

---

## A FRAMEWORK FOR SELECTION OF TEST METHOD AND TEST INTERVAL FOR SAFETY CRITICAL VALVES IN SITUATIONS WITH LIMITED DATA

**E.B. Abrahamsen**

University of Stavanger, Norway  
e-mail: eirik.b.abrahamsen@uis.no

**W. Røed**

•  
Proactima AS, Norway  
e-mail: wr@proactima.com

### ABSTRACT

In this paper we present a practical framework that can support the decision on test methods and test intervals for safety critical valves in situations where limited relevant historical data is available. The framework is based upon a systematic review of the valve functions and associated failure modes, as well as properties of the environment that the valve is located in, and evaluation of this knowledge in qualitative expert workshops. The main application area for the framework is valves located upstream or downstream of large gas transport pipeline segments. The framework is however general and can be applied to smaller hydrocarbon segments as well.

### 1 INTRODUCTION

This paper focuses on periodically tested items as part of passive safety critical systems. For such systems a failure can only be revealed during test or if the safety system is activated. There are many examples of such systems, but for the purpose of this paper, we consider safety critical valves. These valves are supposed to close and sectionalize the main process in case of an emergency situation at a process plant. It is vital that the safety critical valves are functioning on demand, as the consequences of failure could be significant, for example with reference to economy, production assurance, safety and emissions to the environment.

The paper is inspired by studying large land-based valves on the Norwegian gas distribution network. Many of these valves are unique in terms that there are no, or only a few, other valves that can be considered ‘similar’. In addition to this, these valves were originally not designed to be tested at all, and have therefore only occasionally been tested. As a result, there is very little relevant test information that can be used to determine test methods and test intervals for some of the above-mentioned valves. But how can we in an appropriate way decide upon test methods and test intervals for safety critical valves when just limited relevant data is available? This question is the starting point for this paper.

Traditionally, decisions upon test methods have not been well-documented. It is, however, our impression that the decision on test method has primarily been made, with respect to the operational disadvantages of performing the test, for example in terms of shut-down time, rather than on a systematic review of the pros and cons of each alternative. It is also a challenge that tests are performed without considering the functions of the valve, i.e. the reason that they are installed. As discussed in Røed et. al. (2008), this may for example result in testing for potential leakage in one direction, while the most critical valve function is that the valve should not leak in the opposite direction.

For determination of test intervals a number of methods within reliability theory exist. See for example Rausand and Høyland (2004) and Aven (1991). Such methods are all based upon the availability of historical data considered relevant for the future performance of the system in focus. For most systems considered by reliability engineers, relevant data is available, making the above-

mentioned methods relevant. But for the safety critical valves which are the starting point for this paper, little or limited historical data is available. The question may then be asked: ‘How can test intervals be decided upon when no relevant data is available?’. This is certainly a challenge, but decisions on test intervals are also made in situations with no or limited data. Some key questions are then: What kind of information should be used to determine test intervals when no or limited data is available? And how should we take this information into consideration to ensure that decisions are made with the best knowledge available?

It may be argued that due to lack of test data we should implement short test intervals to be on the safe side. This is, however, not necessarily a good strategy since most of the test methods have operational disadvantages when the test is performed. This is not limited to economic loss in terms of reduced production assurance only. We see that some test methods also imply flaring/venting of large amounts of greenhouse gases. Serious accidents when performing a test may also occur. All in all, the above information means that there are many aspects that have to be taken into consideration when decisions are made for both test methods and test intervals.

This paper proposes a framework that takes the above-mentioned aspects into consideration, to suggest a maintenance regime for safety critical valves. The framework is based upon performing several steps, where initially the most appropriate test method is selected, and thereafter the test interval is decided upon for the selected test method. The framework is relevant in situations where little or no relevant historical data is available, but there is still some qualitative knowledge of the system in focus.

## 2 A FRAMEWORK FOR SELECTION OF TEST METHOD AND TEST INTERVAL FOR SAFETY CRITICAL VALVES

The proposed framework consists of three main steps: A, B and C, as illustrated in Table 1. In the first step, the activities to be carried out are prepared, including system understanding, identification of relevant consequence dimensions, identification of the functions that the valve is installed to achieve and potential failure modes. In the second step the most critical failure mode is identified. This is important as no test method can control all the failure modes of the valve. The third step of the framework is carried out in order to select a test interval for the test method selected in Step B.

