Co-funded by the
Erasmus+ Programme
of the European Union

TAME

Training Against Medical Error

561583-EPP-1-2015-1-KZ-EPPKA2-CBHE-JP

D4.2, Evaluation plan

Deliverable number	D.4.2
Action Number	
Delivery date	May 2016
Status	Final
Authors	St George's University of London (SGUL)

Co-funded by the Erasmus+ Programme of the European Union

TABLE OF CONTENTS

1. VERSION HISTORY

Version Number	Date	Changes made by	Details
0.1	20/11/15	Jonathan Round SGUL	Initial draft of overview
0.11	23/11/15	Luke Woodham	Additional material added
0.2	19/03/16	Luke Woodham	Additions for early draft
0.3	04/05/2016	Luke Woodham	Additions
0.4	09/05/2016	Luke Woodham	Complete draft for review by partners
1.0	18/12/2018	Luke Woodham	Changed from draft to final

2. INTRODUCTION

Medical Error has been widely identified as a key cause of preventable adverse events in clinical practice (Institute of Medicine, 1999; Makary & Daniel, 2016), but there is potential to reduce the incidence of error by providing appropriate training (Alberti, 2001). Despite this, relatively little emphasis is currently placed on the need to teach avoidance of error in global undergraduate medical curricula.

Simulation offers a particular opportunity for learners to make errors in a safe environment, and to learn by reflecting on these. When errors during such simulation exercises are approached in an educational environment that permits reflection upon negative emotions associated with error, learner awareness of the possibility of errors is raised, and allows them to take responsibility for adapting their practice to avoid making the same or similar mistakes in future (Ziv, Ben-David, & Ziv, 2005). Some experts contend that education should actually induce error in learners as a formative learning experience, and that targeted feedback when learners make errors can help them construct knowledge in ways that aid learners' retention of knowledge and ability to apply that knowledge to real-life circumstances. (Eva, 2009).

Interactive Virtual Patients (VPs), which are web-based simulated patient scenarios have demonstrated the beneficial learning effect of decision making (Ellaway, Poulton, & Jivram, 2015; Poulton et al., 2014), and the flexibility, scalability and low cost of this form of low-fidelity simulation make them well suited for use with medical students (Ellaway, Poulton, Smothers, & Greene, 2009). A model for creating VPs that

leverage this decision-making capability to specifically train students in error avoidance and awareness has been devised at St George's, University of London, based upon a framework for understanding medical errors in practice (Vaughan, Bate, & Round, 2012).

The aim of the TAME project is to develop VP resources based upon this model, and to adapt and implement them in the curricula at the institutions in the project consortium.

2.1. Evaluation Aims and Objectives

This evaluation plan describes the aims, stakeholders and methods used to evaluate the work of the TAME project. The TAME project evaluation will involve a series of activities throughout the lifecycle of the project designed to evaluate all aspects of the project, drawing upon the experience of the project consortium.

The evaluation plan will be approved by the project consortium, and form the basis of the Evaluation Report Deliverable 4.4.

The key aims of the evaluation include:

- Provide a summary of progress of project
- Capture useful research outcomes from the work
- Capture any unintended outcomes
- Disseminate best practice
- Capture experiences of stakeholders
- Disseminate findings
- Inform future work and evaluation of similar projects.

2.2. Interventions

A key project output to be evaluated is the Virtual Patient cases designed to train against medical error. The nature of the educational interventions is fully described in deliverable D1.2 Repurposed paediatric cases in English. The model outlined in this document will also be used as the basis of the VP resources to be created and described in Deliverable D2.1 Paediatric cases modified.

2.2.1. Philosophical Approach

The design of the VP interventions and the teaching strategy adopted by the project is based upon the following philosophical assumptions:

1. Medical and Healthcare Error is a large cause of mortality and morbidity.
2. Medical error is complex, but can be simplified and categorized. A descriptive classification

identifies 10 basic types of error (Vaughan et al., 2012), while a cognitive taxonomy interprets the underlying cognitive processes (Zhang, Patel, Johnson, & Shortliffe, 2004).

