Medical Informatics Europe '97
C. Pappas et al. (Eds.)
IOS Press, 1997

Evaluation Plan for a Cardiological Multi-Media Workstation (I4C Project)

J.W. van der Hofstede[a], A.B.W.M. Quak[a], A.M. van Ginneken[b], P.W. Macfarlane[c], J. Watson[c], P.R.H. Hendriks[d], C. Zeelenberg[a]

[a] Division of Technology in Health Care, TNO Prevention and Health, Leiden, The Netherlands
[b] Department of Medical Informatics, Erasmus University Rotterdam, The Netherlands
[c] Department of Medical Cardiology, Royal Infirmary, Glasgow, Great Britain
[d] Software Engineering Research Centre, Utrecht, The Netherlands

Abstract. The goal of the I4C project (Integration and Communication for the Continuity of Cardiac Care) is to build a multi-media workstation for cardiac care and to assess its impact in the clinical setting. This paper describes the technical evaluation plan for the prototype.

1. Introduction

I4C (Integration and Communication for the Continuity of Cardiac Care) is a project within the Fourth Framework of the Health Telematics Programme of the European Union. The goal of I4C is to build a multi-media workstation for cardiac care, to demonstrate its use in a variety of clinical settings in different European countries and to assess its impact on cardiac clinical care.

The I4C workstation integrates data of many types from many sources into a virtual record, providing each user with views of the data tailored to match both user preference and the requirements of the particular aspects of clinical care relevant to the patient and care-giver. The workstation is intended to improve the efficiency and efficacy of cardiac care and to promote the exchange of patient data (for example, to support evidence-based medicine) by providing a consistent interface to all elements of the record, eliminating the need to access multiple systems and ensuring that all relevant information is available.

The virtual record contains:

- Administrative data (name, address, date of birth etc.)
- Computer patient record (past and current history, physical findings, diagnoses, problems and medication)
- Laboratory data
- ECG and Holter data
- ECHO data
- Angiographic data

The location of the various elements of the virtual record will depend on local conditions. In general, the administrative data and part or all of the computer patient record will be retrieved from the central hospital information system while the other data (laboratory, images) will come from dedicated laboratory systems. The ability to handle data from legacy systems is essential to this design. Data entered during a session will be stored in the patient record. In practice, response requirements will require local storage of some or all of the elements of the record. The workstation will then make use of the appointment list if available, retrieving the necessary items for scheduled patients in advance.

I4C is divided into two phases.

- Phase I - the development of a multimedia-workstation for cardiac care. This part of the project will result in the demonstration and technical evaluation of the workstation. The first phase of the project will end in 1997.
- Phase II - transformation of the demonstrator into a mature product, the clinical use and technology assessment of the workstation in cardiac practice in several hospitals in Europe, comparing the efficiency of cardiac care with and without the help of the I4C workstation.

This paper discusses the plan for a technical evaluation of the I4C workstation at the end of phase I. The QUINT methodology for quality assessment has been chosen for its emphasis on practical quality characteristics and on their quantification, enabling a more objective discussion of (software) quality. Quantifying the target levels of quality characteristics allows the developers to test intermediate results.

2. Technical evaluation of a multimedia workstation in a test situation

In phase I, a demonstrator will be developed; hence the evaluation will be limited to technical properties and performance.

In assessing a product, there are in general four groups of requirements:

- Process
- Internal products, i.e. aspects of the product that are important for the developers (e.g. in case of software: code, documentation)
- External products, i.e. aspects of the product that are important for the end-users (e.g. in case of software: executables, manuals)
- Resources

Citing Pfleeger et.al.: "Software engineering standards are heavy on process, and light on (external) product, while other engineering standards are the reverse. That is, software engineering standards reflect the implicit assumption that using certain techniques and processes, in concert with "good" tools and people, will necessarily result in a good product. Other engineering disciplines have far less faith in the process; they insist on evaluating the final product in their standards." [Pfleeger, 1994]