Table 1. Framework for selection of test method and test interval

<b>STEP A – Initial planning activities</b>
Clarify boundaries of the system and corresponding segments
Identify relevant consequence dimensions
Identify the valve’s functions and corresponding failure modes
Collect relevant historical test results
<b>STEP B – Selection of test method</b>
Identify the most critical failure mode
Select test method for the most critical failure mode
<b>STEP C – Selection of test interval</b>
Identify potential causes of each of the failure modes
Describe the valve’s operating conditions
Evaluate the test interval

We suggest that Step B and parts of Step C are carried out in a multidisciplinary group. In the following we explain in more detail each step of this procedure. Each of the steps is explained by use of a realistic, although simplified, case study. The steps A, B and C are presented in Sections 2.1, 2.2 and 2.3, respectively.

## 2.1 Step A: Initial planning activities

In order to support decisions it is important that the boundaries of the analysis system are defined. It is also important to achieve an understanding of the boundaries of the corresponding hydrocarbon inventories, since this information is important when the consequences of a failure are going to be described.

To illustrate our ideas we will use the system as shown in Figure 1, including four valves and two pressurized hydrocarbon inventories. In this example, attention will be given to deciding upon the test method for valve 1 (V1) in this system.

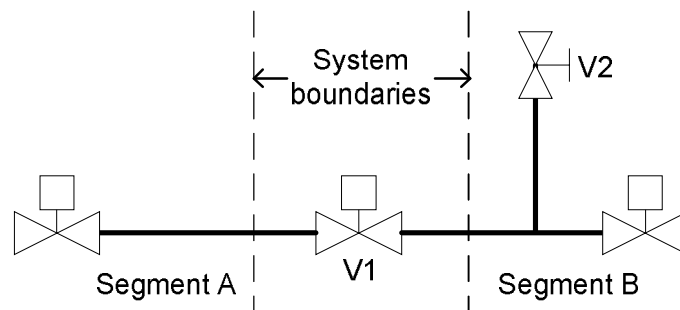


Figure 1: System description (simplified).

The test method decision will depend on several consequence dimensions such as production assurance, safety of personnel, costs and environmental aspects. The framework described in this paper is based on an analysis of each consequence dimension separately. Then the problem of weighting the consequences against each other, as part of the analysis, is avoided. We believe that it is important to leave the weighing of the consequence dimensions for the decision maker, and have not included this in the analysis. Arguments for our view can be found for example in Aven (2003). As part of the planning activity the functions of the valve in focus should also be determined. A function may be seen as a reason for the valve to be installed. Examples of valve functions are:

- F1: Isolate in case of leakage in segment A
- F2: Isolate in case of leakage in segment B
- F3: Isolate in case of maintenance of valve 2

It is important to ensure that the valve function formulations reflect the ‘real’ reasons that the valve is installed, since this may affect subsequent parts of the assessment process. An example: Consider F1 in the list above. What does ‘isolate in case of leakage in segment A’ mean? Does the valve have to isolate the segment completely, or is it sufficient that a large leak through the valve is prevented? Perhaps the valve function should be re-phrased to ‘Reduce leakage through valve in closed position to maximum X kg/s’? The way the valve function is formulated is important, since in the first case, the method may end up with the recommendation of a differential pressure test, while in the latter case, a partial stroke test may be sufficient. To ensure that the valve functions are correctly phrased, we could use control questions, for example whether complete isolation is needed or whether achievement of the function within a longer period of time is sufficient.

Next, the failure modes of the valve should be determined. Examples of failure modes are:

- FM1: The valve does not close on demand
- FM2: Leakage through closed valve from segment A to segment B
- FM3: Leakage through closed valve from segment B to segment A

The failure modes are closely related to the valve functions since, if a failure mode occurs, one or several valve functions are lost. The relationship between the valve functions and the failure mode should be described. An example of such a relationship is shown in Table 2.

Table 2. Relationship between valve functions and failure modes.