3. Errors will occur at specific and predictable points in a patient episode.
4. Medical error requires multiple approaches to combat its effects to address these different causes. These range from system based approaches to practical training to cognitive exercises.
5. Educational strategies are able to address many types of error, firstly by making healthcare professionals aware of the issues, types of error and when they will occur; further theoretical practice of using this information in case based discussions will enable the learnt information to be integrated into clinical thought processes and habits.
6. Tutorials based around real cases are the most effective way of changing behavior and therefore outcomes.
7. Triggering errors in learners can represent an effective learning strategy (Eva, 2009)

3. METHODOLOGY

The evaluation will focus on capturing the experience of key stakeholders in the project to identify and assess its impact and effectiveness.

The methodology used to develop the evaluation plan is based upon the framework described by Frechtling et al (Frechtling Westat, Frierson, Hood, & Hughes, 2002), and builds upon a published model for course evaluations (D. A. Cook & Ellaway, 2015).

The project evaluation will be primarily summative, to assess the effectiveness of the project and its outcomes at the end of the lifetime of the project, or at the end of a project stage. However, there will be formative evaluation components to monitor project progress and effectiveness on an ongoing basis, and these will feed back into the project upon their completion to be carried forward and acted upon during subsequent project activities. The evaluation will use mixed-methods, with quantative methods used to gather feedback from the large learner population, and qualitative methods to gather more in-depth opinions from smaller tutor and partner groups.

3.1. Conceptual Model for TAME

The table below shows a conceptual model for the project, describing the resources, activities and goals of the project. This model forms the basis for identifying the key stakeholders and evaluation questions.

The inputs are the resources that are available at the start of the project. The short-term outputs represent the project deliverables, and can be directly related to the project activities. The long term outputs represent the project's overarching objectives, which will only be realised after the lifetime of the project,

and can therefore not be summatively evaluated as part of the evaluation strategy.

Table 1 Conceptual Model of TAME project

Inputs	Activities	Short-term outputs	Long term outputs
Partner Country Universities (PCUs)	Planning	D1.1 Curriculum plan	Modern curricula based around PBL
Programme Countries (PCs)	Developing VP resources	D1.2 Repurposed paediatric cases in English	Sustainable learning centres
Project Coordination Group (PCG) and Project steering group	Planning	D1.3 Training Plan, staff trained	Improved student learning
Existing curricula and error teaching in PCUs		D1.4 Documented assessment strategy	A range of localised VP and PBL resources
Proposed model for creating error VPs	Implementing Paediatrics VPs	D2.1 Paediatric cases modified	Group of trained PBL tutors
Existing VP PBL cases		D2.2 Cases tested, implemented	
EC Funding		D2.3 Assessment instruments	
	Creating and Implementing additional VPs	D3.1 Training Plan, writing new cases, staff trained	
		D3.2 New cases created	
		D3.3 Cases tested,	

			implemented		
			D3.4 Assessment instruments		
		Evaluation and Quality	D4.1 QC Plan		
			D4.2 Evaluation Plan		
			D4.3 QC Report		
			D4.4 Evaluation Report		
		Dissemination and Sustainability	D.5.1 Project website		
			D5.2 Dissemination events		
			D5.3 Publications		
		Project Management	D6.1 Signed consortium agreement		
			D6.2 Project management plans		
			D6.3 Project reports		

3.2. Key Stakeholders

Having established a conceptual model for the project which summarises the key inputs, activities and outputs, the key stakeholders for the evaluation process can be identified.

Table 2 Table showing identified key stakeholders

Stakeholders	**Persons/spokespersons for each audience**	**Audience's key values, interests, expectations**
Learners	Students at each of the PCUs	Learner experience, learner performance, impact upon workload
Tutors	Tutor staff at each of the PCUs	Training requirements, impact upon workload, learner performance
Tutor Trainers	Project team members at PCs and PCUs	Training strategies, training requirements, documentation requirements
Case writers (creators and adapters)	Project team members at PCUs	Creation and adaptation of resources
Course teams (Curriculum Planning and	Project team members at	Resource requirements (time, rooms,

Course Management)	PCUs	equipment), learner performance
Project Consortium and funding body	Members of the project team, European Commission	Project completion, project progress, effective decision making, project monitoring procedures

3.3. Key Evaluation Questions

Based upon the stakeholder analysis in section 3.2, and the activities and outputs identified in the conceptual model of the project, the following are proposed as key evaluation questions for each individual stakeholder group.

