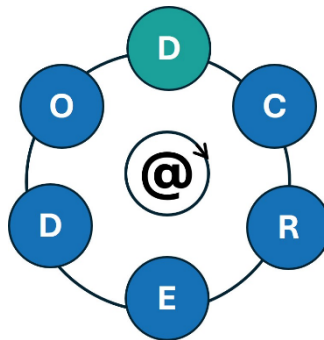# D-CREDO

Digital Health Technologies-Augmented Clinical Reasoning Education

2024-1-PL01-KA220-HED-000247790

D4.2 Evaluation toolset report

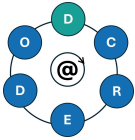|	
Deliverable number	D.4.2
Delivery date	September 30, 2025
Status	final
Authors	Tetiana Shchudrova (BSMU), Andrzej Kononowicz (JU), Inga Hege (Instruct gGmbH), Iryna Popova (BSMU), Nataliia Davydova (BSMU)

Document Revision Table

Version	Date	Author(s)	Organization	Comments
01	07.07.2025	Tetiana Shchudrova	BSMU	First draft
		Elske Ammenwerth Andrzej Kononowicz	UMIT TIROL JU	Add Quality Checklist for Deliverables
02	05.08.2025	Andrzej Kononowicz	JU	Update and edit draft Add Learning analytics
03	06.08.2025	Iryna Popova Nataliia Davydova	BSMU	Process feedback, edit and update draft
04	04.09.2025	Inga Hege	Instruct	Update and edit draft
05	01-15.09.2025	Tetiana Shchudrova	BSMU	Process feedback, edit and update draft
06	20.09.2025	Andrzej Kononowicz	JU	Update and edit draft
1.0	29.09.2025	Tetiana Shchudrova	BSMU	Finalized

Summary

Objectives: The EU project D-CREDO (Digital Health Technologies-augmented Clinical Reasoning Education) aims to enhance the skills of health professions students and faculty in using digital health tools. We plan to develop high-quality learning units (LUs) alongside virtual patients (VPs) that will prepare students for the responsible use of digital tools in clinical reasoning (CR). The main objective of developing an evaluation toolset is to establish a means of assessing the quality and impact of the developed LUs.

Approach: A structured collaborative approach was adopted to facilitate the development of the evaluation toolset. In order to ensure diverse contributions and consensus, the working group comprised representatives from each of the D-CREDO partner institutions. Five major activities were identified: the development of a quality checklist for the project deliverables; the selection of a set of assessment methods and tools aligned with the learning objectives (LOs); the development of quality checklists for the project LUs and VPs; the development of tools for evaluating the learning experience; and the selection of learning analytics tools to monitor the quality of LUs developed in the D-CREDO project. The process involved, wherever possible, collecting existing evaluation methods and tools from the literature and previous CR research and education projects, and adapting them to the project's specific objectives. The process was guided by regular collaborative discussions and iterative feedback within the working group and consortium, with subsequent refinement of the evaluation tools to ensure consensus and alignment with the project's goals.

Results: A comprehensive evaluation toolset has been developed, including the following:
- Quality checklist for project deliverables (25 criteria);
- Quality checklist for monitoring LU development (8 criteria);
- Quality checklist for the formal review of LUs (31 criteria)
- Quality checklists for didactic (40 criteria) and content (11 criteria) review of the VPs;
- Set of assessment methods and tools to measure the expected learning outcomes;
- Satisfaction questionnaires for learners (14 questions) and instructors (26 questions);
- A list of selected learning analytics metrics available in the project's learning management systems Moodle and CASUS.

Conclusion: The evaluation toolset presented in this deliverable will serve as a framework for the upcoming activities in two project work packages: WP3 (Development of the student and educator LUs) and WP4 (Pilot study planning and evaluation).
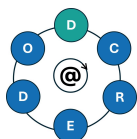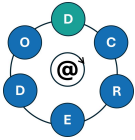
Table of Contents

1 Introduction

The aim of this deliverable is to present a set of evaluation tools to ensure the quality of the project's outputs by monitoring the development process of the LUs and VPs and measuring the expected educational outcomes and learner satisfaction. Our plans include adopting or developing assessment methods and tools that are aligned with the LOs, as well as quality checklists for the project's LUs and VPs. We also intend to create evaluation surveys that measure learner and educator experience, satisfaction and the perceived impact on learning outcomes. Additionally, we will select learning analytics metrics that contribute to quality improvement by monitoring indicators of learning engagement and clinical reasoning success in the LUs and VPs hosted by the project's technical infrastructure. Wherever possible, the development process will involve collecting existing evaluation methods and tools from literature and previous clinical reasoning research and education projects, adapting them to the project's specific objectives. Regular collaborative discussions and iterative feedback within the working groups and consortium will guide the process, with subsequent refinement of the evaluation toolset to ensure consensus and alignment with the project's goals.

This deliverable builds upon the work carried out in WP2 (D2.1. Literature Review Report; D2.2 Learning Objectives Report and D2.3 White Paper) and WP4 (D4.1 Report on Targeted Needs Assessment), as well as the outcomes of the EU-funded projects iCoViP (https://icovip.eu/) and DID-ACT (https://did-act.eu/).

2 Quality criteria

The D-CREDO consortium decided to guide the development of the evaluation toolset based on following quality criteria:

The quantitative indicators:

- a satisfaction questionnaire for students
- a satisfaction tool for educators

The qualitative indicators:

- recommended assessment methods for each LO
- constructive alignment of LOs and assessment methods
- quality monitoring strategy for content of LUs and VPs
- selection of the evaluation tools informed by the D-CREDO evaluation framework
  (as set out in the White Paper)
- selection of digital learning activity monitoring metrics
- implementation of the evaluation tools in a digital system
- final evaluation tool report

3   Quality checklist for deliverables

This checklist is to be used for the final review of all public deliverables of D-CREDO. Authors are asked to use this checklist when writing their deliverable, submitting it to the project partner leads for formal review no later than two weeks before the official deadline.

Checklist

1. The name of the file should include D-CREDO along with the Deliverable number and short title (e.g. D2.1-D-CREDO-Rapid-literature-review.pdf).

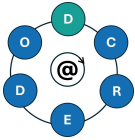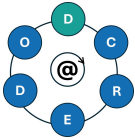2. The title page strictly follows the D-CREDO Deliverables template. It includes the full names and affiliations of all authors, correct delivery date, status changed to "Final" and displays the EU disclaimer.

3. The D-CREDO and EU logos are placed in the header at the top of every page.

4. The second page contains the document revision table, which tracks the development and quality management of the Deliverable, including details on the final review.

5. The third page presents a structured summary designed to be clear and understandable for readers outside the D-CREDO consortium. The summary is written in a concise, accessible style suitable for social media postings, avoiding abbreviations and emphasizing the relevance of the results. The summary is structured (Objectives, Approach, Results, Conclusion)

6. The fourth page features an up-to-date table of contents (TOC) structured to help readers navigate the Deliverable. The TOC avoids unnecessary abbreviations or overly specific terms, maintains a professional and well-organized layout, and ensures each chapter includes at least two subchapters. Appendices are listed with clear names.

7. Starting from the introduction, each page is numbered, beginning with "1."

8. All chapter and subchapter titles use Word's "chapter" formatting and are automatically numbered.

9. All cross-referencing to chapter and subchapters are properly linked using Word's cross-referencing feature.

10. Font type and size are used consistently using the "standard" format style of Word.

11. Text alignment is consistent throughout the Deliverables.

12. All bullet points and numbered lists are formatted consistently throughout the Deliverables.

13. Figures are consecutively numbered and include clear, informative titles below the Figure. Numbering is applied using Word's automatic "Insert Figure" feature. Figures are referenced in the text using automatic cross-referencing to ensure consistency. They are large enough to be legible.

14. Tables are consecutively numbered with clear, informative titles above the Table. Word's automatic "Insert Table" feature is used for numbering, and tables are referenced in the text using automatic cross-referencing for consistency.

15. Paragraphs are spaced (e.g., 3 pt) to enhance readability. This spacing is applied consistently throughout the Deliverable.

16. All external links (URLs) in the document are checked for correctness and properly linked using Word's cross-linking feature.

17. If references are included, they follow a clear and consistent citation style (e.g., APA or Harvard) in both in-text citations and the reference list.

