Appendix A. Operationalization into Action Items

Consistent with the conceptual framework in Section 3, this section describes the practical implementation of measures. The major effort lies in realizing Sections 3.2, 3.3, 3.4, 3.5, and 3.6 — the implementation of the concept per lecture.

The measures described in Section 3 have been translated into tangible action items for each lecturer to ensure implementation. Examples include:

- **ARIVA structuring (Sec. 3.4):** mandatory inclusion of (i) motivational slide ('What we will learn today?' or 'Learning objectives'), (ii) takeaway slide (≤ 3 tangible takeaways aligned to LO-1–LO-4), and (iii) reflective problem slide.
- **Content alignment (Sec. 3.3):** Alignment of content to LO-1–LO-4.
- **Exam redesign (Sec. 3.7):** Short-answer problems aligned to takeaways, ideally one per takeaway. Submission of two to three short-answer problems 24h before class.
- **Early availability of slides (Sec. 3.6):** providing lecture material at least 24h pre-class.
- **Topic focusing (Sec. 3.3):** limiting to two cases or stories, reducing acronyms, and defining on-slide explanations.
- **Lecturer remediation (Sec. 3.2):** coaching on teaching style, interaction with multimedia, directing student engagement, constructive alignment principles, and learning objectives.

The full action-item lists per lecture are documented in Appendix D, ensuring that conceptual measures are operationalized at the level of concrete actions, which is a critical step given the industry experts' limited teaching expertise. The assessment and the targeted measures, phrased as specific action items, are discussed with the lecturers in a personal meeting. The documentation is provided to the lecturers. The lecturers are invited to consult the organizers for a review before the actual lecture. The review is based on the revised/novel lecturing presentation. Also, the exam problems of winter term 2024/2025 are reviewed and suggestions provided to the lecturers on the level of understanding. The lecturers are asked to develop problems on the levels of understanding or apply.

A.1 Onboarding of Novel Lecturers

All novel lecturers for the lecture series of the subsequent term are consulted on the expectations defined in Sections 3.1, 3.3, and 3.4. Particularly they are advised to:

- Read the learning objectives LO-1–LO-4 and reflect on what knowledge to share to align with these. They shall think of a case-story based on their work to convey this knowledge.
- Focus their lecture on a maximum of three, if possible two takeaways. To focus, ask yourself: 'What is the essential learning that I can convey, for the students to make better risk-informed decisions?'
- Write a takeaways slide. State the takeaways of your lecture at a tangible level. Place this slide at the end of the presentation.
- Include a motivational slide entitled 'What we will learn today?' or 'Learning objectives'. Derive the Learning objectives or statements on this slide from the takeaways.
- Build your presentation in between these two slides. Use cases and visualizations from your work to convey the knowledge, not abstract pictures and processes. Always consider that the students want to understand how to solve the case and the problems you encounter and how to use the theoretical knowledge they have acquired.
- Following the slide summarizing the takeaways, include a slide motivating reprocessing the presentation and discussing with you. Thereby, invite students to take your role in making a risk-informed decision on the subject, given the information provided during the lecture, i.e., present another, simplified case or some aspects of a case. Start the task with: 'Imagine you are in my position, and you intend to ...'

ISSN 3006-6301, Vol. 3 (No. 3) 2026

- State an exam problem in the format of a short question for each takeaway. Write a maximum of three problems on the taxonomy level of apply for LO-1 and understand for LO-2–LO-4. Submit these problems together with the presentation a day before your lecture.

These recommendations are discussed with the lecturers in a personal meeting ahead of time to their scheduled lecture date, and communicated in written form immediately after the meeting. Thereby, they are requested to approach the organizers with a concept for their lecture and questions to be discussed preferably in another personal meeting.

Appendix 2

ISSN 3006-6301, Vol. 3 (No. 3) 2026

Appendix B. Ratings for Lectures, Lecture Material, and Lecturer Performance

The degree to which a lecturer fulfills the criteria C1–C4 is rated ex post each lecture. C1–C4 are:

- 0: Not fulfilled
- 1: Partially fulfilled
- 2: Fully fulfilled

Thereby, the early availability of lecture material according to Section 3.6 is part of the assessment.

The degree to which a lecture promotes the learning objectives LO-1–LO-4 is rated ex post each lecture. Three levels are defined:

- 0: Does not promote
- 1: Partially promotes
- 2: Fully promotes

Given the lecture material, the degree of alignment to the ARIVA framework is rated against the following three levels. The lecture either:

- 0: Does not align
- 1: Does partially align
- 2: Fully aligns

with the ARIVA framework. Thereby, if a lecture were to include the three structural elements described in Section 3.4, full alignment to the ARIVA framework is justified.