- **Learners**
  - Does the use of error VPs affect learner performance and knowledge relating to medical error?
  - Do the error VPs provide an effective and engaging learner experience?
  - Do the error VPs represent an appropriate workload for inclusion into existing curricula?
- **Tutors**
  - Was the training provided to tutors to facilitate error VP sessions sufficient?
  - Did the use of error VPs significantly impact upon the workload of tutors?
- **Tutor Trainers**
  - Was the training provided to tutors effective?
  - How was the training delivered to tutors?
  - What training documentation was required?
- **Case writers (creators and adapters)**
  - What were the significant challenges when creating/adapting the VP cases?
  - What skills and input were required to create/adapt the cases?
- **Course teams (Curriculum Planning and Course Management)**
  - What resources were required to deliver the error VP cases effectively?
- **Project consortium and funding body**
  - Has project met key milestones/performance indicators?
  - Has project remained in budget?
  - Have project management tasks (decision making, reporting, communication) been carried out effectively?
  - Have project dissemination activities been targeted effectively, and future opportunities identified?

Co-funded by the
Erasmus+ Programme
of the European Union

4. DATA COLLECTION AND EVALUATION INSTRUMENTS

4.1. Data Collection

Two key difficulties have been identified in the data collection and data analysis stages of the project; the geographical separation of the partners and the barriers caused by language differences represent a significant challenge. The proposed solution to these obstacles is to perform all data collection using online tools, using an approach adopted in previous similar projects.

Evaluating the project from the perspective of Learners and Tutors requires the gathering of data from a relatively large sample of participants. For this reason it is intended that the feedback relating to these groups will be collected using an online questionnaire, delivered using the Survey Monkey tool. The questionnaire questions will be initially written in English, and will predominantly be closed-ended and represented using modalities such as Likert scales and multiple choice questions. This will limit the impact of the language barrier by minimising the number of free-text responses required. Where required, translations will be provided either in-line in the survey instrument, or in alternative localised survey instruments. Data collected using localised and translated instruments will be combined into a single dataset for analysis.

Evaluating the project from the perspective of tutor trainers, case adapters and course teams involves collecting data from a smaller sample of participants. These participants will have a higher level of engagement in the project and it is anticipated that they will have both a working knowledge of English and be able to provide detailed responses with a deeper analysis and understanding of their experiences being required. Given these circumstances, the natural approach to data collection from these groups would be in-person semi-structured interviews and focus groups. However, due to the language difficulties this is not deemed to be feasible. Instead, an electronic form (i.e. PDF or Survey Monkey) will be used to gather data from participants in a "written interview" format. Questions will be written in English, and will be predominantly open-ended in nature, requiring free-text entry. Where possible, participants will be asked to respond in English for consistency of data analysis, and content analysis techniques will be used to codify the responses. If translation is required, a nominated member of the project team at the PCU with good English language skills will be asked to assist the participant in both understanding the question and providing a response in English suitable for analysis.

Development of the questionnaires and questions for online semi-structured interviews will be completed with collaboration of all project partners. The responsible partner for the evaluation activity will develop a first draft for questionnaire/written interview question stem, and solicit feedback from other partners. These evaluation instruments will build upon existing work, and where possible make use of existing

validated instruments.

4.2. Data Analysis

Appropriate data analysis techniques will be determined depending upon the final design of the required evaluation instruments.

Survey results will be aggregated using Survey Monkey online tools. Statistical tests appropriate to data sets and sample sizes will be used to establish whether results allow statistically significant conclusions to be drawn from responses.

Responses to online semi-structured interviews will be codified using thematic-analysis techniques, allowing for further identification of common themes and trends in the interview responses. Based upon identified commonalities, a compound interview response illustrating a general view will be synthesized from the individual responses.

The planned collected datasets to be used for analysis are described in Annex 1.

4.3. Evaluation Instruments

The evaluation instruments will be created using a series of collaborative documents, such as Google Docs. An initial draft will be proposed by the deliverable lead, and all partners will be invited to contribute modifications before a final version is agreed.

Based upon the identified key stakeholders and interests, 7 evaluation instruments will be created.