18. Bold text may be used for emphasis but should be applied sparingly.

19. The text has been checked for spelling and grammar errors. Automatic correction should remain enabled.

20. The qualitative and quantitative performance indicators should be checked against the D-CREDO project proposal. The Deliverable should discuss the extent to which these KPIs have been achieved, providing relevant data and explanations where necessary.

21. All annexes are clearly labelled and referenced in the main text.

22. If numerous abbreviations are used, a list of abbreviations (e.g., placed before the introduction) can aid readability. Abbreviations and full terms must be used consistently throughout the Deliverable.

23. All comments in the document and all track modes should be removed before exporting the document to PDF.

24. The final PDF version of the Deliverable should be sent to Instruct for upload on the project website. Additionally, it must be archived on Google Drive for internal documentation and future reference.

25. The Deliverable lead prepares a short summary of the Deliverable for a blog post and LinkedIn release and sends it to UMIT within two weeks after the Deliverable is completed.

4 Assessment methods and tools

Following Kern's six-step curriculum development approach, we created a comprehensive set of 26 learning objectives - 19 for student LUs and 7 for educator LUs (see [D2.2 Learning objectives report](#)), that integrate digital health tools into clinical reasoning education for both students and teachers. The LOs are categorised systematically according to Bloom's Taxonomy level, target audience, relevance across the health professions, applicable categories of the D-CREDO digital health tools, and alignment with the DID-ACT clinical reasoning themes. The LOs serve as a foundation for the design of individual LUs in WP3 and guide planning for their evaluation and quality assurance in WP4.
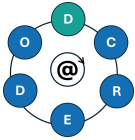
In the project white paper, 'Teaching Clinical Reasoning Enhanced by Digital Tools' ([D2.3 White paper](#)), we present a conceptual and methodological framework to support the integration of digital technologies into the teaching and assessment of CR. Our aim is to provide health professions educators with an evidence-based approach to bridging the gap between traditional CR training and the digitally enhanced environment of contemporary healthcare. Drawing on key learning theories such as cognitive load theory, experiential learning, distributed cognition and reflective practice, this framework promotes instructional strategies such as blended learning, technology-enhanced learning and case-based learning. Our pedagogical approach emphasizes interactive, evidence-based methods that foster engagement, stimulate prior knowledge, enable new knowledge application, and enhance the transfer and retention of clinical reasoning skills in a digitally mediated environment.

The white paper provides targeted guidance on LU design, applying the principle of constructive alignment to ensure coherence between intended outcomes, instructional methods, and assessment strategies. Within WP4, this alignment enables the selection of assessment formats that accurately measure learner achievement, encourage meaningful engagement and promote the responsible and effective use of digital health tools in clinical reasoning.

In light of the inherent complexity of CR and the evolving concept of digitally augmented CR, our objective was to develop a catalogue of assessment methods and tools that accurately reflect the intended learning outcomes.

CR assessment methods

The selection and alignment of assessment methods with specific learning objectives was driven by the programmatic assessment strategy, which supports the implementation of assessment-for-learning and assessment-as-learning practices. Assessment-as-learning is defined as a situation in which learning and assessment are intertwined to encourage self-regulated learning behaviour (Swan Sein et al., 2021). Examples include self-assessment, portfolios, reflection exercises, and peer assessment. Assessment-for-learning (formative assessment) emphasises feedback and reflection to guide and monitor student progress and inform instruction (Biggs & Tang, 2007, Ruczynski, 2024). It includes discussion-based, observation-based and peer- and self-assessments, as well as written and product-based assessments. Performance-based assessments include scenario-based exercises, role plays, practical presentations or demonstrations (e.g. effective use of a tool or software), and portfolio creation.

We analysed the catalogue of clinical reasoning assessment methods (Daniel et al., 2019) and selected relevant methods from three categories: non-workplace-based assessments (e.g. multiple-choice questions, extended matching questions, key feature examinations and script concordance tests); assessments in simulated clinical environments (e.g. technology-enhanced simulation); and workplace-based assessments (e.g. direct observations, oral presentations and written notes). In addition to conventional CR assessments, we examined innovative digital and AI-powered assessment tools to explore their potential for providing immediate, automated feedback and supporting dynamic, learner-centred evaluation strategies (Owan et al., 2023, Saputra et al., 2023).

A VP is an effective and well-established interactive learning tool that enables students to practise CR in a safe and controlled environment which mimics real patient encounters (Kotwal et al., 2021, Plackett et al., 2021). In the context of D-CREDO, integrating VPs into LUs offers a variety of possibilities. A key advantage of VPs is that they can be integrated with the digital health technologies of interest, such as electronic health records (EHRs), clinical decision support systems (CDSSs) and AI-generated content, in order to create an authentic digital environment for learning and practising CR enhanced with digital tools. Furthermore, the implementation of the automated feedback mechanisms into VPs can strongly reinforce learning effectively by prompting reflection and guiding refinement and improvement (García-Torres et al., 2024, Jay et al., 2025).

We plan to design VPs (extending from the [iCoViP collection](#)) combined with CR concept maps, in which findings, tests, differential diagnoses and treatment options can be documented and connected to each other (Hege et al., 2017). According to Mayer et al. (2025), combining concept map activities with VPs can reinforce their educational effect on CR outcomes. Therefore, VPs combined with concept maps could serve as an engaging learning activity and a formative assessment method, providing tailored feedback on CR processes and encouraging reflective thinking during and after the case. Meanwhile, case reports could be used for peer review and debriefing.

CR assessment tools

Validated CR assessment tools (see Table 1) provide structured, ready-to-use instruments (e.g. questionnaires, rubrics and checklists) that are often psychometrically tested and designed to reliably measure CR or its components (Covin et al., 2020; Ilgen et al., 2012; Thammasitboon et al., 2018). Unlike more general assessment methods such as multiple-choice questions (MCQs), script concordance tests (SCTs) or VPs, these tools are specific instruments that can be implemented 'off-the-shelf' with established scoring guidelines. Integrating these validated tools into LUs enables alignment with programmatic assessment strategy, ensuring consistent and reliable measurement across LOs and contexts, as well as facilitating longitudinal progress tracking, allowing for both feedback and reflection (Goldszmidt et al., 2013; Lasater, 2007; Peterson et al., 2022).
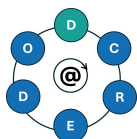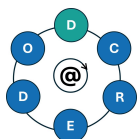
Table 1. Overview of validated CR assessment tools

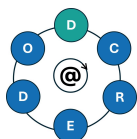| Tool	Purpose	Format	Target group	Reference
REACT Rapid Evaluation Assessment of Clinical Reasoning Tool	To measure five dimensions of CR process: data collection, interpretation, management, communication and reflection during urgent patient care	3-level rubric (5 dimensions)	Medical students, residents	Peterson et al., 2022
IDEA Interpretive summary, Differential diagnosis, Explanation of reasoning, Alternatives	To rate students' reporting, diagnostic reasoning, and decision-making skills based on new patient admission notes.	15-item instrument (4 domains)	Medical students	Baker et al., 2015
ART Assessment of Reasoning Tool	To assess five domains of reasoning: hypothesis-directed data gathering, articulation of a problem representation, formulation of a prioritized differential diagnosis, diagnostic testing aligned with high-value care principles and metacognition	3-point scale (5 domains)	Medical students, residents	Thammasitboon et al., 2018
DTI Diagnostic Thinking Inventory	To measure two aspects of diagnostic thinking: the degree of flexibility in thinking and the degree of knowledge structure in memory.	41-item self-report questionnaire	Medical students, clinicians	Bordage et al., 1990
LCJR Lasater Clinical Judgment Rubric	To assess clinical judgment and decision-making based on four phases of Tanner's (2006) Clinical Judgment Model - noticing, interpreting, responding, and reflecting.	4-level rubric (11 dimensions described in four developmental levels)	Pre-licensure nursing students	Lasater, 2007
CDI Clinical Data Interpretation test	To evaluate development of CR skills by testing ability to interpret clinical data	72-item multiple-choice question instrument	Medical students	Williams et al., 2011 Covin et al., 2019

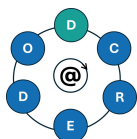| SSAR Summary Statement Assessment Rubric	To assess summary statements as a marker for clinical reasoning using a five-component rubric that includes the following: factual accuracy, appropriate narrowing of the differential diagnosis, transformation of information, use of semantic qualifiers, and a global rating.	5-component rubric	Medical students	Smith S et al., 2016 Covin et al., 2020
PNS Patient Note Scoring rubric	To measure three dimensions: documentation, justified differential diagnosis (DDX), and workup	patient note scoring rubric (3 dimensions)	Medical students	Park et al., 2013 Covin et al., 2020
CRT Clinical Reasoning Task checklist	To assess CR performance in four broad categories: framing the encounter, diagnosis, management, and self-reflection.	24 reasoning tasks checklist	Students, residents, clinicians	Goldszmidt et al., 2013 Covin et al., 2020

Table 2 presents a summary of the assessment methods and tools linked to the D-CREDO LOs.

