CHALMERS
UNIVERSITY OF TECHNOLOGY

Facilitating Engineering Process Alignment, Quality, and Productivity through Cross-Functional Standardization

Development of a Framework for Structuring Engineering Activities in a Multi-Site Context at SKF

Master’s thesis in Industrial Design Engineering & Mobility Engineering

ALFRED HALLIN
LINUS FÄLDT

DEPARTMENT OF INDUSTRIAL AND MATERIAL SCIENCE

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2026
www.chalmers.se

Master’s thesis 2026

Facilitating Engineering Process Alignment, Quality, and Productivity through Cross-Functional Standardization

Development of a Framework for Structuring Engineering Activities in a Multi-Site Context at SKF

ALFRED HALLIN
LINUS FÄLDT

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Industrial and Materials Science
Chalmers University of Technology
Gothenburg, Sweden 2026

Facilitating Engineering Process Alignment, Quality, and Productivity through Cross-Functional Standardization

Development of a Framework for Structuring Engineering Activities
in a Multi-Site Context at SKF

ALFRED HALLIN

LINUS FÄLDT

Supervisor: Peter Hammersberg, Industrial and Material Science

Examiner: Peter Hammersberg, Industrial and Material Science

Master’s Thesis 2026

Department of Industrial and Material Science

Chalmers University of Technology

SE-412 96 Gothenburg, Sweden

Telephone + 46 31 772 1000

Cover: CARB Toroidal Roller Bearing © SKF Group (2026)

Printed by Chalmers Reproservice

Gothenburg, Sweden 2026

Abstract

Ensuring consistency, traceability, and comparability in complex engineering activities remains a key challenge in global organizations, particularly when practices evolve locally across multiple sites. This thesis investigates these challenges within the context of bearing life testing, where variations in workflows, documentation, and data management limit the effective reuse and interpretation of test results over time. The study was conducted as a single case study at SKF, combining document analysis, expert interviews, and iterative concept development based on a Six Sigma-inspired methodology. The objective was to develop a structured framework that improves transparency, supports evaluation, and enables systematic reuse of life testing knowledge while maintaining flexibility for practical engineering.

The result of this work is a multi-layered framework consisting of three core elements: a modular structure describing the life testing process, a maturity perspective enabling evaluation of test completeness and quality, and a data management perspective supporting structured data management and long-term knowledge retention. Across three iterative concepts, the framework evolves from a descriptive reference model into an applied and decision-support-oriented system, incorporating deliverable-based gating, maturity evaluation, and data integration.

The findings demonstrate that the primary limitation in engineering testing is not technical capability, but the lack of a shared structure for describing, evaluating, and reusing knowledge from activities. By introducing a standardized yet flexible framework, the study shows how transparency, consistency, and comparability can be improved across sites. The proposed approach supports both operational decision-making and long-term organizational learning by making implicit practices explicit and enabling systematic continuous improvement by tracking performance over time.

Acknowledgements

This thesis has been a highly rewarding and enriching experience, not only through the knowledge gained, but also through the opportunity to step into an engineering area in which we initially had limited prior experience. The learning process has been both challenging and inspiring, especially through exposure to real-world complexity. We are particularly grateful for the opportunity to collaborate in a global context and to engage with professionals across different parts of the world, all of whom generously shared their time, knowledge, and perspectives throughout this journey.

We would like to express our sincere gratitude to SKF for providing us with the opportunity to carry out this thesis, and for the trust placed in us to explore a long-requested and important area. A special thank you goes to Elin Liden and Luca Lemma for their support, encouragement, and for enabling this work to take shape from the very beginning.

We would also like to express our gratitude to Manlio Becchetti and Claudio Savarese for their continuous support and willingness to share their expertise throughout the project. Their deep knowledge of the testing process and their guidance during the concept development and evaluation phases at site level have been invaluable. Furthermore, we would like to thank Albrecht Nestle, Stefan Weigand, and Hugo Boffy for their insightful advice and for broadening our understanding within Performance Prediction, and how our work can be further strengthened within this area.

We are especially thankful to Peter Hammersberg for his support and guidance as supervisor throughout the entire process. His trust in our approach and his continuous engagement have helped us navigate the project in a thoughtful and structured way. His ability to provide nuanced perspectives on both the purpose of our work and how to approach it has been crucial in shaping our thinking.

Table of Contents

List of Acronyms

AI	Artificial intelligence
DMAIC	Define, measure, analyze, improve, and control
LPMS	Lab process management system
PoC	Proof of concept
SOP	Standard operating procedure

1. Introduction

Engineering activities play a central role in ensuring the performance, reliability and quality of products. In global organizations, these activities are often distributed across multiple sites and involve complex interactions between technical processes, documentation practices and data management. Ensuring consistency in how such activities are structured, executed and evaluated is therefore essential to enable reliable results and effective collaboration across the organization.

This chapter introduces the research conducted as part of a master's thesis within the Industrial and Materials Science department at Chalmers University of Technology. The study is carried out as a single case study within an industrial context, focusing on how engineering activities can be structured and standardized to improve consistency, traceability, and comparability across an organization. Bearing life testing is used as a representative case to explore these challenges in practice.

1.1 Background

Engineering activities in global organizations are often distributed across multiple sites, teams, and technical environments. While these activities may pursue similar objectives, how work is structured, executed, and documented can evolve locally over time. Differences in tools, workflows, and practices emerge as a result of varying constraints, experience levels, and historical development at each site. Although such local adaptations can be beneficial in specific contexts, they can introduce unintended variation at an organizational level.

This variation presents several challenges. When activities are not structured or documented in a consistent manner, it becomes difficult to compare outcomes across sites, reuse existing knowledge, and ensure that results are interpreted in the same way. Over time, this can lead to reduced transparency, ineffective information flows, and an increased reliance on individual experience rather than shared organizational knowledge (Carlile, 2004). In such environments, critical knowledge may remain implicit and difficult to transfer, making onboarding of new personnel more complex and limiting the organization's ability to scale or coordinate activities globally.

To address these challenges, organizations often seek to introduce some form of standardization. However, standardization in complex engineering environments is not a straightforward task. While strict standardization can improve consistency, it may also reduce flexibility and limit the ability of engineers to adapt to local conditions or specific technical requirements. Conversely, leaving processes fully decentralized may preserve flexibility but reduce comparability and control. This creates a fundamental need to balance consistency and flexibility, where activities are aligned at an appropriate level without constraining necessary local variation (Goel et al., 2023).

A key aspect of such alignment is the establishment of a shared structure for how work is described and understood. Rather than prescribing identical execution, effective standardization often focuses on defining common reference points, such as process structures, expected outcomes, and shared terminology. This enables different parts of an organization to communicate about work using a common language, supporting collaboration and making it possible to identify differences in performance and practice in a structured way. In addition to structural alignment, there is also a need to improve how activities are evaluated and how information is managed over time. Without clear criteria for assessing completeness or quality, it becomes difficult to determine whether a given activity has been performed at a compliant quality level.

This thesis explores these challenges in the context of engineering testing activities, using bearing life testing as a representative example of a complex, multi-stage process. Life testing involves coordinated activities related to planning, execution, analysis, and reporting, often performed across different sites and by different stakeholders. As such, it provides a suitable context for studying how variation emerges and how it can be managed through structured approaches.

Rather than focusing on technical optimization of individual tests, the work aims to address how testing activities are structured, documented, and evaluated within an organizational context. The objective is to investigate how a standardized framework can support improved traceability, comparability, and knowledge reuse, while maintaining the flexibility required for practical engineering work. By using life testing as a case example, the study seeks to develop insights that are applicable beyond a single domain and relevant to standardization efforts in complex engineering environments more broadly.

1.2 Aim

The aim of this thesis is to explore how complex engineering activities are structured and executed within an organizational context, using bearing life testing as a representative case. The study seeks to describe the current state and identify gaps related to consistency, traceability, and comparability of practices. Based on these findings, the purpose is, to develop a structured framework that supports global alignment of activities while enabling local flexibility.

1.3 Research Questions

Based on the identified challenges related to consistency, traceability, and flexibility in engineering activities, the following research questions were defined:

I. How are engineering activities currently structured, executed, and documented at the selected reference site?

II. How can insights from the current-state analysis be used to structure and develop a standardized framework for complex engineering activities?

The first research question addresses the absence of an overview of how engineering activities are currently structured, executed, and documented, aiming to establish an understanding of current practices.

The second research question focuses on transforming these insights into a structured framework that supports alignment and comparability, while enabling continuous improvement through evaluation and feedback mechanisms.

1.4 Limitations

This study focuses on analysing and structuring engineering activities within the context of bearing life testing at a single reference laboratory. The scope is limited to understanding how activities are organized, executed, and documented, with the purpose of developing a structured framework that supports improved traceability, consistency, and comparability. The work does not include the development or modification of test rigs, sensors, or measurement hardware. Instead, the focus is on the structural and organizational aspects of engineering work, rather than technical optimization of testing methods or performance.

The study was conducted within a limited time frame, which constrained the depth of analysis and the extent of validation that could be achieved within the scope of the thesis. In addition, the project involved an initial learning phase with the life testing process and its technical characteristics, which shaped the scope and direction of the investigation.

Access to detailed internal data and documentation was limited due to confidentiality requirements, and the study did therefore not include any major systematically collected datasets for quantitative analysis. As a result, the development of the framework is primarily based on available documentation, reported results, and expert input, rather than extensive empirical validation. This approach is consistent with the aim of the thesis, which is to develop a structured standardization framework rather than to perform statistically driven evaluation of technical performance.

A full global implementation of the developed framework is outside the scope of this work. Instead, the study provides a foundation intended to support future standardization, continuous improvement, and potential rollout across additional sites.

2. Theoretical Framework

Standardization is frequently used to manage the complexity of large organizations. When standardized structures are absent or weak, practices tend to evolve unstructured, resulting in difficulties comparing performance. This chapter defines standardization as a coordinating mechanism and explains how it supports visibility, control, and collaboration in organizations. These theoretical principles form the basis for the analysis and conclusions presented later in the report.

2.1 Standardization

This section introduces standardization as a coordination mechanism used to structure how work is performed across organizational contexts. It clarifies what standardization entails, the role of external standards, and how different degrees of standardization can be applied in practice.

2.1.1 Definition and Scope

Standardization is described by the Corporate Finance Institute (2019) as the process of creating protocols to guide the development of goods or services. As such, standardization represents a structured approach to defining how activities are to be performed across different contexts. While applicable in many business areas, standardization has been widely adopted in production processes to improve efficiency, quality and consistency (Salunke et al., 2025, p. 3158). Standardized procedures are therefore typically documented and designed for repeatable use, providing a consistent reference for how work is performed and evaluated.

2.1.2 Regulatory Standards

In industrial settings, standardization plays an important role in ensuring that activities are performed in a predictable and controlled manner, particularly where external laws and regulations influence how work is carried out. Regulatory standards define formal requirements for how processes are documented, executed, and evaluated, providing a common reference for compliance across organizations and industries (ISO, 2015).

Quality management standards are an example of such regulatory standardization. They explicitly rely on documented and repeatable procedures to ensure consistent outcomes. The ISO 9001:2015 standard for quality management systems illustrates how standardization is operationalized through defined processes, measurement requirements, and regular audits, making quality management an established and widely recognized application of standardization (ISO, 2015). Beyond internal alignment, certification to such standards also establishes a shared reference, enabling organizations within and across industries to communicate and compare quality expectations using a common and external terminology.

2.1.3 Degrees of Standardization

The benefits of standardization can be realized through various means and at different organizational levels. Uniform approaches to documentation, workflows, communication, resource usage, policies, performance metrics and customer interaction can all contribute to more efficient and transparent operations (SixSigma.us, 2024b). At the same time, not all activities require the same level of restrictiveness. Some tasks demand detailed instructions while others rely more on operator judgement, which requires an understanding of how much standardization is required to provide the greatest benefit.

This variation is reflected in business process standardization literature through the concept of a *Master process*, which serves as a common baseline for how a process should be performed (Goel et al., 2023). Rather than enforcing identical execution in all situations, the master process allows for adjustments where needed with justification. Differences are expected to be explained by factors such as cultural differences, resource availability, or local constraints, highlighting that effective standardization is less about strict uniformity and more about finding an appropriate balance between consistency and flexibility.

Continuing on the concept of a master process, Goel et al. (2023) emphasize the role of modularization as the subdivision of documented processes into “meaningful and suggestive subprocesses,” allowing each part of the process to be described at an appropriate level of detail. This supports both transparency and manageability, particularly in complex processes where not all activities require the same degree of specification. Importantly, what constitutes meaningful modularization depends on the underlying process logic and organizational structure, meaning that different subprocesses may be viewed, structured, and detailed differently depending on the user perspective and the context it’s used within.

2.2 Visibility

Standardization not only structures work execution but also enables visibility into how processes perform over time. This section explains how standardized practices support performance overview, consistent measurement, and interpretation, forming the basis for continuous improvement. Visibility is presented as a key mechanism through which organizations can learn from execution rather than simply enforce predefined procedures.

2.2.1 Performance Overview

One of the main purposes of standardization is to make organizational performance visible over time. When work is performed in a consistent way, results from different periods, teams, or locations can be collected and compared. This makes it easier to identify trends and deviations that would otherwise remain hidden.

Beyond supporting compliance, standardization enables organizations to improve performance while using resources more efficiently (Salunke et al., 2025, p. 3159). In contrast, inconsistent ways of working often lead to uneven quality outcomes, which in turn result in unnecessary use of time, material, and effort. By reducing uncontrolled variation, standardized practices help create a clearer overview of how operations perform and where improvement efforts should be directed.

2.2.2 Measurement and Interpretation

Achieving a meaningful performance overview depends on the ability to measure results in a consistent and interpretable manner. Without standardized ways of performing and evaluating work, organizations struggle to determine whether indicators such as productivity, quality, or customer satisfaction are improving or declining over time (ISO, 2015). Measurements become difficult to compare, as results are influenced by subjective interpretations and differing evaluation criteria across individuals, departments, or periods.

Standardization addresses this challenge by establishing shared references for both execution and interpretation. When activities are performed according to common procedures and evaluated using defined criteria, performance differences can be identified more reliably and attributed to their underlying causes rather than to inconsistencies in measurement. This consistency is particularly critical for quality control, which relies on distinguishing meaningful variation from noise based on data that is both comparable and trustworthy.

2.2.3 Continuous Improvement

Standardized measurement and performance visibility provide the foundation for continuous improvement. When processes are executed and evaluated in a consistent manner, organizations can track performance trends over time and assess whether changes lead to measurable improvements. This relationship is frequently emphasized in continuous improvement literature, which highlights that standards enable measurement, and that measurement is necessary for systematic improvement (Kaizen Institute, n.d.). Documented baselines make it possible to compare conditions before and after adjustments, supporting fact-based evaluation.

In practice, continuous improvement is not driven by performance data alone. Interpreting measurement results and identifying suitable improvement actions requires an understanding of the underlying work system, as data indicators by themselves do not explain the causes of unwanted outcomes. Deming (1986) argues that managing solely on results creates “management without knowledge,” emphasizing that improvement depends on understanding how the system produces those results rather than reacting to the numbers themselves. From this perspective, input from those performing the work becomes essential, as experiential knowledge provides important context that may not be captured by measuring alone. Combining standardized performance measures with user insights supports organizational learning about how processes function in practice, enabling organizations to understand which aspects of a process benefit from increased standardization and which requires greater flexibility (SixSigma.us, 2024a).

While governance is necessary to ensure that standards are applied consistently and that performance measurements remain meaningful, business process standardization literature stresses that such governance should not eliminate context dependent adaptation in daily work (Goel et al., 2023). From this perspective, standards function as evolving reference points rather than fixed rules. Continuous improvement enables organizations to iteratively adjust standards based on measured performance and user feedback, working toward an appropriate level of standardization that supports both performance objectives and acceptance in practice.

2.3 Collaboration

Beyond enabling visibility and improvement, standardization also plays a central role in supporting collaboration within global organizations. When activities are standardized and documented in a consistent manner, they function as a shared language that allows different parts of the organization to communicate about work using common structures, terms, and expectations. This shared language reduces reliance on local interpretation and tacit knowledge, enabling discussions about processes and performance to be grounded in defined references.

By providing a common basis for describing how work is performed, standardization enables comparability across organizational units, sites, or teams. Differences in execution, outcomes, or performance can be identified and discussed using the same underlying process structure, supporting alignment without requiring identical local conditions (Goel et al., 2023). In this way, collaboration is enabled not through uniformity, but through the ability to understand and explain variation relative to a shared standard.

The use of standardized representations further supports organizational learning by making experiences from one context understandable and relevant in another. When improvements, deviations, or challenges are described using a common process language, knowledge can be transferred more effectively across boundaries (Carlile, 2004). So, rather than each site independently developing local solutions, standardization allows improvement initiatives to be communicated, evaluated, and adapted collaboratively. This enables learning to occur at an organizational level, supporting continuous improvement through shared understanding rather than isolated local optimization (Deming, 1986).

3. Methodology

This chapter describes the methodological approach used to develop a structured and standardized framework for complex engineering activities. The methodology combines iterative concept development, structured data collection, and continuous stakeholder involvement. This chapter explains how data was gathered, how analyses were performed and how the concepts were developed and evaluated.

3.1 Research Design

The research design combines a Six Sigma-inspired improvement structure with iterative concept-development cycles and a reference-site case study. This integrated design establishes both the analytical foundation and the empirical environment needed for developing a standardized life testing framework. The use of a reference site provides a consistent operational context for mapping current practices, validating assumptions, and supporting iterative refinement. Together, these elements enable an understanding of the current situation while supporting a progressive concept-development process.

3.1.1 Six Sigma

The project is based on the Six Sigma DMAIC (define, measure, analyze, improve, and control) structure, which serves as a structured and data-driven framework for breaking projects into well-defined phases and a smoother progression from problem definition to verifying that the implemented solution delivers the expected effect on the system (Lean Outside the Box, 2025). Due to the project's goal of standardization, Six Sigma techniques and tools have supported a data-driven mindset throughout the development of the concepts. The method is also suitable because the project extends beyond the time and resources available for the thesis, and the *Control* phase of DMAIC provides a framework for long-term monitoring by other stakeholders to ensure that results are achieved and that correct action is taken when necessary.

3.1.2 Reference Site Selection

A single reference site was selected for mapping current life testing practices and establishing a baseline for the development of the standardized framework. The chosen site was SKF's life testing facility in Airasca in Italy, which was prioritized due to its recognized expertise in roller-bearing testing, its established procedures for calibration and documentation, and its wide range of test types. These characteristics made the Airasca facility representative of mature life testing operations and suitable as a foundation for identifying standardization needs.

In addition to providing insight into existing workflows, the Airasca site offered access to experienced SKF engineers and specialists who contributed to the evaluation of preliminary concepts and helped clarify procedural questions throughout the project. Because the site was not located near the thesis writers, collaboration and information gathering were conducted through online meetings and interviews.

Although involving multiple SKF sites could have broadened the comparative scope, the decision to work with a single site enabled consistent data collection and allowed iterative improvements to be traced over time. At the time of the thesis, no global standard existed for life testing practices, making a single-site baseline a practical starting point before extending the framework to additional SKF locations in future work. Focusing on one reference site also reduced the need for widespread coordination and communication across many engineers at different sites, enabling more efficient interaction, faster feedback, and tighter alignment during the development phase.

3.1.3 Iterative Concept Development

The development of the standardized life testing framework followed an iterative concept-development process structured around three successive phases:

- **Phase One - Initial definition**, where early assumptions and structural ideas were formulated to establish the overall direction and scope of the framework.
- **Phase Two - Refinement**, where insights from data collection and stakeholder feedback were used to adjust and clarify the structure, content, and feasibility of the concept.
- **Phase Three - Finalization**, where the concept evolved into a coherent and practically implementable framework suitable for validation and future application.

Early concept versions were treated as provisional prototypes that could be evaluated, challenged, and improved in subsequent cycles. Iteration allowed each stage of development to incorporate new findings from internal document review interviews with the reference site, ensuring that the evolving framework remained aligned with real operational needs. Iteration also served as a mechanism for exploring the boundaries of the defined scope. By repeatedly cycling between prototyping, feedback and analytical insights, the process revealed gaps that were not apparent in early drafts and supported the exploration of alternative approaches to emerging challenges. Iterative approaches are recognized for enabling such structured exploration and for improving solution quality by integrating user feedback and insights across repeated refinement cycles (Kelley, 2022).

By allowing the concept to evolve through sequential improvements rather than a single up-front design, the research design ensured that the final framework was both iteratively evidence based and practically feasible. This approach contributes to a solution that’s robust, scalable, and responsive to user needs, reflecting the strengths of iterative development in complex engineering environments.

3.2 Data Collection Methods

A combination of qualitative and quantitative data collection methods was applied to obtain both a broad perspective of the subject and in-depth knowledge where needed. The quantitative methods helped create an understanding of areas where detailed information about current life testing practices and the requirements for standardization were required. This combination of methods is widely recognized for providing a more comprehensive understanding of complex research environments (Isom, n.d.).

Data collection also followed a top-down and bottom-up approach over the course of the project. Initial activities were primarily top-down, focusing on understanding the overall structure, workflow, and main components of the life testing system. As understanding increased, data collection became progressively more bottom-up, shifting towards detailed investigation of specific processes, practices, and limitations. This combined approach supported both a holistic system understanding and detailed insight into critical elements of the testing environment (Lykins, 2023).

3.2.1 Interviews

Semi-structured interviews were conducted with test engineers, local laboratory managers, performance prediction specialists, and site managers. This interview format was chosen due to its balance between structure and flexibility, combining predetermined questions with the possibility for spontaneous follow-up questions to clarify uncertainties or explore topics in greater depth. Such flexibility makes semi-structured interviews well suited for obtaining in-depth information while maintaining alignment with the study's objectives (Ruslin et al., 2022).

Interviews were conducted throughout the project and evolved in scope and level of detail as the understanding of the testing environment increased. Early interviews addressed the overall life testing process, while later interview rounds focused more on key areas of current routines, including documentation, calibration, uncertainty estimation, failure-analysis practices, which could reveal perceived bottlenecks and other disturbances within the system.

In addition to supporting early concept development, interviews were also used during concept-review sessions to improve both the quality and relevance of stakeholder input. Since interview responses directly influenced the following concept iterations, ensuring consistency and clarity in these discussions was essential. These sessions were therefore based on the structure of: what was the previous version, what has been done since last time, and what we expect to do for the following iteration.

During the early stages of the project, interview responses were documented manually by the interviewers and summarized during the interview sessions. In later stages, interviews were transcribed and summarized using Artificial Intelligence (AI) by Microsoft Teams' built-in tool (Microsoft, n.d.). The resulting transcriptions were validated by comparison with interviewer notes to identify inconsistencies or potential misinterpretations by the AI.

3.2.2 Document Review & Analysis

Document reviews and analysis were conducted to establish a structured understanding of life testing practices and to identify inconsistencies across documentation, uncertainty estimation, and test reporting. The reviewed material covered the progression of a life testing process from foundational technical knowledge and applicable international standards to finalized test results.