Table 4 Proposed evaluation instruments

Evaluation Instrument	Type of Data to be collected	Notes
E1.1 Learner Experience Survey	Survey (SurveyMonkey)	Based upon validated instruments produced for evaluating VP effetciveness (Huwendiek et al., 2014), affected states (Thompson, 2007), mental strain (Borg & Borg, 2001) and self-efficacy (Bandura, 2006)
E1.2 Learner Motivation Survey	Survey (SurveyMonkey)	Based upon Motivated Strategies for learning Questionnaire (Pintrich, Smith, Garcia, &

		McKeachie, 1991)
E2.x Learner Assessment instrument at PCU	Performance Data	All assessment instruments will adhere to and be described by the assessment strategy documented in Deliverable D1.4. An individual assessment instrument will be created at each PCU site that is tailored to their specific teaching and curriculum. Where multiple PCUs share an assessment structure instruments can also be used at multiple PCUs.
E3 Tutor Experience Survey	Survey (SurveyMonkey)	Based upon an instrument developed for the ePBLNet project (ePBLNet, n.d.)
E4 Case writer (adapter and creator) Written Interview Questions	Written Interview (PDF form)	Based upon an instrument developed for the ePBLNet project (ePBLNet, n.d.)
E5 Tutor Trainer Written Interview Questions	Written Interview (PDF form)	Based upon an instrument developed for the ePBLNet project (ePBLNet, n.d.)
E6 VP Case Implementation – Written Interview Questions	Written Interview (PDF form)	Based upon additional material developed for the eViP project (De Leng, Huwendiek, Donkers, & EViP, 2009)

5. EVALUATION ACTIVITIES

5.1. Summary of Activities

Table 3 Table showing list of TAME evaluation activities

Activity	Evaluation area	Stakeholder/Source of information	Data collection methods/Evaluation Instruments	Timing	
A1	Learner Experience of Paediatric cases	Students at each of the PCUs	E1.1 Learner Experience Survey E1.2 Learner Motivation Survey	Sep-Dec 2016	Jan-May 2017
A2	Learner performance relating to Paediatric cases	Students at each of the PCUs	Randomised Controlled Trial at PCU sites E2.x Learner Assessment instrument at PCU	Sep-Dec 2016 - Timing to be confirmed based upon PCU assessment schedule	Jan-May 2017
A3	Tutor experience of paediatric cases	Tutors at PCUs	E3 Tutor Experience Survey	Nov-Dec 2016	Jan-May 2017
A4	Case adapters experience of paediatric cases	Case adapters at PCUs	E4 Case writer (adapter and creator) Written Interview Questions	Oct-Dec 2016	Jan-May 2017
A5	Tutor trainer experiences of training tutors at PCUs	Tutor trainers (project team members at PCs and PCUs, additional tutor trainers at PCUs)	E5 Tutor Trainer Written Interview Questions	Oct-Dec 2016	Jan-May 2017
A6	Learner Experience of PCU-selected cases	Students at PCUs	E1.1 Learner Experience Survey	Sep 2017-Apr 2018	Apr-Aug 2018

Commented [1]: Timings all need to be agreed and reviewed by all partners

A7	Tutor experience of PCU-selected cases	Tutors at PCUs	E3 Tutor Experience Survey	Sep 2017-Apr 2018	Apr-Aug 2018
A8	Case writer experience of PCU-selected cases	Case writers at PCUs	E4 Case writer (adapter and creator) Written Interview Questions	Sep 2017-Dec 2017	Apr-Aug 2018
A9	Implementation of PCU-selected cases	Course teams at PCUs, project team members at PCUs	E6 VP Case Implementation – Written Interview Questions Quality Report	Apr 2018	Apr-Aug 2018
A10	Project progress	Project team	Project progress reports, Quality report	Ongoing-Aug 2018	Ongoing-Aug 2018

