

Research Findings from an Industrial Trial of a Traceability Assessment and Implementation Framework

Gilbert Regan
Regulated Software Research Centre
(RSRC)
Lero @DKIT
Ireland
gilbert.regan@dkit.ie

Derek Flood
RSRC
Lero @DKIT
Ireland
derek.flood@dkit.ie

Fergal Mc Caffery
RSRC
Lero @DKIT
Ireland
fergal.mccaffery@dkit.ie

ABSTRACT

Software systems are becoming increasingly complex. Within safety critical domains such as medical device software, this increasing complexity is placing growing demands on manufacturers who must ensure their software not only meets functional requirements but is also safe and reliable. However, the Food and Drugs Administration who regulate medical device software in the United States report a significant increase in recalls between years 2003 and 2012 and have cited software difficulties as one of the frequent causes of recalls. Furthermore a recent analysis of traceability documentation submitted to the Administration has revealed that the traceability data was incomplete, incorrect, and conflicting in many cases. This is problematic as traceability plays an important role in the development of safe and reliable software. In this paper we present the validation, through industry trial, of a traceability assessment and implementation framework which we have developed to assist medical device organizations implement traceability in an efficient and regulatory compliant manner. Our findings show that implementation of the framework within two organizations improved their traceability process and that both organizations found the framework to be both useful and usable.

CCS Concepts

• **Software and its engineering** ~Traceability *Software verification and validation*

Keywords

Process Assessment, Roadmap, Software Process Improvement, Medical Device

1. INTRODUCTION

Traceability, which is the ability to establish and use the links (or traces) between source artefacts and target artefacts [7].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org. ICSSP'16, May 14-15 2016, Austin, TX, USA

Copyright is held by the owner/author(s). Publication rights licensed to ACM.

Traceability is important as developers need to ensure that their software conforms to customer expectations, functional and regulatory requirements. Additionally, traceability is important for impact analysis as it assists developers to visualize how a change in any component can affect the software system. Therefore an effective traceability process supports the development of quality software and reduces the possibility of delays, recalls, or even complete project failures. Despite this traceability is problematic as *'most existing software systems lack explicit traceability links between artefacts'* [8].

Organizations that produce safety critical software must ensure their software is safe, reliable and available. Safety critical software can be described as software, which if fails could cause damage to the environment, injury or death [22]. Within the safety critical medical device domain, medical devices must frequently interface with other equipment, connect to hospital and laboratory information systems, and work in high-stress situations. The increased demands on such devices has resulted in increased software complexity and has created formidable development challenges for their manufacturers [13]. A medical device can be described as any article intended for use in the diagnosis, care, treatment or prevention of disease that does not achieve its intended purpose by medicines [6]. A recent report by the United States Food and Drugs Administration (FDA) finds that the annual number of medical device recalls increased by 97 percent (from 600 to 1200) between years 2003 and 2012 and lists software difficulties as one of the most frequent causes of recall [3]. An analysis of medical device recalls in the EU reports similar figures to that of the United States [2], indicating a growing and significant problem. As in other domains, traceability implementation within the medical device domain is challenging and has been highlighted by a recent analysis of the traceability documents submitted to regulators in the United States as part of the medical device approval process. This analysis has revealed that the traceability data was incomplete, incorrect, and conflicting in many cases [1].

To determine the challenges to implementing traceability effectively, the authors conducted a literature review in addition to interviewing two medical device organizations [14, 17]. As a result a number of challenges were identified including a lack of guidance on what traceability to implement and how to implement it. To assist medical device organizations overcome this challenge the authors have developed a traceability assessment and implementation framework (TAIF). The framework consists of a traceability assessment model (TAM) and a traceability roadmap (TR). The TAM identifies any gaps in an organizations

implementation of traceability best practices [18] and/or medical device standards' traceability requirements[19], and the TR provides the pathway for an organization to plug those gaps. An initial validation of the TAIF was completed through expert review [15, 16]. Further validation of the model has been completed through its implementation in two medical device organizations. In this paper we present the results of that implementation and evaluate the ability of the TAIF to assist medical device software organizations implement effective traceability and put them on the path to regulatory compliance.