Table 2. Assessment methods and tools

Category / digital tool	Learning objective *At the end of the course students are able to…*	Assessment methods	Tools
General	explain the potential benefits of digital technologies in clinical reasoning and list their strengths and limitations	Group discussion / Debates Oral presentation Quiz (MCQs, EMQs, SAQs) Script Concordance Tests (SCT) Written assignment / Essay Concept maps Case (VP, clinical scenario, AI-generated case) E-Portfolio Observation Peer review Peer- and self-assessment (checklists, sample solutions) Reflective assignment (essay / structured report) Feedback (automated, peer-, tutor-)	DTI CDI PNS CRT SSAR ART
	discuss the ethical and legal aspects of using digital technologies in the clinical reasoning process		
	evaluate the validity and reliability of the output of digital technologies in the clinical reasoning process		

	evaluate the use of digital technologies in various clinical settings considering factors such as timing, workload, workflow, and integration in the healthcare team			ART REACT CRT LCJR
	make and justify clinical decisions based on data from digital technologies			IDEA, ART, DTI, CDI CRT, PNS REACT
	explain the meaning and value of the output of digital technologies in an understandable manner appropriate to the target group			DTI CDI SSAR PNS
LLM	apply basic principles of prompt engineering to effectively use LLMs for their clinical reasoning process	Practical presentation / demonstration (prompt engineering) Observation Case (VP, clinical scenario, AI-generated case) Reflective assignment (essay / structured report) Peer- and self-assessment (prompts, prompt-generated outputs) Feedback (automated, peer-, tutor-)		CRT ART SSAR PNS DTI IDEA
	evaluate potential influences on their own clinical reasoning process when using LLMs	Group discussion / Debates Reflective assignment (case analysis with bias identification tasks; comparative reasoning exercises) Peer- and self-assessment (prompt-generated outputs)		DTI CRT, ART, LCJR, IDEA, SSAR, PNS
AI for image analysis	use the output of the AI-generated image analysis to reflect on their own diagnostic process, such as making the differential diagnoses	Quiz (MCQs, EMQs, SAQs) Case (VP, clinical scenario, AI-generated case) Reflective assignment (comparative reasoning exercises)		CDI, SSAR DTI, PNS CRT, ART
	evaluate the impact of AI-generated imaging on clinical decision-making compared to conventional diagnostic methods	Group discussion / Debates Case Reflective assignment (comparative reasoning exercises) Feedback (automated, peer-, tutor-)		CDI, ART, LCJR PNS, CRT DTI

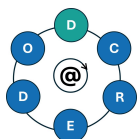|			
CDSS	use the CDSS effectively and responsibly in the clinical reasoning process	Quiz (MCQs, EMQs, SAQs) Key feature cases (KFCs) Concept maps Case	REACT ART, CRT CDI, DTI
EHR	analyze and document patient data within the EHR and create management plans	Case (VP, clinical scenario, AI-generated case) Practical presentation / demonstration Observation Written assignment Oral presentation Concept maps Peer- and self-assessment (checklists, sample solutions) Feedback (automated, peer-, tutor-)	PNS, CRT IDEA, ART SSAR, LCJR
	create management plans collaboratively with the healthcare team within the EHR		REACT CRT, IDEA, PNS, LCJR
	monitor patient outcomes over time and adjust their strategies accordingly within the EHR		ART, CRT, REACT, LCJR
Telehealth	explain when to use telehealth methods for consultation and remote diagnosis considering different contexts	Group discussion / Debates Case (VP, clinical scenario) Reflective assignment (essay / structured report) Feedback (peer-, tutor-)	DTI, ART CRT LCJR CDI, SSAR, PNS
	reflect on how the clinical reasoning process differs in a telehealth setting compared to in-person clinical settings, identifying the unique challenges and opportunities posed by virtual consultations		SSAR, LCJR DTI, CDI, PNS, CRT
	conduct a simulated telehealth visit considering the boundaries of remote consultation	Simulation / Role-play (telehealth visit) Observation & Debriefing Feedback (peer-, tutor-) Peer- and self-assessment (checklists)	REACT, LCJR, CRT IDEA, PNS
mHealth	integrate mHealth technologies into shared decision-making, taking into account patient preferences and context	Simulation / Role-play Case (VP, clinical scenario, KFCs) Group discussion / Debates Observation Feedback (peer-, tutor-)	REACT ART CRT LCJR
	describe how mHealth apps and wearables can be used in routine patient care, for instance continuous patient monitoring and timely intervention	Group discussion / Debates Case Written Assignment Oral Presentation Concept maps	CDI DTI CRT SSAR

5   Quality checklists for LUs

To ensure the pedagogical rigour and instructional coherence of the D-CREDO LUs, we have developed quality checklists based on well-established frameworks for course design and evaluation of didactic materials. The checklists guide the systematic review of LUs across the following key dimensions: alignment with blueprint and learning outcomes; instructional strategies; content quality; learner engagement; and assessment design. Drawing on the Quality Matters Higher Education Rubric (Quality Matters, 2023) and the OLC Quality Scorecard Suite (Online Learning Consortium, 2023), the checklists emphasise alignment, accessibility, and learner support. They also incorporate validated indicators from the system for evaluating didactic materials in online education (Marciniak & Rivera, 2021) and are informed by Kern's six-step approach to curriculum development in medical education (Kern et al., 2009).

Quality checklist for monitoring LU development

To ensure the collaborative, high-quality development of LUs, a set of guiding quality criteria has been established. These criteria support the creation of structured, blueprint-driven LUs and ensure that all LUs are pedagogically coherent, appropriately scoped and aligned with project educational goals. This checklist (see Table 3) is designed to guide the creation of LUs and facilitate the formative peer review process. When assessing the following aspects, reviewers should refer to the LU Blueprint and LU Description workbook.

Table 3. Quality checklist for monitoring LU development

| Criterion	Yes / No / Partial	Comments
The LU is developed in accordance with the approved blueprint and includes all relevant specifications and detailed content.		
The LU description Excel workbook uses the official project template and contains all required worksheets.		
The LU title clearly reflects its core content and is understandable for the intended learner audience.		
The target audience, learner level, prerequisites, and estimated workload are clearly defined.		
The overarching and specific learning objectives are clearly defined and aligned with the LU's content and instructional activities.		
The LU is explicitly linked to the relevant D-CREDO digital health tool(s).		
The instructional design clearly separates synchronous and asynchronous phases, with all instructional steps and events described in sufficient detail.		
The LU aligns with the overarching course goals and is connected to other units, thereby reinforcing a cohesive and cumulative learning experience.		

Quality checklist for the formal review of LU

To support the consistent development and refinement of LUs within the D-CREDO project, a structured quality checklist has been designed for formal review (see Table 4). Organised into thematic sections, the checklist reflects the key elements of instructional design and implementation. It provides specific quality criteria to ensure that each LU is logically structured, easy to navigate and aligned with pedagogical best practice.
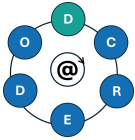
The criteria focus on three core areas: the clarity and structure of the LU; the relevance, accuracy and accessibility of the learning content and resources; and adherence to ethical standards, such as copyright compliance and inclusivity. The checklist also emphasises the importance of creating engaging learning experiences that activate prior knowledge, encourage application of knowledge, and support its long-term retention. Assessment design is also a key component, promoting a variety of methods, including peer and self-assessment, to ensure learners have multiple well-aligned opportunities to demonstrate their progress.