The degree to which constructive alignment (CA) is fulfilled is rated based on the alignment between the lecture content, the stated learning outcomes, and the associated exam problems with the taxonomy levels prescribed in Section 3.1. Three levels are defined:

- 0: Not fulfilled
- 1: Partially fulfilled
- 2: Fully fulfilled

Appendix C. Detailed Examination Analysis

C.1 Winter Term 2024/2025

The students' performance in the exam is plotted as violin plots in Figure C.2. Thereby, the problems are clustered per lecture, to identify potential systematics. Note that the credits of each problem are normalized by the maximum possible credits of that problem. The normalized median and mean scores of problem i are evaluated. A high mean ($\gtrsim 0.80$) suggests the problem is of low difficulty, while a low mean ($\lesssim 0.30$) indicates it is difficult to respond.

Complementarily, the item–total discrimination index r_i computed as the Pearson correlation between the problem score and the total test score excluding that problem is evaluated and visualized in Figure C.3. r_i measures how well the problem differentiates between high- and low-performing students. Positive values of $r_i \gtrsim 0.30$ are desirable, indicating that high-performing students do better and low-performing students do worse on this problem. Negative values potentially indicate ambiguity of the problem or misalignment to the lecture.

The following systematics and anomalies are observed:

Systematics:

- Problem format is the primary driver of discrimination. Every problem achieving $r_i \geq 0.30$ is a short-answer or multi-step question. Most single-choice problems (Problems 2, 3, 5, 6, 22) and multiple-choice problems (Problems 1, 4, 12, 14, 17, 18) fall below the threshold — noted as a structural limitation.
- Lecture 2 (Problems 3–5) shows uniformly low to negative discrimination. Problem 5 is too difficult and yields negative r_i, a signal of ambiguity or content misalignment. The lecture received zero constructive alignment rating and the lowest ARIVA rating in the series.
- Lecture 7 (Problems 16–18) shows uniformly poor performance and low discrimination. Given the lecture's absent ARIVA structure, lack of focus, and zero constructive alignment rating (see Table 3), this pattern reflects a lecture-level failure rather than individual problem design issues.
- Short-answer problems tend to discriminate consistently better, independent of subject matter and lecture quality.

Anomalies:

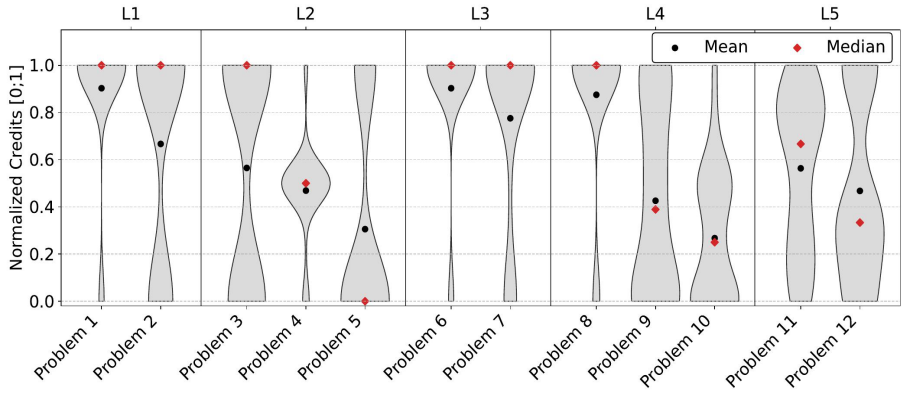
- Problems 5 (L2) and 22 (L10) show negative discrimination. For Problem 5, ambiguity in a single-choice item on dense probabilistic content is the likely cause. Problem 22 is rated too easy.
- Problem 10 (L4) is a short-answer problem with low discrimination, contrary to the general format trend. Lecture 4's acronym overload and unclear focus suggest the content is poorly communicated rather than the format inadequate.
- L1 versus L8 contrast. Both lecturers received strong competence ratings, yet L8 (Problems 19–21) discriminates consistently better than L1 (Problems 1–2). L1's over-specificity on regulatory standards likely produces a bimodal response pattern — students either know the specific standard or do not — suppressing discrimination.
- Within-lecture inconsistency in L10. Problem 22 shows negative discrimination alongside Problems 23–24 with good discrimination, indicating uneven problem quality rather than a lecture-level issue.