Failure modes \ Functions	FM1: Does not close on demand	FM2: Leakage through closed valve from segment A to segment B	FM3: Leakage through closed valve from segment B to segment A
F1: Isolate in case of leakage in segment A	X	X	
F2: Isolate in case of leakage in segment B	X		X
F3: Isolate in case of maintenance of V2	X		X

Additionally, relevant information about the valve in focus should be obtained and described. For example, dependencies between the valve in focus and other systems should be determined. Examples of such dependencies include: Does failure to open or close the valve result in shut down of other process plants? To what extent may testing of the valve be carried out in periods that corresponding systems are shut down? The above may largely affect production assurance and economic loss, in the first case if a valve failure occurs, and in the latter case during testing.

Finally, historical test results for the valve in focus and similar valves should be collected. As presented in the introduction to the paper, our focus is on situations where no or limited relevant data is available. In the case of relevant data being available, a data-driven approach would be a natural starting point.

## 2.2 Step B: Selection of test method

Most safety critical valves have several failure modes, and the most critical failure mode should be identified in order to select a suitable test method. The most critical failure mode can be identified by using a traditional FMECA (Failure Modes, Effects and Criticality Analysis). But there is a need for clarification (a guideline) for how such an analysis should be carried out in order to identify the most critical failure mode for a safety critical valve. Such a guideline is presented in Section 2.2.1. In Section 2.2.2 we describe how to decide upon test method based on the information received from the analysis in Section 2.2.1.

### 2.2.1 Identification of the most critical FM

In the proposed framework the criticality for each failure mode is based on an evaluation of what the consequences may be if a failure mode occurs, and an evaluation of the probability of a failure mode occurring.

A description of the activities in the second part of the framework (Step B) is given in the following.

*Evaluation of potential consequences*

In this step of the approach we evaluate the potential consequences of each failure mode identified in Step A. Each failure mode may result in a number of scenarios with different consequences. Due to this, the assignment process has to be based upon a systematic approach. In the suggested approach the potential consequences of a failure mode are not assigned directly. Instead, attention is paid to the potential consequences when the *valve functions* are lost. Thereafter, the evaluations can be transferred to failure modes by using the relation between valve functions and failure modes obtained in Step A.

In the suggested procedure a qualitative approach, based on a multidisciplinary meeting and expert judgment, is used to assign the consequences of losing each valve function. Attention is paid to the consequences if a best-case scenario occurs, the consequences if a worst-case scenario occurs and the expected consequences. Evaluation of consequences in best and worst cases is important, and of special interest, taking into consideration that the consequences of losing a valve function can be very different in many situations. For example, in the case of an external leak from a pipeline to the atmosphere, followed by loss of the isolation function (F1 or F2), many consequences are possible, ranging from an unignited leak to a large explosion escalating to other equipment.

The expert group decides whether or not the consequences of losing a valve function are high, medium or low, based on an overall evaluation of the consequences for worst- and best-case scenarios as well as the expected consequences. An evaluation of the probability of being in the best and worst state is certainly important when deciding upon the consequence category. The categorization process should be based on some guidelines or criteria to ensure consistency. For each of the consequence dimensions (safety, production, environment, etc.) it should be defined what is meant by low, medium and high consequences.

A simplified version of a questionnaire that can be used in the evaluation of consequences is given in Table 3.

Table 3. Questionnaire used in the evaluation of potential consequences of losing valve functions.

Functions	SAFETY						PRODUCTION						ETC.
	Cons. given loss of funct. on dem						Cons. given loss of funct. on dem						
	Worst case	Best case	Ex-pected	L	M	H	Worst case	Best case	Ex-pected	L	M	H	
<b>F1: Isolate in case of leakage in segment A</b>	Text	Text	Text	X			Text	Text	Text			X	
<b>F2: Isolate in case of leakage in segment B</b>	Text	Text	Text		X		Text	Text	Text		X		
<b>F3: Isolate in case of maintenance of V2</b>	Text	Text	Text	X			Text	Text	Text		X		

The next step is to transfer the evaluations to failure modes by using the relation between valve functions and failure modes obtained in Step A. This is done by combining the information in Table 2 and Table 3. For example, from Table 2 we see that all valve functions (F1-F3) are lost if the valve does not close on demand (FM1 occurs). The safety consequences of losing F1, F2 and F3 are, as seen from Table 3, low, medium and low, respectively. We then consider the safety consequences if the valve does not close on demand equal to the worst consequence category for the relevant valve functions. This means that the safety consequences for failure mode 1 are within the medium category. All the results of this process are given in Table 4.

Table 4. Consequence assignment category for each failure mode.