The remainder of this paper is structured as follows: Section 2, outlines current assessment models' relationship to traceability. Section 3 outlines the methodology used to evaluate the roadmap while Section 4 presents the results of the evaluation. Section 5 provides for a discussion of the results while Section 6 concludes the paper.

2. RELATED WORK

A literature review was conducted by the authors to determine what other traceability assessment models or traceability roadmaps were available in the general, safety critical or medical device domains.

2.1 Roadmaps

A Framework for Requirements Traceability [12] has been proposed which provides guidelines to assist with implementing traceability. This framework advocates the use of 'Value Based Requirements Tracing' which may not be an option if full tracing is a requirement of the customer or the medical device development process standards. Additionally this framework does not fulfil the requirements of a traceability implementation roadmap as defined in the introduction to this paper.

'Software Traceability: A Roadmap' [23] presents a roadmap of research and practices related to software traceability and identify issues that are still open for further research. While this paper does highlight barriers to implementing traceability (and areas for further research) it does not provide a traceability implementation roadmap.

A number of medical device standards implementation roadmaps have been developed [4, 5]. These roadmaps provide a pathway for medical device organizations to implement particular medical device standards for risk management, usability design, and quality management. However these roadmaps only provide traceability requirements from any one standard and there is no detail on how these requirements might be achieved.

2.2 Traceability Assessment

To assist organizations to improve their traceability researchers have published work on the assessment of traceability. Med-trace [10] is a lightweight traceability assessment method (containing 8 stages) to assist medical device organizations to improve their software development traceability process. However, this assessment method is 'lightweight' in nature (as described by the authors), does not assess any of the best practices for implementing traceability, and its development was not based on any recognized international standard e.g. ISO 15504 [9].

An approach to automate the assessment of traceability has been presented in response to 'the gap between what is described by guidelines and what is implemented in practice' [20]. This approach, which is not specific to any particular domain, involves

automatically parsing project data to capture trace information. While this approach automatically assesses the standards' requirements for traceability it does not assess the implementation of traceability best practices.

There are a number of process assessment models which provide common frameworks for assessing software process capability. These models include ISO/IEC 15504 SPICE, Automotive SPICE [14], SPICE 4 SPACE [15], and CMMI [21]. These frameworks assess processes such as the software design process, software construction process, software testing process etc. However, the frameworks do not include a dedicated traceability assessment process. The frameworks do include elements of traceability assessment but it is spread out across a lot of processes and the traceability requirements are sometimes difficult to interpret.

3. INDUSTRY VALIDATION

In order to meet the aim of this study which is the evaluation of the TAIF, a longitudinal study of two medical device organizations was performed over a six month period. For the study the Adept process appraisal method [11] was employed as it could provide a comprehensive assessment of an organizations' traceability compliance without being too onerous in terms of resources that the organization need to provide.

An initial assessment was conducted in both organisations using the TAM which highlighted areas for improvement. A findings report was delivered to the organisation along with the TR which provided a pathway for the organisations to improve their traceability. A follow up assessment conducted six months later (again using the TAM) indicated what improvements were made. Additionally, after this follow up assessment, two representatives from each organization were interviewed about their experience of using the TAIF and how beneficial they found it to be. The interviews were semi-structured in nature as this provided a means to address any issues or points that may arise during the interview that may not have been predicted during the generation of the questions. The interview questionnaire was divided into two main parts. Part A examined the value and usefulness of the assessment experience to the organization. Part B examined the usefulness and usability of the roadmap in implementing the recommendations that had resulted from the initial assessment.

3.1 Participants

Organization 'A' is a small medical device software company, founded in 2002, based in Ireland. The company has a total of ten employees which include one programmer, one software tester and one quality assurance person. The company produces medical device software with a software safety classification of B, meaning non-serious injury is possible.