Table 4. Quality checklist for the formal review of LU

Criterion	Yes / No / Partial	Comments
OVERALL DESIGN		
The LU starts with the general information section		
The LU description outlines its aims, scope and structure		
The LU description specifies teaching formats and methods suitable for the stated learner group		
The overarching and specific learning objectives of the LU are clearly stated		
The prerequisites and/or prior knowledge required are stated		
The estimated time/workload for each activity or chapter is appropriately indicated		
The LU has a clear and logical structure and uses accessible, learner-focused language		
The LU contains a bibliography with appropriate recommended sources		
The technology and tools used in the LU support the learning objectives and promote learner engagement and active learning		
The LU includes opportunities for learners to provide feedback on the unit's technical performance and content clarity to support iterative improvements.		
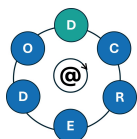
LEARNING CONTENT AND ACTIVITIES		
Learning content and activities are appropriate to the target group of learners and to the learner's level of competence		
Learning content and activities are aligned with the relevant CR themes and D-CREDO digital health tools		
Learning content and activities are organized in a clear, sequential flow to enhance comprehension		
Learning content and activities are constructively aligned with learning objectives and assessment methods		
Learning content and activities emphasize practical application of digital health tools in CR		
Learning content and activities are accurate, up-to-date, reflect current guidelines and practice and have appropriately cited references		
Learning content incorporates strategies like case-based learning (storytelling) and real-world examples to maintain learners' motivation		
Learning content is presented in a legible, visually attractive (but not overloaded) manner and includes engaging multimedia resources		
The LU consists of varied teaching/learning formats to accommodate diverse learning preferences		
The LU offers a variety of active learning activities and resources aligned with LOs (e.g. quizzes, case-based scenarios, VPs, gamification)		
Learning content reflects diversity, equity, and inclusion, taking into account cultural, social and ethical considerations in medical practice		
The learning content of the LU respects copyright laws		
ASSESSMENT AND FEEDBACK		
The assessments are clearly aligned with the learning objectives and measure their achievement		
The assessment instructions explain the purpose and requirements of the assignment		
The assessment process and marking/grading strategy are explained		

The LU incorporates formative, low-stakes assessments to encourage learner engagement		
The formative assessment is integrated into the learning materials and case-based activities such as working on relevant VPs		
The LU includes assessments with practical, real-word applications		
The LU provides learners with opportunities to track their learning progress and receive timely feedback		
The LU supports learner metacognition through encouraging self-reflection activities		
The LU includes collaborative learning activities and assessments that reflect workplace practices		
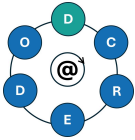
6   Quality checklists for VPs

To ensure the pedagogical soundness, contextual relevance and technical integrity of the virtual patients (VPs) developed within the D-CREDO project, we created structured quality checklists for formal, didactic and content reviews. Based on the comprehensive guidelines and review tools developed by the iCoViP project (https://icovip.eu/), these checklists have been adapted to align with D-CREDO's focus on digital health technologies in clinical reasoning and the project's specific objectives. The checklists aim to support the consistent, high-quality design and integration of VPs into LUs, thereby fostering deliberate practice and the achievement of learning goals.

These checklists underwent several stages of editing, discussion and validation. The initial framework was created by a small working group in WP4 and discussed and edited by all consortium partners in the shared file. The edited versions were then discussed and upgraded at the D-CREDO consortium meeting in Munich. This included brainstorming sessions and reviews from different perspectives and backgrounds. All comments and ideas were considered in the final editing of the framework presented in this document. The quality checklists now focus exclusively on the needs and aims of the D-CREDO project, providing a relevant and coherent assessment tool.
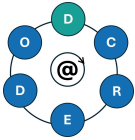
Quality checklist for the formal and didactical review of the VP

Metadata

- ☐ The metadata is complete and correct: the VP name has the correct prefix, the language and licence are selected, the author's names provided
- ☐ The VP corresponds to the outline in the blueprint
- ☐ The VP is aligned with the LU's learning objectives

Structure and text

- ☐ The VP has an appropriate number of cards (5-10)
- ☐ The cards are well organised, with a reasonable quantity of aspects/ information on each
- ☐ The cards are labelled according to their content
- ☐ Direct speech is used for the conversation and a 'hide-and-reveal' mode is used for the dialogue throughout the case
- ☐ The 'Introduction of the Patient/First Impression' card provides information about the scenario, the learner's professional role, the patient's demographics, and the patient's key symptoms/presenting complaint
- ☐ The 'History Taking' card presents a dialogue between the patient and the healthcare professional
- ☐ The 'Physical Examination' card provides a description (audio, image) of the physical examination findings and enables students to interpret them (can be accompanied by a question).
- ☐ Students are asked to compose a summary statement on card 3 (after the physical examination)
- ☐ The 'Tests and Examinations' cards present test results and image material wherever possible and allow students to interpret the findings (can be accompanied by questions).

- ☐ Students are then asked to make a final diagnosis using the concept mapping tool and consider treatment options on the final 'Tests and Exams' card
- ☐ The final diagnosis is revealed only after the students have made their decision in the concept map.
- ☐ The next card includes providing a diagnosis to the patient and discussing treatment options with them. It concludes by describing what happened next (e.g. the patient was discharged, died, etc.).
- ☐ The final card provides one to three main, relevant references for further information (preferred sources are open access)
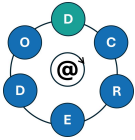
Multimedia material

- ☐ Card 1 shows a matching patient image in a clinical setting
- ☐ The patient image matches the patient description in the text
- ☐ All clinically relevant media material, such as images, videos, and audio files, are included, showing the described findings/pathologies
- ☐ Information about the licence is entered for each multimedia item

Questions

- ☐ The appropriate question types are used
- ☐ The questions do not repeat what students should do in the concept map
- ☐ Questions prompt students to interpret multimedia data (e.g. images, audio)
- ☐ Questions about interpretation of test results include a 'normal' or 'no pathological findings' option
- ☐ The questions and answers follow didactic principles, such as avoiding negations and providing clear, comparable options and reasonable discriminators (incorrect answers)
- ☐ The solutions are correctly marked
- ☐ Answer comments explain the correct and incorrect solutions (in the case of interpreting images, an image with illustration arrows may be provided)

Digital health tools

- ☐ The digital health tool is integrated into the VP. If so, please identify the type:
  - includes a built-in digital health tool (e.g. a CDSS)
  - linked to a simulation of the digital health tool (e.g. an EHR interface)
  - presents AI-augmented content (e.g. AI-analyzed X-ray, ECG or CT scan)
  - includes mHealth data (e.g. data from a patient's wearable device)
  - includes the use of LLMs for CR augmentation or reflection on the CR process
  - simulates telemedicine consultation
  - other, please specify …
- ☐ The tool is integrated in such a way that students have to interact with it and/or elaborate on the results
- ☐ Clear instructions on what the student is supposed to do with the integrated tool.
- ☐ The tool can be easily accessed and no payment is required

Concept mapping tool (if integrated in the VP)

- ☐ It has a reasonable number of findings, including the key symptom
- ☐ It has a reasonable number of differentials
- ☐ It has a reasonable number of tests
- ☐ It has a reasonable number of treatment options
- ☐ It has a reasonable number of connections
- ☐ Includes working diagnosis, must-not-miss and ruled-out diagnoses, and negations (if appropriate)
- ☐ There is a clear final diagnosis
- ☐ Progress is in sync with case progression

References / Links

- ☐ They are current, i.e. no older than 10 years (preferably 5 years or less)
- ☐ They are freely accessible (i.e. not behind a paywall or only accessible through educational institutions)
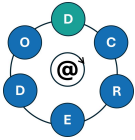
Quality checklist for content review of the VP

Relevance

- ☐ The diagnostic and management processes described are based on current guidelines
- ☐ Any deviations from the guidelines are explained (e.g. in an expert comment), with appropriate references included to emphasise the implications for clinical reasoning
- ☐ The details of the diagnostic and management processes are relevant to learners' CR competence level
- ☐ The integrated/linked digital health tool is relevant and informative for the VP

Quality

- ☐ Appropriate language and medical terminology are used consistently across the case
- ☐ The multimedia material is provided for all relevant findings and clearly shows the described findings and pathologies
- ☐ The content describes a realistic clinical case and is accurate from a clinical point of view
- ☐ The integrated digital health tools are valid, reliable, and up to date
- ☐ The proposed use of the digital health tools corresponds to how clinicians could support their clinical reasoning
- ☐ The concept map includes all the relevant key findings, differential diagnoses, tests, treatment options, and connections
- ☐ The provided references and links are relevant to the symptoms/diagnoses in question, and are suitable for students

7 LUs evaluation tool (satisfaction questionnaires)

The LUs evaluation tool (satisfaction questionnaires) is based on the Evaluation of Technology-Enhanced Learning Materials (ETELM) tool, which forms part of a broader framework developed by Cook and Ellaway (2015) for the systematic assessment of technology-enhanced learning in medical education. The ETELM tool was designed based on a theoretical model developed by the authors, informed by the analysis and reuse of items from several established evaluation tools (Cook & Ellaway, 2015). This tool comes in three versions: ETELM-LP (Learner Perceptions), ETELM-LP-S (Learner Perceptions – Short Form) and ETELM-IP (Instructor Perceptions), developed to evaluate courses from the perspectives of learners and instructors. The instrument was validated using a sample of potential users, including educators and students. The authors permit adaptation of the tool to context and specific situational needs.