5.2. Timeline of Activities

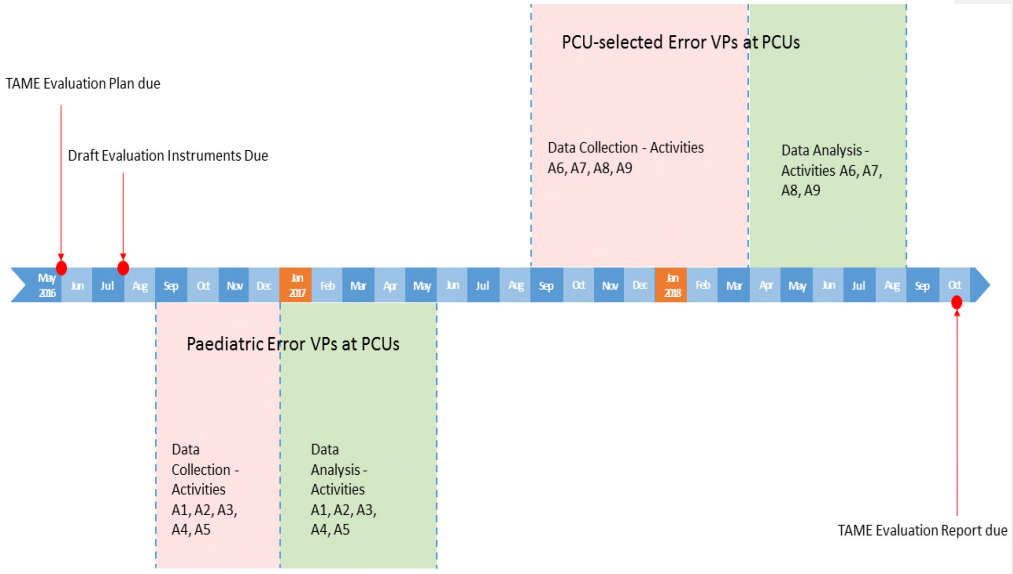

Figure 1 Timeline of TAME evaluation activities

5.3. Evaluation Activity Details

5.3.1. Evaluation Activity A1 – Learner experience of paediatric cases

Participants	Students at each of the PCUs
Instrument(s)	E1.1 Learner Experience Survey E1.2 Learner Motivation Survey
Key Evaluation Questions	Do the error VPs provide an effective and engaging learner experience? Do the error VPs represent an appropriate workload for inclusion into existing curricula?

This survey activity aims to capture learner experiences of the paediatric error VPs at all PCUs. Learners participating in this study will also be required to self-report on measures of self-efficacy, mental strain, affected states and motivation.

Activity A1 will be conducted in consort with a cluster pseudo-randomised controlled trial study methodology, allowing a comparison to be made between those who receive a standard VP intervention (control group), and those who receive an error VP intervention. Each PCU will be asked to identify the number of available students/teaching groups that will participate in the trial at their institution. These clusters will be pseudo-randomised into the control and intervention groups; full randomisation will not be possible due to the need to ensure an equal spread of control and intervention groups at each PCU.

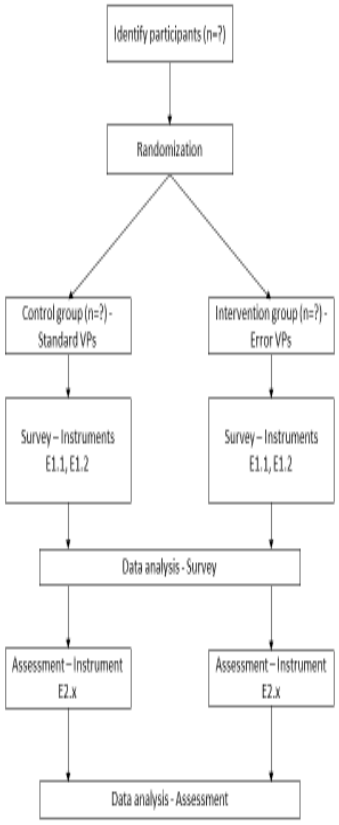

Figure 2 Study Flow Diagram

The survey instruments E1.1 and E1.2 will be delivered to all participant learners at PCUs after they have received the intervention. E1.1 will be completed by each participant after completing each VP, while E1.2 will be administered after all VP cases have been completed.

Data collected using E1.1 and E1.2 will be merged across all PCUs into a single dataset for analysis, possible since the learner experience will be standardised and controlled across all sites.

5.3.2. Evaluation Activity A2 – Learner performance relating to paediatric cases

Participants	Students at each of the PCUs
Instrument(s)	E2.x Learner Assessment instrument at PCUs
Key Evaluation Questions	Does the use of error VPs affect learner performance and knowledge relating to medical error?