At the time of the project, no formally documented end-to-end procedure for life testing existed at the case company. As a result, the document review focused on mapping and interpreting the available, but partially fragmented, documentation that described individual aspects of the process. This documentation provided an initial structural framework, while much of the detailed process knowledge was captured through interviews with practitioners involved in life testing activities.

The case company provided access to internal documentation related mainly to technical aspects of bearing life and life testing. Due to confidentiality and information security requirements, access was limited to test reports, and data collection and document analysis were therefore conducted within these defined boundaries. The review began with documentation addressing bearing life principles and failure-mode classification, establishing a baseline technical understanding. This was followed by an analysis of relevant ISO standards related to calibration, laboratory operations, and quality management.

Subsequent reviews focused on operational material, including internal documentation related to test request handling and report structures. These sources supported the mapping of how life testing activities are initiated, planned, and executed in practice, while historical test reports were examined to understand how results, uncertainties, and conclusions are documented and communicated across stakeholders.

3.2.3 Use of AI in Report writing

AI-based tools have been used as support during the writing process of this thesis. The use of AI was limited to language refinement, structural suggestions, and improvement of clarity in written text. All concepts, evaluations, and conclusions presented in this report are the result of the authors' own work, based on collected data and expert input. The use of AI is therefore considered a supporting tool for communication and presentation, rather than a contributor to the research outcomes.

3.3 Analytical Approach

The analytical approach of this study focused on systematically interpreting collected empirical information by comparing observed life-testing practices with established theoretical frameworks. Literature and standards were used to define how testing activities, documentation structures, and uncertainty estimation are intended to function. This analytical baseline was then applied to the current practices at the reference site, allowing differences, gaps, and improvement opportunities to be identified in a methodical manner.

3.3.1 Visual Information Mapping

To manage and interpret the large volume of qualitative information collected during the data-collection phase, a visual information-mapping approach was applied. Qualitative inputs from interviews, document reviews, and observations were decomposed and structured into visual diagrams created in the collaborative modelling environment *Miro* (Miro, n.d.).

Three diagrams were primarily used to organize workflows, documentation handovers, and structural relationships between stakeholders and processes. Breaking complex systems into structured visual elements supported comparison across information sources and reduced interpretive ambiguity. The use of visual decomposition aligns with established Lean and Six Sigma practices, where diagram-based structuring, such as tree diagrams, is used to analyze complex processes by separating them into manageable and comparable components (Lean Outside the Box, 2026).

3.3.2 Literature and Standards Review

International standards (e.g., ISO/IEC 17025, ISO 15243) and internal SKF guidelines were reviewed to establish the analytical reference framework for the study. These sources clarified expectations related to calibration procedures, uncertainty estimation, documentation practices, laboratory operation, and standardized failure-mode classification.

In addition to supporting the analytical evaluation of current practices, these standards were used as reference criteria during concept development. Since compliance with such standards is essential for laboratory certification and alignment with accepted industry practices, they defined boundary conditions for feasible and implementable solutions.

3.3.3 Current-State Analysis

Using the reference site described in Section 3.1.2, current life testing practices were analyzed by integrating interview findings and document-review insights with the visual representations created during the mapping stage. The mapped workflows and documentation structures were compared with the analytical baseline defined by standards and guidelines. This comparison enabled systematic identification of deviations between intended and actual practices, supporting the development and refinement of concepts.

3.4 Validation of Findings

Validation activities were conducted to ensure that each iteration of the developed concept addressed user needs and aligned with both theoretical expectations and operational constraints. Validation was not treated as a final, isolated step, but as an integrated activity throughout the project, improving each concept development through repeated feedback.

3.4.1 Feedback Sessions

Validation through feedback sessions was carried out continuously with engineers and technical specialists involved in life testing activities. These sessions focused on reviewing intermediate concept versions, clarifying and refining definitions of minimum requirements, and assessing the feasibility and practicality of proposed concepts within the laboratory environment. Insights from feedback sessions were incorporated after each concept development cycle, enabling progressive refinement of both structure and content. This iterative validation approach ensured that concept development remained closely aligned with stakeholder expectations and practical constraints rather than entirely on theoretical assumptions.

3.4.2 Proof of Concept

The final validation step was conducted through a proof of concept (PoC), in which the developed framework was illustrated using example life testing data. The purpose was not to derive the concept from empirical data, but to demonstrate how it can be applied in a realistic context and to highlight its potential benefits.

PoC is a recognized approach for assessing feasibility and practical by showing how a conceptual solution functions when used in practice (Grant, Bugge, & Wells, 2020). This enabled evaluation of the framework's clarity, internal consistency, and usefulness as an analytical and decision-support tool. While not constituting full implementation, the PoC provides a realistic illustration of how the framework can be used in practice and supports its relevance for future application.

4. Pre-study of the Life Testing Process

This prestudy provides a consolidated overview of the bearing life testing process as it is currently performed across sites. As part of this work, the existing life testing practices have been mapped to create a structured understanding of the overall process, as this hasn’t been thoroughly done at the case company before. This mapping was not only conducted to describe current practices, but also to create a foundation for analysis and improvement. As discussed in Chapter 2, achieving standardization and performance improvement requires an understanding of how processes operate in practice. In line with Deming’s (1986) perspective that improvement depends on understanding how a system produces its results, rather than only evaluating outcomes, the mapping enables identification of where standardization can support the life testing activity.

The purpose is not to detail individual methods or laboratory procedures, but to summarize the overall workflow, main responsibilities, and key decision points that structure bearing life testing from initiation to final documentation. Detailed descriptions of practices, tools, and methodological considerations are presented separately in Appendix A. To support readability and reflect how life testing is organized in practice, the process is described through three overarching phases: testing preparation, test execution, and test analysis & documentation. Together, these phases describe how bearing life tests are planned, conducted under controlled laboratory conditions, and transformed into reusable engineering knowledge.

4.1 Testing Preparation

The life testing process begins with a preparation phase in which the motivation, scope, and feasibility of a potential test are established. During this phase, inputs originating from customer needs or internal development activities are analyzed through structured dialogue and technical evaluation to define a clear test purpose that reflects both performance expectations and laboratory capabilities, see Figure 1. Responsibility for the preparatory work is assigned to a designated test engineer, and feasibility is assessed with respect to technical requirements, available resources, and organizational priorities.

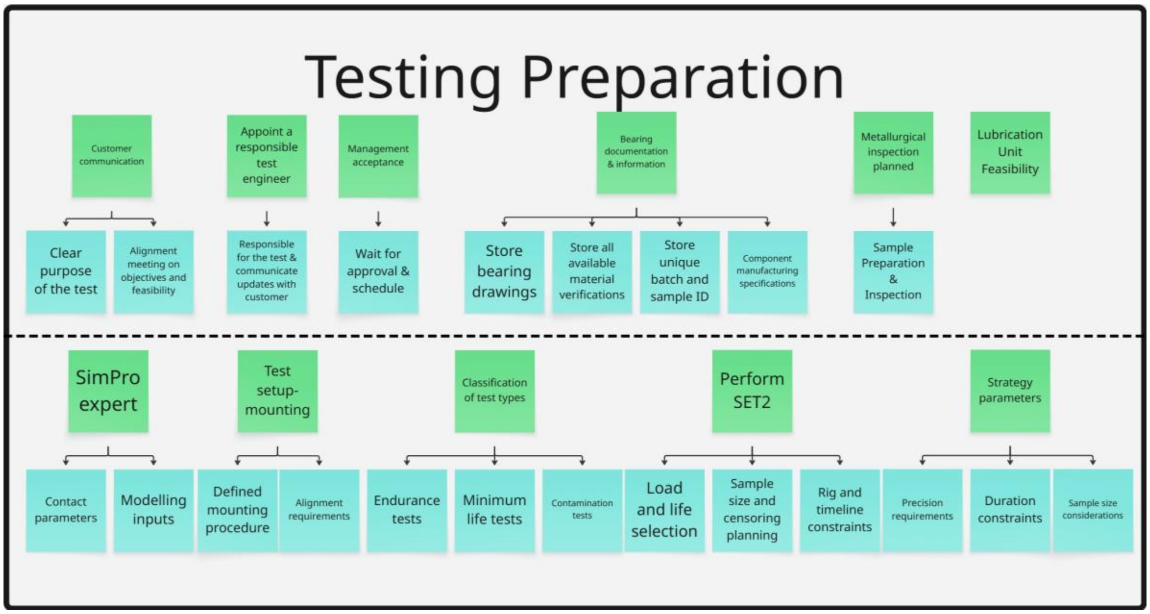


Figure 1: Overview of the testing-preparation phase.

As part of the preparation activities, the bearings used for testing are identified and documented to ensure traceability and representativeness. Bearings are selected to reflect standard production conditions, and each unit is defined through batch information, drawings, and available material documentation. Where relevant to the test objective, metallurgical inspections and sample preparation activities are planned in advance to support later interpretation of test results and failure mechanisms.

The preparation phase further includes the definition and validation of the test environment and strategy. Activities within test preparation are typically developed in parallel and iteratively, as decisions in one area directly influence and constrain the others. Bearing simulation modelling is performed to evaluate internal contact conditions, stress levels, lubrication specifications, and overall feasibility of the intended operating parameters. Based on these assessments, applied loads, boundary conditions, lubrication specifications, and operating limits are proposed to achieve the desired fatigue environment while maintaining a stable and realistic test setup. Test categories, such as endurance testing, minimum life testing, or contamination testing, are outlined using statistical planning tools to balance sample size, expected test duration, required precision, and practical constraints related to rig availability and timelines.

The preparation phase concludes with the formalization of the defined scope, conditions, and feasibility assessments into a test request that marks the transition from the testing preparation to the execution phase, where detailed planning and resource commitment take place.

4.2 Test Execution

Once the test is approved and prepared, the life testing process proceeds to the execution phase. This phase establishes the experimental framework under which the life test is performed and focuses on ensuring stable, controlled, and reproducible operating conditions throughout the test duration.

Test execution involves disciplined laboratory practices aimed at minimizing external disturbances and measurement uncertainty, see Figure 2. Measurement systems are calibrated to ensure accurate monitoring of critical parameters such as load, speed, temperature, vibration, torque, and lubrication flow. As highlighted in expert discussions, calibration is a critical prerequisite for reliable test results:

> *“Calibration is the first activity that can ensure that the results are reliable.”*
> - Global Testing Quality Expert, see Appendix B.

While calibration is typically performed through accredited external partners, the availability of internal calibration capabilities was also identified as a valuable complement, supporting verification, increased flexibility, and improved understanding of measurement uncertainty within the testing environment. Cleanliness requirements are applied to bearings, tooling, and lubricants to reduce the risk of contamination-induced damage and to ensure that observed failures are due to rolling-contact fatigue rather than incorrect test preparation and execution.

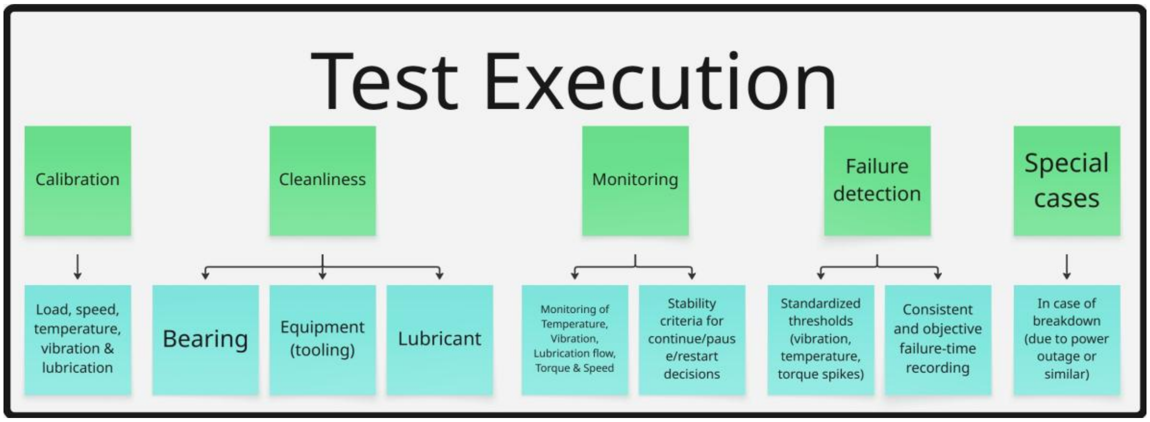


Figure 2: Overview of the test-execution phase.

Throughout the test, operating parameters are continuously monitored. Predefined stability criteria support objective decisions on whether a test can continue as planned or must be paused, adjusted, or suspended due to deviations. Failure-detection criteria are defined in advance in the preparation phase and are consistently applied to ensure repeatable and unbiased identification of bearing end-of-life. These criteria are based on predefined thresholds set at the start of each test, covering aspects such as abnormal vibration, temperature variations, and torque responses. If, for example, the vibration level exceeds its threshold, the test is automatically terminated.

Procedures are also established to handle special situations such as power outages or other unexpected interruptions. These procedures provide guidance on continuation, data validity, and potential rerunning of tests, ensuring consistency in decision-making and protection of data integrity.

4.3 Test Analysis and Documentation

After completion of the experimental phase, the life testing process proceeds to test analysis and documentation. The purpose of this phase is to evaluate the outcome of the test, understand the observed bearing behaviour in comparison to simulations results, and ensure that results are documented in a clear, traceable, and reusable manner.

Test analysis includes examination of bearing condition and damage using established microscopy and documentation practices, see Figure 3. Observations such as surface damage, crack initiation points, wear patterns, discoloration, or thermal effects are recorded and interpreted in relation to the applied test conditions. This supports a consistent understanding of how and why bearings reached the end of their operational life under the designed test environment.

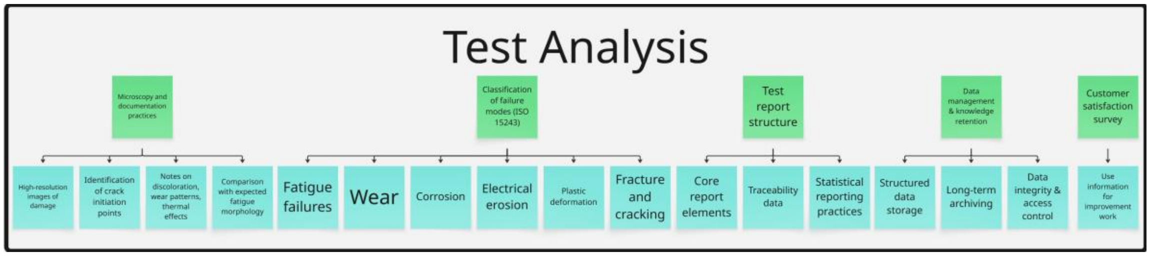


Figure 3: Overview of the test-analysis and documentation phase.

Observed damage is described and categorized using ISO15243 failure-mode descriptions, enabling failures to be reported in a consistent and comparable manner across different tests (ISO, 2017). The objective is to reduce reliance on subjective judgements and to ensure that test outcomes can be interpreted using a common technical language. Detailed procedures for damage evaluation and failure description are provided in Appendix A.

In parallel with damage evaluation, reporting activities ensure that test results are formally documented. Structured test reports summarize the test purpose, sample information, applied test conditions, observed results, and statistical evaluations. Supporting material such as operational logs, modelling inputs, calibration records, and analysis files are stored together with the report to maintain traceability and support future reference.

In addition to current reporting practices, SKF is developing a centralized data platform. The platform is intended to support long-term storage, structured organization, and improved accessibility of bearing life testing data across different testing activities and sites. At the time of this work, the platform is under development and has not yet been fully implemented as an operational standard. Its ongoing development nevertheless highlights the increasing importance of consistent documentation, traceability, and standardized data structures within the life testing process. The documentation practices described in this chapter are therefore aligned with the type of structured data handling that future tools such as the centralized data platform are intended to support.

Currently, life testing data is managed across multiple systems, where test requests are handled and tracked within the internal *lab process management system* (LPMS), while finalized reports and supporting documentation are stored in the internal documentation repositories. This separation reflects the existing workflow but also highlights the need for a more integrated solution. The data platform is therefore intended to consolidate these data flows into a unified structure, enabling improved traceability between test requests, execution data, and final reporting

4.4 From Process Overview to Concepts

Chapter 4 has outlined the bearing life testing process as applied in practice, covering preparation, execution, and analysis. As discussed in Sections 2.1–2.2, the absence of standardized structures often leads to variations in how practices are performed, making it difficult to compare performance, ensure consistent quality, and create organizational visibility. Standardization is therefore not only a tool for alignment, but also a prerequisite for enabling consistent measurement, interpretation, and continuous improvement. This creates a need to translate the observed process into a more structured and shared approach.

5. Results

This chapter presents the results of the study based on expert interviews, analysis of current life testing practices, and iterative concept development. The results are structured around three progressively refined concepts that together address the standardization of structural, operational, digital and organizational aspects of bearing life testing. Building on the findings presented in Chapter 4, the life testing process has been translated into a more structured and shared approach. The objective is to establish a common reference, or a "master process", as described by Goel et al. (2023), that supports consistency and transparency while still allowing for necessary local adaptations.

5.1 Key Findings from Expert Interviews

The expert interviews consistently indicate that SKF's life testing capability is technically strong at individual sites, but that these capabilities are not fully leveraged across the organization. Interviewees describe a situation where tests are generally planned and executed with high engineering competence, yet where differences between sites, documentation practices, and information flows limit comparability and retention of knowledge. A recurring theme is the absence of a common, shared structure for discussing life testing activities:

> *"The problem is not that people don't know how to run tests. The problem is that if you move from one site to another, there is no common reference for what has actually been done and why."*
>
> - Global Testing Technical Expert, see Appendix B.

Another recurring observation concerns the reliance on individual experience rather than structured documentation. While experienced engineers often know what to look for, this knowledge is not always explicitly captured:

> *"A lot of things are done almost automatically by experienced people, but they are not written down anywhere. If that person is not available, the knowledge is simply not visible."*
>
> - Global Testing Technical Expert, see Appendix B.

Together, these findings highlight that the primary challenge is not technical execution of tests, but rather harmonization, traceability and long-term retention of life testing knowledge. These insights form the empirical basis for the three concepts presented below.

5.2 Concept 1

Concept 1 represents the first major result of this study and is directly grounded in the life testing process described in Chapter 4. While Chapter 4 provided a descriptive overview of how bearing life tests are currently planned, executed, and documented in practice, Concept 1 translates this understanding into a structured and explicit reference model.

The purpose of Concept 1 was therefore not to introduce new procedures or define requirements, but to synthesize existing life testing practices into a coherent structure that makes roles, activities, and information flows visible and comparable across different testing locations.

From the outset, Concept 1 was developed around four perspectives: Modular Structure, Maturity rating, Digital & Data, and Process. Together, these perspectives were used to organize the insights gained from Chapter 4, identify common patterns across sites, and explore how consistency and comparability could be improved without constraining local engineering practices.

5.2.1 Modular Structure

The modular structure developed in Concept 1 is the result of synthesizing the life testing process described in Chapter 4 into a set of distinct yet connected building blocks. The structure applies the process-modularization principle discussed in Section 2.1.3, based on the approach described by Goel et al. (2023). While Chapter 4 provided a process-level description of how life testing activities are planned, executed, analyzed, and documented in practice, Concept 1 translates this understanding into an explicit and structured reference model.

Based on how life testing activities are organized in practice, seven modules were identified to cover the full life testing lifecycle: *Test Motivation, Test Sample, Test Conditions, Test Strategies, Experimental Requirements, Failure Analysis,* and *Reporting & Data Management*, see Figure 4. These modules reflect how life tests are discussed, planned, and performed within the reference site, rather than representing a newly imposed or theoretical structure. The primary intent was therefore not to redesign the process, but to make existing practices explicit and comparable across different testing locations.

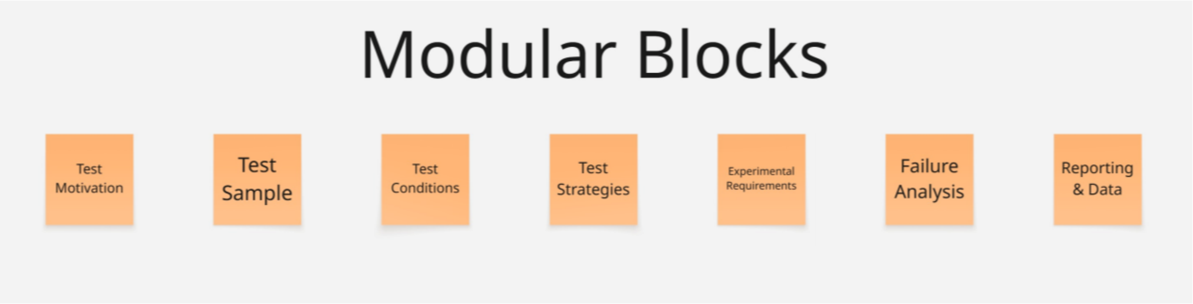


Figure 4: Representation of the seven modular blocks in bearing life testing.

While Figure 4 presents the modular structure as a set of distinct building blocks, it does not capture how these modules interact over time. To illustrate the dynamic flow of activities, Figure 5 presents a timeline representation of the life testing process. In this representation, the modules are arranged along a horizontal timeline showing the overall progression from initial motivation to final reporting and data management. However, the process is not strictly sequential. In particular, the modules related to T*est samples*, T*est conditions*, and T*est strategies* are often developed in parallel and iteratively. These activities are closely interconnected, and decisions in one area frequently influence and require adjustments in the others. The timeline representation therefore complements the modular view by combining a structured overview with the flexibility needed to reflect how life testing activities are performed in practice.

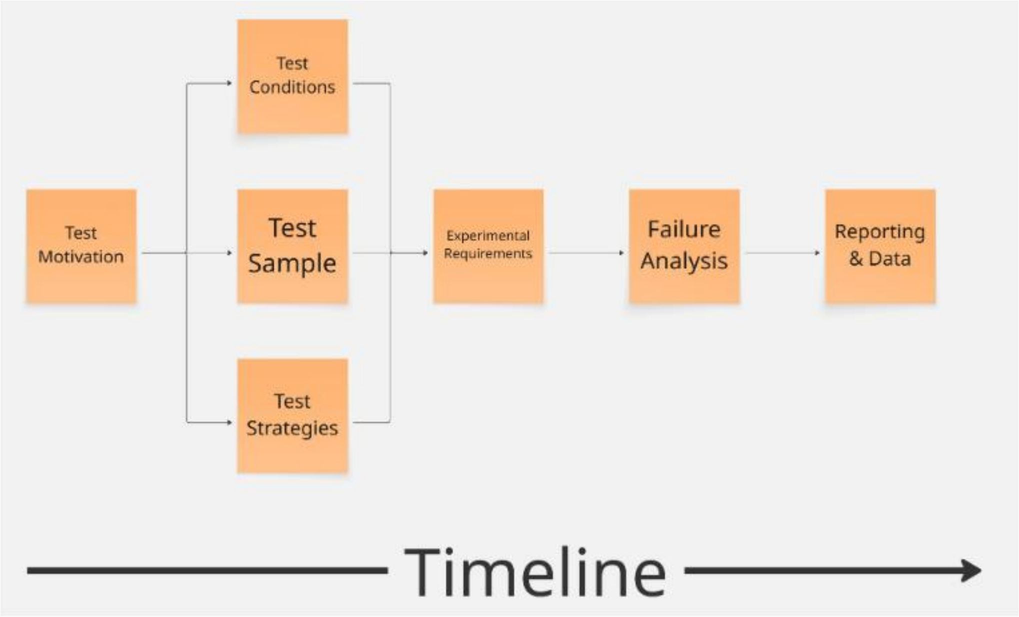


Figure 5: Bearing test flow timeline.