Organization 'B' is a small Product Development & Design Engineering company focused on the Medical Device and Life Science market and is based in Ireland. The company, which was formed in 2007, employs 14 individuals with skills in mechanical, hardware and software engineering. The company are a third party supplier of software to medical device companies and have recently been accredited with IEC 62304 certification.

4. RESULTS

The results of the initial traceability implementation assessment and follow up assessment are presented in section 4.1. Section 4.2 presents the results of evaluating the assessment experience while section 4.3 presents the results of the roadmap evaluation. The data presented in section 5.1 was collected through document analysis and interviews as part of the Adept method while the data presented in section 5.2 and 5.3 was collected through interviews.

4.1 Processes Assessment

The results of the initial assessment and final assessment are presented in tabular format in Table 1 and for the purpose of clarity the results are also presented as percentages in Table 2.

Table 1. Best Practices and Traceability Links Implementation- Initial and Final Assessment

Initial Assess	Best Practices and Trace Links Implemented	Not Implemented	Final Assess
	Best Practices		
	Company policy on traceability		
	Traceability improvement communication method		B
	Traceability Information Model		
	Traceability standard operating procedure		
	Resources to implement traceability unavailable		B
	Appropriate techniques not deployed		B
	Risk Management Traceability Requirement		
	Link hazardous situation to software item		B
	Link software item to software cause		
B	Link software cause to risk control measure		B
B	Link RCM to risk control verification		A,B
B	Link Hazard to assessment of residual risk		B
	SDLC Traceability Requirement		
B	Link software req. to system req. and their source		A,B
	Link software req. to source code (where required)		
B	Link soft. arch. design to soft. detailed design		B
B	Link soft. detailed design to source code		B
B	Link soft arch. design to test		A,B
B	Link soft detailed design to test		B
B	Link source code to test		B
	Change Management Traceability Requirement		
B	Link change request to problem report		A,B

Table 1, presents the traceability best practices and traceability requirements that were *not* implemented by either organisation 'A' or 'B' during the initial assessment (Column 1.) and during the final assessment (Column 3). For example, both organizations 'A' and 'B' had no traceability improvement communication method in place at the initial assessment. However, at the final assessment organisation 'B' had met this requirement. This means that organisation 'B' implemented a traceability improvement communication method between the initial and the final assessments while organisation A had not. Table 2 indicates the percentage achievement of the outcomes for each process. For example, organisation 'A' had fully achieved three (or 75%) of the outcomes in the initial assessment and in the final assessment they had fully achieved all four outcomes (i.e. 100%). As indicated in Table 1 the purpose of this outcome is to implement traceability between each change request and the problem report ID from which it originated. The organization achieved this by adding a 'source of change request' section to their change control form.

Table 2. Results of Initial Assessment and Final Assessment

	Best Pract.		SDLC Trac		RM Trac		CM Trac	
	Init.	Final	Init.	Final	Init.	Final	Init.	Final
Org. A	5%	5%	22%	44%	40%	53%	75%	100%
Org. B	10%	41%	89%	89%	75%	88%	100%	100%

The results indicate an improvement in all processes except for organization A's Best Practice process which remained at 5% of full implementation and organization B's SDLC process and CM process (which had already been 100% complete).

4.2 Evaluation of the Assessment Experience

To fully evaluate the TAIF it was necessary to appraise the participating organizations experience of undergoing the assessment. Therefore, questions were formulated to ascertain the participating organizations perspective with regard to the value and usefulness of the assessment to the organizations.

Within the assessment, all questions asked the users to respond using a five point Likert scale. In some cases answers ranged from strongly disagree (1) to strongly agree (5), or very bad (1) to very good (5). Additionally, the organizations were asked to provide a rationale for their decision.

4.2.1 Value of Assessment

With regard to the value of the assessment, questions were asked during the final assessment to determine:

- The organization's understanding of the assessment process and its results;
- The organization's understanding of the purpose of the activities that took place as part of the assessment;
- The organization's perspective on whether the assessment was worth the time and effort expended?
- The organization's view on how good or bad the assessment results characterize their traceability strength or weaknesses.

The results are presented in Figure 4.