Failure modes	SAFE-TY			PROD-UCT.			ETC.
	L	M	H	L	M	H	
FM1: Does not close on demand		X				X	
FM2: Leakage through closed valve from A to B	X					X	
FM3: Leakage through closed valve from B to A		X		X			

*Evaluation of the probability of a failure mode occurring*

The next step is to assign the probability of each failure mode occurring. Since the starting point for the paper is that little or no historical relevant data is available, we suggest that the probability of a failure mode occurring is revealed in a multidisciplinary meeting by use of expert facilitation. This can be carried out in the same meeting as the one in which the consequences are assigned. The procedure proposed uses a classification scheme with three categories: high, medium and low. The categorization should be based on some criteria to ensure consistency.

In some situations it is required to understand the potential causes of each of the failure modes in order to express the probability of the failure mode occurring. This information is revealed during Step C of the method. In case this information is required to assign the probability category, we suggest that parts of Step C are carried out before Step B is finalized.

In the evaluation of the probability of a failure mode occurring, there is a need for taking the uncertainties into consideration. We may for example consider the probability of a failure mode occurring within the category medium, but because of high uncertainties we say that the probability is considered to be within the high probability category. This procedure is justified in Abrahamsen and Røed (2010). The point is that the evaluations on the probabilities are based on some background information. In mathematical terms you may look at the probability of failure mode occurring as  $P(\text{failure mode occurring} | K)$  where K is the background information and knowledge. The background knowledge covers historical system performance data, system performance characteristics and knowledge about the phenomena in question. Assumptions and presuppositions are an important part of this information and knowledge. The background knowledge can be viewed as frame conditions of the analysis, and the assigned probabilities must always be seen in relation to these conditions. Thus, different analysts could come up with different values, depending on assumptions and presuppositions. The differences could be very large. Hence, uncertainty needs to be considered, beyond the assigned probability number (Aven, 2008). Similar ideas are found underpinning approaches such as the risk governance framework (Renn, 2008) and the risk framework used by the UK Cabinet Office (Cabinet Office, 2002).

*Criticality evaluation*

Deciding on the criticality for each failure mode is based on the information in the consequence assessment process and on the assigned probability of a failure mode occurring. The categorization of the criticality should be based on some guidelines to ensure consistency. One possible way to categorize the criticality of the failure modes is given in Table 5.

Table 5. Criticality categories of failure modes.

ASSIGNED CATEGORIES		RESULT
Consequence given that the failure mode occurs on demand	Probability of failure mode occurring	Criticality category
Medium	High	<b>High</b>
High	Medium	
High	High	
Low	Low	<b>Medium</b>
Low	Medium	
Medium	Low	
Low	High	<b>Low</b>
Medium	Medium	
High	Low	

The idea is to allocate the failure modes into the three criticality categories, based on the probability and consequence assignments. The most critical failure mode should be brought forward to Step C of the framework. Since the assignment process is carried out for the consequence dimensions separately, the failure mode considered the most critical may differ from one consequence dimension to another. In other words, the failure mode considered the most critical with regard to production assurance may not be the same as the one considered the most critical with regard to safety of personnel. We will come back to this in the discussion; cf. Section 3. In some cases, by following the suggested procedure, several failure modes will have equal criticality categorization. We will revisit this situation as well in the discussion; cf. Section 3.

### 2.2.2 Selection of test method for the most critical FM

Selection of test method is based upon identification of alternative test methods, evaluation of potential consequences, evaluation of the reliability of each test method and evaluation of the overall performance of each test method. These activities are presented in the following.

#### *Identification of alternative test methods*

All the possible ways the most critical failure mode can be tested should be listed. Examples of test methods are:

- Pressure differential across the valve
- Cavity test
- Partial stroke test
- ...

Note that only those test methods able to test the failure mode classified as the most critical in the first part of the proposed framework should be included.

#### *Evaluation of the potential consequences of each test method*

Before a test is carried out, we do not know what the consequences of performing the test will be for the different consequence dimensions. We may predict that performing the test will result in a number of hours'/days' loss of production, a certain amount of gas flaring (resulting in CO<sub>2</sub> emission to the environment) and no harm to the personnel performing the test, but we should also describe other potential outcomes. Some test consequences depend on factors outside the analysis

object. For example, in the case where the test can be performed when the facility at the other end of the pipeline is shut down, the consequences in terms of costs may be reduced. In many cases, the outcome of performing a test is associated with uncertainty. The test may for example take a longer or shorter time, or it may, in the worst case, result in an ignited external gas leak exposing the test personnel to fire or explosion.