As described in Activity A1, Evaluation Activity A2 will also be conducted as part of a randomised controlled trial, and will use the same allocations and interventions as activity A1.

Following exposure to the control and intervention conditions, participants will be asked to complete an assessment. These assessment instruments E2.x will be unique to each partner site, as they will be tailored to the individual partner curriculum as described in D1.4 Assessment Strategy. Scores on these assessments will then be compared and tested using hypothesis testing techniques, to identify whether there is a significant impact upon learning performance resulting from the error VP intervention.

Due to the differing assessment instruments at each PCU, datasets from each site will be kept separate for analysis. Attempts at meta-analysis to generalise findings across a larger population may be made if appropriate.

5.3.3. Evaluation Activity A3 – Tutor experience of paediatric cases

Participants	Tutors at PCUs
Instrument(s)	E3 Tutor Experience Survey
Key Evaluation Questions	Was the training provided to tutors to facilitate error VP sessions sufficient? Did the use of error VPs significantly impact upon the workload of tutors?

This survey activity aims to capture tutor experiences of the paediatric error VPs at all PCUs. The survey instruments E3 will be delivered to all participant tutors at PCUs after they have received the intervention. Data collected from each PCU will be merged into a single dataset for analysis.

5.3.4. Evaluation Activity A4 – Case adapters experience of paediatric cases

Participants	Case adapters at PCUs
Instrument(s)	E4 Case writer (adapter and creator) Written Interview Questions
Key Evaluation Questions	What were the significant challenges when adapting the VP cases? What skills and input were required to adapt the cases?

This written interview activity is designed to capture experiences of those adapting the SGUL paediatrics error VPs for local use at the PCU. The instrument will be sent out to purposefully selected individuals who have been identified as having adapted cases. Data collected from each PCU will be merged into a single dataset for analysis.

5.3.5. Evaluation Activity A5 – Tutor trainer experiences of training tutors at PCUs

Participants	Tutor trainers (project team members at PCs and PCUs, additional tutor trainers at PCUs)
Instrument(s)	E5 Tutor Trainer Written Interview Questions
Key Evaluation Questions	Was the training provided to tutors effective? How was the training delivered to tutors? What training documentation was required?

This written interview activity is designed to capture experiences of those training tutors at the PCU. The instrument will be sent out to purposefully selected individuals who have been identified as having been involved in training tutors at PCUs. Data collected from each PCU will be merged into a single dataset for analysis.

5.3.6. Evaluation Activity A6 – Learner experience of PCU-selected cases

Participants	Students at PCUs
Instrument(s)	E1.1 Learner Experience Survey
Key Evaluation Questions	Do the error VPs provide an effective and engaging learner experience? Do the error VPs represent an appropriate workload for inclusion into existing curricula?

This survey activity will use the same instruments as activity A1, but will instead capture learner experiences of the PCU-selected cases. This activity will not be run as a randomised controlled trial, and all learners will receive the designed error VP intervention.

Due to the differing interventions used at each PCU, data collected from each partner will be kept as separate datasets for analysis and will not be merged.

5.3.7. Evaluation Activity A7 – Tutor experience of PCU-selected cases

Participants	Tutors at PCUs
Instrument(s)	E3 Tutor Experience Survey
Key Evaluation Questions	Was the training provided to tutors to facilitate error VP sessions sufficient? Did the use of error VPs significantly impact upon the workload of tutors?

This survey activity will use the same instruments and be run in the same way as activity A2, but will instead capture tutor experiences of the PCU-selected cases.

Due to the differing interventions used at each PCU, data collected from each partner will be kept as separate datasets for analysis and will not be merged.

5.3.8. Evaluation Activity A8 – Case writer experience of PCU-selected cases

Participants	Case writers at PCUs
Instrument(s)	E4 Case writer (adapter and creator) Written Interview Questions
Key Evaluation Questions	What were the significant challenges when creating/adapting the VP cases? What skills and input were required to create/adapt the cases?

This written interview activity will use the same instruments and be run in the same way as activity A**4**, but will instead capture tutor experiences of the PCU-selected cases. Data collected from each PCU will be merged into a single dataset for analysis, since the data recorded will be qualitative in nature.