According to Trienekens [Trienekens, 1995], the available methods and techniques for improving the quality of software can be characterised in two different ways:

- a requirements-based classification: assessment of (external) product vs. process
- a goal-based classification: assurance vs. improvement

The following table shows the best known methods for software quality improvement (a discussion of the different methods is beyond the scope of this paper) in agreement with this classification:

	quality assurance	**quality improvement**
(external) product	ISO/IEC 9126 QUINT	SLA/PLA ITIL
process	ISO 9000-3 TickIT ANSI/IEEE Std 1012	CMM Bootstrap SPICE

The ISO 900x standards and the CMM (Capability Maturity Model) method assess predominantly the process while ISO/IEC 9126 is concerned with quality assurance of the product.
For the I4C project there are two principal reasons for applying a method that assesses the final software product and pays less attention to the process and the internal products:

- since the workstation is to be extended and customised by centres across Europe, each with their own 'informatics culture', it would be difficult to assess all these processes
- the I4C workstation has been built by prototyping, the final specifications will be fine-tuned during further development and customisation.

In other words, we will test the quality of the workstation for the end-users rather than the quality of the process of design and implementation. The international standard applicable to this type of evaluation is ISO/IEC 9126 "Information technology - Software product evaluation - Quality characteristics and guidelines for their use". The QUINT method is an extension to ISO/IEC 9126. QUINT stands for QUality in INformation Technology and is described in Dutch in [van Zeist, 1996a] (an English version will be marketed in 1997 [van Zeist, 1997]), and in English [Paulussen 1995] [van Zeist, 1996b]. The QUINT methodology involves more quality characteristics than does ISO/IEC 9126 and, more importantly, for each quality characteristic QUINT describes one or more indicators together with their measurement protocols. We have therefore decided to use the QUINT method to evaluate the I4C workstation quality.

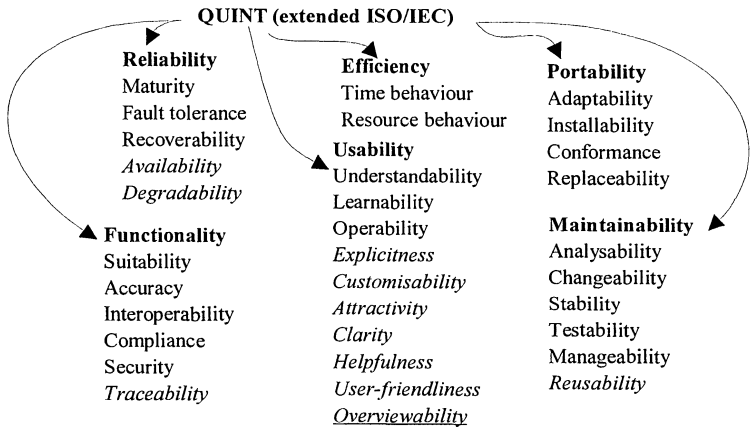

Figure 1. Six groups of quality characteristics. Extensions to the ISO/IEC 9126 are in italic, our addition is underlined.

3. Implementation of the QUINT method in the I4C project phase I

Using the QUINT method, there are three steps in writing the test plan:

1. Specification of quality requirements
    - selection of relevant and important characteristics
    - specification of characteristics with indicators
2. Specification of effort
    - planning of effort needed for testing

3. Testing
    - building the tests
    - performing measurements
    - analysing the results

The choices for all quality characteristics are discussed in the test plan. If any quality characteristic is discarded, the reason is explicitly stated.