The 2024/2025 exam insufficiently differentiates competences. The dominant limitation is structural: single-choice and multiple-choice formats do not generate sufficiently reliable discrimination in this context. For Lectures 2, 7, and 9, low discrimination reflects not only suboptimal problem format but

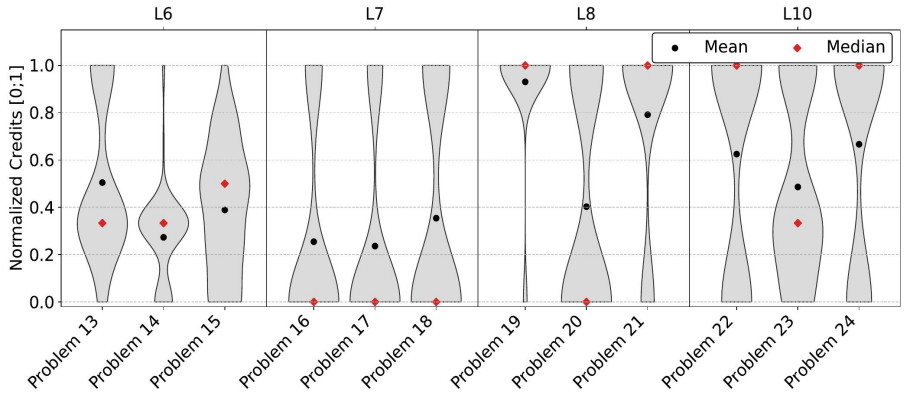
also lecture-level failures in structure and constructive alignment, which prevented students from extracting coherent, examinable knowledge. The mean 56%, coincident mean and median, and $\alpha_{norm} = 0.65$, noted in Section 2, are consistent with these observations. The exam covers the breadth of the subject matters, addresses different levels of difficulty and is fair, yet, it does not reliably reflect the students' individual potentials.

In summary, the following systematics and anomalies are observed:

- Systematic: The three problems of Lecture 2 indicate a low or even negative discrimination.
- Systematic: The two problems of Lecture 5 indicate an acceptable discrimination. The student performance is good.
- Systematic: All three problems of Lecture 7 have been answered poorly and the discrimination is low.
- Anomalies: Problems 5 and 22 indicate a negative discrimination.
- Anomalies: The single choice type problems 1 and 6 show a poor discrimination.
- Systematic: Short question type problems discriminate well.

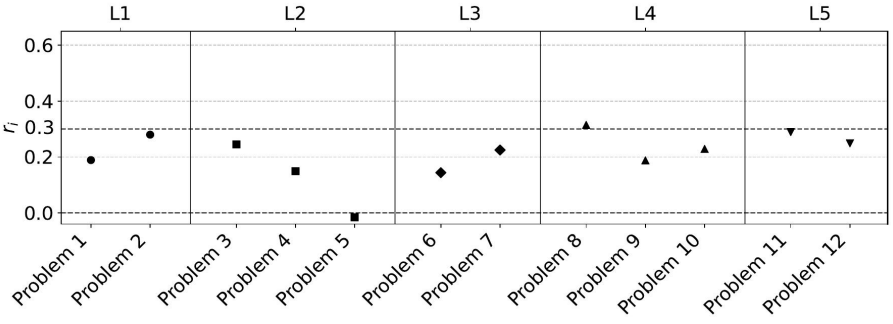


(a) Normalized scores for lectures L1–L5

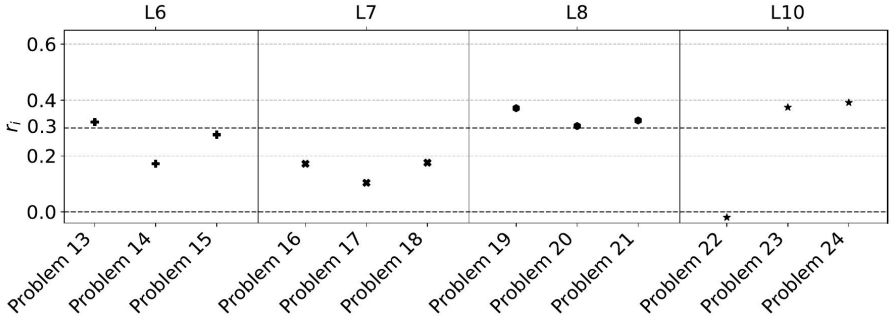


(b) Normalized scores for lectures L6–L10

Figure C.2: Violin plots of normalized scores per problem, split by lecture groups for exam of winter term 2024/2025.