In the suggested procedure a qualitative approach, based on a multidisciplinary meeting and expert judgment, is used to assign the potential consequences of performing each of the test methods. The consequence category is based on an overall evaluation of the consequences for worst- and best-case scenarios and the expected consequences in the same way as mentioned during identification of the most critical failure mode. These evaluations can be carried out in the same meeting as mentioned above. Table 6 presents a simplified version of the questionnaire.

Table 6. Questionnaire used to assign the potential consequences of performing each of the

Test methods	SAFETY						PRODUCTION						ETC.
	Consequence of performing test						Consequence of performing test						
	Worst case	Best case	Ex-pected	L	M	H	Worst case	Best case	Ex-pected	L	M	H	
<b>T1: Pressure diff. across the valve</b>	Text	Text	Text			X	Text	Text	Text			X	
<b>T2: Cavity test</b>	Text	Text	Text		X		Text	Text	Text		X		
<b>T3: Partial stroke test</b>	Text	Text	Text	X			Text	Text	Text	X			

alternative tests.

*Evaluation of the reliability of each test method*

Another important aspect to take into consideration is the reliability of each test method. In many cases there are great differences between test methods in how much the test results can be trusted. While a positive outcome of one test method gives high trust that the valve will work as intended on demand, a positive outcome of another test method only gives an indication that this is the case. For example, if a test including pressure differential across the valve tells us that there is no leak, this result may be trusted more than the result from a partial stroke test telling us that the valve can be turned. The latter test method gives no confidence that the valve is not leaking in the closed position. We suggest that three qualitative categories are used for the reliability assignment in line with the former assignments: high, medium and low reliability; ref. Table 7 below.

Table 7. Categorization of test reliability.

Test methods	Test reliability			
	Descr- iption	L	M	H
<b>T1: Pressure diff. across the valve</b>	Text			X
<b>T2: Cavity test</b>	Text		X	
<b>T3: Partial stroke test</b>	Text	X		

*Evaluation of the overall performance of each test method*



The overall performance of each test method is made based on the reliability and consequence assignments. The categorization should be based on some guidelines to ensure consistency. One possible way to categorize the overall performance of each test method is given in Table 8.

Table 8. The overall performance/quality of test methods.

ASSIGNED CATEGORIES		RESULT
Reliability	Consequence of performing the test	Overall performance
H	L	High
H	M	Medium
H	H	
M	L	
M	H	
L	L	
M	H	
L	M	Low
L	H	

### 2.3 Step C: Selection of test interval

In Step C a decision about test interval for the test method in Step B is made. This is determined by first considering the potential causes of each of the failure modes revealed in Step A. Next, the valve’s operating conditions are described. A test interval suggestion is then revealed by combining the above-mentioned information. In practice, this is done by expert facilitation in the previously mentioned multidisciplinary group consisting of personnel with thorough understanding of the valve/system in focus. The different activities in Step C are presented in the following.

#### *Identify potential causes of each of the failure modes*

The expert group should systematically go through each of the failure modes revealed in Step A, and identify potential causes in terms of failures within the valve. To carry out this activity, detailed knowledge of the valve in focus is needed. As mentioned in the introduction to the paper, many of the large safety critical valves are unique, having no, or just a few, comparable items. The differences are primarily related to technical details within the valves, and for each valve in focus it is vital to understand and describe the design of the valve. To reveal this information, it may be necessary to study in detail information from the vendor delivering the valve, and this means that parts of the study may be carried out in the planning phase (as part of Step A). We will not in this paper go into detail on valve design, since the results will only be valid for one particular valve, and since the framework is the key focus of the paper. We do, however, emphasize that for most large safety critical valves, each failure mode may have a number of potential causes.

Each of the potential causes should be described qualitatively, based on discussions in the group. In the case of there being redundancy within the design, e.g. double seals, the extent to which this redundancy can be tested should be discussed. In the case where it is not possible to test that redundant components are working as they should - a situation that is common for safety critical valves - it is recommended not to take the redundancy into consideration when deciding upon the test intervals. In the opposite case, potential dependencies and common cause errors should be discussed. In most cases there will be a list of potential causes of each of the failure modes. For the failure mode ‘leakage through the valve in closed position’, potential causes can for example be ‘seats in wrong position’, ‘damaged seals’ and ‘leak behind the valve’s seats’, etc.