The original ETELM instrument was adapted for use in the DID-ACT project (https://did-act.eu/). It was successfully used there to evaluate the Student Clinical Reasoning Curriculum from the perspectives of students and instructors, as well as Train-the-Trainer courses on clinical reasoning from the perspectives of participants and facilitators. In D-CREDO, we have decided to use the same adapted versions of the ETELM instrument, making only minor modifications to the demographic section. Depending on the evaluation setting, the questionnaires will be administered either in paper form or via a web-based survey tool. For the online questionnaire, we will use the D-CREDO project's LimeSurvey web application installation maintained by Instruct. For the purposes of the D-CREDO project, the surveys will be available in English, Polish and Ukrainian.

Learning unit evaluation (Learner Perceptions)

Thank you for participating in this evaluation survey. Your feedback is essential to help us improve the quality and effectiveness of the learning units developed within the D-CREDO project. This questionnaire invites you to reflect on your experience with the learning unit you have just completed. It includes both scaled and open-ended questions to capture your perceptions of engagement, instructional quality, assessment alignment, and learning outcomes.
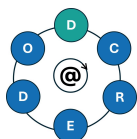
The survey is based on an adapted version of the *Evaluation of Technology-Enhanced Learning Materials: Learner Perceptions (ETELM-LP) evaluation tool by Cook DA, Ellaway RH. Med Teach. 2015;37(10).*

Which learning unit are you evaluating?

Free text field

Which institution do you work/study at?

- Jagiellonian University
- Erasmus MC
- UMIT Tirol
- BSMU
- Other (please specify)

What educational programme do you relate to?

- Medicine
- Nursing
- Other (please specify)

What is your primary role/roles at your institution? (multiple answers possible)

- Student
- Teacher
- Researcher
- Resident/intern doctor
- Other (please specify)

How many years of experience in healthcare education (excluding years of study) do you have? (faculty only)

- less than 2 years
- 2-5 years
- 6-10 years
- more than 10 years

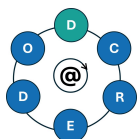**What is your current year of study ?** (students only)

- 1 - 2 - 3 - 4 - 5 - 6 - internship
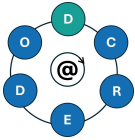
What is your gender?

- Female
- Male
- Prefer not to disclose

Please indicate your agreement with each of the statements below. Answer questions based on your experiences, unless otherwise noted. If you strongly agree or disagree with any item (ratings of 1 or 7), we kindly ask you to provide a brief explanation in the comment section at the end of the survey.

No	Question	Strongly disagree						Strongly agree
Learning activities								
1	The educational activities encouraged engagement with learning unit materials / content.	1	2	3	4	5	6	7
2	The educational activities promoted achievement of the learning unit objectives.	1	2	3	4	5	6	7
3	Educational activities encouraged interaction and collaboration with other participants	1	2	3	4	5	6	7

4	The learning unit effectively blended online and face-to-face elements.	1 2 3 4 5 6 7
Assessment/Feedback		
5	Assessments (e.g. tests and self-assessments) were appropriate for the learning unit objectives, content, and activities.	1 2 3 4 5 6 7
6	I had sufficient opportunity to assess and reflect upon my learning progress.	1 2 3 4 5 6 7
7	I received adequate feedback on my learning progress.	1 2 3 4 5 6 7
8	I received adequate support for any questions or concerns I had about my learning.	1 2 3 4 5 6 7
Impact		
9	This learning unit will change my practice of clinical reasoning.	1 2 3 4 5 6 7
10	The overall quality of this learning unit was excellent.	1 2 3 4 5 6 7
11	The overall effectiveness of the instructor(s) was excellent.	1 2 3 4 5 6 7
Free text comments		
12	How could the quality of the learning unit be improved? What would you change, remove, or add? Please describe and explain.	
13	Overall, what elements of this learning unit most contributed to your learning, excitement and engagement as a learner? Please describe and explain.	
14	Further comments:	

Learning unit evaluation (Instructor Perceptions)

Thank you for participating in this evaluation survey. Your feedback is essential to help us improve the quality and effectiveness of the learning units developed within the D-CREDO project. This questionnaire invites you to reflect on your experience with the learning unit you have just completed. It includes both scaled and open-ended questions to capture your perceptions of engagement, instructional quality, assessment alignment, and learning outcomes.

The survey is based on an adapted version of the *Evaluation of Technology-Enhanced Learning Materials: Instructor Perceptions (ETELM-IP) evaluation tool by Cook DA, Ellaway RH. Med Teach. 2015;37(10).*

Which learning unit are you evaluating?

Free text field

Which institution do you work/study at?

- Jagiellonian University
- Erasmus MC
- UMIT Tirol
- BSMU
- Other (please specify)

What educational programme do you relate to?

- Medicine
- Nursing
- Other (please specify)

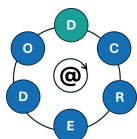**What is your primary role/roles at your institution?**

- Teacher
- Researcher
- Curriculum Planner/Manager
- Other (please specify)

How many years of experience in healthcare education (excluding years of study) do you have?
- less than 2 years
- 2-5 years
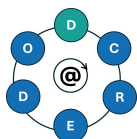- 6-10 years
- more than 10 years

What is your gender?

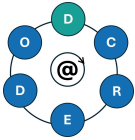- Female
- Male
- Prefer not to disclose

Please indicate your agreement with each of the statements below. Answer questions based on your experiences, unless otherwise noted. If you give any item a 1 (strongly disagree) or 7 (strongly agree), please be sure to comment on this at the end of the survey.

No	Question	Strongly disagree						Strongly agree
Learning activities								
1	Instructions provided a good introduction to the course (e.g., participants did not have obvious questions about how to get started, where to find various course components, how to obtain technical support if needed, etc.).	1	2	3	4	5	6	7
2	Course objectives were relevant to participant needs.	1	2	3	4	5	6	7
3	Navigation of the technology-based components of the course was logical, consistent, and efficient.	1	2	3	4	5	6	7
4	The course technologies and media supported the learning objectives.	1	2	3	4	5	6	7
5	This course required that participants possess inappropriately high computer skills. (negative answer is better; if agree, please explain below)	1	2	3	4	5	6	7
6	The educational activities encouraged participants' engagement with course materials / content.	1	2	3	4	5	6	7
7	The educational activities promoted participants' achievement of the course objectives.	1	2	3	4	5	6	7
8	I was able to contribute a personal presence / personal touch during the course's development and/or delivery.	1	2	3	4	5	6	7
9	Educational activities encouraged participants' interaction and collaboration. [Omit if there was not a collaborative element in the course]	1	2	3	4	5	6	7
10	The course effectively blended online and face-to-face elements. [Omit if there was not a face-to-face element in the course]	1	2	3	4	5	6	7
11	Face-to-face activities contribute meaningfully toward achieving the course learning objectives. [Omit if there was not a face-to-face element in the course]	1	2	3	4	5	6	7
12	Assessments (e.g. tests and self-assessments) were appropriate for the course objectives, content, and activities. [Omit if there were no assessment opportunities in the course]	1	2	3	4	5	6	7

13	Learner assessments and provision of feedback proceeded smoothly (i.e., no unforeseen problems). [Omit if there were no assessment opportunities in the course]	1 2 3 4 5 6 7
14	I plan to use learner feedback to improve the course.	1 2 3 4 5 6 7
15	The course will be easy to maintain and deliver again.	1 2 3 4 5 6 7
16	It will be easy to re-use all or part of the course materials in other, future courses.	1 2 3 4 5 6 7
17	I had access to needed tools during course development and delivery.	1 2 3 4 5 6 7
18	I had significant computer / technical problems while developing or delivering this course. (negative answer is better; if agree, please explain below)	1 2 3 4 5 6 7
19	I received adequate support for any technical issues encountered while developing and delivering this course.	1 2 3 4 5 6 7
20	I was able to provide adequate support to students for questions or concerns about their learning.	1 2 3 4 5 6 7
21	The course was a good use of time and resources.	1 2 3 4 5 6 7
22	The overall quality of this course was excellent.	1 2 3 4 5 6 7
The most helpful part of feedback is specific suggestions to improve the course. **Please take a few minutes to share your ideas.**		
23	What could have been done to improve the process of course development and implementation?	
24	How could the quality of the course be improved? What would you keep the same? What would you change, remove, or add? Please describe and explain.	
25	Overall, what elements of this course most contributed to your excitement and engagement as an instructor? What could have been done to improve your engagement? Please describe and explain.	
26	Is there anything else you'd like to share about your experience in this course? If so, please describe and explain.	

