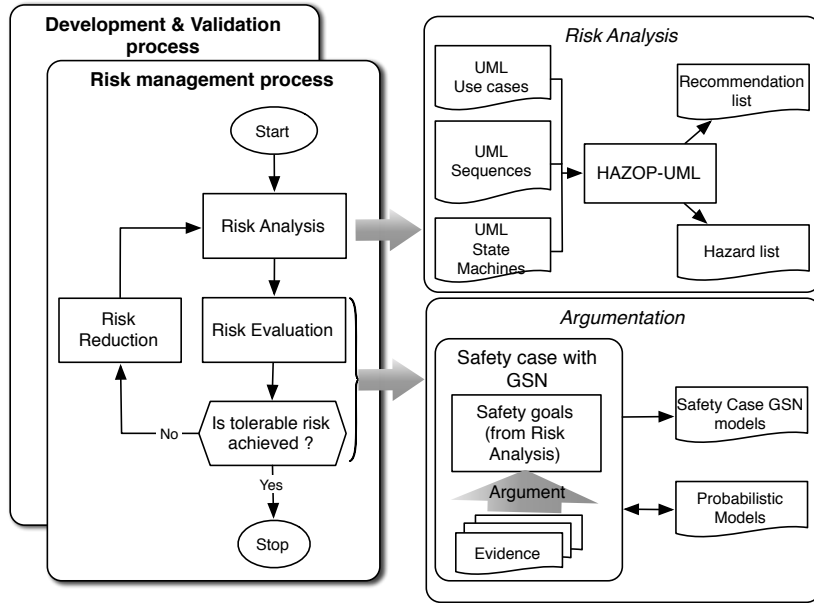


Fig. 2. Risk management and safety case in our safety analysis process

cope with those issues, we base our approach on the generic iterative risk management process, sharing models with the development process. We chose several modeling and analysis techniques presented in Figure 2.

Our risk analysis approach is based on a re-interpretation of HAZOP guidewords [8] in the context of certain UML models [9]. A similar approach has been followed in some previous studies considering UML structural diagrams [10]–[12] and dynamic diagrams [13]–[17]. We actually extend the results of those studies, focusing only on use case, sequence and state machine diagrams, in order to explore deviant behaviors during operational life. In particular, in this paper, we show through the application of the HAZOP-UML method to the robotized rollator how the three types of diagrams complement each other to identify relevant hazards.

During the development of the MIRAS robot, a list of recommendations was issued at each step of the risk analysis process: both UML modeling *per se* and the HAZOP analysis gave rise to general recommendations to enhance safety. Recommendations were fed back to the system developers at the user level (e.g., a new procedure for sitting on a chair), the specification level (e.g., the first prototype did not include

an integrated seat), and the design level (e.g., a heartbeat mechanism to regularly check the state of the robot and send an alarm to the medical staff in case of robot failure).

Once hazards have been identified, the argumentation with a safety case is carried out using the Goal Structuring Notation (GSN) [18]. This approach is based on expert judgement and strongly depends on expertise level. But, as mentioned in [19], a safety case argumentation can be built with probabilistic models. In the case study presented in this paper, we show how Markov models can be used to support the argumentation with GSN.

In the following sections, we illustrate our approach by applying it to the MIRAS robot and we summarize the main lessons learned.

III. RISK ANALYSIS

A. Hazard identification

The risk analysis starts with the description of the system with UML, focusing on use case, sequence and state machine diagrams. Our approach requires use cases to be described