### *Describe the valve's operating conditions*

The valve's operating conditions should be described by use of a brainstorming process. It is important that all aspects that may affect the valve's performance are revealed. We suggest that check lists with guide words are used as part of the brainstorming process in order to reveal as many relevant operating conditions as possible. Examples of guide words are 'internal', 'external', 'temperature' and 'corresponding equipment'. Examples of relevant operating conditions that may be revealed by use of expert facilitation in combination with guide words are harsh weather, vibration, sand in the HC flow, hydrate build-up and regular pigging.

### *Evaluate the test interval*

In this activity, the relationship between operating conditions and potential causes of each of the failure modes should be revealed and discussed, and this information should be used to evaluate the test interval. We suggest that the activity is carried out by performing the following steps:

- Reveal the relationship between operating conditions and potential causes of valve failure
- Discuss time to occurrence for the relationships revealed in the step above
- Evaluate the test interval based on the above information

These activities are discussed below:

Firstly, the relationship between the operating conditions and the potential causes of valve failure should be revealed. In most cases this can be carried out in the work group. A simple way to reveal the relationship is to consider the lists of operating conditions and potential causes, and draw lines between the elements on the lists. In this way, a simple influence diagram is made showing which of the operating conditions may cause a valve error.

The next step is to discuss time to occurrence for the relationships revealed in the step above. In order to cover this in a proper manner, in-depth knowledge of the system in focus is required. In some cases, the work group will have sufficient information to assign approximate time to failure. In other cases additional knowledge will be required, for example technical design information and other information from vendors.

In most cases, assigning time to failure is not a simple process, due to lack of complete understanding of design, degradation mechanisms and how such mechanisms are affected by the operating conditions. Due to this, there are, in most cases, considerable uncertainties. These uncertainties should somehow be included in the discussion during the assignment process. This can be taken into consideration by assigning the expected time to failure as well as best-case and worst-case values.

The last step is to evaluate the test interval based on the above information. In principle, the shortest of the above assigned times to failure should be chosen. Due to the uncertainties there is, however, a need to make an overall decision taking all the above-mentioned aspects into consideration.

In cases where a test interval less frequent than once a year is decided upon, we recommend that an annual review is carried out in order to reveal if there is any new information that may affect the test method and test interval evaluations. For example, suppose we have revealed that a valve may have been damaged during a pigging operation during the last year. Then the test method and test interval evaluations should be updated, taking the new information into consideration. In most cases, however, the annual review will conclude that no relevant new information is gained, and that the maintenance programme can be followed as scheduled. The annual review will not be particularly time consuming since it can be carried out by a few persons, and in many cases can be carried out for several valves at a facility at the same time.

### 3 DISCUSSION AND CONCLUSION

The qualitative framework suggested in this paper is particularly relevant in situations where the safety critical items are unique and limited historical test data is available. In such situations, the ability to learn from the past through relevant test experience is limited, and alternative sources of knowledge should be considered. The framework shows how expert knowledge can serve as an alternative source of information. The framework is based on a structured approach where key evaluations can be carried out in a multidisciplinary work group during one or a few days. Due to the limited need for resources, the framework can work as an alternative to more arbitrary - and in some cases unstructured - approaches. Since the evaluations carried out are documented in work sheets, the framework also provides documentation of test method decisions subsequent to the work group meeting. This documentation can also serve as transfer of knowledge from the valve experts, having key knowledge of the systems in focus, to other staff.

A key aspect of the framework is, however, that additional knowledge is gained each time a test is carried out. Since this knowledge is taken into consideration in the framework, the test method considered the best in one case may not be considered the best next time the framework is applied. Suppose, for example, that the framework is applied to a safety critical valve resulting in a differential pressure test being considered the best. Thereafter, this test is carried out. Then, next time the framework is applied, the test results and other information gained during the test are taken into consideration. The result may be that another test method is considered as the best next time. This dynamic characteristic of the framework makes it possible over time to build insight on the properties and performance of the valve in focus.