(a) r_i for lectures L1–L5



(b) r_i for lectures L6–L10

Figure C.3: Item-total discrimination index r_i per problem, split by lecture groups for exam of winter term 2024/2025.

C.2 Winter Term 2025/2026

Figure C.4 shows the credit distribution for the 55 students taking the exam in winter term 2025/2026. Maximum achievable credits are 75.5, with a 20% overhang giving a practical ceiling of 63 credits (see Section 3.7). The highest achieved score is 58, i.e., 92% of the ceiling, confirming appropriate overhang calibration with no ceiling effect. The mean and median scores are 31 and 32 credits, respectively. The standard deviation is 12 credits. The distribution is near-symmetric with a skewness of 0.19 and excess kurtosis -0.38. The internal consistency is indicated with $\alpha_{norm} = 0.85$.

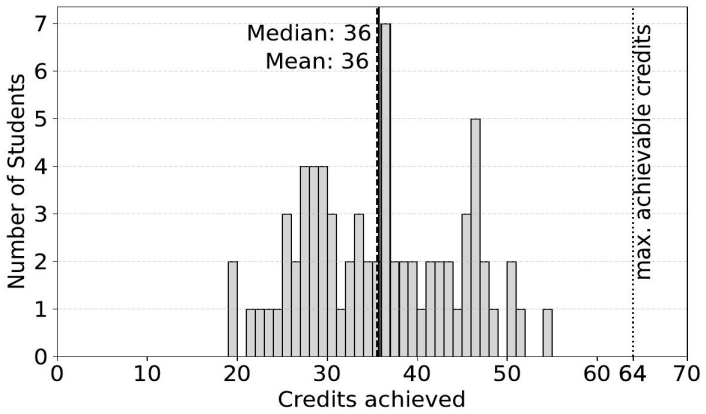


Figure 1: Distribution of Credits per Student (Winter Term 2024/2025).

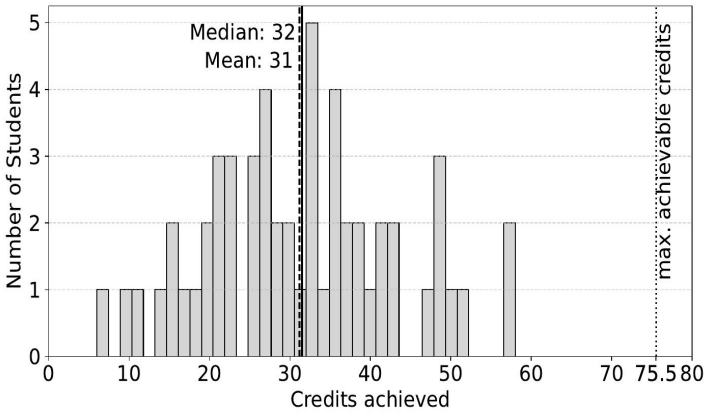


Figure C.4: Distribution of Credits per Student (Winter Term 2025/2026).

The per problem score distributions and discrimination indices (Figures C.5 and C.6) reveal the following systematics and anomalies:

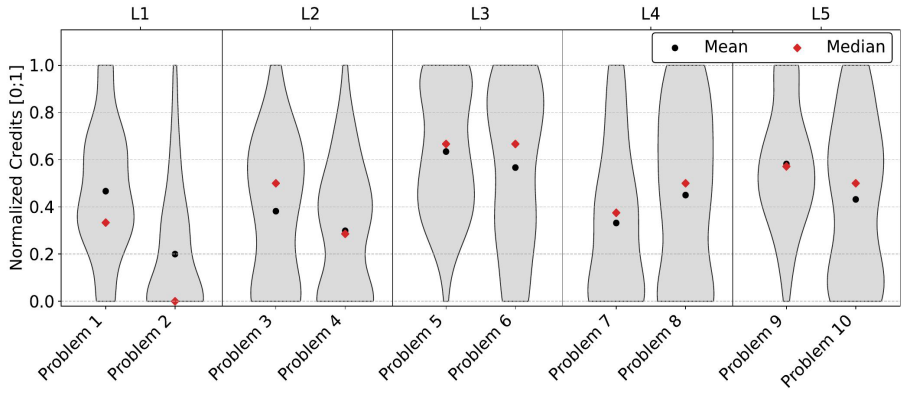
Systematics:

- Short-answer format achieves near-universal discrimination. With 19 of 20 problems in short-answer format, almost all meet $r_i \geq 0.30$. The single exception is Problem 12 (L6), also the only single-choice problem, replicating the 2024/2025 pattern and confirming format as the binding constraint on discrimination.
- Lecture 7 (Problems 13–14) is consistently very difficult with high discrimination. RaSa received strong teaching ratings and substantially improved the lecture content; the difficulty reflects the inherent difficulty of the subject matter of emerging technology risk assessment, not poor instruction. Both problems are valid.
- Lecture 1 (Problems 1–2) is the second most difficult cluster, attributable to the subject matter of aerospace safety management.
- Lecture 3 (Problems 5–6) represents optimal calibration: moderate difficulty, good to very high discrimination.
- Problems 2 and 13, though very difficult and often left blank, show valid discrimination. High performers answer them substantially better than low performers, confirming genuine difficulty rather than ambiguity. Their presence raises overall exam demand without harming validity. Yet, these indicated that the content relevant for these problems, and relevant in view of the lecturers, is not well understood. Hence, the lectures need to improve in conveying the content.

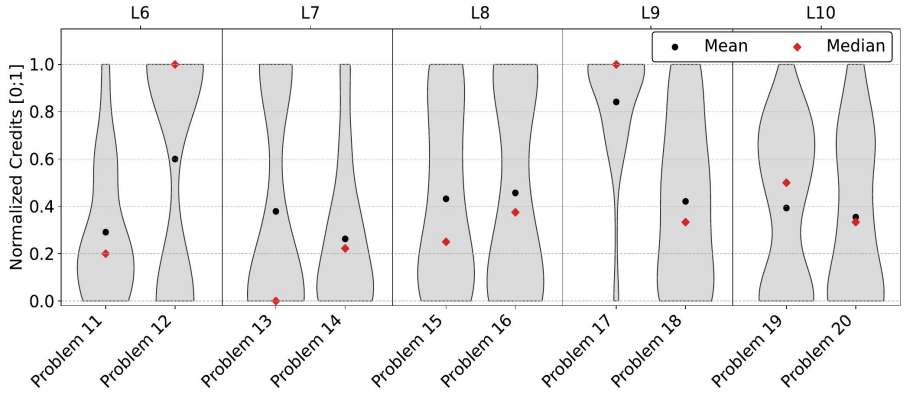
Anomalies:

- Problem 12 (L6) is the only problem with low discrimination and the only single-choice item. Crucially, it is identical to Problem 2 of the 2024/2025 exam, where it achieved higher discrimination. This provides the clearest available evidence that single-choice format should be eliminated regardless of content quality.
- Problem 17 (L9) is easy with low discrimination, contrasting with Problem 18 from the same lecture, which achieves the highest r_i in the entire exam. The anomaly is attributable to within-lecture problem calibration. Note that Problem 17 focuses on taxonomy level 1, remembering.

ISSN 3006-6301, Vol. 3 (No. 3) 2026

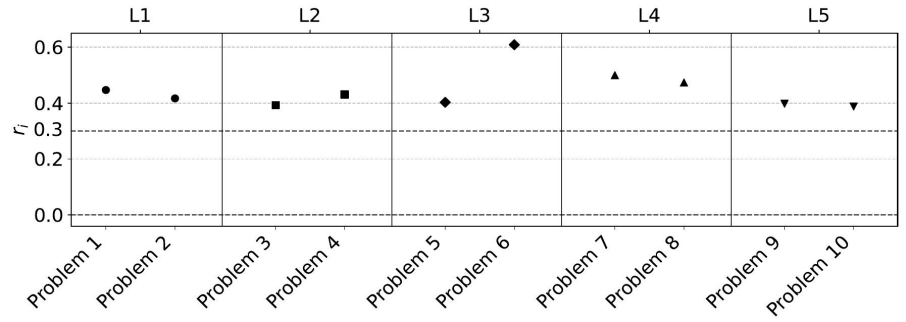


(a) Normalized scores for lectures L1–L5

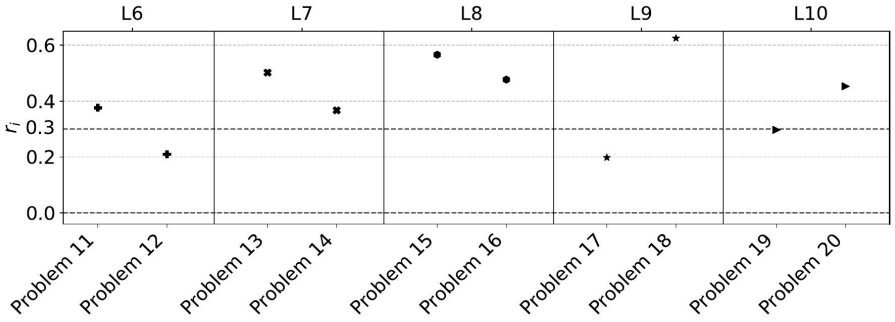


(b) Normalized scores for lectures L6–L10

Figure C.5: Violin plots of normalized scores per problem, split by lecture groups for exam of winter term 2025/2026.