The decision to structure the life testing process into seven modules was driven by several considerations. First, the modules provide sufficient coverage of the end-to-end life testing workflow without dividing the process into overly detailed or task-specific steps. Second, each module corresponds to a meaningful stage at which key decisions are taken, activities are performed, and information is generated or handed over. This balance between completeness and clarity enables focused discussion of responsibilities, deliverables, and interfaces between stages, while preserving the flexibility required for local adaptations.

An overview of the seven modular blocks and their initial sub-blocks is presented in Figures 4, while the complete workflow is provided in Appendix C. Together, these figures provide a visual mapping between practical laboratory activities and the modular structure, showing how the seven modules span the preparation, execution, and analysis phases of life testing without introducing additional procedural requirements.

The preparation phase modules, corresponding to the activities described in Chapter 4.1, are detailed in four sub-blocks, as per Figure 6.

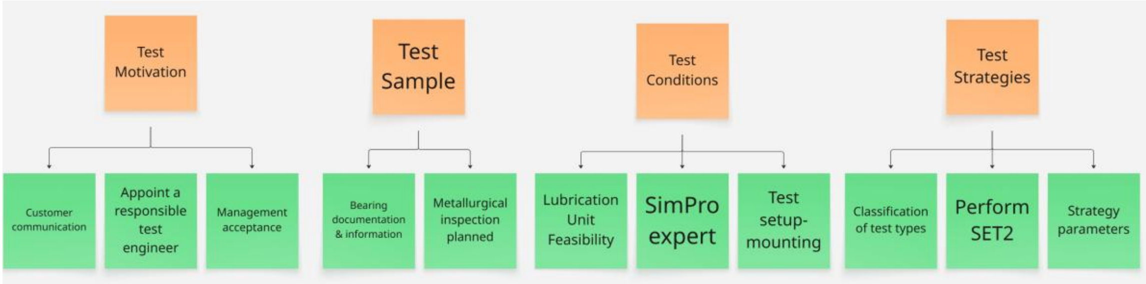


Figure 6: Modular representation of life testing preparation activities.

The execution phase module, corresponding to the activities described in Chapter 4.2, as per Figure 7.

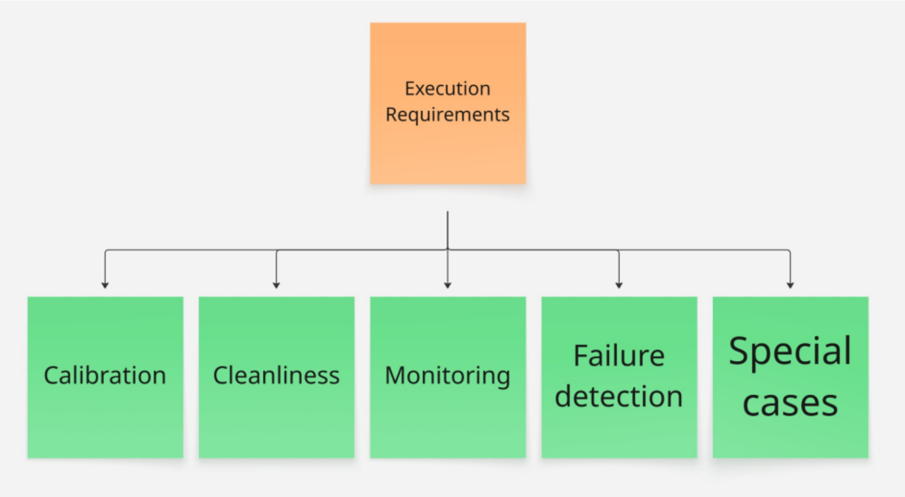


Figure 7: Modular representation of life testing execution activities.

The test analysis and documentation phase module, corresponding to the activities described in Chapter 4.3, are detailed in two sub-blocks, as per Figure 8.

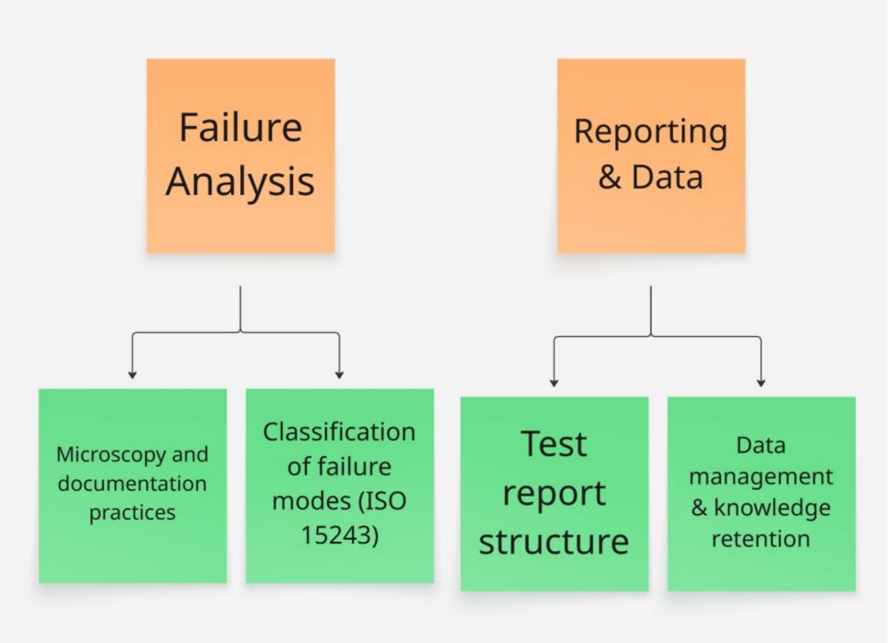


Figure 8: Modular representation of life testing analysis and documentation activities.

Each module is described through a corresponding set of *Deliverables* that clarify what types of activities, decisions, and documentation typically occur at that stage of a life test. Importantly, these deliverables were formulated to support transparency, to make implicit practices explicit, and to establish a shared vocabulary for discussing life testing activities within the organization. A complete overview of the deliverables associated with each module in Concept 1 is provided in Appendix C.

By decomposing life testing work into modular blocks with associated deliverables, Concept 1 establishes a clear structural foundation for describing and comparing life testing activities across different contexts. This modular structure provides the common reference against which additional perspectives, such as maturity rating, digital and data handling, and process formalization, are applied in the subsequent concepts to Concept 1.

5.2.2 Maturity Perspective

The Maturity element in Concept 1 evaluates how well each site fulfills the defined requirements within each modular block. The purpose is not to create competition between sites, but rather to support a structured and transparent approach to continuous development. This includes identifying improvement areas, guiding investment decisions, strengthening site capabilities, standardizing minimum requirements, and enabling long-term tracking of progress.

To achieve this, the maturity assessment provides a clear overview of the current state of each site and what is required to reach full readiness for different types of life testing. Concept 1 adopts a flexible approach by introducing different scoring methodologies to be considered and refined over time. Several scoring system options have been identified to support this development:

- **Three-level scale** (e.g., underperforming, standard, best practice)
  Provides a straightforward and easily applicable approach for introducing maturity assessment. Its simplicity facilitates alignment across sites, although its limited number of levels restricts the ability to capture incremental improvements.
- **Five-point numerical scale**
  Enables a more detailed differentiation of performance and supports tracking of progression over time. This approach requires clearer definitions and alignment in comparison to the three-level scale to ensure consistent interpretation and application.
- **Weighted multi-criteria scoring system**
  Represents a more advanced approach in which individual criteria are assigned different levels of importance based on expert input. This allows for a more accurate reflection of key performance drivers, but also introduces increased complexity in design, implementation, and governance.

To balance simplicity and long-term robustness, a phased rollout strategy is recommended. Initially, a basic three-level scale can be introduced to establish alignment and ease of use. As maturity and consistency improve, the system can evolve into a five-point scale to better differentiate performance. In a later stage, once sufficient data and organizational alignment are achieved, weighted scoring can be implemented to provide a more comprehensive and tailored evaluation model, as illustrated in Figure 9.

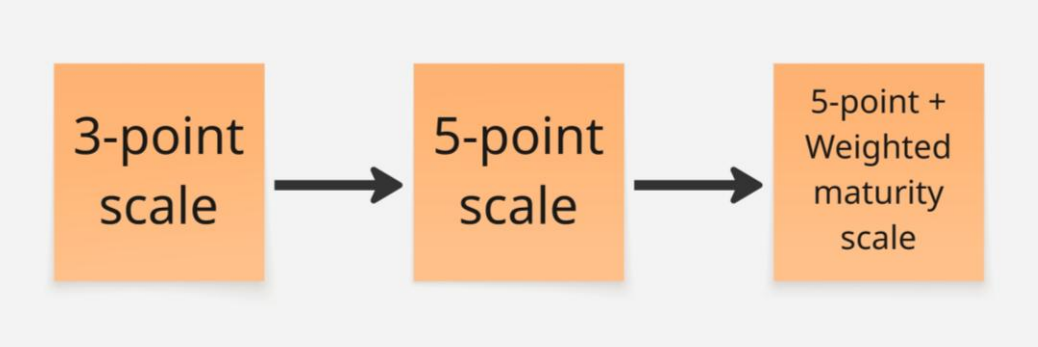


Figure 9: Concept 1: Rollout strategy for maturity rating.

In addition to supporting technical evaluation, this staged maturity approach can also be used to support structured reporting and follow-up at site level. By aggregating maturity results across modular blocks, it becomes possible to provide a clear and transparent overview of performance in periodic reports, such as quarterly or annual reviews. This enables visualization of which areas perform well and where improvement efforts should be prioritized, supporting both operational alignment and long-term development. This stepwise approach ensures that the maturity model remains practical to implement while still allowing for continuous refinement over time.

5.2.3 Digital & Data

Concept 1 also addressed data management as a fundamental part of the standardization, focusing on how information is generated, managed, and reused across the life testing process. The analysis considered what types of information are created throughout the process and how this contributes to traceability, transparency, and long-term knowledge retention. This includes modelling data, test configurations, operational logs, raw life data, failure documentation such as microscopy analysis, calibration certificates, and final reporting outputs.

Rather than specifying digital systems or tools, Concept 1 focuses on clarifying what information exists and how it flows between modular blocks. This ensures that data is consistently collected, traceable, and structured in a way that supports reuse. The intention is to enable a unified approach to data handling, where information can be stored and accessed through a common reference, such as the centralized data platform, without constraining future system choices. From a process perspective, Digital & Data enables a closed learning loop across the life testing workflow. Similarly, failure analysis is strengthened from comparisons with legacy data, and that validated results are continuously fed back into the shared data environment, enabling continuous improvement.

By connecting all stages of the process through consistent data structures and flows, Digital & Data supports the development of a sustainable knowledge base that strengthens over time. This not only improves traceability and transparency, but also enables more informed decision-making and continuous improvement across sites.

5.2.4 Process

The process SOP (Standard operating procedure) perspective was included in Concept 1 to explore how standardized operating procedures and process maps can support consistent execution across sites. The analysis focused on identifying where procedures already exist, where variation occurs, and how increased process clarity could improve alignment without prescribing detailed local execution methods.

SOPs and process maps were therefore considered key enablers of standardization. They provide a clear definition of workflows within each modular block, reducing variation between operators and testing locations while improving documentation, traceability, and audit readiness. In addition, structured procedures support onboarding and facilitate knowledge transfer, creating a more stable foundation for future improvements, including automation and digitalization.

Process maps further complement SOPs by visualizing how the modular blocks interact throughout the life testing process. This enables a clearer understanding of dependencies between activities and helps identify bottlenecks, inefficiencies, and opportunities for improvement across the testing chain. Rather than enforcing strict uniformity, the SOP perspective in Concept 1 promotes a balanced approach to standardization, where common process structures act as a shared reference while still allowing necessary adaptations to local conditions.

5.2.5 Summary of Concept 1

Rather than directly reflecting how work is currently performed, Concept 1 represents a synthesis of observed practices into a more explicit and structured framework. This includes the introduction of a modular structure, maturity evaluation, digital & data perspective, and process SOP perspective, which together provide new ways of describing, comparing, and analyzing life testing activities across sites. The intent of Concept 1 is therefore not to define final requirements or enforce standardized execution, but to establish an initial conceptual foundation that makes variations visible and enables more structured discussions around performance, data handling, and process alignment.

In line with the theoretical framework, this can be understood as a first step toward increased standardization by creating a shared reference for how life testing activities can be described and evaluated. The concept balances structure and flexibility, allowing existing practices to be interpreted within a common framework without constraining local engineering decisions. Concept 1 serves as a foundation for further development. By introducing structure and terminology, it enables subsequent concepts to build on a clearer and more aligned understanding of the life testing process, supporting future work related to evaluation models, governance, and system integration.

5.3 Concept 2

Concept 2 builds upon the foundation established in Concept 1 by extending the modular structure into a more operational and governing framework. Through the introduction of gated Deliverables, Maturity rating evaluation, and strengthened data governance, the concept defines clearer expectations for how life testing activities are executed, evaluated, and reused.

5.3.1 Modular Structure

The modular structure applied in Concept 2 builds directly upon the framework established in Concept 1 and is therefore not redefined in terms of structure or scope. Instead, the same seven module blocks are used as the underlying foundation for further development.

In Concept 2, the modular structure is extended from a descriptive reference model to a more operational framework by introducing a more condensed interpretation of Concept 1 deliverables, see Appendix D. While Concept 1 perceives deliverables as a set of checkboxes, Concept 2 treats them as formal progression criteria. This means that each module is associated with a set of deliverables that must be fulfilled before the test process is allowed to proceed, effectively transforming the modular structure into a gated system.

Through this shift, the modular design moves to actively governing the execution of the testing process. The structure thereby ensures that critical aspects of test preparation, execution, and analysis are consistently addressed, while maintaining alignment with the established process logic introduced in Concept 1. At the same time, the modular approach retains its flexibility, allowing individual modules and their associated deliverables to evolve without altering the overall structure.

In addition to this reinterpretation, Concept 2 introduces targeted extensions within the existing modules to strengthen knowledge reuse and evaluation transparency. Within the *Test Motivation* module, a mandatory step for "Review of previous tests", highlighted by pink boxes, are introduced, see Figure 10. This ensures that historical data, prior configurations, and earlier findings are systematically considered during test preparation. By formalizing this activity, the framework supports more consistent use of existing knowledge and reduces the risk of overlooking relevant prior work.

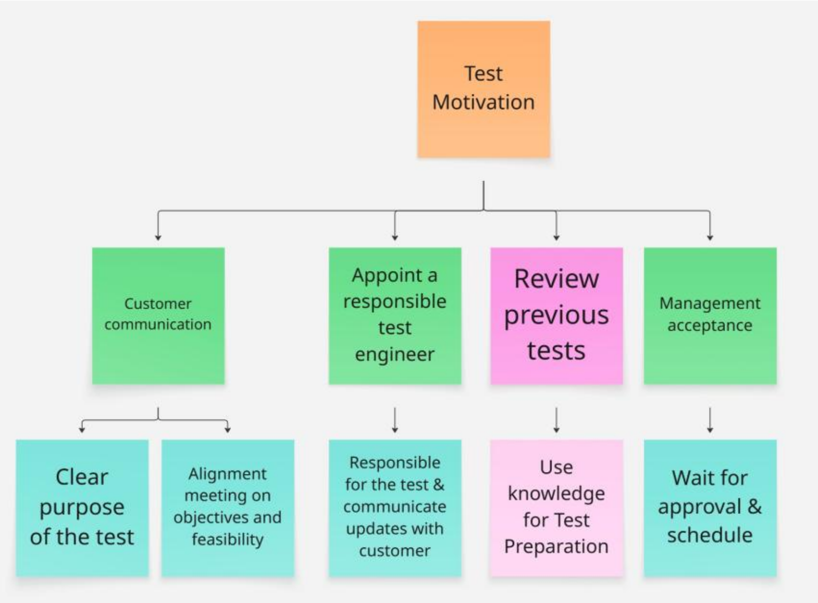


Figure 10: Additional activity for "Review previous tests".

A complementary addition is made within the Reporting & Data Management module through the inclusion of a Maturity Assessment, also highlighted in pink boxes, see Figure 11. This enables completed tests to be evaluated based on their fulfillment of the defined deliverables and assigned an overall maturity level. The assessment is documented together with the test results, allowing engineers to quickly assess completeness, traceability, and overall quality when reusing test data.

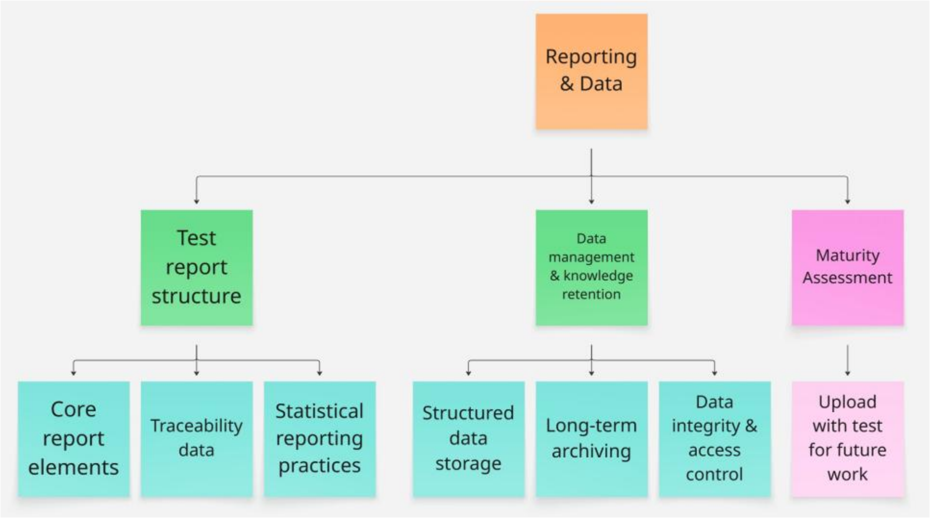


Figure 11: Additional activity for Maturity Assessment.

Together, these additions establish a feedback loop across the modular structure, where the maturity assessment of completed tests directly informs the review of previous tests in upcoming preparation phases. Through this integration, the modular structure evolves from a linear representation into a circular framework that supports continuous learning and more consistent reuse of validated test knowledge. At the same time, the modular approach preserves its inherent flexibility. Individual modules and their associated criteria can evolve as practices develop without requiring changes to the overall structure or compromising the continuity of the framework.

5.3.2 Maturity Perspective

The maturity framework presented in Concept 2 represented a continued development of the modular structure and evolves into a central focus of the study. While Concept 1 introduced maturity as a conceptual perspective, Concept 2 advanced it into a more applied approach aimed at supporting transparency, quality, and continuous improvement within life testing.

The foundation of the maturity rating framework is closely linked to the modular structure. By dividing the testing workflow into distinct blocks with clearly defined deliverables as seen in Appendix D, the framework enables a structured evaluation of how well each stage of the process is executed. Each block, governed by its deliverables, provides a natural basis for assessing completeness, traceability, and overall quality. This need for consistency and comparability was also highlighted by experts:

> *"For traceability reasons and for future comparability purposes, the simulations are run using standardized methodology, so that all labs produce the same output."*
>
> - Global Testing Quality Expert, see Appendix B.

The purpose of the maturity framework was therefore to evaluate how well life tests and their associated documentation fulfilled the defined deliverables across the modular structure. Rather than functioning as a ranking system, maturity was intended to create a structured understanding of strengths, gaps, and development potential at both test level and site level. This intention was emphasized in the concept evaluation interviews:

"Maturity should not be a ranking. It should help you understand where you are strong and where you need support."

- Global Testing Quality Expert, see Appendix B.

At this stage, however, the framework was not finalized. Instead, it was iteratively developed based on insights from expert interviews, exploratory testing, and ongoing work related to scoring structures and weighting of deliverables. Multiple approaches were explored, including both simpler maturity levels and more advanced models incorporating weighted criteria depending on their importance.

A key aspect of the maturity framework was its practical application in the use of the upcoming centralized data platform when reviewing previous tests. In some cases, engineers relied on their own historical test data when planning follow-up tests, particularly those intended to confirm or revise earlier findings. The importance of reusing historical knowledge was also emphasized in the interviews:

"When a new flow starts, you start from what was built in the past."

- Global Testing Technical Expert, see Appendix B.

In such situations, engineers needed a simple way to determine whether a test was complete enough to be trusted and qualitative enough to be used as a reference. The maturity framework was therefore intended to differentiate between tests with missing mandatory requirements, those that met a minimum acceptable standard, and tests that represented best-practice execution with high completeness and traceability. Without such differentiation, there was a risk that incomplete or poorly documented tests are used as references, which could lead to incorrect conclusions or unnecessary analysis effort. By making maturity visible across the modular structure, the framework supported more efficient and reliable reuse of existing test knowledge.

Another central part of the work concerned the development of the scoring structure. Several alternative approaches were explored, including simple maturity levels as well as more advanced methods where deliverables were weighted based on their relative importance. This reflected the understanding that not all deliverables contributed equally to test quality, traceability, or decision-making, and therefore needed to be prioritized accordingly. Particular focus was placed on identifying which deliverables should be considered mandatory and how different elements should be prioritized within the overall evaluation. While these aspects were explored during Concept 2 through discussions with experts and preliminary data collection, no final definitions were established at that stage.

To support such implementation across sites, careful consideration was also given to how the maturity framework should be introduced and scaled in practice. Building on the structured baseline introduced in Concept 1, the rollout strategy for the maturity framework was therefore refined to better reflect practical impact and ease of adoption. As per Figure 12, an initial three-level scale is maintained as a simple entry point to establish alignment and usability between sites.