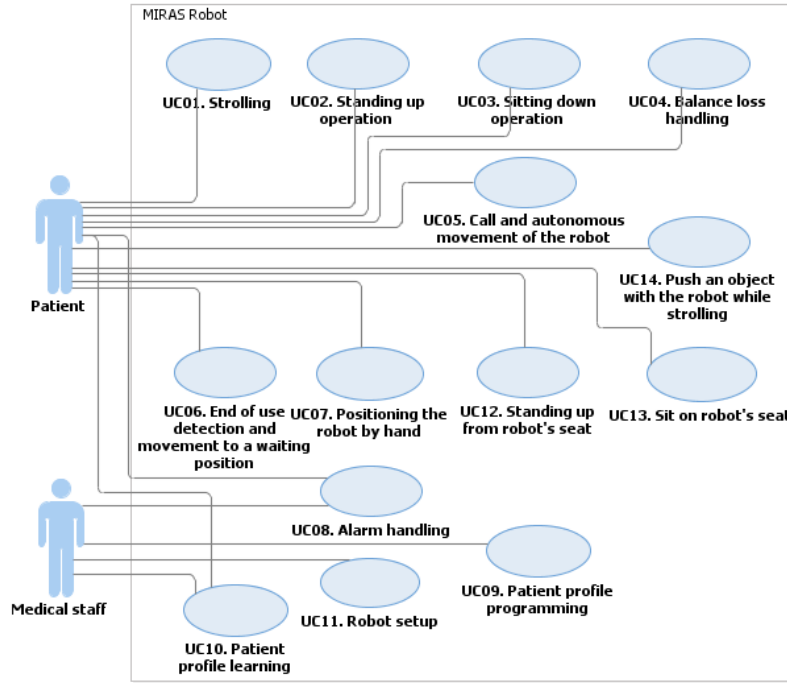


Fig. 3. Complete use case diagram of MIRAS project

Use case name	UC02. Standing up operation
Abstract	The patient stands up with the help of the robot
Precondition	The patient is sitting down The robot is waiting for the standing up operation Battery charge is sufficient to do this task and to help the patient to sit down The robot is in front of the patient
Postcondition	The patient is standing up The robot is in admittance mode
Invariant	The patient holds both handles of the robot The robot is in standing up mode Physiological parameters are acceptable

Fig. 4. Example use case description for deviation analysis

textually, with 3 required fields: preconditions, postconditions, and invariants. Invariants are conditions that must hold throughout all the use case execution. The complete use case diagram of the MIRAS project is shown in Figure 3 and an example of a textual description is presented in Figure 4. For each UML use case, which can be understood as an objective for users, there is at least one sequence diagram describing a scenario. The sequence diagrams of interest to us are “context sequence” diagrams, also called “system sequence” diagrams, i.e., scenarios described only considering human actors in the system environment and the system itself (and not all the components of the system) as in Figure 5.

Similarly, we consider state machines that specify the system-level behavior, and not the internal states of system components. From our experience, these diagrams are the most suitable for describing human-robot interactions at the start of the development process. The limitation to only these three diagrams has two major benefits. First, we have experienced that they are easily understandable by non UML experts, coming from the robotics and medical domains. Second, the use

of just 3 types of system-level models limits the combinatory explosion of hazard identification using HAZOP in the next step. Indeed, our case study led to the definition of 14 use cases, 15 sequence diagrams, and a state machine with 10 states. This produces a considerable amount of data but it is manageable with simple tools.

The second step is to identify hazards that can arise from the use of the robot. HAZOP-UML adapts the HAZOP [8] method to analyze deviations of the UML diagrams. According to the UK Defence Standard 00-58 [20], HAZOP analysis is the systematic identification of every deviation of every *attribute* of every *entity*. Each deviation is a potential hazard that can lead to a harmful event. We adapted the guideword lists to apply them to attributes of use case, sequence and state machine diagrams. Due to space limitation the lists of guidewords are not presented in this paper but can be found in [1].

Each resulting deviation defines a line of a HAZOP table. The analyst then establishes the effect at the use case level, and the result in the real world. The other columns of the table guide the analyst to establish a severity level, to deduce requirements and otherwise make remarks on that deviation. The guidewords also integrate human error models, that are analyzed in well-identified scenarios of use, showing also system response, which is not the case in many human error analysis methods (see [21]).