This LUs evaluation survey enables correspondents to systematically assess how instructors perceive the design, implementation and impact of educational materials. It provides insight into the usability and pedagogical value of the materials, as well as the practical barriers to their implementation. It also complements the learner-focused survey, providing a comprehensive evaluation of the courses.

8 Learning analytics

Introduction

In the D-CREDO project, learning analytics is understood as the collection and analysis of automatically captured behavioural data, reflecting learners' interactions with educational activities encoded in the technical educational infrastructure of the project.

The technical educational infrastructure of D-CREDO comprises of:

1. The project learning management system Moodle where the D-CREDO LUs are hosted as individual courses. The Moodle system for the D-CREDO project expands the platform deployed for the DID-ACT project and has been updated to version 4.5. The project Moodle installation is available here: https://d-credo.instruct.eu

2. Virtual patient cases are hosted at the CASUS platform: https://player.casus.net

3. Linked digital health tools activities (AI in image processing, LLMs, mHealth, EHRs, telehealth).
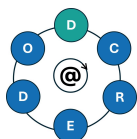
The goal of learning analytics in D-CREDO is to contribute to quality improvement by monitoring identified indicators of learning engagement and clinical reasoning success in the LUs and VPs hosted by the project's technical infrastructure.

We adopted the definition of engagement as "the emotional, behavioural, cognitive, or social energy and effort students direct towards learning" (Bond et al., 2023; Bergdahl et al., 2024). In our analyses, we focus mainly on the behavioural aspects of engagement, as this is the component most readily captured in the Moodle learning management system.

The cognitive component is in our setting related to indicators of clinical reasoning success recorded by the virtual patient system CASUS. Clinical reasoning is defined in the D-CREDO project as "health professionals thinking and acting in assessment, diagnostic, and management processes in clinical situations, taking into account the patient's specific circumstances and preferences" (Huesmann et al., 2023). The operationalisation of the clinical reasoning construct in the virtual patients is achieved through such measures as diagnostic accuracy, the presence of specific elements and connections added to concept maps constructed by students while solving the VPs, or textual summary statements of the patient's problems.

The emotional component of engagement is captured through satisfaction questionnaires, following the ETELM model, as described in the previous chapter of this report.

Finally, as D-CREDO is a blended learning curriculum in which the digital components are primarily used to facilitate asynchronous guided learning in preparation for synchronous classes, the social aspects are mainly implemented in the classroom. Therefore, we acknowledge that the social component of engagement is often not applicable for monitoring through learning analytics measures in this project.

Development process

Overall design

The learning analytics plan was devised by a small working group consisting of representatives from three partner institutions: Instruct, JU, and UMIT. The collaboration was carried out through a series of online meetings, consultation activities with the consortium (including a consensus-building activity during a face-to-face meeting), and asynchronous research.

The group proceeded with a top-down approach, starting from the general aims of learning analytics through specification and prioritization of analytics questions, selection of high-level indicators, down to concrete data points recorded in our learning environments.

We grouped the identified data points in two types of reports: 1. Real-time/Formative Reports - based on standard reporting tools, supporting immediate or regular semi-automatic feedback for educators. 2. Summative Reports - more open and tailored to specific educator needs and potentially requiring custom on-demand scripting.

The learning analytics workshop

The card sorting game was a needs analysis and prioritisation activity carried out during the second face-to-face meeting of the D-CREDO consortium on May 28, 2025, in Munich. Its purpose was to support decision-making regarding what kinds of questions could be posed about student activities within the learning unit (on Moodle) and their interactions with the associated virtual patients (on CASUS).

Participants, working in small groups, arranged five pre-prepared learning analytics questions printed out on flashcards in order of relevance. The most relevant placed on top, indicating a stronger desire to see that question answered during the course. Participants were also invited to suggest additional questions that they believed would be useful and realistically answerable based on system logs. Finally, if they felt any of the existing questions should be rephrased, they were asked to explain how and why by writing the revised version directly on the card.

The 5 starter cards (questions) were:

C1. **How long** do my students engage with the activities & materials provided in my course?

C2. **How often** do my students engage with the activities & materials provided in my course?

C3. **How regularly** do my students engage with the activities & materials provided in my course?

C4. **How successfully** do students engage with the activities & materials provided in my course?

C5. How does engagement of my students on given activities & materials **develop over time** in the course?

In total 21 project members participated in the meeting (Fig. 1). From that 16 participate face-to-face and 5 online. They formed in total 5 working groups (G1-G5): 4 in-person groups and 1 on-line group.
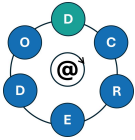
Fig. 1. Card sorting activity during the second F2F meeting in Munich

The results, including the ranks and comments, were recorded in a spreadsheet. To summarise the voting, each card received a score calculated as 6 minus its rank (meaning each score was in the range of 1 to 5). One group decided to assign the same rank to all cards, as all questions appeared to them to be equally interesting and reasonable, and needed to be considered simultaneously in order to gain a clearer overall picture.

The results for the rest of the groups are presented in figure 2.
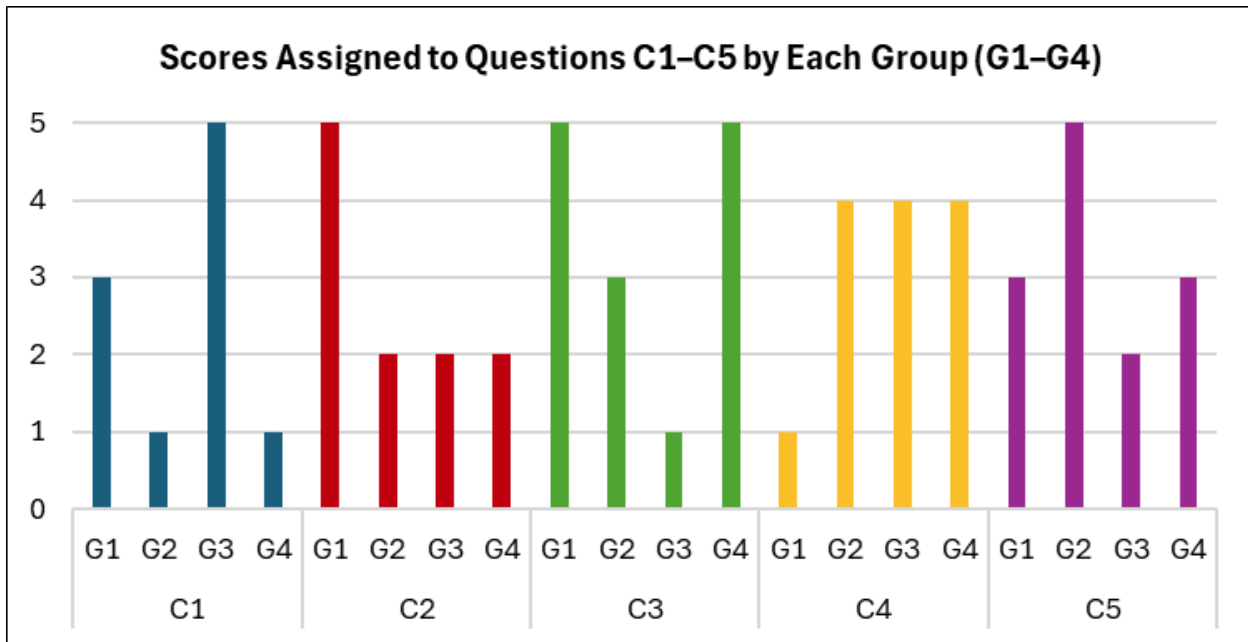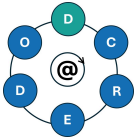
Fig 2. Scores Assigned to Questions C1–C5 by Each Group (G1–G4)

The total scores of the groups for the questions were as follows: C1 = 10 pts; C2 = 11 pts; C3 = 14 pts; C4 = 13 pts; C5 = 13 pts. This confirms the view that it was difficult to select the most relevant question. For instance, the question that received the highest score (C3), which concerned the regularity of learning, was ranked as the most relevant by two groups, but placed last by another.