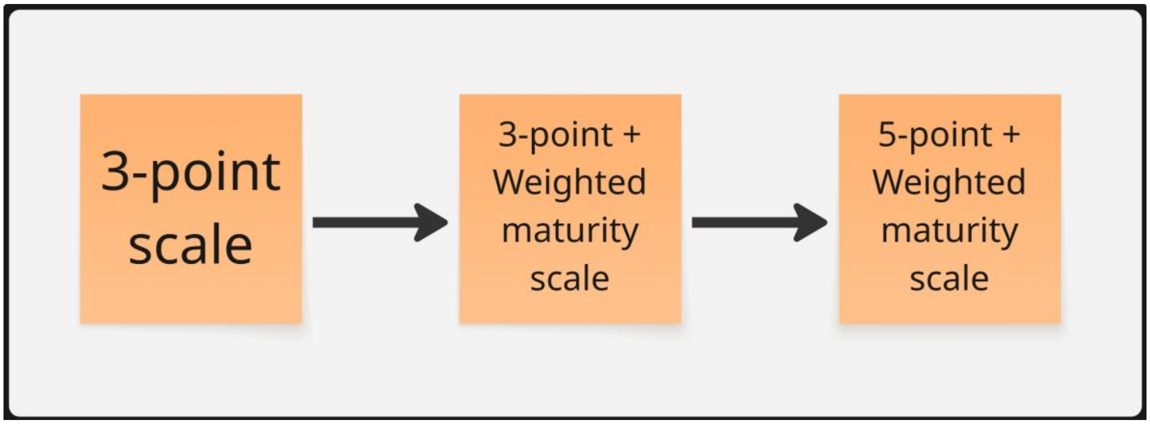


Figure 12: Concept 2: Adjusted rollout strategy for maturity rating.

However, rather than first increasing the number of maturity levels, the next step focuses on introducing weighted scoring to prioritize deliverables based on their relative importance. This adjustment was made as prioritization was found to have a greater influence on evaluation quality than an increased number of grades. Once sufficient experience and alignment are achieved, a five-point scale can then be introduced to provide additional grading that provides more information for tracking improvement over time.

5.3.3 Digital & Data

The Digital & Data block in Concept 2 builds upon the foundation established in Concept 1 but extends its scope by emphasizing governance, usability, and long-term value of life testing data. While Concept 1 primarily focused on identifying data types and describing how information flows across the testing process, Concept 2 further develops this perspective by defining how data should be structured, managed, and reused.

Digital & Data defines the principles for how life testing information is handled across the entire lifecycle and across all modular blocks. The objective is to ensure that data is consistently collected, traceable, and reusable, thereby supporting continuity over time, alignment between sites, and preservation of engineering knowledge.

Rather than treating test data as isolated documentation, this block positions it as part of a shared knowledge base that can support decision-making, comparison, and continuous improvement. This includes enabling engineers to systematically reuse historical data when planning new tests, validating results, or understanding failure mechanisms. The block therefore establishes common expectations for:

- Structured and consistent data collection across sites
- Full traceability from test motivation to final reporting
- Centralized storage and accessibility of life testing data
- Long-term retention of knowledge beyond individual projects or engineers
- Reuse of historical data in simulations, planning, and analysis

To support this, Digital & Data defines the scope of information included in the life testing framework. This covers all relevant data generated before, during, and after testing, including modelling and simulation data, test plans and configurations, operational logs, raw life data, microscopy results, calibration certificates, failure analysis documentation, and final reports. All data is linked to a unique test instance and treated as lifecycle-based information, ensuring traceability and consistency over time.

A central element within this block is the data platform as discussed in Chapter 4.3, which functions as the common knowledge and information infrastructure. The platform provides a shared environment for storing and accessing life testing data, structured according to the modular workflow. This enables clear traceability between requirements, execution, results, and conclusions, while also supporting controlled access, ownership, and long-term preservation of information. Rather than being defined as a specific IT system, the centralized data platform represents a conceptual backbone that ensures consistent data handling across sites.

By connecting historical data with ongoing and future test activities, Digital & Data enables a closed learning loop within the life testing process. This means that previous test results can directly inform new test designs, that simulations and requirements can be based on validated historical data, and that failure mechanisms and lessons learned become systematically reusable.

Through this, the Digital & Data block strengthens the organization's collective knowledge base and reduces dependency on individual experience. It also creates a foundation that supports the maturity framework, as the availability, quality, and traceability of data directly influence how effectively tests can be evaluated, compared, and improved over time.

5.3.4 Process

In Concept 1, Process (SOP) was considered as a separate element intended to support standardization through defined procedures and structured workflows. However, in Concept 2, this element was intentionally removed as an independent component. This decision reflects the evolution of the overall framework, where standardization is no longer treated as dependent on a single process layer, but instead emerges from the combined structure of Modular, Digital & Data, and Maturity Rating. By integrating standardization into these core components, the need for a separate SOP layer is removed. Introducing such a layer on top of the framework would risk creating redundancy and unnecessary complexity, without adding further clarity to how testing activities are structured and evaluated

5.3.5 Summary of Concept 2

Concept 2 defines a structured yet flexible framework intended to improve the execution, evaluation, and reuse of life testing within the organization. The scope of the concept covers the entire life testing process, from initial test motivation to final reporting and long-term data utilization, while ensuring alignment across sites and engineering functions.

The primary intent of Concept 2 is not to prescribe detailed procedures, but to establish a consistent system of expectations built upon modular structure, data governance, and maturity rating evaluation. By combining these elements, the concept aims to strengthen traceability, increase transparency, and enable systematic reuse of knowledge without constraining local engineering practices. Furthermore, Concept 2 seeks to support both operational efficiency and long-term capability development. At the operational level, it ensures that tests are conducted with sufficient completeness and quality to be reliable and comparable.

At the organizational level, it enables continuous improvement by making performance and gaps visible across both individual tests and sites.

5.4 Concept 3

Concept 3 focuses on the practical application of the developed framework by integrating the modular structure, maturity rating perspective and Digital & Data into a unified setup. The concept builds on previous work, but shifts from development to use by introducing aggregation, visualization and demonstration of how maturity rating information can support both operational and strategic decision-making.

5.4.1 Modular Structure

The modular structure applied in Concept 3 remains unchanged from the foundation established in Concept 1 and further operationalized in Concept 2. The same seven modules are used and continue to define the overall structure of the life testing process.

Rather than introducing structural changes, Concept 3 builds upon the existing framework by leveraging its established modular logic. The structure is therefore maintained as a stable backbone, ensuring continuity and consistency across concepts while enabling further development in other areas.

5.4.2 Maturity Perspective

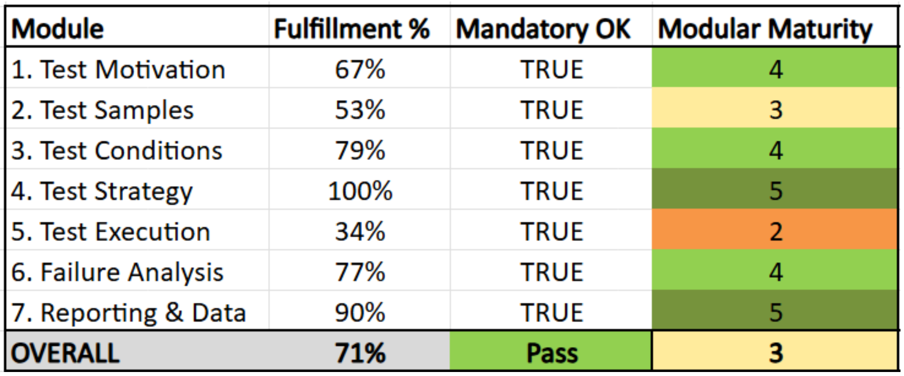
Concept 3 extends the *Maturity rating* introduced in Concept 2 by moving from a conceptual framework to a more applied and demonstrable solution. While previous concepts focus on maturity at test level and the structure of evaluation, Concept 3 introduces aggregation logic, site-level visualization and practical implementation through an example.

At test level, maturity rating is evaluated per modular block and summarized into an overall test maturity rating score. The aggregation follows two governing principles. First, all defined deliverables are currently considered mandatory for a test to be valid as these are minimum requirements. This decision was established based on feedback from the reference site in Airasca, where it was emphasized that all core elements must be fulfilled to ensure traceability and reliability of test results. If any mandatory deliverable is missing, the test is classified at the lowest maturity level.

While all deliverables are currently treated as mandatory for validation purposes, the visualizations presented in the following sections incorporate additional weighted deliverables to provide a more informative representation of maturity rating concept. This approach enables differentiation between varying levels of performance and illustrates how the maturity rating framework can evolve over time. The weighted representations further reflect how maturity rating assessment is intended to be applied in the future, where site-based deliverables are weighted according to their relative importance, allowing more critical aspects of test execution to have a greater influence on the overall maturity evaluation.

This approach reflects the current set of deliverables defined in this study, see Appendix E, which represent a baseline for acceptable test execution. However, the weighting principle introduced in Concept 2 is not removed. Instead, it is intended to be applied in future development, as additional or more differentiated deliverables are defined. In such cases, weighting can be used to reflect the relative importance of different criteria, allowing sites to adapt prioritization based on local context and experience while maintaining a consistent overall evaluation framework.

Second, the overall maturity rating score of a test is constrained by its weakest modular block, such that the total maturity rating level cannot exceed the lowest module maturity rating by more than one level. For example, as shown in Figure 13, the average modular maturity is close to *4*. However, since the *Test Execution* module is rated at level *2*, the overall maturity is limited to level *3*. This rule ensures internal consistency and prevents high overall scores from masking critical weaknesses within individual modules.

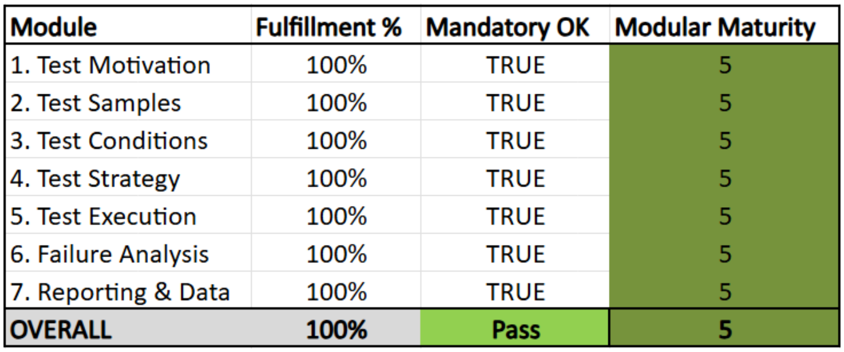


Module	**Fulfillment %**	**Mandatory OK**	**Modular Maturity**
1. Test Motivation	67%	TRUE	4
2. Test Samples	53%	TRUE	3
3. Test Conditions	79%	TRUE	4
4. Test Strategy	100%	TRUE	5
5. Test Execution	34%	TRUE	2
6. Failure Analysis	77%	TRUE	4
7. Reporting & Data	90%	TRUE	5
OVERALL	**71%**	**Pass**	**3**

Figure 13: Illustration of the rule limiting overall test maturity based on the weakest modular block.

From a practical perspective, test-level maturity rating primarily supports technical review and validation. It provides a structured way to assess whether an individual test is complete, reliable and suitable for reuse as a reference, thereby improving decision-making during test planning and reducing the risk of relying on incomplete or insufficient data. In addition, test-level maturity provides visibility into performance over time by enabling comparison between multiple test instances. This makes it possible to assess whether test quality is consistent or varies significantly across similar tests, highlighting potential instability in execution or documentation practices. While test-level maturity ensures the validity of individual tests through strict evaluation rules, the site-level perspective reflects performance across multiple tests, allowing variation while still providing an overall indication of capability.

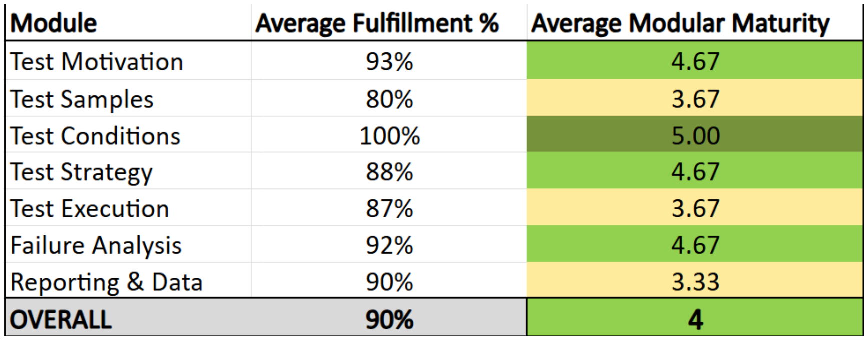
To validate and illustrate the approach, an Excel document was developed in which maturity rating can be visualized at both test level and site level. At test level, individual test instances are assessed and assigned an overall maturity rating score, providing a direct indication of test completeness and usability for comparison or reuse, see Figure 14.



Module	Fulfillment %	Mandatory OK	Modular Maturity
1. Test Motivation	100%	TRUE	5
2. Test Samples	100%	TRUE	5
3. Test Conditions	100%	TRUE	5
4. Test Strategy	100%	TRUE	5
5. Test Execution	100%	TRUE	5
6. Failure Analysis	100%	TRUE	5
7. Reporting & Data	100%	TRUE	5
OVERALL	**100%**	**Pass**	**5**

Figure 14: Example of test-level maturity assessment for an individual test.

Building on this, the maturity rating logic enables aggregation across multiple tests within a site. By summarizing test-level maturity results, a site-level overview is obtained, where each module is represented by an aggregated performance and maturity rating level, see Figure 15. This enables identification of systematic strengths, recurring weaknesses and overall capability across modules.



Module	Average Fulfillment %	Average Modular Maturity
Test Motivation	93%	4.67
Test Samples	80%	3.67
Test Conditions	100%	5.00
Test Strategy	88%	4.67
Test Execution	87%	3.67
Failure Analysis	92%	4.67
Reporting & Data	90%	3.33
OVERALL	**90%**	**4**

Figure 15: Example of site-level maturity overview based on multiple test results.

At site level, maturity rating supports a broader, strategic perspective. Rather than focusing on individual tests, the aggregated results provide a basis for decision-making related to resource allocation, investment priorities and distribution of testing activities. Sites with higher maturity ratings in specific modules can be better positioned to handle more advanced or critical testing tasks, while identified gaps can guide targeted improvements and capability development. And lower-maturity sites can be assigned more standardized and recurring tests, allowing them to build experience while more advanced capabilities are developed.

In addition to the static site-level overview, the Excel also enables visualization of maturity rating development over time. By tracking results across multiple test series, it is possible to monitor how sites evolve and improve their testing practices, see Figure 16. This introduces a performance over time perspective to the framework and supports continuous improvement based on measurable progress.

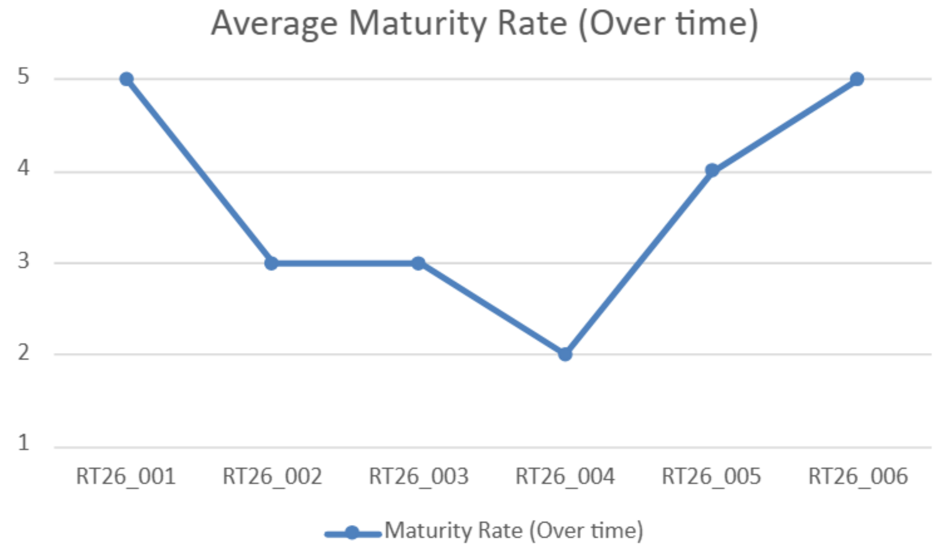
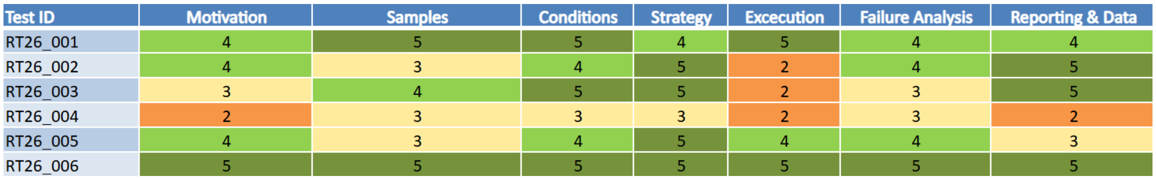


Figure 16: Example of maturity development over time across multiple test series.

To help assess progress over time at a more detailed level, a heatmap visualization is introduced, see Figure 17. In this representation, the seven modular blocks are displayed along the vertical axis, while individual tests are arranged horizontally in chronological order. Each cell represents the maturity rating for a specific module within a given test, allowing for a clear overview of development patterns over time. This enables identification of both consistent performance and areas of improvement or fluctuation within specific modules, thereby supporting continuous monitoring of progression. For example, as illustrated in Figure 17, there is notable variation across modules, where *Motivation* and *Execution* exhibit lower scores and greater fluctuation, while *Conditions* and *Strategy* demonstrate consistently higher performance levels.



Test ID	Motivation	Samples	Conditions	Strategy	Excecution	Failure Analysis	Reporting & Data
RT26_001	4	5	5	4	5	4	4
RT26_002	4	3	4	5	2	4	5
RT26_003	3	4	5	5	2	3	5
RT26_004	2	3	3	3	2	3	2
RT26_005	4	3	4	5	4	4	3
RT26_006	5	5	5	5	5	5	5

Figure 17: Heatmap illustrating progression of module maturity rating over time.

Together, these elements extend the maturity rating framework from an evaluative concept to a practical decision-support tool, supporting both detailed validation of individual tests and high-level understanding of site capabilities and development potential over time.

5.4.3 Digital & Data

The Digital & Data perspective in Concept 3 builds upon the foundation established in Concepts 1 and 2, but shifts focus from data structure to practical usage and integration. While earlier concepts defined how life testing information is generated, structured and stored, Concept 3 emphasizes how this data is utilized together with the maturity rating framework to support decision-making. All life testing data, including: modelling inputs, test configurations, operational logs, raw life data, failure analysis and final reporting, remains structured and stored within a shared environment such as the centralized data platform, ensuring traceability and accessibility across the entire testing lifecycle. This establishes a consistent data backbone that enables both reuse and comparison of test results across sites.

A key development in Concept 3 is the integration of maturity rating evaluation directly into this data environment. Instead of treating data and maturity rating as separate elements, maturity rating is applied as an additional layer that enables immediate interpretation of stored test information. In its current implementation, each test is evaluated based on defined mandatory deliverables and assigned a validation status (pass/fail), providing a clear indication of whether the test meets the baseline requirements for traceability and reliability. This allows engineers to quickly identify which test results are valid and suitable for reuse without detailed manual review. As the framework evolves, additional deliverables and weighting criteria can be introduced, enabling a more advanced maturity rating score (e.g. 1–5) to differentiate test quality beyond basic validity. This development is intended to remain flexible, allowing sites to define weighted criteria based on local priorities while maintaining a consistent overall structure.

Through this integration, the Digital & Data perspective enables efficient filtering and selection of relevant test data. Engineers can directly identify tests that are both valid and of sufficient quality, supporting faster decision-making during test planning and reducing reliance on manual evaluation of historical data. In combination with the site-level aggregation introduced in the maturity rating framework, this also enables data to be analyzed beyond individual tests. Performance can be assessed across sites and over time, supporting identification of trends, strengths and areas for improvement. By combining structured data management with integrated maturity rating evaluation, the Digital & Data perspective transforms the data platform from a passive data repository into an active decision-support system. This enables both operational use of test data and strategic insights into testing performance, supporting more consistent and informed engineering decisions across the organization.

5.4.4 Summary of Concept 3

Concept 3 builds on the previously developed framework by moving from structure and evaluation towards practical application. By integrating modular structure, maturity rating, and data management, the concept provides a standardized and consistent way of interpreting life testing activities across sites. Through aggregation, visualization and integration with shared data environments, maturity rating results can be applied at both test level and site level. This enables not only validation of individual tests, but also consistent comparison of performance, identification of systematic differences, and structured follow-up of development over time. By standardizing how test quality and performance are evaluated and communicated, Concept 3 supports both operational decision-making and strategic alignment across the organization.

6. Discussion

The results of this study demonstrate that the main challenges in complex engineering activities are not primarily related to technical capability, but to how activities are structured, documented and evaluated across an organization. While individual sites often exhibit strong engineering expertise, the lack of a shared structure limits transparency, comparability and effective reuse of knowledge. This highlights that the core issue is not how tests are performed, but how testing activities are interpreted and managed at an organizational level.

From a positive perspective, the developed framework provides a structured approach that enables a more consistent way of describing and evaluating testing activities. The introduction of a modular structure clarifies how work is organized, while the maturity rating perspective enables systematic evaluation of completeness and quality. Together with the Digital & Data component, the framework supports improved traceability and makes it easier to interpret and reuse test data across sites. When applied correctly, the maturity rating framework also provides a transparent way of making implicit practices visible, which supports alignment and shared understanding across the organization. This aligns well with the theoretical perspective presented in Chapter 2, where standardization is described as a mechanism for enabling visibility, collaboration and continuous improvement.

It should also be emphasized that while the framework is developed within the context of bearing life testing, its underlying structure is intended to be applicable to complex engineering activities more broadly. However, its application in other contexts would require a similar pre-study to map and understand the specific process, enabling the definition of a suitable modular structure and the identification of corresponding deliverables. Furthermore, the framework is particularly suited for recurring engineering activities, where knowledge can be accumulated, reused, and continuously refined over time. In contrast, its application to non-repetitive tasks is more limited, as much of the value of the framework depends on the ability to establish a structured and continuously evolving knowledge base.

However, the introduction of such a framework also presents several challenges. One limitation is the potential increase in perceived administrative workload. Defining deliverables, documenting activities and maintaining structured data requires additional effort, particularly during initial implementation. In early stages, the maturity rating evaluation may be experienced as time-consuming, as both criteria and workflows are still being established and internalized. But, as experience increases and routines become more familiar, the process should become more efficient and less time-intensive. Without proper integration into existing systems and workflows, the framework nevertheless risks being perceived as an added layer of complexity rather than a supporting tool. In this context, it is equally important that users perceive the framework as a tool that supports and simplifies their daily work, for example by enabling more effective reuse of existing knowledge and reducing the need for repeated analysis, rather than as an additional administrative task to be performed for each test.

This highlights the importance of digital integration. When maturity rating evaluation and documentation are embedded into systems such as the centralized data platform, the additional effort can be reduced while significantly increasing the value of the framework. In such cases, structured data becomes directly usable for analysis, comparison and decision-making, transforming documentation from a static requirement into an active asset that supports both operational work and long-term knowledge development.

Another important consideration is the balance between standardization and flexibility. While the framework aims to create a shared structure, overly strict definitions may limit the ability of engineers to adapt to local conditions, resources and test setups. At the same time, insufficient structure may lead to continued variation and reduced comparability. The results therefore reinforce that effective standardization requires a balanced approach, where alignment is achieved without constraining necessary local adaptation. This balance is also not static. As the framework is implemented and used over time, it can be adjusted based on observed effects and practical experience. This allows the level of standardization to evolve, enabling organizations to refine where stricter alignment is needed and where flexibility should be maintained.