5.3.9. Evaluation Activity A9 – Implementation of PCU-selected cases

Participants	Course teams at PCUs, project team members at PCUs
Instrument(s)	E6 VP Case Implementation – Written Interview Questions Quality Report
Key Evaluation Questions	What resources were required to deliver the error VP cases effectively?

This written interview activity is designed to capture experiences of those implementing cases at the PCUs. The instrument will be sent out to purposefully selected individuals who have been identified as having been involved in implementing cases at PCUs. Additional data relating to case implementation and its success will be collected as part of quality procedures during the project, so this analysis may also include data from D4.3 Quality report.

Data collected from each PCU will be merged into a single dataset for analysis, since the data recorded will be primarily qualitative in nature.

5.3.10. Evaluation Activity A10 – Project progress

Participants	Project team
Instrument(s)	Project progress reports, Quality report
Key Evaluation Questions	Has project met key milestones/performance indicators? Has project remained in budget? Have project management tasks (decision making, reporting, communication) been carried out effectively? Have project dissemination activities been targeted effectively, and future opportunities identified?

This evaluation task will be completed by the project team as part of the final evaluation and project reports. It will be based upon analysis of key indicators (e.g. project remaining in budget, completion of deliverables, completion of project, number of dissemination activities), project progress reports and anecdotal experiences shared by project partners. It will be conducted on an ongoing basis, linking with quality procedures, and summarised in the final evaluation report. Key outputs will include recommendations of best practice for future projects.

6. REFERENCES

Alberti, K. G. (2001). Medical errors: a common problem. *BMJ (Clinical Research Ed.)*, *322*(7285), 501–2.

Bandura, A. (2006). Guide for constructing self-efficacy scales. *Self-Efficacy Beliefs of Adolescents*, 307–337.

Borg, G., & Borg, E. (2001). A new generation of scaling methods: Level-anchored ratio scaling. *Psychologica*, (October).

Cook, D. a, Thompson, W. G., & Thomas, K. G. (2011). The Motivated Strategies for Learning Questionnaire: score validity among medicine residents. *Medical Education*, *45*(12), 1230–40. http://doi.org/10.1111/j.1365-2923.2011.04077.x

Cook, D. A., & Ellaway, R. H. (2015). Evaluating technology-enhanced learning: A comprehensive framework. *Medical Teacher*, *37*(10), 961–970. http://doi.org/10.3109/0142159X.2015.1009024

De Leng, B., Huwendiek, S., Donkers, J., & EViP. (2009). *eViP Deliverable 5.2: Published evaluation instrument for Virtual Patients and for learning and teaching activities with Virtual Patients*.

Ellaway, R. H., Poulton, T., & Jivram, T. (2015). Decision PBL: a 4-year retrospective of using virtual patients in problem-based learning. *Medical Teacher*, *in press*.

Ellaway, R. H., Poulton, T., Smothers, V., & Greene, P. (2009). Virtual patients come of age. *Medical Teacher*, *31*(8), 683–4.

ePBLNet. (n.d.). ePBLNet Project. Retrieved from www.epblnet.eu

Eva, K. W. (2009). Diagnostic error in medical education: where wrongs can make rights. *Advances in Health Sciences Education : Theory and Practice*, *14 Suppl 1*, 71–81. http://doi.org/10.1007/s10459-009-9188-9

Huwendiek, S., De Leng, B. a, Kononowicz, A. a, Kunzmann, R., Muijtjens, A. M. M., Van Der Vleuten, C. P. M., … Dolmans, D. H. J. M. (2014). Exploring the validity and reliability of a questionnaire for evaluating virtual patient design with a special emphasis on fostering clinical reasoning. *Medical Teacher*, 1–8. http://doi.org/10.3109/0142159X.2014.970622

Institute of Medicine. (1999). *To err is human: building a safer health system*. (L. T. Kohn, J. M. Corrigan, & M. S. Donaldson, Eds.). National Academy Press, Washington. http://doi.org/10.1017/S095026880100509X

Makary, M. A., & Daniel, M. (2016). Medical error—the third leading cause of death in the US. *BMJ*, *2139*(May), i2139. http://doi.org/10.1136/bmj.i2139