In the MIRAS project, the analysis of 397 possible deviations led to the identification of 16 hazard classes. Table I presents the main hazardous situations of the system, and the number of occurrences of the considered hazards when iden-

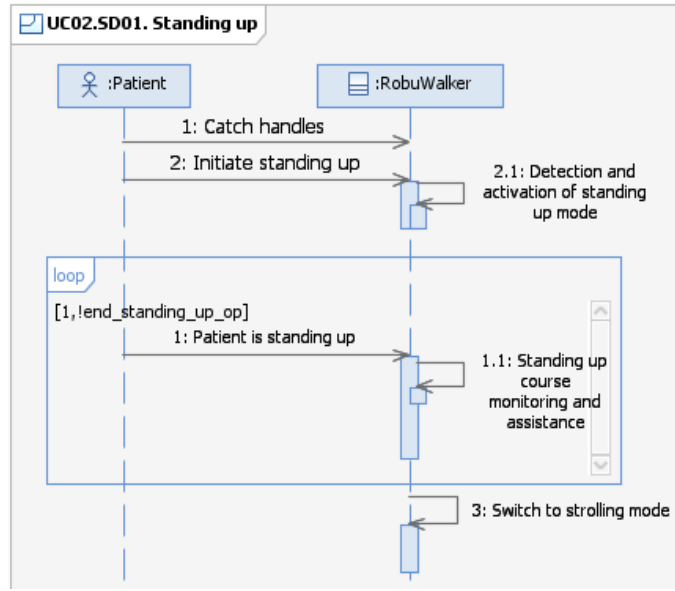


Fig. 5. Sequence diagram for UC02 ‘Standing Up Operation’

TABLE I. HAZARDS AND THEIR OCCURRENCE IN PHA AND HAZOP-UML ANALYSES

Num	Description	PHA	HAZOP-UML		
			UC	Seq.	State Machine
HN1	Incorrect posture of the patient during robot use	2	4	3	4
HN2	Fall of patient due to imbalance not caused by the robot		29	27	30
HN3	Robot shutdown during its use	1	2		5
HN4	Patient falls without alarm or with a late alarm		11	13	32
HN5	Physiological problem of the patient without alarm or with a late alarm		15	10	
HN6	Fall of the patient due to imbalance caused by the robot	10	51	37	10
HN7	Failure to switch to safe mode when a problem is detected. The robot keeps on moving		8		
HN8	Robot parts catching patient or clothes	3	5	4	
HN9	Collision between the robot (or robot part) and the patient	2	14	14	
HN10	Collision between the robot and a person other than the patient		5	14	2
HN11	Disturbance of medical staff during an intervention		1		
HN12	Patient loses his/her balance due to the robot (without falling)	11	1	70	1
HN13	Robot manipulation causes patient fatigue	12	1	53	21
HN14	Injuries of the patient due to robot sudden movements while carrying the patient on its seat			3	
HN15	Fall of the patient from the robot seat	2	10	12	
HN16	Frequent false positive alarms (false alarm)			3	

tifying risks with a Preliminary Hazard Analysis (PHA) based on use of checklists and brainstorming with domain experts, and with each of the three types of diagrams investigated with HAZOP-UML. For instance, HN11 (disturbance of medical staff during an intervention) has only been encountered during the analysis of the UML use case deviation analysis. HN5 (physiological problem without alarm or late alarm), which is a quite obvious hazard, was not mentioned during brainstorming workshops, but was identified 15 times during the analysis of use cases, and 10 times during the analysis of sequence diagrams. We observe that HAZOP-UML covers all hazards identified during the PHA. We also observe that use case and sequence diagram analyses were complementary (both identified hazards that the other did not). The state machine analysis did not identify any new hazards, although it did find new deviant behaviors that induced already identified hazards.

B. Hazard severity and likelihood levels

By definition, risk estimation should consist in estimating the severity and probability of occurrence of each potential harm. For that, analysts need to use probabilistic or ordinal scales. Some standards define such scales as examples, but there are none for robotic systems used in medical applications. Hence we collaborated with the doctors of three hospitals involved in the MIRAS project to establish a severity ranking scale that is suitable for the application context of the assistive robot considered in our study. For severity ranking, we first adapted a scale presented in [22], and asked the doctors to estimate the severity level of the identified hazards. This led to a redefinition of levels as presented Figure 6, adding an important dimension, which is the loss of confidence in the robot. Even if this is not directly related to safety, the psychological impact on the patients and the medical staff is