4. Work done so far

We classified the quality characteristics (see figure 1) as follows:

- *Usability* (which includes the users' perception of time response) is the most important quality aspect in the I4C project., The aim of the demonstrator is to prove the value of a multimedia workstation for cardiologists and this aim will only be realised if cardiologists feel comfortable with the workstation.
- The specifications for the project have been presented as a description of basic functionality, data dictionaries and scenarios. Each scenario describes the standard procedures for a specific type of patient. These scenarios should be easy to follow using the workstation. Since the level of abstraction of the specification is that of a description of the functionality, the demonstrator is largely being developed by prototyping. This places heavy demands on *maintainability*, making it a factor of great importance.
- The *suitability* (e.g. the proper implementation of previously described scenarios) ranks highest of the functionality characteristics.
- The three characteristics, *reliability, efficiency* and *portability* are of great importance for a mature product but are of less importance for a demonstrator. For the I4C demonstrator, it is far more important to pay attention to implementing the desired functionality in a very usable product with extremely flexible software.

We described the relevant quality characteristics with one or more associated measurement protocols for each characteristic selected. These are described in a chronological test plan, taking two days of testing per test site. The description is such that the developers of the workstation can test individual parts of the software according to this plan. Specific code to perform the tests was discussed and implemented.

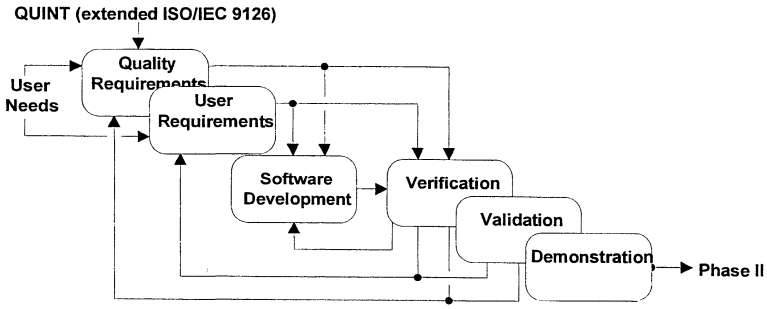

Figure 2. From user and quality requirements, the project will iterate toward the demonstration at the end of phase I.

During the evolution of the workstation, functional specifications may be added or reviewed and the quality characteristics and measurement methods revised (see figure 2). We further prioritised the individual quality characteristics by roughly defining acceptable and target levels. This helps the developers to judge their intermediate results against these goals.

5. Future work

Prior to the final evaluation of the demonstrator, we will make the necessary measurements in order to state realistic minimal levels and challenging target levels for the selected quality characteristics.
At the end of 1997, the demonstrator will be evaluated in several European hospitals. The results will be described in a test report. Deviations from the original plan will be documented and discussed.
After passing the evaluation stage, the demonstrator will be transformed into a more mature product. The evaluation of that version of the workstation will have a similar usability and functionality to that of the demonstrator; in contrast, the evaluation at that stage will concentrate primarily on reliability and on efficiency and portability characteristics.
Only when the product meets the predefined criteria can the clinical stage of technological assessment be initiated. We expect to reach that stage in 1998.

6. Conclusions

Application of the QUINT method with its emphasis on practical quality characteristics and metrics has made awareness of quality issues in software development a fundamental aspect of the I4C project and facilitated objective assessment of quality.

References

R.M.C. Paulussen, The QUINT Approach to the Specification of Software Quality. 1st World Congress for Software Quality.

S.L. Pfleeger, N.Fenton and S.Page, Evaluating Software Engineering Standards. IEEE Computer, September 1994, 71-79.

J. Trienekens, Afnemers van IT produkten verlangen garantie. Automatiserings Gids, **8**, friday 24 february 1995, 15.

B. van Zeist, P. Hendriks, R. Paulussen en J. Trienekens, Kwaliteit van softwareprodukten. Kluwer Bedrijfswetenschappen, Deventer, Holland, 1996.

R.H.J. van Zeist and P.R.H. Hendriks, Specifying Software Quality with the Extended ISO Model. In: M. Bray, M. Ross, G. Staples (editors), Software Quality Management IV Improving Quality, Mechanical Engineering Publications Limited, London and Bury St Edmunds, UK, 1996, pp. 145 - 159.

B. van Zeist, P. Hendriks, R. Paulussen and J. Trienekens, Quality of Software Products - Presentation of a Framework and Practical Experiences. To be published.