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. (1991). *A manual for the use ofthe Motivated*

Strategies for Learning Questionnaire (MSLQ). Ann Arbor, MI.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & Mckeachie, W. J. (1993). Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). *Educational and Psychological Measurement*, *53*, 801–813. http://doi.org/10.1177/0013164493053003024

Poulton, T., Ellaway, R. H., Round, J., Jivram, T., Kavia, S., & Hilton, S. (2014). Exploring the efficacy of replacing linear paper-based patient cases in problem-based learning with dynamic Web-based virtual patients: randomized controlled trial. *Journal of Medical Internet Research*, *16*(11), e240. http://doi.org/10.2196/jmir.3748

Thompson, E. R. (2007). Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology*, *38*(2), 227–242. http://doi.org/10.1177/0022022106297301

Vaughan, S., Bate, T., & Round, J. (2012). Must we get it wrong again? A simple intervention to reduce medical error. *Trends in Anaesthesia and Critical Care*, *2*(3), 104–108. http://doi.org/10.1016/j.tacc.2012.01.007

Watson, D., & Clark, L. A. (1988). Development and Validation of Brief Measures of Positive and Negative Affect : The PANAS Scales, *54*(6), 1063–1070.

Zhang, J., Patel, V. L., Johnson, T. R., & Shortliffe, E. H. (2004). A cognitive taxonomy of medical errors. *Journal of Biomedical Informatics*, *37*(3), 193–204. http://doi.org/10.1016/j.jbi.2004.04.004

Ziv, A., Ben-David, S., & Ziv, M. (2005). Simulation based medical education: an opportunity to learn from errors. *Medical Teacher*, *27*(3), 193–9. http://doi.org/10.1080/01421590500126718

APPENDIX 1 – LIST OF EVALUATION DATASETS

Dataset	Evaluation Activity	Instrument Source	Partner	Approx. Date
DS1.1	A1 (Learner experience)	E1.1	All PCUs	Q4 2016
DS1.2	A1 (Learner experience)	E1.2	All PCUs	Q4 2016
DS2.1	A2 (Learner performance)	E2.1	Assessment responses from ZSMU	Q4 2016
DS2.2	A2 (Learner performance)	E2.2	Assessment responses from BSMU	Q4 2016
DS2.3	A2 (Learner performance)	E2.3	Assessment responses from KSMU	Q4 2016
DS2.4	A2 (Learner performance)	E2.4	Assessment responses from AMU	Q4 2016
DS2.5	A2 (Learner performance)	E2.5	Assessment responses from PCU5	Q4 2016
DS2.6	A2 (Learner performance)	E2.6	Assessment responses from PCU6	Q4 2016
DS3	A3 (Tutor experience)	E3	All PCUs	Q4 2016
DS4	A4 (Case adaptor)	E4	All PCUs	Q4 2016
DS5	A5 (Tutor trainer)	E5	All PCUs	Q4 2016
DS6.1	A6 (Learner experience)	E1.1	Survey responses from ZSMU	Q4 2017
DS6.2	A6 (Learner experience)	E1.1	Survey responses from BSMU	Q4 2017
DS6.3	A6 (Learner experience)	E1.1	Survey responses from KSMU	Q4 2017
DS6.4	A6 (Learner experience)	E1.1	Survey responses from AMU	Q4 2017
DS6.5	A6 (Learner experience)	E1.1	Survey responses from PCU5	Q4 2017
DS6.6	A6 (Learner experience)	E1.1	Survey responses from PCU6	Q4 2017
DS7.1	A7 (Tutor experience)	E3	Survey responses from ZSMU	Q4 2017
DS7.2	A7 (Tutor experience)	E3	Survey responses from BSMU	Q4 2017

DS7.3	A7 (Tutor experience)	E3	Survey responses from KSMU	Q4 2017
DS7.4	A7 (Tutor experience)	E3	Survey responses from AMU	Q4 2017
DS7.5	A7 (Tutor experience)	E3	Survey responses from PCU5	Q4 2017
DS7.6	A7 (Tutor experience)	E3	Survey responses from PCU6	Q4 2017
DS8	A8 (VP Implementation)	E5	All PCUs	Q4 2017