Severity Levels	
Level	Description
Catastrophic	Leads to patient's death
Critical	Leads to permanent deficiency or an injury putting in jeopardy patient's life
Serious	Leads to an injury (a) requiring intervention of health professional or (b) causing loss of patient's confidence in the system (with possible psychological impact)
Minor	Leads to a temporary injury (a) not requiring intervention of health professional or (b) causing medical staff to have less confidence in the system
Negligible	Causes annoyance or inconvenience

Fig. 6. Severity levels

Likelihood levels	
Level	Occurrence frequency
Extremely Frequent	~ once a week
Frequent	~ once a month
Probable	~ once every 6 months
Occasional	~ once a year
Remote	~ once every 10 years
Improbable	~ once every 100 years
Incredible	less than once every 100 years

Fig. 7. Frequency of occurrence levels

of great importance.

The second dimension of risk, frequency, was addressed at the same time as risk acceptance levels. We proceeded as follows: we defined three levels of acceptance according to the ALARP principle (which states that risk must be reduced to a level that is As Low As Reasonably Practicable [23]): *unacceptable* (risk cannot be justified except in extraordinary circumstances), *tolerable* (tolerable only if further risk reduction is impracticable or if its cost is grossly disproportionate to the improvement gained) or *acceptable* (negligible risk). Considering the patients' pathologies (some cannot talk or can hardly formulate structured sentences), it was not possible to ask them for risk acceptance criteria, so we decided to ask the doctors contributing to the study to assess for each hazard, the frequency (according to the scale defined in Figure 7) at which the investigated hazard could be considered acceptable, tolerable or unacceptable when using a single robot in their hospital. An example of a doctor response is given in Figure 8.

This led us to define precisely both likelihood levels and risk levels. There were some inconsistencies between the assessments of different doctors and hospitals, and also for a single doctor, sometimes estimating 2 hazards with the same severity level, but with different acceptability levels. We had several meetings and iterations, and finally reached a consensus. The result is presented in Figure 9. In particular, Figure 9 shows the severity level of the hazards listed in Table I. We chose to not present the occurrence frequencies per hour, because this was confusing when discussing with doctors, particularly when their estimation was at the boundary between two levels. It is for this reason that we have so many levels, compared to examples given in standards (3 or 5 levels in examples in [22]). It is noteworthy that the occurrence frequencies in Figure 7 are estimated with respect to calendar time. However, the average cumulated time of use of the MIRAS robot is estimated to be two hours per day. This leads us to consider that catastrophic events are "acceptable" at a target occurrence rate of "incredible". This rate is one event every 100 years, which, for 2 hours of use per day, is equivalent to 10^{-5} catastrophic events per hour. This is much more than in other safety critical applications, (usually around 10^{-7} catastrophic failures per hour is considered acceptable).

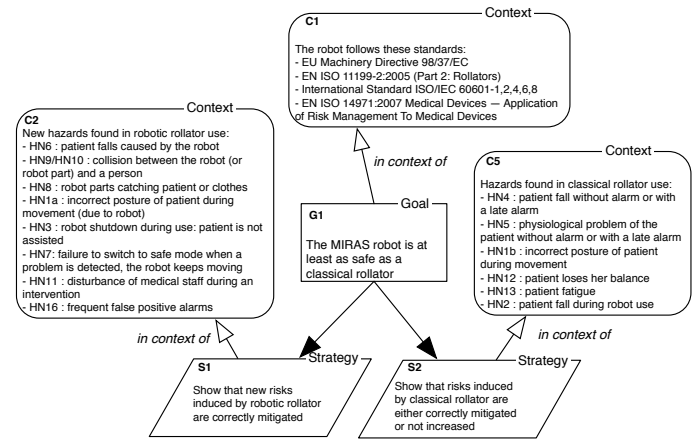


Fig. 10. The High Level Goal Structure of MIRAS robot

This over-estimation of the acceptable catastrophic failure rate can be explained by the fact that the estimation was carried out for a whole hospital service without considering the number of robots. If the number of robots deployed in the service is about 100 then the corresponding catastrophic failure rate per robot would be around 10^{-7} . Furthermore, we believe the doctors unconsciously performed a risk/benefit analysis when answering this question, which led them to be more tolerable with risks. For the considered study, we used their initial assessment because the objective was to use a single robot during clinical evaluation, in three hospitals.