Several potential risks should also be acknowledged. One concern is the dependency on user adoption. The framework relies on consistent use of defined structures and evaluation criteria, and its effectiveness is therefore closely linked to how well it is accepted and applied in practice. Resistance to change, differences in working culture and varying levels of experience between sites may affect implementation. Another risk is that maturity rating evaluation could be interpreted as a performance ranking tool rather than a support for improvement. If used incorrectly, this may create unintended competition or discourage transparent reporting of incomplete tests. It is therefore important that the maturity rating framework is applied as a mechanism for understanding and development rather than control.

In addition, there is a risk related to how deliverables are assessed in practice. When maturity rating evaluation relies on manual input, there is potential for overestimation of completeness, where deliverables are marked as fulfilled without being fully completed. This may occur unintentionally due to unclear criteria, or intentionally to present results in a more favorable way. Such behaviour would reduce the reliability of the maturity rating framework and undermine its purpose as a tool for transparency and comparison.

At the same time, several aspects of the framework reduce this risk in practice. Since test documentation and maturity rating evaluations are logged and visible over time, results can be reviewed and compared by other engineers and stakeholders. This transparency makes it more difficult to misrepresent the status of a test without being identified. In addition, the framework introduces a clearer sense of responsibility at site level, where completed tests are expected to meet defined standards before being considered valid and reusable. This shifts the focus from simply completing activities to ensuring that they are performed with sufficient quality.

To further mitigate risks related to interpretation and evaluation, clear definitions of deliverables, shared interpretation guidelines and, where possible, integration with objective data sources are important. This emphasizes that the effectiveness of the framework depends not only on its structural design, but also on how it is applied, monitored and maintained in practice.

An additional benefit of the framework is its ability to provide insight at site level. By making maturity rating visible across multiple tests and modules, it becomes possible to identify site-specific strengths and weaknesses. This creates opportunities for improved collaboration, where knowledge and best practices can be shared between sites. From an organizational perspective, such visibility can also support more informed decision-making regarding how tests, resources and expertise are distributed across the organization. Finally, the framework's validation has been limited to a single reference site, which may affect its general applicability. Differences in infrastructure, competence and organizational context across other sites may require adjustments to ensure practical usability.

Overall, the study demonstrates that a structured framework combining modularization, evaluation and data management has strong potential to improve how engineering activities are managed and interpreted. At the same time, its success depends on careful implementation, user acceptance and integration with existing systems, highlighting that standardization is not only a technical challenge, but also an organizational one.

7. Conclusions

This thesis has addressed the challenge of achieving consistency, traceability and comparability in engineering testing activities within a global organizational context. By using bearing life testing as a representative case, the study has demonstrated how variation in workflows, documentation and evaluation practices can limit the effective use and interpretation of test results across an organization. The findings show that the primary challenge is not the absence of technical capability, but the lack of a structured and shared approach for how testing activities are described and evaluated.

To address these challenges, a structured framework was developed through an iterative concept development process. The framework combines a modular structure, maturity rating evaluation and data principles to establish a common reference for managing testing activities. Rather than prescribing specific procedures, the approach focuses on defining clear structures, expectations and evaluation criteria that enable alignment while maintaining flexibility for local adaptation.

The results indicate that such a framework improves transparency by making testing activities more explicit and comparable across sites. It further supports consistent interpretation of test results and enables more efficient reuse of historical data, thereby strengthening the organization’s ability to leverage existing knowledge. In addition, the integration of maturity rating evaluation and data management provides a practical link between execution, evaluation and continuous improvement.

At a broader level, the study highlights that standardization in complex engineering environments is not achieved through strict procedural control, but through establishing a shared structure for interpretation and communication. This aligns with the theoretical perspective presented in Chapter 2, where standardization is described as a mechanism for enabling visibility, collaboration and improvement.

Overall, the proposed framework provides a foundation for improving how complex engineering activities are managed, interpreted and developed over time. While its successful implementation depends on factors such as user adoption and system integration, the study demonstrates that a structured and flexible approach to standardization can support both operational consistency and long-term organizational learning.

8. Recommendations

While this study establishes a structured framework for improving consistency and standardization of engineering activities, several areas are recommended for further development and implementation. A key recommendation is the implementation and evaluation of the framework across multiple sites. Since this study is based on a single reference site, broader application is necessary to assess how the framework performs under varying conditions, organizational structures and levels of maturity. Such implementation would also provide insight into how the balance between standardization and flexibility is maintained in practice.

It is further recommended to refine the maturity rating framework, particularly regarding the definition and weighting of deliverables. While the current approach establishes a baseline through mandatory criteria, the introduction of weighted elements could improve differentiation and better reflect the relative importance of different activities. At the same time, clear guidelines and evaluation methods should be developed to reduce subjectivity and ensure consistent application across sites.

Another important recommendation concerns integration with digital systems such as the upcoming centralized data platform. Embedding the framework into existing data environments would reduce administrative effort and enable automated data capture, validation and analysis. This would strengthen the reliability of maturity rating assessments and increase the practical value of the framework as a decision-support tool.

Finally, it is also recommended to establish clear governance structures and define responsibilities across roles such as test engineers, site managers, and central functions. Defining ownership of different parts of the framework is critical to ensure consistent application and long-term sustainability, and to prevent the system from losing effectiveness after initial implementation. Practical guidelines should define responsibilities related to evaluation, validation and follow-up of maturity rating assessments within and across sites. It is also recommended to support implementation through training, clear communication of purpose, and alignment of expectations, to ensure that the framework is perceived as a support tool rather than an administrative requirement.

Reference list

Carlile, P. R. (2004). *Transferring, translating, and transforming: An integrative framework for managing knowledge across boundaries*. Organization Science, 555-568. http://www.jstor.org/stable/30034757

Deming, W. E. (1986). *Out of the crisis*. MIT Press.

Goel, K., Bandara, W., & Gable, G. G. (2023). Conceptualizing business process standardization: A review and synthesis. Schmalenbach Journal of Business Research, 75, 195–237. https://doi.org/10.1007/s41471-023-00158-y

Grant, A., Bugge, C., & Wells, M. (2020). *Designing process evaluations using case study to explore the context of complex interventions evaluated in trials*. Trials. https://doi.org/10.1186/s13063-020-04880-4

Isom, E. (n.d.). *Multiple and mixed methods research*. FHSU Digital Press. https://fhsu.pressbooks.pub/socialresearch/chapter/multiple-and-mixed-methods-research/

International Organization for Standardization. (2015). *ISO 9001:2015 quality management systems - Requirements*. https://www.iso.org/standard/62085.html

International Organization for Standardization. (2017). *ISO 15243:2017 rolling bearings—Damage and failures—Terms, characteristics and causes*. https://www.iso.org/standard/59619.html

Kelley, B. (2022). *Incorporating user feedback into iterative design. Human-Centered Change and Innovation*. https://bradenkelley.com/2022/07/incorporating-user-feedback-into-iterative-design/

Lean Outside the Box. (2025, January 21). *DMAIC: A proven framework for process improvement.* https://leanoutsidethebox.com/dmaic/

Lean Outside the Box. (2026, February 6). *Tree diagrams in Lean Six Sigma: How to Break Down Complex Problems*. https://leanoutsidethebox.com/tree-diagrams/

Lykins, T. (2023). *What's the difference between top-down and bottom-up processing*. Fischer Institute. https://fischerinstitute.com/difference-between-top-down-and-bottom-up-processing/

Microsoft. (n.d.). *Start, stop, and download live transcripts in Microsoft Teams meetings*. https://support.microsoft.com/en-us/office/start-stop-and-download-live-transcripts-in-microsoft-teams-meetings-dc1a8f23-2e20-4684-885e-2152e06a4a8b

Miro. (n.d.). *What is Miro? Get to know our AI innovation workspace*. https://miro.com/what-is-miro/

Ruslin, R., Mashuri, S., Abdul Rasak, M. S., Alhabsyi, F., & Syam, H. (2022). *Semi-structured interview: A methodological reflection on the development of a qualitative research instrument in educational studies*. IOSR Journal of Research & Method in Education, 12(1), 22–29. https://www.iosrjournals.org/iosr-jrme/papers/Vol-12%20Issue-1/Ser-5/E1201052229.pdf

SixSigma.us. (2024a). *Deming's 14 points: Transforming business through quality management*. https://www.6sigma.us/six-sigma-in-focus/demings-14-points/

SixSigma.us. (2024b). *Ultimate guide to process standardization: Strategies for operational excellence*, 3158-3163. https://www.6sigma.us/business-process-management-articles/process-standardization-for-operational-excellence/

SKF. (2026). *CARB toroidal roller bearing – cross-sectional view* [Image]. https://cdn.skfmediahub.skf.com/api/public/0901d196802e5cb3/png_highpreview_800/0901d196802e5cb3_png_highpreview_800.png

Appendix A: SKF Current Test Flow

This appendix provides a description of the life testing process as it is currently performed at a SKF testing site. While Chapter 4 presents a consolidated overview of the process divided into three main phases (testing preparation, test execution, and test analysis & documentation), this appendix expands on the underlying practices, methods, and technical considerations that support each phase. The structure of this appendix follows key functional areas identified during the prestudy.

A1. Testing Preparation

The life testing process begins with a preparation phase in which the foundation for the test is established. This phase focuses on defining the purpose, scope and feasibility of the test, as well as ensuring that all necessary inputs are properly defined before execution. Through structured dialogue, technical evaluation and planning activities, this phase ensures that the test is both technically justified and practically executable. It includes the clarification of requirements, definition of samples and operating conditions, and development of appropriate test strategies. Together, these activities create a consistent and well-defined basis for the subsequent execution of the life test.

A1.1 Initiation and Responsibility Assignment

Life testing at SKF is initiated through a structured process intended to ensure that each test is technically justified, feasible and aligned with both customer expectations and internal priorities. Rather than starting directly from a predefined test configuration, the process progresses through defined stages that transform an initial request into a clearly motivated and approved life test. This chapter describes how life tests are currently initiated and aligned in practice before detailed test preparation begins.

A1.1.1 Customer Proposal and Responsible Test Engineer

The starting point for a life test is a formal proposal from a customer. This request originates from an external customer or from internal development, verification or benchmarking activities. At this stage, the proposal captures the underlying motivation for testing rather than a fully specified technical setup. Its primary function is to formally initiate the process and ensure that the need for testing is visible, documented and subject to evaluation.

Upon receipt of the test proposal, a responsible test engineer is appointed. This appointment represents a key transition from a general request to an owned technical task. The responsible test engineer becomes accountable for coordinating the life test throughout its entire lifecycle and serves as the main contact for technical and communicative continuity.

A1.1.2 Customer Dialogue

Following the assignment of responsibility, a structured dialogue with the customer is initiated. This dialogue aims to clarify expectations, boundary conditions and operational requirements associated with the requested life test. Initial customer wishes are translated into technically interpretable parameters, while laboratory capabilities and constraints are assessed.

During this phase, the responsible test engineer evaluates whether the requested conditions can realistically be achieved and whether the resulting test would generate meaningful and defensible results. If necessary, adjustments to scope, operating conditions or test duration are discussed to align technical relevance with practical feasibility.

In addition, potential risks, limitations, and uncertainties associated with the proposed test setup and its execution are identified and considered. This ensures that technical challenges, boundary conditions, and possible limitations in the interpretation of results are understood early in the process. This alignment step is essential to ensure that both parties share a common understanding of what the life test can demonstrate and, equally important, what it cannot.

A1.1.3 Test Purpose

Based on the clarified requirements and feasibility assessment, the purpose of the life test is formally defined. At this stage, the test purpose represents a consolidated and agreed description of why the test is performed, what value it is expected to deliver and within which boundaries the results are considered valid. The purpose may involve validating performance under defined operating conditions, generating life data or supporting design and application-related decisions. Establishing a clearly articulated test purpose ensures that the test has a well-defined objective and contributes with meaningful insight.

In connection with defining the test purpose, criteria for evaluating the outcome of the test are established and agreed upon. These criteria clarify how results will be interpreted and what constitutes acceptable or meaningful performance, thereby reducing ambiguity and ensuring a common basis for evaluation among stakeholders. Establishing a clear purpose and agreed evaluation criteria is essential to prevent life tests that consume resources without delivering meaningful insight. This concern was strongly emphasized during the expert interviews:

> *"We need to prevent that a lab runs a test with six bearings for three hours and then claims the life is OK."*
>
> - Global Testing Technical Expert, see Appendix B.

By defining the test purpose only after requirement alignment, SKF reduces the risk of poorly motivated life tests and ensures that subsequent decisions are anchored in a shared technical understanding.

A1.1.4 Management Acceptance

Before a life test transitions from the motivation and alignment phase into detailed preparation, management acceptance is required. This step functions as a final control point to ensure that the proposed test is not only technically justified but also operationally feasible.

Management evaluates the defined purpose and intended value of the test, the agreed purpose and evaluation criteria, as well as identified risks and uncertainties. In addition, practical aspects such as availability of test rigs, resource allocation, staffing and scheduling are considered, together with how the requested life test aligns with broader organizational priorities. Only after this acceptance can the engineering team proceed with detailed preparation activities, such as finalizing test sample orders, defining test conditions and planning execution. Management acceptance therefore acts as a gating mechanism, ensuring that motivated life tests can be realistically executed.

A1.2 Sample Definition and Traceability

Reliable interpretation of life test results requires that tested bearings are both traceable and representative of standard production. Without this fundamental context, observed failures risk being misinterpreted or difficult to compare across tests. Documentation and material understanding therefore form a necessary foundation for subsequent analysis, modelling and failure evaluation.

A1.2.1 Documentation, Traceability and Sample Handling

Traceability is fundamental to any life test campaign. Each bearing must be traceable throughout the entire process, from initial sample delivery to final failure analysis. This traceability is ensured through standardized metadata and controlled documentation practices, which together form the technical identity of each sample bearing. As one of the experts expressed clearly:

> *"Metadata to me is the identity card of the bearing."*
>
> - Global Testing Technical Expert, see Appendix B.

Metadata typically includes bearing designation, production batch, internal markings, supporting drawings and, when available, material certificates. Together, these elements ensure that every bearing can be linked back to its manufacturing origin, enabling correct interpretation of modelling inputs, operational logs and final test results.

In addition to basic identification data, the manufacturing and material history of each sample is documented. This includes relevant production information, material specifications and available certification data. Such information is essential for understanding how material properties, production routes and batch variations may influence test behaviour and failure mechanisms. Maintaining traceability further requires that test samples are formally received and registered upon delivery to the testing laboratory. This step ensures that all samples included in the test campaign are accounted for and correctly linked to their associated documentation before testing begins.

Following registration, each sample and associated test campaign are assigned unique internal identifiers. This includes both a sample ID and a test series ID, ensuring consistent traceability across planning, execution and analysis phases. These identifiers provide the formal linkage between physical samples, test conditions and recorded data, enabling consistent tracking throughout the test lifecycle.

Correct sampling practices are also an essential part of test-sample handling. Bearing samples must reflect standard production rather than selective or exceptional units. This was clearly emphasized during the expert interviews:

> *"The correct way to state it is to have samples which are representative*
> *of standard production. Avoid cherry-picking."*
>
> - Global Testing Technical Expert, see Appendix B.

Ensuring representativeness at this stage helps avoid bias and provides a solid foundation for later statistical evaluations, comparative studies and overall standardization. Rather than showcasing results from only top-performing bearings, the objective is to provide a realistic representation of how bearings perform under the specified conditions.

A1.2.2 Metallurgical Inspections and Material Verification

To ensure that life test results accurately reflect material fatigue behaviour rather than unintended manufacturing effects, selected bearing samples undergo metallurgical inspection and material verification. The purpose of these inspections is not routine documentation but to support correct interpretation of test outcomes when material-related mechanisms are relevant to the test objective or subsequent failure analysis. The expert interviews consistently emphasized that insufficient understanding of the tested material condition increases the risk of misinterpreting results. In particular, failures caused by manufacturing irregularities or material deviations may otherwise be incorrectly attributed to fatigue-driven failure modes. As one expert explained:

> *"If material differences are not understood, you risk attributing failures to fatigue when they come from manufacturing."*
>
> - Global Testing Technical Expert, see Appendix B.

Based on the interviews, metallurgical verification is therefore applied selectively rather than systematically for every life test. It is primarily conducted when the test aims to compare materials, heat treatments or manufacturing routes, or when unexpected failure behaviour requires confirmation that results are not influenced by material differences. For baseline tests, standard life tests or tests involving well-known and controlled materials, full metallurgical characterization may not be necessary.

When performed, metallurgical inspections typically include standardized analyses such as microstructure evaluation, hardness measurements, retained austenite content, residual stress profiles and chemical composition. In addition, surface and subsurface conditions are examined to identify defects or damage that could influence fatigue initiation. These inspections follow established internal routines and are designed to provide sufficient insight to distinguish between material-driven effects and test-induced fatigue behaviour, while avoiding unnecessary cost or lead time.

From a documentation perspective, metallurgical findings are not fully reproduced within the life test report itself. Instead, the detailed analysis is documented in a separate material or metallurgy report, stored within the internal documentation system. The life test report includes a high-level summary of relevant material information and references the corresponding report when applicable.

A1.3. Definition of Operating Conditions

The purpose of Test Conditions is to ensure that bearings operate under controlled, consistent and technically validated operating parameters throughout testing. By defining how loads, constraints and lubrication are applied and verified, this section establishes the requirements for a stable and reproducible testing environment where observed failures can be confidently attributed to rolling-contact fatigue rather than unintended test-setup variations.

A1.3.1 Bearing Modelling Practices

SKF uses the in-house developed SimPro Expert program, where standardized bearing modelling ensures that identical inputs, such as software versions, lead to identical simulation results, regardless of who performs the analysis or where it is carried out. Consistent modelling establishes a shared baseline for defining internal bearing conditions and verifying that planned test parameters are physically feasible. To ensure traceability and reproducibility, modelling activities require that inputs, assumptions and software versions are defined, documented and stored together with the simulation results. Three categories of modelling inputs are especially critical: load application, boundary conditions and geometric tolerances.

A1.3.1.1 Load Application

Load application defines the amount, direction and distribution of external forces acting on the bearing. Even small deviations in load level or orientation can significantly alter internal stress patterns, thereby affecting contact pressure, fatigue behaviour and, therefore, the expected life of the bearing. One expert highlighted why correct load definition is so important:

> *"If the contact pressure is extremely high, the bearing cannot even start.*
> *If it is too low, the test might run for years."*
>
> - Global Testing Technical Expert, see Appendix B.

Loads may be radial, axial or combined, depending on the bearing design and intended use. The applied load should represent the target operating condition while avoiding unrealistic or excessively severe loading. The expert statement highlights a key challenge in bearing life testing: *accelerated testing*. Since bearings are typically expected to operate for years, laboratories increase the severity of test conditions to shorten test times. This helps save both time and resources, but requires engineers to carefully balance test acceleration with result accuracy.

A1.3.1.2 Boundary Conditions

Boundary conditions describe how the bearing is constrained, supported and aligned in the simulation environment. They strongly influence how forces are transmitted through the bearing and directly affect internal contact mechanics. Misalignment or overly rigid constraints can introduce artificial stress concentrations or non-representative behaviour. To ensure consistency, positional and rotational degrees of freedom must be defined in a controlled way so that the simulated stress environment corresponds to what the physical test rig will impose. These definitions also form the basis for how the physical test setup is configured and documented.

A1.3.1.3 Geometric Tolerances

Despite high manufacturing quality, bearings cannot be produced as perfectly identical components, and characteristics such as roundness deviation, waviness, roller profile variation and internal clearance affect load distribution and fatigue progression. Incorporating realistic tolerances in the model ensures that simulations match the behaviour of actual physical samples. During interviews, the importance of following each modelling step precisely was emphasized:

> *"If you miss even one modelling step, you will get a different result*
> *and then you will question everything."*
>
> - Global Testing Technical Expert, see Appendix B.

Standardizing modelling routines therefore ensures reproducibility across test sites and builds confidence that internal bearing conditions reflect the intended test configuration.

A1.3.2 Contact and Stress Parameters

Understanding internal bearing behaviour requires calculating a set of key contact and stress parameters. These parameters define the fatigue-driving environment inside the bearing and must be documented to verify that selected test conditions are representative. The main internal parameters include:

- **Contact pressure**, which determines the local stress at the rolling contact and is a primary driver of subsurface fatigue.
- **Contact ellipse dimensions**, specifying the size and shape of the loaded zone and influencing stress gradients.
- **Subsurface stresses**, representing stress peaks at depths where fatigue cracks often initiate.
- **Film thickness** & **kappa lubrication parameter**, which indicate the lubrication regime and potential for surface distress.
- **Friction torque**, reflecting sliding losses and corresponding heat generation.
- **Load distribution**, describing how the applied load is shared among rolling elements.

These values enable engineers to judge whether the test is operating within a realistic fatigue state or deviating into unintended conditions. One expert highlighted why these internal parameters matter:

> *"Understanding the contact parameters helps you make*
> *sense of what you see after the test."*
>
> - Global Testing Technical Expert, see Appendix B.

Documenting these parameters in the test report provides transparency and supports correct interpretation of failures.

A1.3.3 Lubrication Feasibility

Before testing begins, it is essential to verify that the lubrication system can sustain the intended operating conditions. Lubrication feasibility ensures that the bearing can operate in a stable thermal and tribological state throughout the test and that failures originate from fatigue rather than lubrication-related issues. Two linked factors must be evaluated:

- **Frictional heat generation**, which increases with load, speed and insufficient lubrication.
- **Cooling capacity**, determined by lubricant flow rate, inlet temperature and the thermal properties of the oil.

If the lubrication system cannot remove heat at a sufficient rate, the bearing may enter thermal runaway, leading to elevated temperatures, oil degradation and premature failure. Test conditions are therefore considered feasible only when frictional heat generation and heat removal are in balance. As one expert noted during the discussions:

> *"You need to ensure that the lubrication system can actually handle the load, otherwise the test is not meaningful."*
>
> - Global Testing Technical Expert, see Appendix B.

This verification ensures that the defined test conditions are both technically feasible and representative of the intended operating environment.

A1.4. Test Planning and Strategy Development

The purpose of Test Strategies is to ensure that life tests are planned and executed in a consistent, statistically reliable and operationally feasible way. By defining how test types are selected, how loads and durations are balanced, and how sample sizes are determined, test strategies provide the framework required to produce meaningful and comparable life test results. This section outlines the principles used to classify test types, apply planning tools and design strategies that reflect the intended fatigue mechanisms while meeting practical constraints.

A1.4.1 Classification of Test Types

A shared classification system is essential for ensuring that engineers across laboratories use consistent terminology and interpret objectives in the same way. Three core test types are commonly used in life testing environments: endurance tests, minimum life tests and contamination tests.