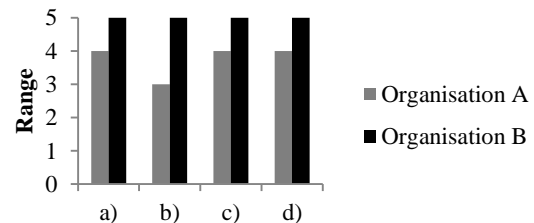


Figure 4. Value of Assessment

As an example, in response to a question on whether the assessment was worth the time and effort expended, organization 'A' responded with a '4' and provided the following rationale: *"Our organization did not have the knowledge or the expertise required. The assessment identified many areas that current practice was lacking. Extensive knowledge and experience would be required to produce the quality of report provided by the assessment....very few people would have this"*

4.2.2 Usefulness of assessment

To evaluate the usefulness of the assessment, questions were formulated to determine the organization's viewpoints with regards to:

- How the assessment helped us better understand what needs to be improved with regards to traceability?

- b) If the assessment improved awareness, buy-in and support for traceability within our organization;
- c) If the assessment model provides real direction for long term traceability improvement?

The results are presented in Figure 5.

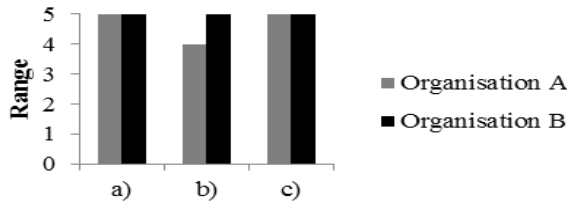


Figure 5. Usefulness of Assessment

As an example, in response to a question on how much the assessment helped the organization better understand what needs to be traced, organization ‘B’ returned a ‘5’ and provided the following rationale: *“We are definitely much more aware of what needs to be traced. Although we were aware that traceability was a requirement of the standards we obviously weren’t aware of how much was required. The assessment has highlighted our deficiencies and as we are a small organization everyone is now very aware”*.

4.3 Evaluation of the Roadmap

To evaluate the roadmap, questions were formulated to determine the participating organization’s perspective of the roadmap with regard to its usefulness and usability. These questions assessed their agreement with:

- a) The overview section is easy to understand;
- b) The correctness of the order of implementation as depicted in Overview section;
- c) The roadmap is useful in practice;
- d) The roadmap is usable in practice;
- e) The roadmap provides sufficient guidance.

The results are presented in Figure 6.

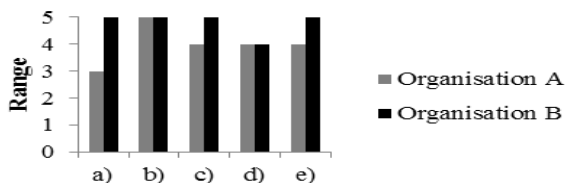


Figure 6. Roadmap - Usefulness and Usability

As an example, in response to a question on how much they agreed that the roadmap provides sufficient guidance on how to go about improving traceability, organization ‘B’ scored this category with a ‘5’ and provided the following rationale: *“The roadmap is comprehensive. It explains what you have to do and is not difficult to comprehend.”*

In addition to the above, a number of questions were asked to determine the organization’s viewpoint with regard to any deficiency they have observed with the roadmap, any suggestions they have for improvement, or what they thought were the major benefits of the roadmap. Organization ‘A’ suggested that it might help if the best practices could be prioritized, with ‘Resources’ as first best practice. The major benefit of the roadmap as far as organization ‘A’ was concerned was that *“it provided them with an expertise that was not available within the organization and that they do not believe many organizations would have such a*

wealth of knowledge on the subject’. The fact that the roadmap detailed all the medical device traceability requirements was a major benefit for organization ‘B’ as they felt that the standards can be difficult to interpret. As an example of the usefulness and usability of the roadmap organization ‘A’ stated that for the CM process they were easily able to link change request to problem report as *“Method B in Appendix C of the roadmap clearly defined how we could easily do this simply by adding a source of change request section to the change request form”*.