IV. RISK EVALUATION

As presented before, the answer to the question "is tolerable risk achieved" should be based on a formal demonstration or at least a well-structured argumentation.

To do this, we used the Goal Structuring Notation (GSN) [18] and studied how HAZOP-UML outputs could be integrated in the argumentation. A first goal to be assessed was to compare the assistive robot to a classic rollator (also called frame walker). If the robot shows higher performance from the safety perspective compared to a traditional robot, the project will be successful. Hence, we have set as top-goal G1 the claim that: "The MIRAS robot is at least as safe as a classical rollator" (Figure 10). This goal is broken down into sub-goals through two strategies: we argue safety claims with respect to, on one hand, risks induced by the robot technology and, on the other hand, risks that are equally relevant to a classic rollator.

We then applied the GSN pattern defined in [24], which simply consists in creating as many subgoals as there are hazards to be addressed. In our case, we had 16 subgoals stated as "Hazard HN_x has been addressed". Then, for all subgoals, we identified various pieces of evidence according to the sub-goal to be solved: test results, estimation of error detection coverage and compensation efficiency, proof of correct implementation of code, failure rate of physical components, compliance with standards, etc. In our case, a list of 44 pieces of evidence to be collected has been identified.

A specific point in our application is that our system has many safety detection-reaction features for monitoring the patient's posture and health. In this context, any failure

Hazard Number	Severity	Frequency						
		Incredible	Improbable	Remote	Occasional	Probable	Frequent	Extremely Frequent
HN3	Critical	Acceptable	Tolerable	Tolerable	Not Acceptable	Not Acceptable	Not Acceptable	Not Acceptable

Fig. 8. An extract of a doctor response for risk evaluation levels elicitation

Catastrophic (HN4, HN6, HN14, HN15)							
Critical (HN3, HN5, HN8, HN10)							
Serious (HN7, HN9, HN12, HN13, HN16)							
Minor (HN1, HN11)							
Negligible							
Severity (Hazards)	Incredible	Improbable	Remote	Occasional	Probable	Frequent	Extremely Frequent
Frequency							

Acceptable
 Tolerable
 Not acceptable

Fig. 9. Elicited risk acceptability matrix, with hazard numbers

of such a feature is considered to be a hazard. For this we developed a pattern that we applied several times during our argumentation. It consists in decomposing the argumentation that a safety monitoring function is acceptable into three subgoals: a) there are no design faults, b) the failure rate of the supporting hardware is acceptable, and c) the monitoring function performance is acceptable (i.e., with an acceptable False Positive and False Negative rates). We used this pattern to argue about the safety of the robot functions designed for the monitoring of patient fall, physiological problems, or posture problems. Should any such undesired behavior be detected, the reaction would be to put the system in a safe state and trigger an alarm.

The robot also implements an imbalance compensation function that detects imbalance of the patient, and compensates with an appropriate movement (backward or forward) of the robot. This function is designed to reduce the consequences of patient imbalance, the worst final effect being a fall (see hazard HN2 in Table I). Of course, we need to argue that such a system effectively reduces the risk, and does not add new risks (such as bad compensation). This feature is analyzed in Figure 11 applying the previous pattern. Acceptability can be argued through the achievement of three goals: i) design faults are avoided or removed (G9.1) using rigorous development methods as suggested, e.g., in the IEC 61508 standard [23]; ii) the compensation system failure rate is shown to be acceptable (G9.2) using, e.g., using fault tree analysis and iii) the compensation function coverage factor (effective detection and compensation) is shown to be acceptable (G9.3) by testing under different patient imbalance scenarios.

To demonstrate G9, for which the acceptability criterion is established through discussions with medical experts, we need to demonstrate that the subgoals are satisfied, and estimate the parameters λ and μ presented in Figure 11. We performed this task using a Markov model [2].