A1.4.1.1 Endurance Tests

Endurance tests, here referring to bearing life tests run to failure, are used to study fatigue behaviour under controlled and well-defined conditions. The tests are commonly continued until a sufficient number of bearings has reached failure to allow statistical evaluation of life parameters such as L10 and β.

In this context, failure does not imply complete bearing destruction. Instead, the test is terminated at the first reliable indication of fatigue damage, such as spalling or other characteristic signs of surface or subsurface fatigue. As emphasized during the expert interviews:

> *"The definition of bearing life is the appearance*
> *of the first evidence of fatigue."*
>
> – Global Testing Technical Expert, see Appendix B.

This approach ensures that the measured life reflects fatigue initiation rather than secondary damage mechanisms caused by continued operation after failure onset. The resulting data is primarily used for comparative evaluation, design validation and assessment of bearing performance under defined conditions.

A1.4.1.2 Minimum Life Tests

Minimum life tests are conducted to verify that a bearing meets a specified minimum life requirement rather than to characterize the full life distribution. In these tests, the objective is to demonstrate compliance with a defined threshold under specified conditions, rather than to generate detailed statistical life parameters.

Unlike endurance tests, minimum life tests are often run for a fixed duration or until a predefined number of operating cycles has been completed. The outcome is then evaluated based on whether failures occur within this time window. As a result, the focus is typically on pass–fail evaluation and qualification rather than on estimation of parameters such as L10 or β. The predefined test length also makes the test duration more predictable, which facilitates planning and scheduling compared to enudrance testing.

Such tests are commonly used for acceptance, qualification or verification purposes, particularly when the intent is to confirm that a bearing design or manufacturing process satisfies minimum performance requirements before release or further application.

A1.4.1.3 Contamination Tests

Contamination tests are designed to assess bearing performance under harsh conditions where the presence of foreign particles or contaminants is expected to influence fatigue behaviour. These tests deliberately introduce controlled contamination into the lubrication system in order to accelerate damage and provoke contamination-induced failure mechanisms. The level and nature of the contamination are defined in a controlled manner to ensure repeatability and meaningful comparison between tests. The expert interviews highlighted that contamination fundamentally alters the nature of the test by imposing severe and continuous surface distress from the outset. As one expert described:

> *"If you introduce contamination, you are essentially hammering the bearing from the start. They will not survive too long."*
>
> - Global Testing Technical Expert, see Appendix B.

As a result, contamination tests are not intended to represent normal operating conditions, but rather to evaluate robustness, lubrication sensitivity, filtration effectiveness and material resistance to surface-initiated fatigue. They are particularly valuable for stress-testing designs and lubrication strategies in applications where contamination exposure is unavoidable.

A1.4.2 Strategy Development Using SET2

Life test strategies are developed by evaluating trade-offs between statistical reliability and practical constraints such as available test capacity, sample availability and acceptable test duration. SET2 is the dedicated and consistently applied simulation tool used during this planning phase. It is used to model and evaluate how different choices of precision, test duration and sample size interact under assumed failure behaviour before any test conditions are finalized. By simulating alternative scenarios, SET2 provides a common basis for decision-making and ensures that all key strategy parameters are defined before test execution.

In addition to sample size and test duration, censoring strategies are defined as part of the test planning. This includes decisions regarding suspension time and handling of censored test data, which are essential for ensuring correct statistical interpretation of the results. Results from SET2 simulations, including assumptions, input parameters and selected strategy configurations, are documented and stored to ensure traceability and enable future review or comparison between tests.

A1.4.2.1 Precision

Precision reflects the statistical confidence in the estimated life distribution and is primarily driven by the number of observed failures. A higher number of failures generally leads to narrower confidence intervals and more reliable life estimates, whereas limited failure data results in high uncertainty. As one expert summarized:

> *"The more bearings you test, the more failures you obtain,*
> *the more precise your results become."*
>
> - Global Testing Technical Expert, see Appendix B.

Achievable precision is therefore often evaluated during the planning phase using the projection tool SET2 to assess whether a proposed test strategy can achieve sufficient statistical confidence.

A1.4.2.2 Duration

Test duration depends on factors such as applied load, speed, lubrication conditions and target life level. Increasing load or speed reduces test duration but may introduce non-representative failure mechanisms, while milder conditions improve realism at the cost of longer test periods. Despite careful planning, test duration remains inherently uncertain:

> *"You never know what will happen until it happens, that's why we*
> *make projections but must always be ready to adjust."*
>
> - Global Testing Technical Expert, see Appendix B.

Projected test duration is therefore commonly assessed in advance using scenario simulations to support realistic scheduling and resource allocation. In this context, the selected test duration also defines the suspension time for tests that are not run to failure, making it a key parameter in both test execution and statistical interpretation.

A1.4.2.3 Sample Size

Sample size directly influences both the achievable precision and overall test feasibility. Too few samples risk producing confidence intervals that are too wide to support meaningful conclusions, while excessively large sample sizes may exceed rig capacity or prolong testing without proportional gains in insight. Sample availability further constrains strategy design, particularly for prototype or aerospace bearings produced in limited quantities. As one expert noted:

> *"Some customers send us five or ten bearings and expect a*
> *statistical conclusion. That is not realistic."*
>
> - Global Testing Technical Expert, see Appendix B.

Sample size selection is therefore typically guided by a combination of statistical considerations and practical constraints, often supported by planning projections such as those provided by SET2.

A2. Test Execution

This chapter outlines the experimental requirements necessary for executing life tests in a controlled, stable and reproducible manner. While test samples, test conditions and test strategies define what is being tested and how the test is designed, test execution defines the experimental requirements under which the test is performed. These requirements establish minimum laboratory standards needed to ensure that the generated life test data is valid, comparable and not influenced by external disturbances, in line with established principles for competent laboratory testing (ISO/IEC 17025).

Before test execution begins, a formal test request is completed and recorded in the internal test management system, ensuring that all required information is documented and approved. In addition, resources, timelines and responsibilities are defined and confirmed to ensure that the test can be executed as planned. Any required tooling or equipment modifications are identified and procured prior to test start.

A2.1 Cleanliness and Laboratory Standards

Cleanliness is one of the most critical experimental requirements in bearing endurance testing. Even small contaminants, such as dust, metal particles or residues from previous tests, can disturb lubrication conditions, trigger surface distress, or introduce failure mechanisms unrelated to fatigue. As one expert explained:

> *"If you have a metal particle or a sand particle in the oil, it will generate damage on the raceway, and then the bearing will fail with a mechanism that has nothing to do with the material performance."*
>
> - Global Testing Technical Expert, see Appendix B.

To minimize such risks, mounting areas must be free from debris, tools and fixtures must be cleaned before use, and bearings must be handled with gloves or dedicated tools. Lubrication units should be flushed before starting new test series to remove any residual contaminants. These practices are consistent with the requirement to control environmental conditions and prevent contamination of test objects in competent laboratory activities according to ISO/IEC 17025.

Strict cleanliness enhances the reproducibility of results and ensures that contamination does not influence test outcomes. These cleanliness and handling requirements must be fulfilled prior to and throughout test execution.

A2.2 Calibration and Measurement Uncertainty

Accurate measurement of load, speed, temperature and lubrication flow is essential for interpreting life test results. Calibration ensures that sensors operate within their specified tolerances and that the recorded values accurately reflect the true operating conditions. Calibration must therefore be performed regularly and documented for traceability. An expert highlighted the resources used to ensure high quality calibration:

"Calibration can be almost 20% of our yearly budget.
That's how important it is."

- Global Testing Quality Expert, see Appendix B.

This is further emphasized by another expert:

"If the sensor is not calibrated, you simply cannot trust the results."

- Global Testing Quality Expert, see Appendix B.

In addition to calibration, measurement uncertainty must be evaluated to understand the reliability of the recorded data. This is particularly important when comparing results across sites or over long periods of historical testing. Calibration activities are supported by documented calibration certificates, which are maintained and traceable to relevant standards. These certificates provide evidence of measurement validity and are essential for ensuring compliance with laboratory quality requirements. The evaluation and documentation of measurement uncertainty is a fundamental part of ensuring result validity in laboratory testing, as required by ISO/IEC 17025. Understanding calibration status and uncertainty is therefore essential for producing life test results that are credible and statistically meaningful.

In practice, calibration is either performed internally or using accredited external laboratories, where calibration certificates include both measured values and associated uncertainty across defined operating ranges. These certificates provide traceable reference values and represent the baseline for measurement reliability within the laboratory. At the same time, internal calibration capabilities may also be applied where available, using dedicated reference equipment and standardized procedures. As highlighted during expert discussions, calibration is not only a formal requirement but a fundamental prerequisite for reliable testing:

"Calibration is the first activity that can ensure that the results are reliable."
- Global Testing Quality Expert, see Appendix B.

In addition, calibration results must always be interpreted together with their associated uncertainty, as even relatively small deviations can influence the validity of test conditions, particularly when operating within narrow tolerance ranges.

Calibration is typically performed across multiple reference points within the operating range of the sensor. For example, when calibrating temperature, measurements may be taken at different setpoints such as 50°C, 70°C, 90°C and 110°C. The recorded values may then deviate from the reference, such as 50.5°C, 70.8°C, 91.2°C and 111.6°C respectively. This illustrates how measurement deviation and uncertainty can vary across the range, often increasing at higher temperatures, which must be considered when evaluating test conditions and tolerances.

The interviews further demonstrated how uncertainty is evaluated in practice, where measured values are compared against reference instruments, and contributions such as repeatability, deviation and reference uncertainty are included in the final estimation. This highlights that calibration is not a single isolated activity, but part of a broader measurement framework where both equipment performance and uncertainty evaluation directly affect the credibility and comparability of life test results across sites.

A2.3 Stability of Operating Conditions

Life tests must run under stable operating conditions to ensure that observed failures reflect true rolling-contact fatigue rather than unintended variations in the test environment. Stability is continuously monitored through defined measurement channels, including temperature, vibration levels, lubrication flow rate, torque and power consumption. These monitoring channels must be active, validated and functioning correctly throughout the test execution.

Stability deviations may indicate misalignment, lubrication starvation, rig malfunction or unexpected contamination. Standardized stability thresholds help determine when a test can continue, when intervention is required, or when the test must be restarted entirely. Continuous monitoring of operating conditions is essential to ensure result validity and reproducibility, in accordance with the requirement to control and document testing conditions. As one expert explained:

> *"You never know what will happen until it happens,*
> *that's why we monitor everything and react quickly."*
>
> - Global Testing Quality Expert, see Appendix B.

In addition, procedures for handling special cases and unexpected disruptions, such as power loss or equipment failure, must be defined, documented and readily available prior to test execution. These procedures ensure consistent handling of interruptions and minimize the risk of invalid or non-comparable results.

A2.4 Criteria for Failure Detection

Failure detection methods must be objective, reproducible and consistent to ensure valid comparison. Common failure indicators include predefined vibration thresholds, sudden temperature increases, torque spikes or abnormal rig behaviour. These thresholds must be defined and validated prior to testing to ensure consistent and reproducible failure detection. Using standardized methods to detect failure ensures that the recorded failure times are comparable and do not depend on operator interpretation. Objective and validated detection criteria are required to maintain statistical consistency and avoid subjective bias, which is consistent with the requirement for validated test methods and objective evaluation of results in laboratory testing. As one expert summarized the importance of objectivity:

> *"Failure detection must be objective, otherwise your*
> *life results become meaningless."*
>
> - Global Testing Technical Expert, see Appendix B.

A2.5 Data Recording and Traceability

During test execution, all relevant operational data are continuously recorded and stored. This includes raw life data, failure times and any censored or suspended test results. In addition, data from monitoring channels such as temperature, vibration, torque and lubrication flow are logged to provide a complete record of test conditions throughout the test duration. Comprehensive data recording ensures full traceability of test outcomes and enables consistent statistical analysis across test series. It also allows test results to be reviewed, compared and verified over time, which is essential for long-term data integrity and knowledge development within the organization.

A3. Test Analysis and Documentation

Following test execution, the life testing process proceeds to analysis and documentation, where the outcome of the test is evaluated and transformed into structured and reusable knowledge. This phase focuses on understanding the observed bearing behaviour, identifying underlying failure mechanisms and ensuring that results are documented in a transparent and traceable manner. Through a combination of technical analysis and structured reporting, this phase enables consistent interpretation of test outcomes, supports comparison across test campaigns and ensures that the results can be effectively communicated and reused in future engineering work.

A3.1 Damage Mechanism Analysis

Failure analysis provides insight into the mechanisms that cause bearing failure and is a critical step in interpreting bearing life test results. While experimental requirements define how a bearing test is conducted, failure analysis explains why the bearing reached the end of its life. By linking observed damage to fatigue processes, lubrication conditions, material behavior, and operational disturbances, failure analysis ensures that test outcomes can be interpreted in a meaningful and comparable way.

A3.1.1 Failure Modes and Classification (ISO 15243)

Bearing failures observed during life testing are primarily associated with rolling-contact fatigue, but may also be influenced or accelerated by corrosion, lubrication, contamination, misalignment, or electrical effects. To ensure consistent interpretation of such damage, ISO 15243 provides a standardized taxonomy for bearing failure modes. Within ISO 15243, fatigue failures are broadly divided into subsurface-initiated and surface-initiated mechanisms. Subsurface-initiated fatigue originates below the raceway surface where cracks initiate at material weak points and propagate toward the surface, eventually producing spalling. This failure mechanism is generally associated with adequate lubrication and minimal surface distress.

Surface-initiated fatigue, in contrast, begins at or near the surface. Typical manifestations include micropitting, grey staining, and micro-spalls, which may evolve into larger damage areas with continued operation. The stress and kinematic environment strongly influences which failure mode dominates. Beyond fatigue, ISO 15243 also defines additional categories such as wear, corrosion, electrical erosion, plastic deformation, and fracture. These mechanisms may act as primary failure modes or as contributing factors that alter local stress or lubrication conditions and thereby accelerate fatigue. Using ISO 15243 as the conceptual framework ensures that observed damage is interpreted consistently and that failure modes are discussed using a common technical language.

A3.1.2 Failure Modes Using Microscopy

Microscopy is the primary tool used to identify and distinguish the failure modes defined in ISO 15243. High-resolution microscopy enables detailed examination of surface damage, crack initiation sites, failure origin locations, subsurface crack propagation paths, and morphological features linked to specific stress and lubrication conditions. The SKF Bearing Damage and Failure Analysis guide emphasizes that microscopic inspection is essential for separating surface-initiated fatigue features, such as micropitting, grey staining, and surface distress, from subsurface-initiated fatigue, where cracks propagate upward toward the raceway.

Microscopy also allows interpretation of raceway path characteristics. Under normal loading and alignment, raceways exhibit symmetric and stable running tracks. Deviations such as skewed, uneven, or discontinuous tracks may indicate abnormal loading, misalignment, or dynamic instability, which correlate with local stress concentrations. Experts empathized the importance of interpreting microscopy findings in the context of internal contact mechanics:

> *"Understanding the contact parameters helps you make sense of what you see after the test."*
>
> – Global Testing Technical Expert, see Appendix B.

A3.1.3 Documentation of Failure Observations

Once failure modes and failure origins have been identified through microscopy, systematic documentation is essential to ensure traceability, repeatability and comparability across test series and laboratories. Microscopy-based documentation involves capturing high-resolution images of relevant damage features such as spalling, micro-cracks and surface distress, together with clear identification and recording of the failure origin and crack initiation location. All imaging and associated observations are stored within internal documentation systems to ensure traceability and future reference. The documented damage morphology is then compared with characteristic features associated with specific ISO 15243 failure modes to support consistent and reproducible classification.

In addition to classification, all observed damage is evaluated in relation to expected rolling-contact fatigue behaviour based on defined test conditions and modelling results. Comparison with SimPro modelling outputs is performed to assess whether the observed failure mechanisms correspond to predicted stress distributions and contact conditions. If deviations are observed, potential contributing factors such as lubrication issues, contamination, misalignment or operational disturbances are identified, documented, and stored. This ensures that alternative failure drivers are considered, and that observed results are not incorrectly attributed solely to fatigue behaviour.

Consistent and thorough documentation enables historical failure cases to be revisited when new bearing designs are evaluated or when recurring damage patterns emerge across multiple tests. It also provides a technical basis for correlating observed damage with test conditions, contact mechanics, and fatigue life results, thereby strengthening the overall interpretation of bearing life test outcomes.

A3.2 Reporting of Test Results

Reporting and data management form the final core component of the life testing procedure. While earlier stages ensure correct samples, controlled test conditions and robust strategies, this stage ensures that the resulting data is traceable, interpretable and comparable over time. A consistent reporting structure combined with reliable data-management practices supports long-term knowledge retention and enables engineers to evaluate test outcomes across sites.

A3.2.1 Core Elements of a Life Testing Report

A complete and approved life testing report must document the full testing context in a transparent and traceable manner, such that the test can be understood, evaluated and repeated at a later stage. The report is prepared according to standardized templates to ensure consistency across tests and laboratories.

The report begins by establishing the project background, including the relevant business division, application or use case, bearing type, project identification and the specific motivation for performing the test. The overall objectives of the test must be clearly stated, providing a reference against which results and conclusions can be evaluated. The report must describe the tested sample population in sufficient detail to ensure full traceability, including bearing type, batch information, manufacturing site, number of bearings and selection criteria. Storage conditions and references to applicable procedures are also included.

Material and geometry details are documented, including material grades, heat-treatment conditions and references to supporting drawings and metallurgical investigations where applicable. The test description section outlines how the test was executed, including test equipment, operating conditions (loads, temperature, lubrication, flow rate), and all relevant parameters. Contact-parameter calculations and modelling results are reported together with SimPro versions and settings used. Operational logs are included in the report to document actual test execution and operating behaviour over time.

Test results are presented through structured documentation of failure and life data. This includes tables identifying each bearing, achieved revolutions, failure or suspension status and failure mode classification according to ISO 15243. Statistical analysis is performed and reported using standardized methods, including Weibull analysis, characteristic life, L10 values, Weibull slope (β), confidence intervals and corresponding plots.

Interpretation and discussion sections highlight key findings, including comparison between observed and predicted results, analysis of failure distributions and evaluation of confidence levels. The conclusions section summarizes the outcome of the test in relation to its objectives and includes identified learnings and recommendations for future tests. The report must clearly state whether predictions were confirmed and explain any deviations. Final technical conclusions are explicitly stated. Test results are formally communicated to the customer as part of the reporting process, ensuring that findings are understood and can support decision-making.

Finally, the report is reviewed, approved and published in relevant internal systems, such as the centralized data platform or internal documentation repositories, ensuring that it is accessible for future reference. Supporting documentation such as metallurgical reports, previous test reports and microscopy images are referenced and included where relevant.

A3.2.2 Data Management and Traceability

In addition to the written test report, effective data management is required to preserve the full value of life testing activities. Reporting alone is insufficient if underlying data and supporting documentation are not stored, traceable and accessible over time.

All relevant data, including raw life data, operational logs, calibration certificates, modelling inputs and analysis files, are archived together with the corresponding test report. Data is stored using standardized naming conventions, metadata structures and access control mechanisms to ensure that it can be retrieved and correctly interpreted. Access to stored data is controlled to prevent unauthorized modification while still allowing engineers to review and analyze historical results. Changes to approved reports or datasets must be clearly documented and traceable.

Once the report and associated data have been finalized and approved, the corresponding test request in the test management system is formally reviewed, approved and closed. As part of the post-test process, a satisfaction survey is issued to relevant stakeholders to capture feedback on the test execution, reporting quality and overall outcome for continuous improvement purposes.

Long-term traceability enables engineers to link test results back to conditions, assumptions and execution details, supporting comparisons across tests and evaluation of new designs against historical data. By combining standardized reporting with disciplined data-management practices, the life testing organizations of SKF can ensure that the investment associated with life testing continues to generate value beyond the original test. Effective data management transforms individual tests into long-term knowledge assets that support informed engineering decision-making.

Appendix B: Expert Interview Quotes

This appendix presents selected quotes from the expert interviews conducted as part of the study. The purpose is to support and illustrate the findings presented throughout the report, particularly those discussed in Chapter 5. The quotes included are not complete transcripts, but representative examples selected to reflect recurring patterns, challenges, and practitioner perspectives identified during the analysis. The quotes are grouped according to key areas corresponding to the life testing process and the main findings of this study.

B1. Test Motivation and Purpose

The interviews highlight recurring concerns regarding insufficiently defined test purposes and lack of alignment between stakeholders. Several experts emphasized the risk of conducting tests without a clear objective, which may lead to misleading conclusions and inefficient use of resources. These findings are reflected in Section 5.1, where lack of structure and shared understanding is identified as a key challenge.

"We need to prevent that a lab runs a test with six bearings for three hours and then claims the life is OK."
– Global Testing Technical Expert

"The problem is not that people don't know how to run tests. The problem is that if you move from one site to another, there is no common reference for what has actually been done and why."
– Global Testing Technical Expert

"A lot of things are done almost automatically by experienced people, but they are not written down anywhere. If that person is not available, the knowledge is simply not visible."
– Global Testing Technical Expert

B2. Sample Selection and Traceability

A strong emphasis was placed on traceability and representativeness of test samples. Experts highlighted how insufficient documentation or biased sampling can significantly affect the interpretation of test results. These insights support the importance of standardized documentation and traceability discussed in Chapter 4.

"Metadata to me is the identity card of the bearing."
– Global Testing Technical Expert

"The correct way to state it is to have samples which are representative of standard production. Avoid cherry-picking."
– Global Testing Technical Expert

"If material differences are not understood, you risk attributing failures to fatigue when they come from manufacturing."
– Global Testing Technical Expert

B3. Test Conditions and Modelling

The interviews consistently emphasized that defining correct test conditions is critical for obtaining meaningful results. Experts stressed that small deviations in modelling assumptions or operating parameters can significantly affect test outcomes. These findings are directly linked to the modelling practices and operating condition definitions described in Appendix A.

"If the contact pressure is extremely high, the bearing cannot even start. If it is too low, the test might run for years."
– Global Testing Technical Expert

"If you miss even one modelling step, you will get a different result and then you will question everything."
– Global Testing Technical Expert

"Understanding the contact parameters helps you make sense of what you see after the test."
– Global Testing Technical Expert

"You need to ensure that the lubrication system can actually handle the load, otherwise the test is not meaningful."
– Global Testing Technical Expert

B4. Test Strategy and Statistical Considerations

The importance of well-designed test strategies was a recurring topic throughout the interviews. Experts highlighted the trade-offs between precision, test duration, and sample size, as well as the challenges associated with accelerated testing. These insights support the discussion on strategy development in Chapter 4.

"The more bearings you test, the more failures you obtain, the more precise your results become."
– Global Testing Technical Expert

"You never know what will happen until it happens, that's why we make projections but must always be ready to adjust."
– Global Testing Technical Expert

"Some customers send us five or ten bearings and expect a statistical conclusion. That is not realistic."
– Global Testing Technical Expert

"If you introduce contamination, you are essentially hammering the bearing from the start. They will not survive too long."
– Global Testing Technical Expert

"The definition of bearing life is the appearance of the first evidence of fatigue."
– Global Testing Technical Expert

B5. Test Execution and Laboratory Practices

B5.1 Cleanliness and Handling

Cleanliness was highlighted as a critical factor influencing test validity. Experts emphasized that external contamination can introduce failure mechanisms unrelated to material performance, thereby invalidating results.