As mentioned in Section 2, there may be situations where several failure modes are considered equally critical. In such cases, all test methods testing one or several of the critical failure modes revealed in the initial part of Step B should be considered when the test method is decided upon. In such situations, the relationship between the failure modes and the test methods obtained in the initial part of Step B should be taken into consideration in the overall evaluation of test method performance in the last part of Step B.

One aspect of the framework is that the consequence dimensions such as, for example, production assurance, safety, environmental impact and economy are treated separately. This increases the number of evaluations that will have to be carried out during the multidisciplinary meeting. Certainly, it would also be possible to combine the consequence dimensions, and thus, to obtain a less comprehensive procedure. We do, however, argue that separating these dimensions is necessary to provide sufficient decision support: when a result is obtained, we need to know whether the test method suggested is the 'best' with regard to production assurance or safety. The above means that the method may result in several test methods being recommended, regarding different consequence dimensions. This is not a problem since there should always be a step from decision support on the one hand, and the decision and implementation, on the other hand.

Some readers would perhaps claim that the framework is not sufficiently founded on reliability theory. To this, the answer is that alternative methods founded on reliability theory have a common need for sufficient relevant test data. We do indeed agree that in the case of such test data being available, traditional reliability methods can and should be applied. But the purpose of the framework presented in this paper is to obtain decision support in situations where no or only limited data is available. And for such situations, the framework presented can provide additional decision support, and can serve as an alternative to more arbitrary - and in many cases unstructured - approaches. After all, decisions on test methods and test intervals are also made in situations with limited data available.

As argued in Abrahamsen and Røed (2010), decision processes regarding safety related items should not be mechanistic. There is always a need for broad evaluations taking all relevant aspects into consideration before decisions are made. This is also the case for decision situations

regarding choice of test method and test interval: Although the framework could serve as a tool to decide upon a test method and a test interval in a mechanistic manner, it should not be used in such a way. There is always a need for broader reflections before a decision is made. This emphasizes that all assumptions made during the evaluations should be provided to the decision maker. And in the case of there being something about the result that can be questioned, further evaluations should be carried out to obtain further decision support.

The method presented in this paper concludes on one test method and one corresponding test interval for one individual valve. This is of course a simplification, since a complete maintenance programme may consist of several test methods with one test interval for each test method. It is also a challenge that only one individual valve is considered, since to ensure that the maintenance programme is manageable in practice, it may be argued that the valves should be put into groups with one dedicated combination of test methods/ intervals for each group. It is believed that the method presented can be further developed in the future to take these aspects into consideration. For the time being, the focus has, however, been on a simplified challenge considering one individual valve, one test method and one associated test interval. By gaining experience with such a simplification, we will be able to acquire information that can be used to develop more sophisticated methods in the future.

## ACKNOWLEDGEMENT

The authors are grateful to the Ramona research project and its contributors for financial support, valuable discussions and access to relevant information that has been a source of inspiration to the method presented in the paper. We would also like to thank Terje Aven at the University of Stavanger for many useful comments and suggestions to an earlier version of this paper.

## REFERENCES

1. Abrahamsen EB, Røed W. 2010. A semi-quantitative approach for evaluation of cost-effectiveness of safety measures. *Business Review*, Cambridge 14(2): 134-140.
2. Aven T. 2008. *Risk analysis – Assessing uncertainties beyond expected values and probabilities*. Chichester: Wiley.
3. Aven T. 2003. *Foundations of risk analysis: A knowledge and decision-oriented perspective*. New York: Wiley.
4. Aven T. 1991. *Reliability and risk analysis*. Elsevier, London.
5. Cabinet Office. 2002. *Risk: Improving government's capability to handle risk and uncertainty*, Summary report (The Strategy Unit, London).
6. IEC – International Electrotechnical Commission. 2010. IEC 615108. *Functional safety of electrical/electronic/programmable electronic safety-related systems*. International Electrotechnical Commission, Geneva.
7. Rausand M., Høyland A. 2004. *System reliability theory: Models, statistical methods and applications*. Wiley.
8. Renn O. 2008. *Risk governance: coping with uncertainty in a complex world*. London: Earthscan.
9. Røed W., Haver, K., Wiencke, H.S. and Nøklund, T.E. 2008. *Consequence based methodology to determine acceptable leakage rate through closed safety critical valves*. Proceedings of the European Safety and Reliability Conference (ESREL) 2008.