5. DISCUSSION

The inadequate implementation of the best practice process in both organizations is clearly evident in Table 1. Organization ‘A’ indicated that the reason for this was that they were simply unaware of the best practices and did not have any expertise in this area. Additionally, the reason that they showed no improvement in this area was that they focused their resources towards the other three processes for two reasons: the SDLC, RM and CM processes contained ‘regulatory requirements’ whereas the best practice process did not; their resources were particularly limited during this time period with some personnel leaving the organization and not yet been replaced. Organization ‘A’ has indicated that they wish to achieve 100% implementation across all four processes but with their very limited resources they were not able to indicate when this would happen. While the results from organization ‘B’ indicated a marked improvement in their implementation of the best practices process they are acutely aware that they still have significant work to do in this area. They have indicated that the reason why further improvement has not been achieved is one of resources. However, organization ‘B’ has set itself a goal of achieving 100% implementation across the four process areas and to this extent a new person has been hired with part of his remit being to work towards this goal.

The results of the evaluation of the assessment experience have been very positive as indicated in Figures 4 and 5. Figure 4 point b indicates that organization ‘A’ could have had a better understanding of the purpose of the activities that took place during the assessment. While Step 2 of the Adept method requires the assessment process be explained to staff prior to the assessment, for future assessments a more detailed explanation of the process will be delivered at Step 2. However it is believed that this did not have any effect on the results of the assessment. Another area noted for improvement is that the assessment report could better characterize the organizations traceability strong points, with organization ‘A’ stating: *“although the assessment did expose many areas that we are not covering in current TM template, we are compliant in some areas which maybe could be more evident in the report”*.

While the evaluation of the assessment process has highlighted areas for improvement, the lightweight nature of the assessment was noted as a major benefit with organization ‘B’ stating: *“it was good that the assessment interviews were done in less than 1.5 hours as time is a limited resource in this organization”*.

With regard to the evaluation of the roadmap and as depicted in Figure 6, organization ‘A’ neither agreed or disagreed that the overview section was easy to understand and suggested that some form of color-coding would improve it and make it visually more appealing. On reflection, the authors consider this to be a valid point and so the overview is amended to highlight the SDLC, CM and RM processes (in red) and also the adoption of the best practices (in blue) through these processes. Additionally,

organization 'A' suggested that the 'resource' best practice should come first and that the best practices should be given a priority rating. However, textual information in the overview section indicates that an organization should decide at an organization level, or the project manager should decide at the project level which of the traceability best practices to implement. Factors which will influence these decisions will include the availability of resources and how highly regarded traceability is within the organization. Additionally, the roadmap recommends that the best practices are implemented in the order presented in the roadmap where possible. This order of implementation arose as a result of the roadmap validation through expert review.

6. CONCLUSION

The results of implementing the TAIF within two medical device organizations clearly indicate a significant improvement in almost all traceability processes within both organizations, although none of the processes improved to a level of 100% adherence, the overriding reason given for this is resource issues (mostly time) within the organizations. These resource issues are clearly outside the scope of the TAIF.

An evaluation of the assessment experience and traceability roadmap clearly indicates that both organizations found them to be very useful, usable and generally providing an awareness, knowledge and expertise in traceability that they did not previously have. Additionally, both organizations considered that the assessment provided them with guidance and real direction for long term traceability improvement.

7. ACKNOWLEDGEMENT

This work was supported with the financial support of the Science Foundation Ireland grants 10/CE/11855 & 13/RC/20194 to Lero - the Irish Software Research Centre.