The cards added by the groups included the question: "Do they [the students] initiate discussion about activities and materials (with peers or teachers)?" There was also a suggestion to triangulate the learning analytics scores with both quantitative and qualitative feedback from surveys and in-class observations, which are not captured by learning analytics.

Finally, the individual comments on the cards highlighted signs of confusion about the differences between some of the C2 and C3 (regarding the distinction between "how often" and "how regularly"). One comment noted that the word "successful" (C4) is not precise enough. Another explicitly stated that time spent on an activity (C1) is not the most appropriate indicator of engagement. There was also a remark that C5 applies only to learning units of longer duration.

Selection of electronic learning activity monitoring metrics

The analytic questions identified, prioritised, and discussed during the workshop were analysed in the context of the types of data available in D-CREDO's technical education infrastructure. This process resulted in the identification of the following high level indicators, each operationalised through concreted observable analytic data points available in the learning software (see Table 5 and Table 6).
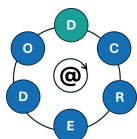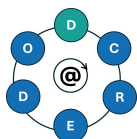
Table 5. Standardized Moodle Metrics to answer our analytic questions

Question	Pedagogical Focus	Moodle Metric	Evaluation	Note
C1: How long do my students engage with the activities & materials provided in my course?	Time on task	Not directly measurable with standardized Moodle metrics/without plug-ins Logs/Activity report: Exportable list (CSV/Excel) with one row per event: date and time, student ID, type of action (e.g., "view", "submitted" …). For each student the first and last timestamp of a session will be compared	Formative & Summative	No direct metric without plug-ins Indicator possible to calculate by comparing log timestamps (first/last actions)
C2: How often do my students engage with the activities & materials provided in my course?	Frequency of participation/ interaction	Logs/Course Participation/Activity Completion/Activity Report/*opt. Forum Report:* Tables with one row per student and columns (e.g., Number of views, number of submissions…) Absolute counts; for example Student1: 15 views, 3 submissions… Forum report shows number of posts and replies by each participant	Formative & Summative	Absolute count of interactions (not time-based) Forum activity report (if forums were used)
C3: How regularly do my students engage with the activities & materials provided in my	Rhythm/ Consistency of engagement	Participants page/Logs aggregated by week/Analytics Overview per student with the date and time of	Formative	Distribution of interactions over time Exclusion of teacher/admin

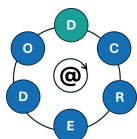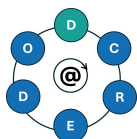| course?		last access Time-series table or chart: how many learners were active and graph of inactive students		activity
C4: How successfully do students engage with the activities & materials provided in my course?	Quality of Engagement/ Goal Completion	Activity Completion/Course Completion/User report/Quiz statistics/Assignment submissions Tables with check metrics "yes/no" for each activity plus date of completion Gradebook view Quiz statistics (average.; max. Scores; % of correct answers) Assignment submissions showing submission times and status ("submitted")	Summative	Grades could also support formative feedback. However, not all assignments can be graded automatically.
C5: How does student engagement with activities & materials develop over time?	Trends over time/ Development	Logs/Activity reports/Analytics/*opt. Forum Report* Time-series chart/table reporting how many views/completions had a student over a longer time. Columns: Date, Total views, Total completions in a time unit. Graphical analysis: increase/decrease in activity across course period	Formative	Formative reports with manual effort Forum activity report (if forums were used)

Table 6. Standardized CASUS Metrics to answer our analytic questions

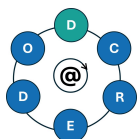| Question	Pedagogical Focus	CASUS Metric	Evaluation	Note
C1: How long do my students engage with the activities & materials provided in my course?	Time on task	Time in minutes spent on different levels (a course of VPs, a specific VP, a card in a VP, or an activity, e.g. quiz question, on a card) by individual students	Formative & Summative	Time on task, is not a very reliable indicator for engagement
C2: How often do my students engage with the activities & materials provided in my course?	Frequency of participation/interaction	Number and timestamps of VP sessions and re-visiting VPs by individual students in given VPs	Formative & Summative	
C3: How regularly do my students engage with the activities & materials provided in my course?	Rhythm/ Consistency of engagement	Total number and timestamps and intervals of VP sessions and re-visiting VPs by individual students over time	Formative	
C4: How successfully do students engage with the activities & materials provided in my course?	Quality of Engagement/ Goal Completion	Scores (quantitative) achieved by individual students or cumulative on the activities in a VP (quiz questions, summary statement, CR concept map) Quality of answers provided by students in these activities	Formative & Summative	
C5: How does student engagement with activities & materials develop over time?	Trends over time/ Development	Number, selected VPs, interactions with VPs, and scores of VP sessions of an individual student over a longer period of time	Formative	

Analysis of suitable plug-ins

The working group considered using plug-ins for Moodle to support implementation of the project reports. Optional components developed by the Moodle community and available for the installed version of the software were analysed. The following plug-ins were reviewed: Learning Analytics API, Engagement, Analytics, Dashboard, and Logstore xAPI (see Table 7).
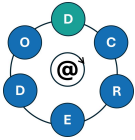
Table 7. Moodle plug-ins

Plugin Name	Function	Indicators	Link
Learning Analytics API	Framework for data-driven analysis.	Prediction of risk of dropping out based on a set of indicators: Activity patterns, course progress, frequency of interaction Comment: These "Predictions" are not helpful in the D-CREDO context (small groups, partly voluntary usage of D-CREDO materials, predictions do not help for quality improvement of LUs and VPs, information split between Moodle and Casus etc.). Besides, training the correct algorithm may be complex.	Analytics API | Moodle Developer Resources
Engagement Analytics	Assesses engagement based on configurable rules and indicators.	Login frequency, forum activity, assignment submissions, time spent in the course Comment: Seems easy to implement, but only has very basic indicators. May be interesting for a "full" module, but probably not for D-CREDO courses, as it is not fine-granular enough to improve LU and VPs.	Engagement Analytics Plugin - MoodleDocs
Dashboard	Allows free setup of the dashboard page layout using blocks.	Not update since 2016 - outdated	Dashboard - Moodle Plugins directory

| Logstore xAPI	When the students view a course, complete a quiz, or grade a student, a log of that activity is recorded in the Logstore. The plugin uses these logs to produce xAPI statements and then sends (emits) them to an LRS (Learning Record Store).	Comment: An API to extract fine-granular data from several tools. Could be interesting to use in aggregation and triangulation of logs from several digital health tools but not for extending the learning analytics functions in	Logstore xAPI - Moodle Plugins directory Moodle Plugins directory: Learning Analytics Log

As none of the analysed plug-ins met our need to implement the analytic question C1 or accelerated the planned report building in C2-C5, we decided, for this stage, to use the standard logging functionality of Moodle and the CASUS reporting tools, with the additional implementation of a custom report for C1 in Moodle.

Regarding the learning analytics of specific digital health tools integrated as examples into the LUs and VPs, it is clear that the selected tools are highly heterogeneous, and there is limited possibility to extend or interfere with their logging functions. For that reason, we decided to record the activation of tools by tracking the use of URLs linking to the tools and LTI connectors. Furthermore, we recommend conducting a qualitative analysis of the outcomes of tool use through assignment results and self-reflection activities in a portfolio format. The specification of research protocols in D4.3 will enable the creation of additional on-demand summative reports with logging functionality for selected tools within the proposed framework of analytic questions C1 to C5.

Conclusions

The D-CREDO comprehensive evaluation toolset provides a robust framework for monitoring and assessing the quality and impact of project outcomes. Through systematic collaboration between partner institutions, we have selected evidence-based assessment methods and tools aligned with learning objectives. We have also created quality checklists to ensure pedagogical rigour, as well as satisfaction questionnaires to capture the experiences of learners and instructors. Integrating these with innovative learning analytics capabilities provides the basis for both formative feedback and summative evaluation of student engagement and achievements. All that leads to the conclusion that the quality criteria for the deliverable as declared in the project proposal has been met completely. The toolset will underpin quality assurance activities in work packages WP3 and WP4, facilitating the development of high-quality learning units and enabling comprehensive evaluation of pilot studies. Combining traditional assessment and evaluation methods with modern learning analytics will advance the field of technology-enhanced clinical reasoning education while maintaining a focus on preparing healthcare professionals to responsibly integrate digital tools into clinical practice.
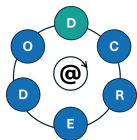
References

Baker EA, Ledford CH, Fogg L, Way DP, Park YS. The IDEA Assessment Tool: Assessing the Reporting, Diagnostic Reasoning, and Decision-Making Skills Demonstrated in Medical Students' Hospital Admission Notes. Teach Learn Med. 2015;27(2):163-73. doi: 10.1080/10401334.2015.1011654.