"If you have a metal particle or a sand particle in the oil, it will generate damage on the raceway, and then the bearing will fail with a mechanism that has nothing to do with the material performance."
– Global Testing Technical Expert

B5.2 Calibration and Measurement Reliability

Calibration and measurement accuracy were identified as fundamental prerequisites for reliable testing. Experts stressed both the importance and the cost associated with maintaining accurate measurement systems.

"Calibration can be almost 20% of our yearly budget. That's how important it is."
– Global Testing Quality Expert

"If the sensor is not calibrated, you simply cannot trust the results."
– Global Testing Quality Expert

"Calibration is the first activity that can ensure that the results are reliable."
– Global Testing Quality Expert

B5.3 Monitoring and Stability

Continuous monitoring of operating conditions was considered essential to ensure test validity and reproducibility. Experts emphasized the importance of proactive monitoring and rapid response to deviations.

"You never know what will happen until it happens, that's why we monitor everything and react quickly."
– Global Testing Quality Expert

B6. Failure Detection and Analysis

The interviews highlighted the importance of objective and standardized failure detection methods to ensure consistent and comparable results. This aligns with the requirements for validated methods discussed in Chapter 4.

"Failure detection must be objective, otherwise your life results become meaningless."
– Global Testing Technical Expert

"Understanding the contact parameters helps you make sense of what you see after the test."
– Global Testing Technical Expert

B7. Data Management, Knowledge Reuse, and Standardization

A central theme in the interviews was the challenge of knowledge reuse and lack of shared structure across sites. Experts emphasized that valuable knowledge exists but is not always captured or made accessible.

"When a new flow starts, you start from what was built in the past."
– Global Testing Technical Expert

"For traceability reasons and for future comparability purposes, the simulations are run using standardized methodology, so that all labs produce the same output."
– Global Testing Quality Expert

"Maturity should not be a ranking. It should help you understand where you are strong and where you need support."
– Global Testing Quality Expert

B8. Calibration and Measurement Uncertainty

Discussions related to calibration and measurement uncertainty revealed both methodological challenges and practical constraints. Experts highlighted that uncertainty information must be readily available and based on reliable, accredited sources.

"Available on request should not mean we start the work when the customer asks – it should already be available."
– Global Testing Quality Expert

"We have to rely on accredited laboratories, because there are audits and controls ensuring the validity of the results."
– Global Testing Quality Expert

"The uncertainty increases with temperature and operating conditions – the environment always matters."
– Global Testing Quality Expert

B9. Practical Constraints and Implementation Considerations

In addition to technical considerations, experts also emphasized practical aspects of implementation, including time constraints, complexity, and the need to balance detail with usability. These insights informed the development of the conceptual framework and its level of detail.

"Some parts need to stay at a guideline level rather than strict rules, especially given the differences between sites."
– Global Testing Quality Expert

"You need to balance how much detail you include – otherwise it becomes too complex to implement."
– Global Testing Quality Expert

Appendix C – Concept 1

This appendix presents the process mapping developed for Concept 1. The figures illustrate the overall life testing workflow and its main subprocesses, providing a structured overview of the testing process from motivation to reporting. In addition, the appendix includes the identified deliverables, presented as a structured checklist to support consistency and practical implementation of the process.

C.1 Process Mapping

Figure C.1: Complete workflow of *life testing* for Concept 1.

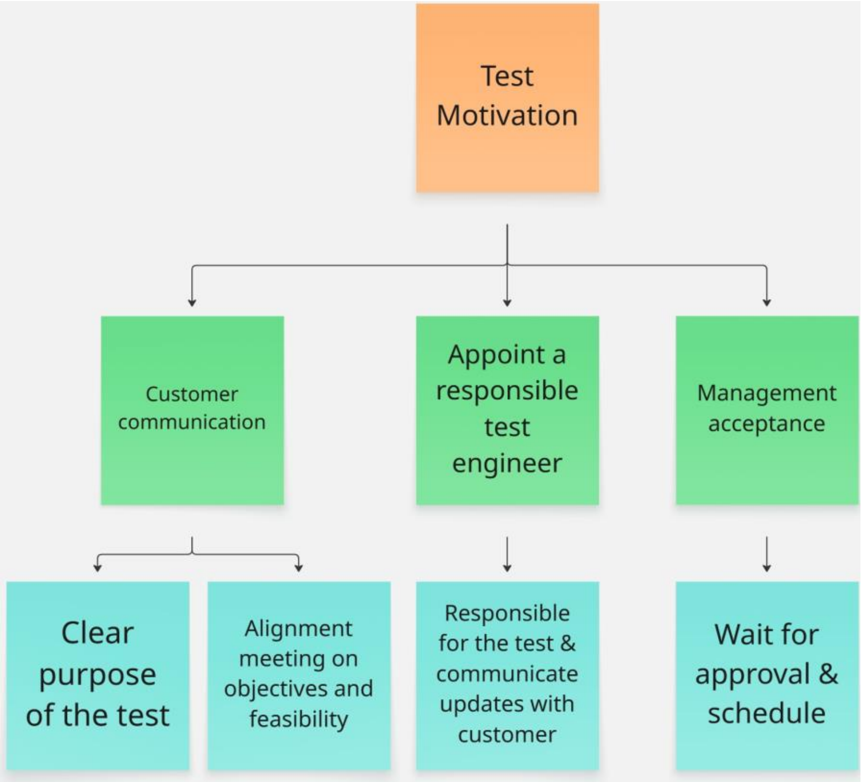


Figure C.2: Complete workflow of *Test Motivation* for Concept 1.

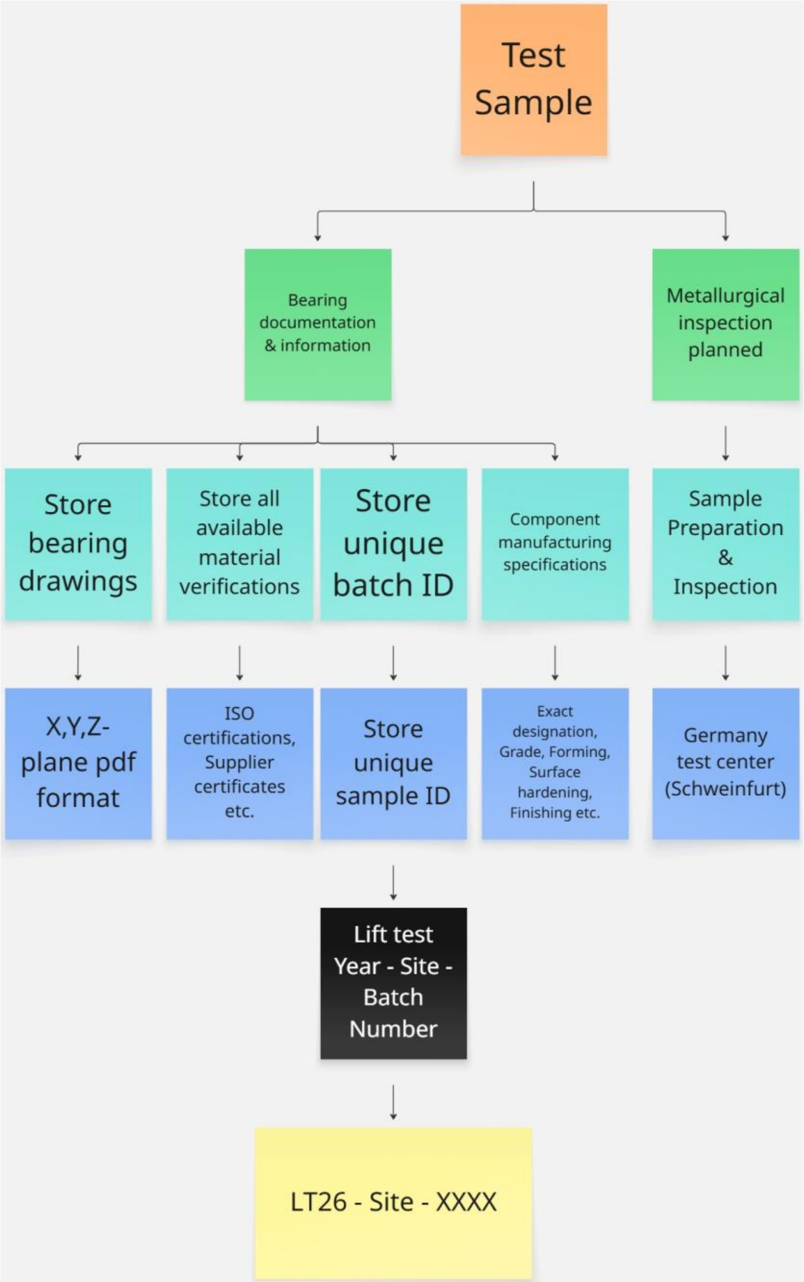


Figure C.3: Complete workflow of *Test Sample* for Concept 1.

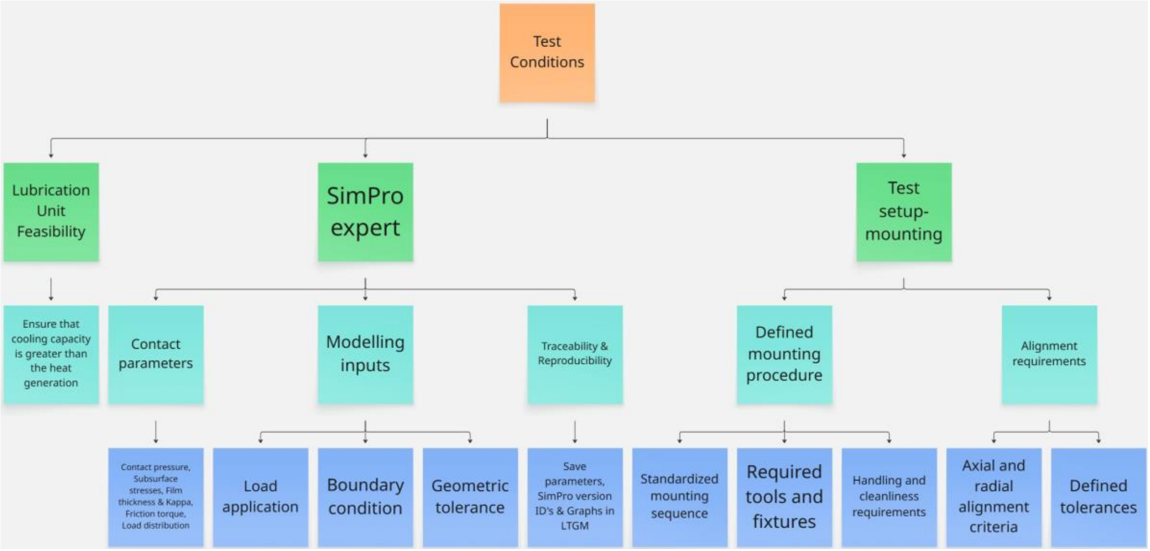


Figure C.4: Complete workflow of *Test Conditions* for Concept 1.

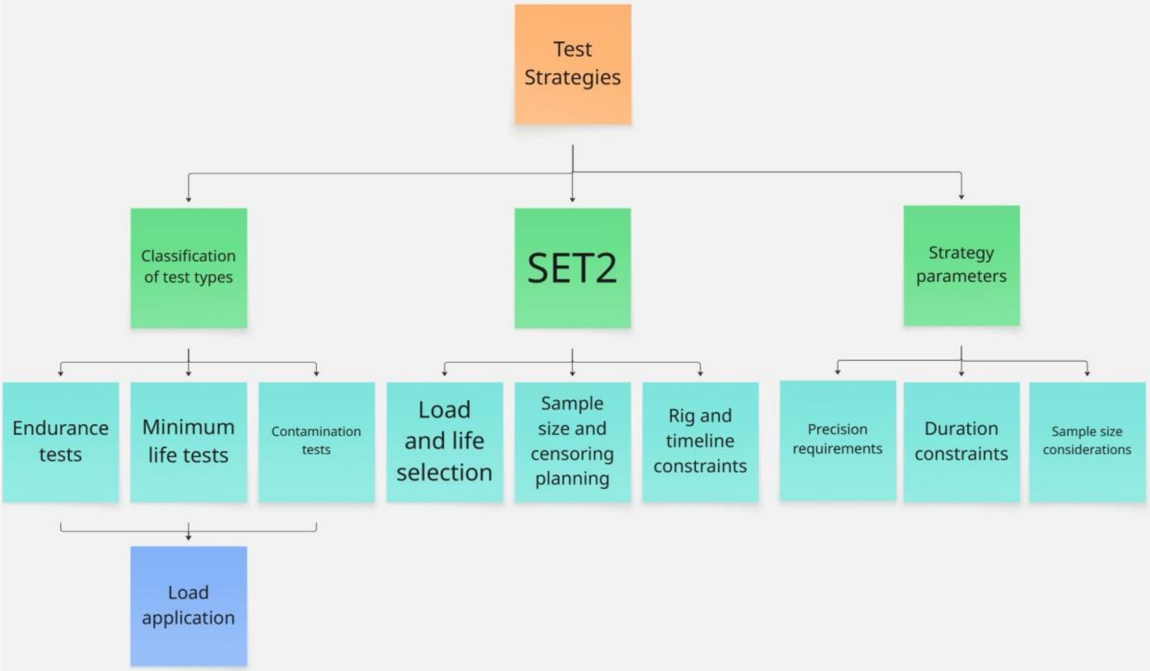


Figure C.5: Complete workflow of *Test Strategies* for Concept 1.

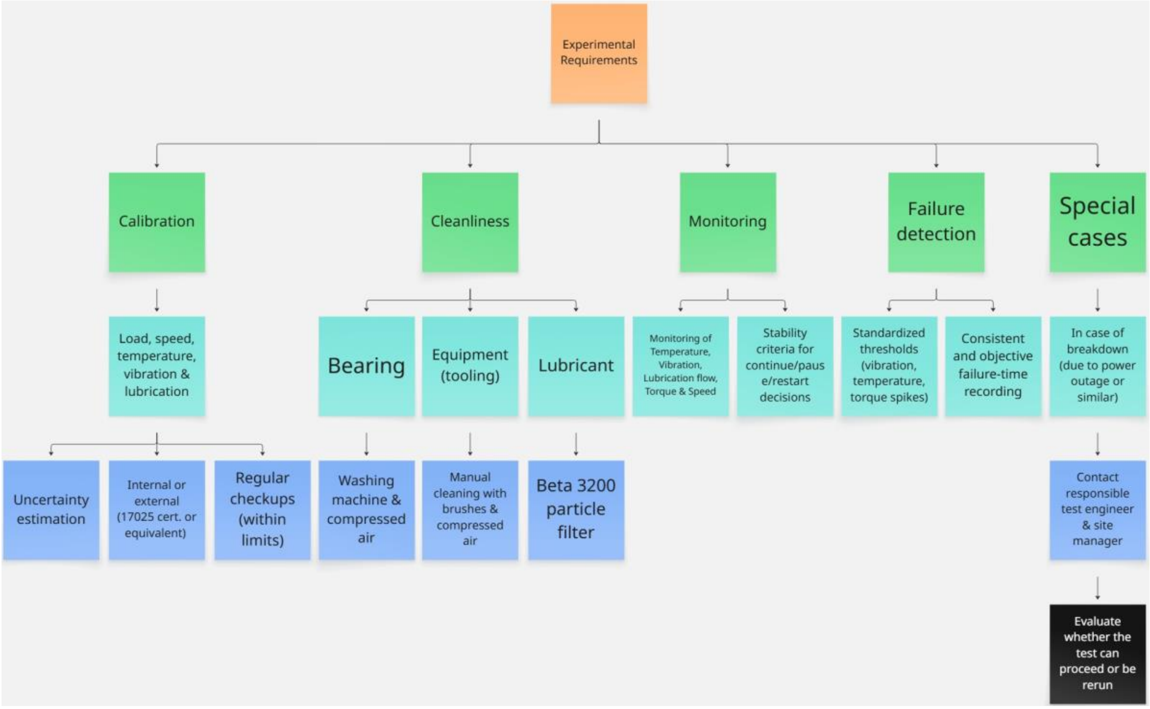


Figure C.6: Complete workflow of *Experimental Requirements* for Concept 1.

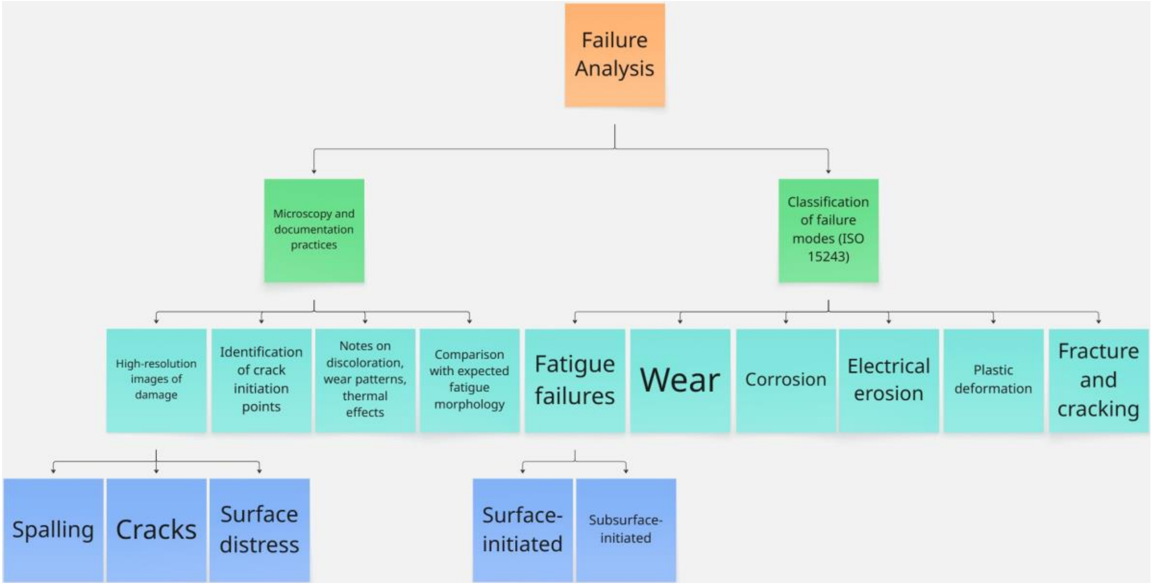


Figure C.7: Complete workflow of *Failure Analysis* for Concept 1.

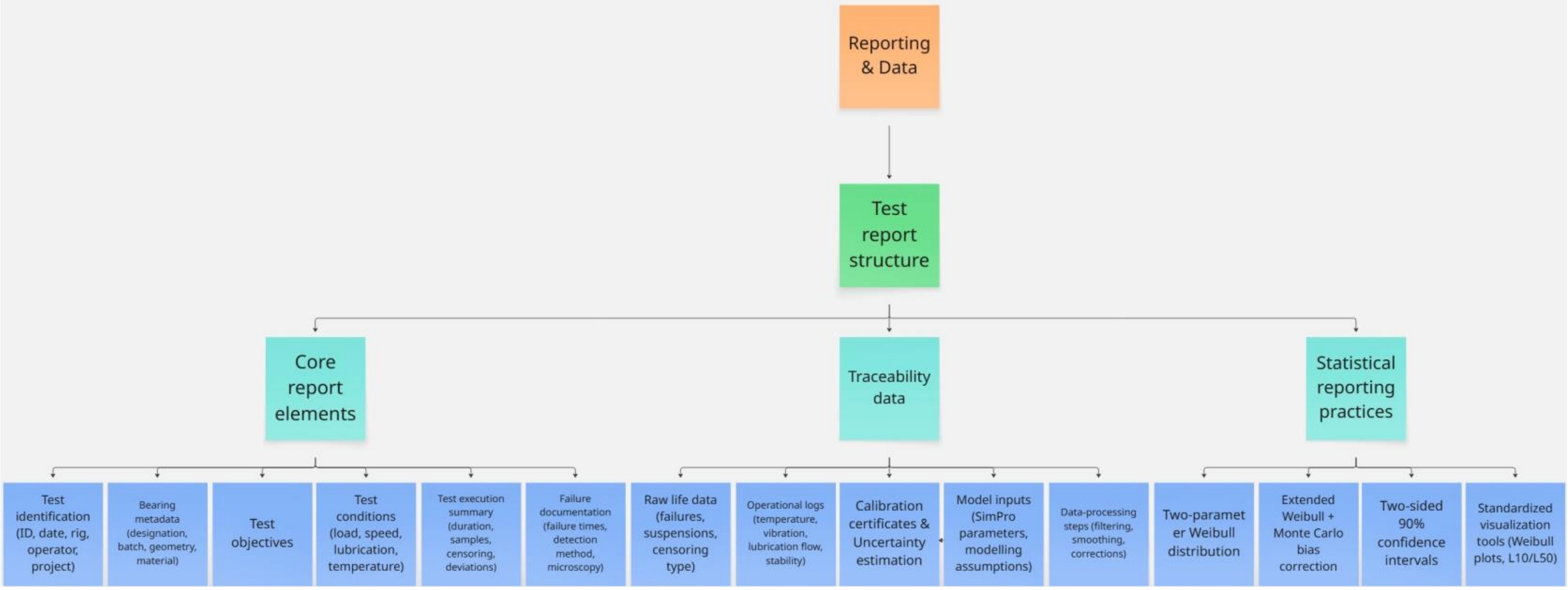


Figure C.8: First part of the workflow of *Reporting & Data* for Concept 1.