8. REFERENCES

- [1] Cleland-Huang, J., Gotel, O., Hayes, J. H., Mäder, P. and Zisman, A. 2014. Software Traceability: Trends and Future Directions. In *Proceedings of the ICSE* (Hyderabad, India, 2014)
- [2] Davis, S., Gilbertson, E. and Goodall, S. 2011 EU Medical Device Safety Assessment - A comparative analysis of medical device recalls 2005-2009. *BCG*.
- [3] FDA. 2014 *Medical Device Recall Reports FY2003 to FY2012*.
- [4] Flood, D., Caffery, F. M. and Casey, V. 2013. A roadmap to implement a quality management system. In *Proceedings of the 6th International conference on Health Informatics (Healthinf 2013)* (2013)
- [5] Flood, D., Caffery, F. M., Casey, V., McKeever, R. and Rust, P. 2015 A roadmap to ISO 14971 implementation. *Journal of Software: Evolution and Process*
- [6] Fries, R. 2013. *Reliable Design of Medical Devices*. CRC Press, 978-1-4398-9491-0, New York, CRC Press, 93
- [7] Gotel, O. and Mader, P. 2012. *Acquiring Tool Support for Traceability*. Springer, Acquiring Tool Support for Traceability, New York, Springer
- [8] Hayes, J. H. and Dekhtyar, A. 2009. Towards Traceable Test-Driven Development. In *Proceedings of the TEFSE'09* (Vancouver, Canada, 2009)
- [9] ISO/IEC. 1997 *15504-SPICE (Software Process Improvement and Capability dEtermination)*. ISO
- [10] McCaffery, F. and Casey, V. 2011. Med-Trace. In *Proceedings of the Software Process Improvement and Capability Determination* (Dublin, Ireland, May 30 – June 1, 2011, 2011). Springer
- [11] McCaffery, F., Richardson, I. and Coleman, G. 2006. Adept – A Software Process Appraisal Method for Small to Medium-sized Software Development Organisations. In *Proceedings of the EuroSPI06* (Finland, 2006)
- [12] Raja, U. and Kamran, K. 2008 *Framework for Requirements Traceability*. Blekinge Institute of Technology
- [13] Rakitin, S. R. 2006 *Coping with Defective Software in Medical Devices*. IEEE.
- [14] Regan, G., Caffery, F. M., Daid, K. M. and Flood, D. 2013. Investigation of Traceability within a Medical Device Organization. In *Proceedings of the SPICE* (Bremen, Germany, June 4 - 6, 2013). Springer
- [15] Regan, G., Caffery, F. M., Daid, K. M. and Flood, D. 2014. The Development and Validation of a Traceability Assessment Model. In *Proceedings of the SPICE* (Vilnius, Lithuania, 4-6 Nov 2014, 2014). Springer
- [16] Regan, G., Flood, D. and McCaffery, F. 2015. The Development and Validation of a Roadmap for Traceability. In *Proceedings of the SPICE* (Gothenberg, Sweden, 2015). Springer
- [17] Regan, G., Mc Caffery, F., Mc Daid, K. and Flood, D. 2012. The Barriers to Traceability and their Potential Solutions: Towards a Reference Framework. In *Proceedings of the 38th Euromicro Conference on Software Engineering and Advanced Applications* (Cesme, Turkey, 5-8 Sept 2012, 2012). IEEE
- [18] Regan, G., Mc Caffery, F., Mc Daid, K. and Flood, D. 2013. Implementation of traceability best practices within the medical device domain. In *Proceedings of the EuroSPI* (Dundalk, Ireland, 2013). DELTA
- [19] Regan, G., Mc Caffery, F., Mc Daid, K. and Flood, D. 2013 *Medical device standards' requirements for traceability during the software development lifecycle and implementation of a traceability assessment model*. Elsevier.
- [20] Rempel, P., Mader, P., Kuschke, T. and Cleland-Huang, J. 2014. Mind the Gap: Assessing the Conformance of Software Traceability to Relevant Guidelines. In *Proceedings of the ICSE 2014* (Hyderabad, India, 31-05-2014, 2014)
- [21] SEI. 2011 *Standard CMMI® Appraisal Method for Process Improvement (SCAMPI) A, Version 1.3: Method Definition Document*. SEI
- [22] Sommerville, I. 2011. *Software Engineering*. Addison-Wesley, ISBN-13: 978-0-13-703515-1, Addison-Wesley

[23] Spanoudakis, G. and Zisman, A. 2004 Software Traceability: A Roadmap. *Handbook of Software Engineering and Knowledge Engineering*.