(a) r_i for lectures L1–L5



(b) r_i for lectures L6–L10

Figure C.6: Item-total discrimination index r_i per problem, split by lecture groups for exam of winter term 2025/2026.

The exam of 2025/2026 is a substantially improved assessment. Near-universal short-answer format, combined with content focused to explicit takeaways, produces reliable competence differentiation across the full performance range. The wider standard deviation, improved α_{norm}, and higher proportion of problems above the discrimination threshold are consistent. The lower mean score is an intended consequence of assessing at higher taxonomy levels; the absence of a ceiling effect indicates it being both demanding and fair.

This indicates a consistent difficulty with the subject matter of lecture 7.

Table C.8: Interpretation of per problem statistics of exam winter term 2024/2025, grouped by lecture.

Problem	Lecture	Interpretation	Type of problem
P1	L1	too easy, low discrimination	Multiple Choice
P2	L1	easy, acceptable discrimination	Single Choice
P3	L2	easy, low discrimination	Single Choice
P4	L2	good, low discrimination	Multiple Choice
P5	L2	too difficult, negative discrimination	Single Choice
P6	L3	too easy, low discrimination	Single Choice
P7	L3	easy, low discrimination	Short Question
P8	L4	too easy, acceptable discrimination	Single Choice
P9	L4	good, low discrimination	Short Question
P10	L4	difficult, low discrimination	Short Question
P11	L5	good, acceptable discrimination	Short Question
P12	L5	difficult, acceptable discrimination	Multiple Choice
P13	L6	difficult, acceptable discrimination	Multiple Choice
P14	L6	difficult, low discrimination	Multiple Choice
P15	L6	good, acceptable discrimination	Short Question
P16	L7	too difficult, low discrimination	Multiple Choice
P17	L7	too difficult, low discrimination	Multiple Choice

P18	L7	too difficult, low discrimination	Multiple Choice
P19	L8	too easy, good discrimination	Single Choice
P20	L8	too difficult, acceptable discrimination	Single Choice
P21	L8	easy, acceptable discrimination	Single Choice
P22	L10	too easy, negative discrimination	Single Choice
P23	L10	good, good discrimination	Multiple Choice
P24	L10	easy, good discrimination	Single Choice

Table C.9: Interpretation of per problem statistics of exam winter term 2025/2026, grouped by lecture.

Problem	Lecture	Interpretation	Type of problem
P1	L1	difficult, good discrimination	Short Question
P2	L1	very difficult, good discrimination	Short Question
P3	L2	medium difficulty, good discrimination	Short Question
P4	L2	difficult, good discrimination	Short Question
P5	L3	low difficulty, good discrimination	Short Question
P6	L3	low difficulty, very high discrimination	Short Question
P7	L4	difficult, high discrimination	Short Question
P8	L4	medium difficulty, high discrimination	Short Question
P9	L5	medium difficulty, good discrimination	Short Question
P10	L5	medium difficulty, good discrimination	Short Question
P11	L6	very difficult, good discrimination	Short Question
P12	L6	easy, low discrimination	Single Choice
P13	L7	very difficult, high discrimination	Short Question
P14	L7	very difficult, good discrimination	Short Question
P15	L8	medium difficulty, very high discrimination	Short Question
P16	L8	medium difficulty, high discrimination	Short Question
P17	L9	easy, low discrimination	Short Question
P18	L9	medium difficulty, very high discrimination	Short Question
P19	L10	medium difficulty, acceptable discrimination	Short Question
P20	L10	difficult, high discrimination	Short Question

Appendix 10

Appendix D. Action-Items on Lecture Level

This section documents the list of Action Items as requested to be implemented by the lecturers of winter term 2025/2026. These can be included as supplementary material.

The following instructions are addressed directly to the lecturers of **L1**:

- Given the change in taxonomy level, rather lecture on the purpose of standards and following these instead of highlighting a particular standard. Highlight the underlying concept of certification of medical devices.
- Focus the lecture on three takeaways / learning objectives instead of five.
- Review the already existing Learning objectives slides and Takeaway slide and align these.
- Please write a short question problem for risk management of medical devices – E.g. you could consider asking for 'When pursuing to market a new medical device, the benefit-risk ratio has to be evaluated. A) What is the meaning of a benefit risk ratio and B) how is it evaluated?'
- State an exam problem in the format of a short question for each takeaway. Submit these problems together with the presentation a day before your lecture.

The following instructions are addressed directly to the lecturers of **L2**:

- Focus your lecture on a maximum of three takeaways. Suggestions, specific to the subject matter are given.
- Build your lecture presentation around these three takeaways by explaining the ideas, concepts, and theoretical background information.
- Include a takeaways slide at the end of the presentation – Phrase the takeaways at a tangible level.
- Include a motivational slide entitled 'What we will learn today?' or 'Learning objectives'. Derive the Learning objectives or Statements on this slide from the takeaways.
- Include a slide motivating reprocessing the presentation and discussing with you with a task such as 'Imagine you are in our position, and you are informed that a new function will be marketed. How do you provide the quantitative analysis for a hopefully positive risk balance for its safety case, given what you have learned today?'.
- State an exam problem in the format of a short question for each takeaway. Submit these problems together with the presentation a day before your lecture.

The lecturer of **L3** is being replaced, this is discussed with the lecturer. The lecturer is involved in selecting lecturers from their organization that fulfill the criteria stated in Section 3.2. Given the analysis of the existing subject matter and slide deck, the following instructions are addressed to and discussed with the new lecturer:

- Focus your lecture on a maximum of three takeaways. Suggestions, specific to the subject matter are given.
- Build your lecture presentation around these three takeaways by explaining the ideas, concepts, and theoretical background information.
- Describe underlying challenges of your work and how to solve these for risk-based decision making.
- Include a takeaways slide at the end of the presentation – Phrase the takeaways at a tangible level. Thereby restate the key messages / core messages of your lecture at a tangible level.
- Include a motivational slide entitled 'What we will learn today?' or 'Learning objectives'. Derive the Learning objectives or statements on this slide from the takeaways.
-

- Following the slide summarizing the takeaways, include a slide motivating reprocessing the presentation and discussing with you. Thereby, invite to the discussion by inviting the students to take your role in making a risk informed decision on the subject, given the information provided during the lecture, i.e. present another, simplified case or some aspects of a case. You could start the task by 'Imagine you are in my position, and you intend to ...'
- State an exam problem in the format of a short question for each takeaway. Submit these problems together with the presentation a day before your lecture.

The following instructions are addressed directly to the lecturers of **L4**:

- Limit the use of acronyms to a minimum. Whenever necessary, include the definition of the acronym on the slide.
- Focus your lecture on a maximum of three takeaways. Suggestions, specific to the subject matter are given.
- Build your lecture presentation around these three takeaways by explaining the ideas, concepts, and theoretical background information.
- Include a takeaways slide at the end of the presentation – Phrase the takeaways at a tangible level. Thereby restate the key messages / core messages of your lecture at a tangible level.
- Include a motivational slide entitled 'What we will learn today?' or 'Learning objectives'. Derive the Learning objectives or statements on this slide from the takeaways.
- Following the slide summarizing the takeaways, include a slide motivating reprocessing the presentation and discussing with you. Thereby, invite to the discussion by inviting the students to take your role in making a risk informed decision on the subject, given the information provided during the lecture, i.e. present another, simplified case or some aspects of a case. You could start the task by 'Imagine you are in my position, and you intend to ...'
- State an exam problem in the format of a short question for each takeaway. Submit these problems together with the presentation a day before your lecture.

Furthermore, during a personal meeting with the lecturers of L4, we pinpoint issues observed during the previous term, remediating the deficiencies in C1–C3. These address e.g. observed presence in the classroom, voice, addressing students more directly, interaction with multimedia material, preparing for particular questions.