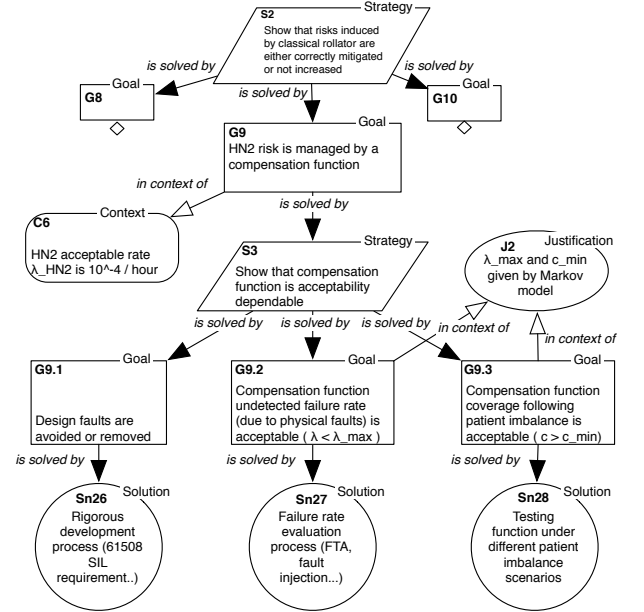


Fig. 11. Safety goal to be achieved by robot: compensate patient's loss of balance in time and in an efficient way

V. SUMMARY AND CONCLUSION

This paper has described our experience with applying model-based safety analysis techniques to an assistive robot. The main contribution has been to show how several complementary techniques can be integrated into a coherent process in which the uncertainties induced by the complexity of such a system are tackled in a pragmatic way.

The method has been applied to a practical case study: a walking assistance robot.

Our method can be evaluated according to four perspectives: **integrability** with the development process, **usability**,

applicability and validity.

Integrability: the UML design models for HAZOP-UML analysis were shared with the development process. This helps to avoid inconsistencies and supports maintainability of the analysis. Indeed, we found it relatively easy to revise the models following design changes, and to trace and update corresponding hazards when necessary. Our approach allowed safety analysis to be carried out not only at the beginning of the development, but also during design refinement. Nevertheless, HAZOP-UML should not be applied to detailed UML models because of the combinatory explosion of the number of deviations.

Usability: the overheads of using this method in the overall process were found to be acceptable. Apart from the modeling that was shared between stakeholders, the main overhead was that induced by the meetings for establishing levels for hazard severity and frequency, and risk acceptability. Safety case construction with GSN could also be time consuming. Indeed, choosing appropriate argumentation strategies strongly relies on the level of available expertise, but the structuring provided by the method facilitated communication with the consortium partners and allowed us to validate our choices. HAZOP-UML and GSN (integrating Markov chains) were reasonably manageable by hand but would be even more so with a tool to assist traceability and result formatting. We have developed an initial prototype of such a tool [1].

Applicability and validity: the first iteration of our risk assessment process led to conclusive evidence that the first prototype of the robot was not safe. Our recommendations for risk reduction were taken into account by the robotics experts and successfully integrated in the second prototype. Operational hazards identified during PHA were all covered by the HAZOP-UML analysis, which also identified additional hazards. Even though it is impossible, by principle, to judge the completeness of *any* hazard identification technique, we are comforted by the fact that no new hazards were discovered during laboratory tests of the robots.

The GSN-based safety case was used to justify safety to the French regulatory authority (AFSSAPS), which qualified the system to perform clinical testing. The tests will last two more years.

ACKNOWLEDGEMENTS

This work was partially supported by MIRAS, a project funded under the TecSan (Technologies for Healthcare) program of the French National Research Agency (ANR) and SAPHARI, a project funded under the 7th Framework Program of the European Union.