Bergdahl N, Bond M, Sjöberg J, Dougherty M, Oxley E. Unpacking student engagement in higher education learning analytics: a systematic review. Int J Educ Technol High Educ. 2024;21(1):63. doi:10.1186/s41239-024-00493-y

Biggs J, Tang C. Teaching for quality learning at university. 3rd ed. Maidenhead: McGraw-Hill Education; 2007.

Bond M, Viberg O, Bergdahl N. The current state of using learning analytics to measure and support K-12 student engagement: a scoping review of engagement in K-12 learning analytics. In: Proceedings of the 13th International Conference of Learning Analytics Research (LAK23); 2023 Mar 13-17; Texas, USA.

Bordage G, Grant J, Marsden P. Quantitative assessment of diagnostic ability. Med Educ. 1990;24(5):413-25. doi: 10.1111/j.1365-2923.1990.tb02650.x.

Bordage G, Grant J, Marsden P. Diagnostic Thinking Inventory.instrument. Open University Centre for Education in Medicine, 1991 [cited 2025 Sep 04]. Available from: https://www.researchgate.net/publication/263733394

Cook DA, Ellaway RH. Evaluating technology-enhanced learning: A comprehensive framework. Med Teach. 2015;37(10):961-70. doi:10.3109/0142159X.2015.1009024

Covin Y, Longo P, Wick N, Gavinski K, Wagner J. Empirical comparison of three assessment instruments of clinical reasoning capability in 230 medical students. BMC Med Educ. 2020;20(1):264. doi:10.1186/s12909-020-02185-3.

Daniel M, Rencic J, Durning SJ, Holmboe E, Santen SA, Lang V, Ratcliffe T, Gordon D, Heist B, Lubarsky S, Estrada CA, Ballard T, Artino AR Jr, Da Silva A, Cleary T, Stojan J, Gruppen LD. Clinical reasoning assessment methods: A scoping review and practical guidance. Acad Med. 2019;94(6):902-12. doi:10.1097/ACM.0000000000002618

García-Torres D, Vicente Ripoll MA, Fernández Peris C, Mira Solves JJ. Enhancing clinical reasoning with virtual patients: a hybrid systematic review combining human reviewers and ChatGPT. Healthcare. 2024;12(22):2241

Goldszmidt M, Minda JP, Bordage G. Developing a unified list of physicians' reasoning tasks during clinical encounters. Acad Med. 2013;88(3):390-397. doi:10.1097/ACM.0b013e31827fc58d

Hege I, Kononowicz AA, Adler M. A clinical reasoning tool for virtual patients: design-based research study. JMIR Med Educ 2017;3(2):e21. doi: 10.2196/mededu.8100
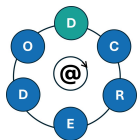
Huesmann L, Sudacka M, Durning SJ, Georg C, Huwendiek S, Kononowicz AA, Schlegel C, Hege I. Clinical reasoning: What do nurses, physicians, and students reason about. J Interprof Care. 2023 Nov 2;37(6):990-998.

Jay R, Sandars J, Patel R, Leonardi-Bee J, Ackbarally Y, Bandyopadhyay S, Faraj D, O'Hanlon M, Brown J, Wilson E. The use of virtual patients to provide feedback on clinical reasoning: a systematic review. Acad Med. 2025;100(2):229-38. doi:10.1097/ACM.0000000000005908

Kern DE, Thomas PA, Hughes MT. Curriculum development for medical education: A six-step approach. Baltimore: Johns Hopkins University Press; 2009.

Kotwal S, Fanai M, Fu W, Wang Z, Bery AK, Omron R, Tevzadze N, Gold D, Garibaldi BT, Wright SM, Newman-Toker DE. Real-world virtual patient simulation to improve diagnostic performance through deliberate practice: a prospective quasi-experimental study. Diagnosis. 2021;8(4):489-96. doi:10.1515/dx-2020-0127

Lasater K. Clinical judgment development: using simulation to create an assessment rubric. J Nurs Educ. 2007;46(11):496-503. doi: 10.3928/01484834-20071101-04.

Marciniak R, Cáliz Rivera C. A system of indicators for the quality assessment of didactic materials in online education. Int Rev Res Open Distrib Learn. 2021;22(1):180-98. doi:10.19173/irrodl.v22i1.5069

Mayer A, Hege I, Kononowicz AA, Müller A, Sudacka M. Collaborative development of feedback concept maps for virtual patient-based clinical reasoning education: mixed methods study. JMIR Med Educ. 2025;11:e57331. doi:10.2196/57331
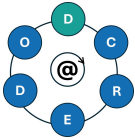
Online Learning Consortium. OLC quality scorecard suite. 2023 [cited 2025 Jul 11]. Available from: https://onlinelearningconsortium.org

Owan VJ, Abang KB, Idika DO, Etta EO, Bassey BA. Exploring the potential of artificial intelligence tools in educational measurement and assessment. EURASIA J Math Sci Tech Ed. 2023;19(8):em2307. doi:10.29333/ejmste/13428

Park YS, Lineberry M, Hyderi A, Bordage G, Riddle J, Yudkowsky R. Validity evidence for a patient note scoring rubric based on the new patient note format of the United States Medical Licensing Examination. Acad Med. 2013;88(10):1552-7. doi:10.1097/ACM.0b013e3182a34b1e.

Plackett R, Kassianos AP, Timmis J, Sheringham J, Schartau P, Kambouri M. Using virtual patients to explore the clinical reasoning skills of medical students: mixed methods study. J Med Internet Res. 2021;23(6):e24723. doi: 10.2196/24723.

Peterson BD, Magee CD, Martindale JR, Dreicer JJ, Mutter MK, Young G, Sacco MJ, Parsons LC, Collins SR, Warburton KM, Parsons AS. REACT: Rapid Evaluation Assessment of Clinical Reasoning Tool. J Gen Intern Med. 2022;37(9):2224-2229. doi: 10.1007/s11606-022-07513-5.

Quality Matters. Higher education rubric for course design. 2023 [cited 2025 Jul 11]. Available from: https://www.qualitymatters.org

Ruczynski LI, Schouwenberg BJ, Custers E, Fluit CR, van de Pol MH. The influence of a digital clinical reasoning test on medical student learning behavior during clinical clerkships. Adv Health Sci Educ Theory Pract. 2024;29(3):935-47. doi:10.1007/s10459-023-10288-x

Sandars J. The use of reflection in medical education: AMEE Guide No. 44. Med Teach. 2009;31(8):685-95. doi:10.1080/01421590903050374

Saputra I, Kurniawan A, Yanita M, Putri EY, Mahniza M. The evolution of educational assessment: how artificial intelligence is shaping the trends and future of learning evaluation. Indones J Comput Sci. 2024;13(6). doi:10.33022/ijcs.v13i6.4465.

Smith S, Kogan JR, Berman NB, Dell MS, Brock DM, Robins LS. The development and preliminary validation of a rubric to assess medical students' written summary statements in virtual patient cases. Acad Med. 2016;91(1):94-100. doi: 10.1097/ACM.0000000000000800.

Swan Sein A, Rashid H, Meka J, Amiel J, Pluta W. Twelve tips for embedding assessment for and as learning practices in a programmatic assessment system. Med Teach. 2020;43(3):300-6. doi:10.1080/0142159X.2020.1789081

Thammasitboon S, Rencic JJ, Trowbridge RL, Olson APJ, Sur M, Dhaliwal G. The Assessment of Reasoning Tool (ART): structuring the conversation between teachers and learners. Diagnosis (Berl). 2018;5(4):197-203. doi: 10.1515/dx-2018-0052.

Williams RG, Klamen DL, White CB, Petrusa E, Fincher RM, Whitfield CF, Shatzer JH, McCarty T, Miller BM. Tracking development of clinical reasoning ability across five medical schools using a progress test. Acad Med. 2011 Sep;86(9):1148-54. doi:10.1097/ACM.0b013e31822631b3.