The following instructions are addressed directly to the lecturers of **L5**:

- Explain the purpose and business of reinsurance for the insurance industry. Thereby, work with an example.
- Include a takeaways slide at the end of the presentation – Phrase the takeaways at a tangible level. Thereby restate the key messages / core messages of your lecture at a tangible level.
- Include a motivational slide entitled 'What we will learn today?' or 'Learning objectives'. Derive the Learning objectives or statements on this slide from the takeaways.
- Following the slide summarizing the takeaways, include a slide motivating reprocessing the presentation and discussing with you. Thereby, invite to the discussion by inviting the students to take your role in making a risk informed decision on the subject, given the information provided during the lecture, i.e. present another, simplified case or some aspects of a case. You could start the task by 'Imagine you are in my position, and you intend to ...'
- Remove the slide entitled 'Questions'.
- State an exam problem in the format of a short question for each takeaway. Submit these problems together with the presentation a day before your lecture.

The lecturer and subject matter of **L6** are replaced.

The following instructions are addressed directly to the lecturers of **L7**:

- Limit the use of acronyms to a minimum. Whenever necessary, include the definition of the acronym on the slide. Do not use different acronyms for the same purpose.
- Use no more than two projects to tell the story of data-informed risk-based decision making. Thereby, clearly highlight the key aspects as takeaways of each project.
- Include a takeaways slide at the end of the presentation – Phrase the takeaways at a tangible level. Thereby restate the key messages / core messages of your lecture at a tangible level.
- Include a motivational slide entitled 'What we will learn today?' or 'Learning objectives'. Derive the Learning objectives or statements on this slide from the takeaways.
- Following the slide summarizing the takeaways, include a slide motivating reprocessing the presentation and discussing with you. Thereby, invite to the discussion by inviting the students to take your role in making a risk informed decision on the subject, given the information provided during the lecture, i.e. present another, simplified case or some aspects of a case. You could start the task by 'Imagine you are in my position, and you intend to ...'
- State an exam problem in the format of a short question for each takeaway. Submit these problems together with the presentation a day before your lecture.

The following instructions are addressed directly to the lecturers of **L8**:

- Focus your lecture on two problems, stories. Define one takeaway for each story. Highlight these takeaways.
- Include a takeaways slide at the end of the presentation – Phrase the takeaways at a tangible level. Thereby restate the key messages / core messages of your lecture at a tangible level.
- Include a motivational slide entitled 'What we will learn today?' or 'Learning objectives'. Derive the Learning objectives or statements on this slide from the takeaways.
- Following the slide summarizing the takeaways, include a slide motivating reprocessing the presentation and discussing with you. Thereby, invite to the discussion by inviting the students to take your role in making a risk informed decision on the subject, given the information provided during the lecture, i.e. present another, simplified case or some aspects of a case. You could start the task by 'Imagine you are in my position, and you intend to ...'
- State an exam problem in the format of a short question for each takeaway. Submit these problems together with the presentation a day before your lecture.

The lecturer and subject matter of **L9** are replaced.

The subject matter of **L10** is replaced. The lecturer of L10 as well as novel lecturers for the lecture series of the subsequent term are consulted on the expectations defined in Sections 3.3 and 3.4. Particularly, the novel lecturers are advised to:

- Focus your lecture on a maximum of three, if possible two takeaways. To focus, ask yourself: 'What is the essential learning that I can convey, for the students to make better risk-informed decisions?'
- Write a takeaways slide. State the takeaways of your lecture at a tangible level. Place this slide at the end of the presentation.
- Include a motivational slide entitled 'What we will learn today?' or 'Learning objectives'. Derive the Learning objectives or statements on this slide from the takeaways.
- Build your presentation in between these two slides. Use cases and visualizations from your work to convey the knowledge, not abstract pictures and processes. Always consider that the students want to understand how to solve the case and the problems you encounter and how to use the theoretical knowledge they have acquired.
- Following the slide summarizing the takeaways, include a slide motivating reprocessing the presentation and discussing with you. Thereby, invite to the discussion by inviting

the students to take your role in making a risk informed decision on the subject, given the information provided during the lecture, i.e. present another, simplified case or some aspects of a case. You could start the task by 'Imagine you are in my position, and you intend to ...'

- State an exam problem in the format of a short question for each takeaway. Write a maximum of three problems. Submit these problems together with the presentation a day before your lecture.

Each lecturer is offered a review of and constructive feedback on the lecture material by the organizers prior to their lecture.

Appendix E. Student Evaluation

Student evaluation results for Winter Terms 2024/2025 and 2025/2026 are provided as supplementary material. The Winter Term 2024/2025 evaluations gathered seven responses; the Winter Term 2025/2026 evaluation involved 29 students. Ratings improved across all assessed dimensions from one term to the next, with the overall grade rising from a mean of 3.0 to 1.9 on a 1–6 scale (1 = very good, 6 = insufficient).