REFERENCES

- [1] D. Martin-Guillerez, J. Guiochet, D. Powell, and C. Zanon, "UML-based method for risk analysis of human-robot interaction," in *Int. Workshop on Software Engineering for Resilient Systems (SERENE2010)*, 2010.
- [2] Q. A. Do Hoang, J. Guiochet, M. Kaaniche, and D. Powell, "Human-robot interactions: model-based risk analysis and safety case construction," in *Embedded Real Time Software and Systems (ERTS2012)*, 2012.
- [3] J. Guiochet, Q. A. Do Hoang, M. Kaaniche, and D. Powell, "Applying existing standards to a medical rehabilitation robot: Limits and challenges," in *Workshop FW5: Safety in Human-Robot Coexistence & Interaction: How can Standardization and Research benefit from each other?*, IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS2012), 2012.
- [4] MIRAS, "Multimodal Interactive Robot for Assistance in Strolling," Project supported by the French ANR (National Research Agency) under the TecSan (Healthcare Technologies) Program (ANR-08-TECS-009-04), <http://www.miraswalker.com/index.php/en>.
- [5] ISO/IEC-Guide51, "Safety aspects - Guidelines for their inclusion in standards," International Organization for Standardization, 1999.
- [6] P. Bishop and R. Bloomfield, "A methodology for safety case development," in *Safety-Critical Systems Symp.*, 1998.
- [7] DefStan00-56, "Defence standard 00-56 issue 4: Safety management requirements for defence systems," Ministry of Defence, UK, 2007.
- [8] IEC61882, "Hazard and operability studies (HAZOP studies) - Application guide," International Electrotechnical Commission, 2001.
- [9] OMG-UML2, "OMG Unified Modeling Language (OMG UML), Superstructure, V2.1.2," Object Management Group, formal/2007-11-02, 2007.
- [10] K. M. Hansen, L. Wells, and T. Maier, "HAZOP analysis of UML-based software architecture descriptions of safety-critical systems," in *Nordic Workshop on UML and Software Modeling (NWUML04)*, 2004.
- [11] J. Gorski and A. Jarzebowicz, "Development and validation of a HAZOP-based inspection of UML models," in *3rd World Congress for Software Quality*, 2005.
- [12] A. Jarzebowicz and J. Gorski, "Empirical evaluation of reading techniques for UML models inspection," *ITSSA*, vol. 1, no. 2, pp. 103–110, 2006.
- [13] P. Johannessen, C. Grante, A. Alminger, U. Eklund, and J. Torin, "Hazard analysis in object oriented design of dependable systems," in *2001 Int. Conf. on Dependable Systems and Networks, Göteborg, Sweden*, 2001, pp. 507–512.
- [14] K. Allenby and T. Kelly, "Deriving safety requirements using scenarios," in *Requirements Engineering, 2001. Proceedings. Fifth IEEE Int. Symp. on*, 2001, pp. 228–235.
- [15] A. Arlow, C. Duffy, and J. McDermid, "Safety specification of the active traffic management control system for english motorways," in *The First Institution of Engineering and Technology Int. Conf. on System Safety*, 2006.
- [16] F. Iwu, A. Galloway, J. McDermid, and T. Ian, "Integrating safety and formal analyses using UML and PFS," *Reliability Engineering and System Safety*, vol. 92, no. 2, pp. 156–170, 2007.
- [17] T. Srivatanakul, "Security analysis with deviational techniques," Ph.D. dissertation, University of York, 2005.
- [18] T. P. Kelly, "Arguing safety – a systematic approach to managing safety cases," Ph.D. dissertation, University of York, 1998.
- [19] P. Bishop and R. Bloomfield, "The SHIP safety case approach," in *The Int. Conf. on Computer Safety, Reliability and Security (SAFE-COMP95)*, vol. 1. Springer, 1995, pp. 437–451.
- [20] DefStan00-58, "HAZOP studies on systems containing programmable electronics," Defence Standard, Ministry of Defence, UK, 2000.
- [21] N. Stanton, P. Salmon, G. Walker, C. Baber, and D. P. Jenkins, *Human Factors Methods: A Practical Guide for Engineering And Design*. Ashgate Publishing, 2006.
- [22] ISO/FDIS14971:2006, "Medical devices - Application of risk management to medical devices," International Standard Organisation, 2006.
- [23] IEC61508, "Functional safety of electrical/electronic/programmable electronic safety-related systems," International Electrotechnical Commission, Ed. 2, April 2010.
- [24] T. Kelly and J. McDermid, "Safety case construction and reuse using patterns," in *16th Int. Conf. on Computer Safety, Reliability and Security (SAFE-COMP97)*, 1997.