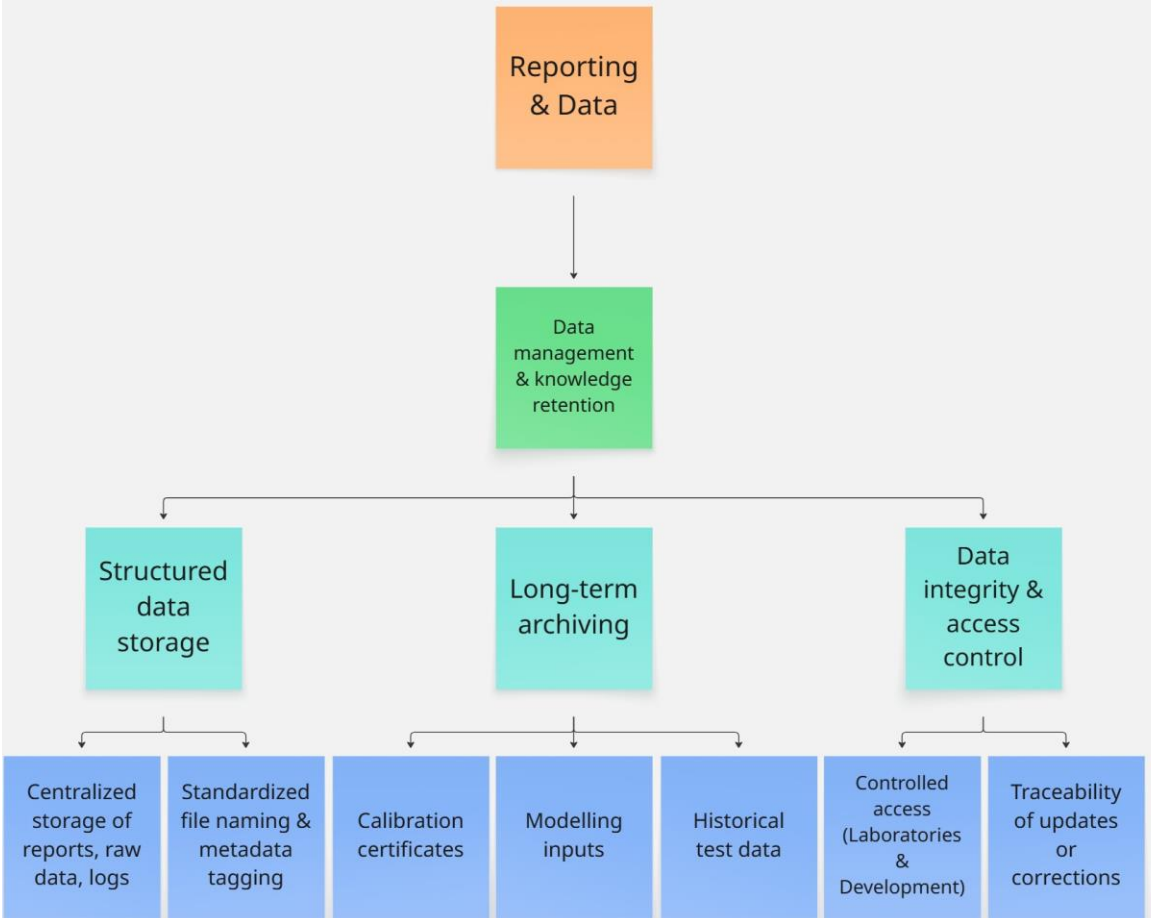


Figure C.9: Second part of the workflow of *Reporting & Data* for Concept 1.

C.2 Deliverables

☐ Testing Preparations

Test Motivation

- ☐ **The test has a clear purpose**
    - ☐ **Customer communication is completed**
        - ☐ Customer need / reason for testing
        - ☐ Expected application or decision supported by the test
        - ☐ Customer understanding of limitations and scope
    - ☐ **Alignment meeting is held**
        - ☐ Test objectives
        - ☐ Feasibility assessment
        - ☐ Scope and boundaries of the test
        - ☐ Agreement on what constitutes “success” or “value”
    - ☐ **A responsible test engineer is appointed**
        - ☐ Overall ownership of the test
        - ☐ Communication with customer and internal stakeholders
        - ☐ Ensuring that requirements and constraints are respected
    - ☐ **Management acceptance is obtained**
        - ☐ The test purpose
        - ☐ The resource allocation
        - ☐ The scheduling in the test plan
        - ☐ Alignment with strategic priorities
    - ☐ **Initial understanding of test progression is documented**
        - ☐ How the test is expected to develop over time
        - ☐ Critical points, risks, and uncertainties
        - ☐ Expected complexity and feasibility issues

Test Sample

- ☐ **The test sample is fully validated and traceable**
    - ☐ Bearing documentation is complete
        - ☐ Bearing drawings (X, Y, Z-plane in PDF format)
        - ☐ Material verification documents
        - ☐ ISO certifications and supplier certificates
        - ☐ Technical metadata relevant for the bearing design
    - ☐ **Unique identification is established**
        - ☐ Unique batch ID
        - ☐ Unique sample ID
        - ☐ Lift Test ID format (e.g., LT26 – Site – XXXX)

    - ☐ Correct linkage to storage, documentation, and future failure analysis
  - ☐ **Component manufacturing specifications are documented**
    - ☐ Material grade
    - ☐ Forming process (e.g., forging, cold forming)
    - ☐ Heat treatment or surface hardening process
    - ☐ Finishing and grinding processes
    - ☐ Dimensional tolerances relevant to performance
  - ☐ **Metallurgical inspection is planned and scheduled**
    - ☐ Scheduling metallurgical analysis (before or after test depending on protocol)
    - ☐ Assigning responsible examiner
    - ☐ Defining inspection scope (microstructure, inclusions, hardness, etc.)
  - ☐ **Sample preparation and inspection are arranged**
    - ☐ Confirmed preparation at Germany Test Center (Schweinfurt)
    - ☐ Verification that cleaning, marking, and initial measurements comply with internal standard

Test Conditions

- ☐ **<u>The test conditions are feasible, well-defined, and fully reproducible.</u>**
  - ☐ **Lubrication capacity is verified**
    - ☐ Cooling capacity > expected heat generation
    - ☐ Verified lubrication stability (flow rate, pressure, viscosity)
    - ☐ Confirmation that the lubrication unit can run the test duration without risk
    - ☐ Documentation that the system configuration is compatible with the test type
  - ☐ **SimPro modelling is completed and all inputs/outputs are archived**
    - ☐ SimPro simulation completed
      - ☐ Contact pressure
      - ☐ Subsurface stress
      - ☐ Film thickness
      - ☐ Kappa
      - ☐ Friction torque
      - ☐ Load distribution
    - ☐ Full traceability of modelling data
      - ☐ Exact modelling inputs
      - ☐ Boundary conditions used
      - ☐ Geometric tolerances
      - ☐ SimPro version
      - ☐ All result graphs, IDs, and centralized data platform uploads

- ☐ **Test setup & mounting procedure is following standards and are reproducible**
  - ☐ Follow standards of mounting procedure
    - ☐ Detailed mounting sequence
    - ☐ Required tools, fixtures, torques
  - ☐ Handling requirements
    - ☐ Handling protocols to avoid contamination or misalignment
  - ☐ Alignment tolerances defined
    - ☐ Axial alignment acceptance limits
    - ☐ Radial alignment limits
    - ☐ Clear documentation of allowable deviations
  - ☐ Traceability
    - ☐ All setup steps documented in a way allowing future repetition

Test Strategies

- ☐ **A finalized and justified test strategy is delivered.**
  - ☐ **Classification of test types**
    - ☐ Endurance test
    - ☐ Minimum life test
    - ☐ Tests under contamination
  - ☐ **SET2 simulation is performed and used for strategic design**
    - ☐ Predicting bearing life for proposed loads
    - ☐ Estimating expected failure distribution
    - ☐ Predicting censoring levels (run-outs)
    - ☐ Determining feasible test duration
    - ☐ Supporting selection of sample size
    - ☐ Evaluating precision of the planned test
  - **A final test strategy is produced by integrating:**
    - ☐ The test taxonomy classification
    - ☐ SET2 simulation results
    - ☐ Insights from previous tests
  - **The final document must explain how decisions were shaped by:**
    - ☐ historical learnings (previous rigs, failures, anomalies)
    - ☐ expected sample size needs
    - ☐ required precision (e.g., 90% CI)
    - ☐ predicted censoring
    - ☐ acceptable duration vs precision tradeoff

When all Testing Preparation points are fulfilled:

- ☐ Review the application

- ☐ Align all parts
- ☐ Get an agreement signed from all parties
- ☐ **Submit a Formal Test Request in the lab process management system**

☐ Test Execution

Experimental Requirements

- ☐ **The test environment is controlled, validated, and continuously monitored.**
  - ☐ **All measurement systems are calibrated and traceable**
    - ☐ Calibration certificates are available and stored
      - ☐ Load
      - ☐ Speed
      - ☐ Temperature
      - ☐ Vibration
      - ☐ Lubrication
    - ☐ Uncertainty estimation is documented
      - ☐ Whether calibration is internal or external (e.g., ISO17025 or equivalent)
      - ☐ Confirmed that uncertainty lies within allowed limits
    - ☐ Regular calibration checkups scheduled
      - ☐ Pre-test verification
      - ☐ Planned intermediate checks if required
      - ☐ Accepted tolerance windows
  - ☐ **Cleanliness and handling procedures are followed**
    - ☐ Bearing, equipment & tooling cleanliness
      - ☐ Bearing is cleaned with washing machine & compressed air and with verification of cleanliness level before assembly
      - ☐ Tooling is cleaned manually with brushes and compressed air
      - ☐ Lubricant filtered with Beta 3200 particle filter
      - ☐ Verification that lubricant meets required purity
  - ☐ **Monitoring, stability criteria, and failure detection are defined**
    - ☐ Monitoring channels active and validated
      - ☐ Temperature
      - ☐ Vibration
      - ☐ Lubrication flow
      - ☐ Torque
      - ☐ Speed
    - ☐ Standardized failure detection thresholds
      - ☐ Vibration thresholds
      - ☐ Temperature thresholds

        - ☐ Torque spike thresholds
        - ☐ Repeatable interpretation of failure time
    - ☐ Special case procedures (during rig breakdown, power supply failure or unexpected shutdown)
        - ☐ Ensure who to contact (responsible engineer / lab manager)
        - ☐ Evaluate criteria's for whether the test can proceed or cancelled

When all Test Execution points are fulfilled:

- ☐ Disassembly and clean the test rigs
- ☐ Prepare bearing for Failure Analysis
- ☐ Check if the calibrations value is the same as before the test
    - ☐ If yes
        - ☐ No further actions are needed
    - ☐ If no
        - ☐ Contact site manager about calibrations are needed
- ☐ **Move on to Test Analysis**

☐ Test Analysis

Failure Analysis

- ☐ **<u>The failure mode is accurately documented and correctly classified.</u>**
    - ☐ **Microscopy and damage documentation is complete**
        - ☐ High-resolution imaging
            - ☐ Spalling
            - ☐ Cracks
            - ☐ Surface distress
            - ☐ Micro-spalls
            - ☐ Wear features
            - ☐ Thermal discoloration
        - ☐ Identification of initiation points
            - ☐ Likely crack initiation zone
            - ☐ Subsurface vs surface origin
            - ☐ Geometry of failure propagation
            - ☐ Associated stress concentration indicators
        - ☐ Evidence of tribological effects
            - ☐ Wear patterns
            - ☐ Discoloration
            - ☐ Localized heat effects
            - ☐ Debris imprints
            - ☐ False brinelling or fretting

- ☐ **Failure mode is classified according to ISO 15243**
  - ☐ Determination of failure category
    - ☐ Fatigue failures
      - ☐ Surface-initiated fatigue
      - ☐ Subsurface-initiated fatigue
    - ☐ Wear
      - ☐ Abrasive
      - ☐ Adhesive
    - ☐ Corrosion
      - ☐ Moisture corrosion
      - ☐ Fretting corrosion
    - ☐ Electrical erosion
    - ☐ Plastic deformation
      - ☐ Overload
      - ☐ Indentation
    - ☐ Fracture and cracking
  - ☐ Comparison with expected fatigue morphology
    - ☐ Whether failure matched SimPro prediction
    - ☐ Whether observed failure matches known mode for this bearing geometry
    - ☐ Assessment of anomalies or mixed failure modes
  - ☐ Documentation of uncertainty or borderline cases
    - ☐ Provide justification
    - ☐ Provide supporting images
    - ☐ Suggest alternative modes
    - ☐ Document needed follow-up analyses

☐ Reporting & Data

- ☐ **<u>A complete, traceable, and standardized test report is delivered.</u>**
  - ☐ **The test report contains all core elements**
    - ☐ Test identification
      (ID, date, rig, operator, project reference)
    - ☐ Bearing metadata
      (designation, batch, geometry, material)
  - ☐ **Traceability data is documented**
    - ☐ Raw life data
      (failures, suspensions, censoring types)
    - ☐ Operational logs
      (temperature, vibration, lubrication flow, stability conditions)

        - ☐ Calibration certificates
          (traceable back to Gate 5)
        - ☐ Modelling inputs
          (SimPro, SET2, modelling assumptions)
    - ☐ **Data processing steps are explicitly documented**
        - ☐ Filtering
        - ☐ Smoothing
        - ☐ Corrections
        - ☐ Any data trimming or noise handling
    - ☐ **Statistical reporting**
        - ☐ Two-parameter Weibull distribution
        - ☐ Extended Weibull + Monte Carlo bias correction
        - ☐ Two-sided 90% confidence intervals
        - ☐ Standardized visualization tools:
            - ☐ Weibull plots
            - ☐ L10/L50 estimates
    - ☐ **Data management and long-term storage are completed**
        - ☐ Archiving includes:
            - ☐ Report PDF
            - ☐ Raw data
            - ☐ Include improvements for future tests in the conclusion
            - ☐ Operational logs
            - ☐ Calibration certificates
            - ☐ Modelling inputs
            - ☐ Microscopy images
        - ☐ File management includes:
            - ☐ Standardized naming (eg. Tests, batches and samples)
            - ☐ Metadata tagging
            - ☐ Controlled access rules
              (Laboratories & Development)

When all Test Analysis points are fulfilled:

- ☐ Send the test report to the customer
- ☐ Have a dialogue with the customer and get at an agreement that the test is fulfilled and done
- ☐ Send out the satisfaction survey
- ☐ Publish everything in the internal data platform for further use
- ☐ Close test

Appendix D – Concept 2

This appendix presents the guideline developed for Concept 2. It outlines the key aspects to be considered throughout the life testing process, from test preparation to reporting. The guideline aims to support a more structured and consistent approach while maintaining flexibility in how different test sites implement and apply the framework.

D.1 Deliverables

1. Test Preparation

1. Test purpose, intended value, and scope clearly defined
2. Success criteria understood and agreed
3. Responsible test engineer assigned
4. Resources, timeline, and responsibilities confirmed
5. Previous relevant tests reviewed
6. Management approval obtained
7. Expected test risks, and uncertainties identified

2. Test Samples

1. Unique Sample & Test ID
2. Manufacturing and material history documented
3. Drawings, certificates, and technical metadata stored
4. Metallurgical inspection performed if necessary
5. Sample preparation completed per ISO 17025

3. Test Conditions

1. Test conditions technically feasible and realistic
2. Lubrication and cooling capacity verified for full test duration
3. SimPro modelling completed and stored
4. All modelling inputs, assumptions, and software versions stored
5. Test setup, mounting, and alignment limits defined and stored

4. Test Strategy

1. Test type, sample size & duration is selected and justified
2. SET2 simulations completed and stored
3. Confidence level and censoring assumptions defined
4. Trade-off between duration, precision, and resources justified

5. Formal test request completed in lab process management system

5. Test Execution

1. Measurement systems calibrated and traceable
2. Uncertainty assessment completed
3. Cleanliness and handling requirements fulfilled
4. Monitoring channels active and validated
5. Failure detection thresholds defined
6. Special-case procedures defined, updated and available (e.g. power loss)
7. Raw life data and censoring stored
8. Calibration certificates traceable

6. Failure Analysis

1. Failure damage stored with high-resolution imaging
2. Failure origin identified and classified
3. Failure mode classified per ISO 15243
4. Comparison with SimPro modelling
5. Other potential causes of failures stored

7. Reporting

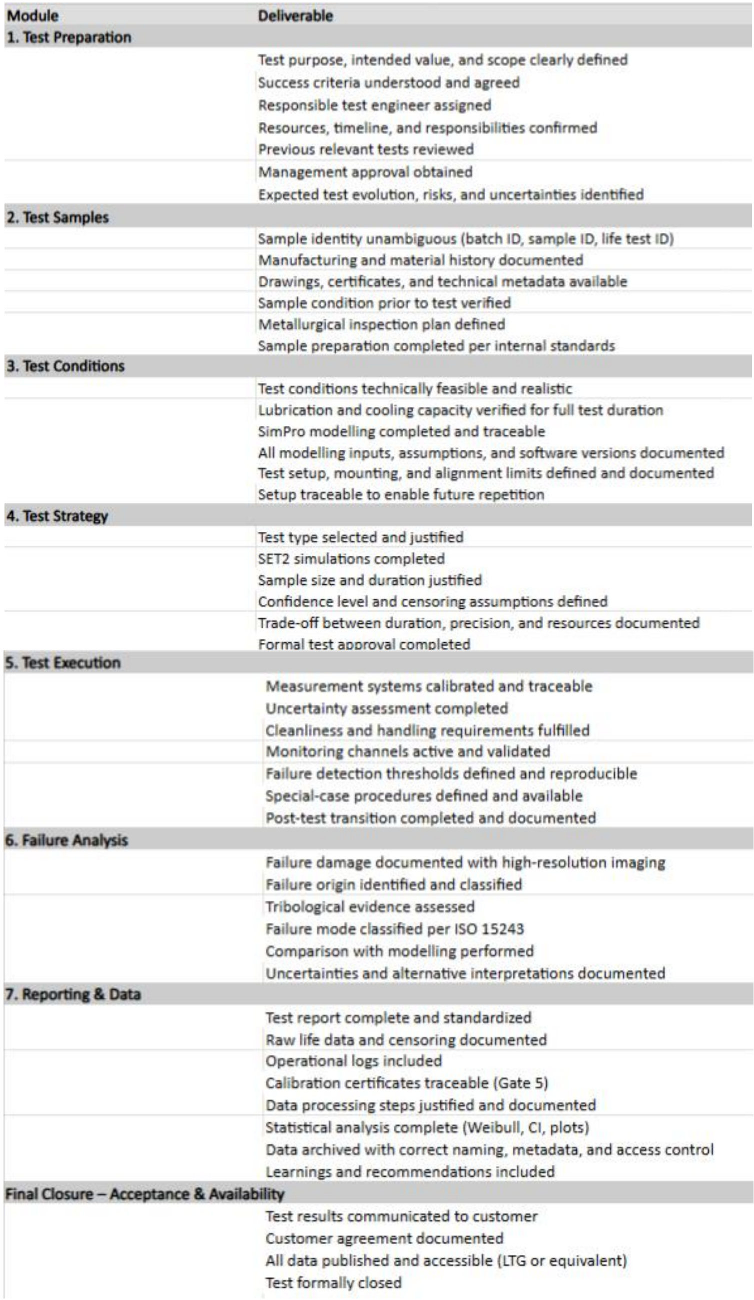
1. Test report complete according to standard template
2. Operational logs included
3. Statistical analysis complete (Weibull, confidence intervals, plots)
4. Data archived with correct naming, metadata, and access control
5. Learnings and recommendations included for future tests
6. Test results communicated to customer
7. Customer agreement stored
8. All data published and correctly accessible (lab process management system, documentation repositories or equivalent)
9. Test formally closed

Appendix E - Concept 3

This appendix presents the guideline and evaluation approach developed for Concept 3. It builds upon the previous concepts by introducing a structured method for assessing maturity and consistency in life testing practices.

E.1 Deliverables – Concept 3

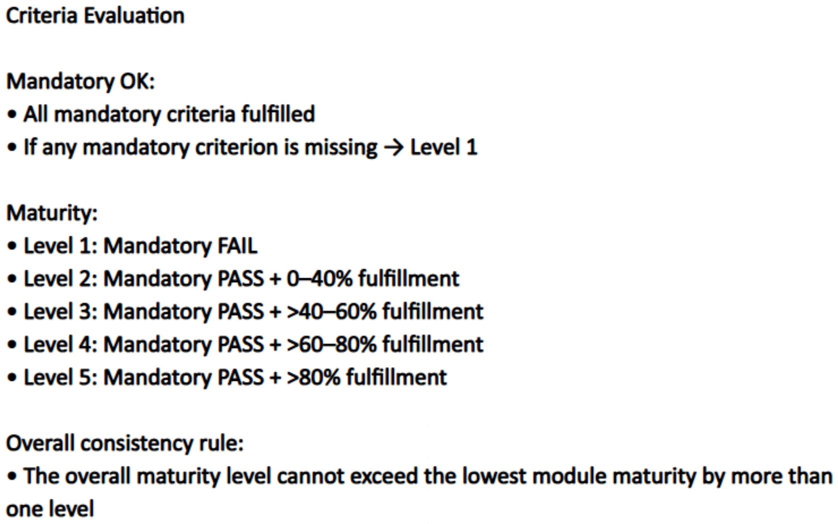
E.1 Structured version of deliverables, this example does not include the mandatory part nor the weighted scale due to confidentiality reasons.



Module	Deliverable
1. Test Preparation	
	Test purpose, intended value, and scope clearly defined
	Success criteria understood and agreed
	Responsible test engineer assigned
	Resources, timeline, and responsibilities confirmed
	Previous relevant tests reviewed
	Management approval obtained
	Expected test evolution, risks, and uncertainties identified
2. Test Samples	
	Sample identity unambiguous (batch ID, sample ID, life test ID)
	Manufacturing and material history documented
	Drawings, certificates, and technical metadata available
	Sample condition prior to test verified
	Metallurgical inspection plan defined
	Sample preparation completed per internal standards
3. Test Conditions	
	Test conditions technically feasible and realistic
	Lubrication and cooling capacity verified for full test duration
	SimPro modelling completed and traceable
	All modelling inputs, assumptions, and software versions documented
	Test setup, mounting, and alignment limits defined and documented
	Setup traceable to enable future repetition
4. Test Strategy	
	Test type selected and justified
	SET2 simulations completed
	Sample size and duration justified
	Confidence level and censoring assumptions defined
	Trade-off between duration, precision, and resources documented
	Formal test approval completed
5. Test Execution	
	Measurement systems calibrated and traceable
	Uncertainty assessment completed
	Cleanliness and handling requirements fulfilled
	Monitoring channels active and validated
	Failure detection thresholds defined and reproducible
	Special-case procedures defined and available
	Post-test transition completed and documented
6. Failure Analysis	
	Failure damage documented with high-resolution imaging
	Failure origin identified and classified
	Tribological evidence assessed
	Failure mode classified per ISO 15243
	Comparison with modelling performed
	Uncertainties and alternative interpretations documented
7. Reporting & Data	
	Test report complete and standardized
	Raw life data and censoring documented
	Operational logs included
	Calibration certificates traceable (Gate 5)
	Data processing steps justified and documented
	Statistical analysis complete (Weibull, CI, plots)
	Data archived with correct naming, metadata, and access control
	Learnings and recommendations included
Final Closure – Acceptance & Availability	
	Test results communicated to customer
	Customer agreement documented
	All data published and accessible (LTG or equivalent)
	Test formally closed

Figure E.1: Shows Concept 3 deliverables.

E.2 Evaluation Criteria for deliverables on test level



Criteria Evaluation

Mandatory OK:

- **All mandatory criteria fulfilled**
- **If any mandatory criterion is missing → Level 1**

Maturity:

- **Level 1: Mandatory FAIL**
- **Level 2: Mandatory PASS + 0–40% fulfillment**
- **Level 3: Mandatory PASS + >40–60% fulfillment**
- **Level 4: Mandatory PASS + >60–80% fulfillment**
- **Level 5: Mandatory PASS + >80% fulfillment**

Overall consistency rule:

- **The overall maturity level cannot exceed the lowest module maturity by more than one level**

Figure E.2: Definition of maturity levels, handling of mandatory criteria, and overall consistency rules applied in the Concept 3 maturity rating aggregation.

DEPARTMENT OF INDUSTRIAL AND MATERIAL SCIENCE
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2026
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY