Sensing, Artificial intelligence, and Edge networking towards Rural Health monitoring (SAFE-RH)

D3.3

Measuring Training Effectiveness with the Kirkpatrick Model

SAFE-RH Project no. 619483-EPP-1-2020-1-UK-EPPKA2-CBHE-JP

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

Contents

List of Tables

1. Introduction

In today's rapidly evolving healthcare landscape, the importance of effective training programs for healthcare professionals cannot be overstated. The Kirkpatrick Model, a renowned framework for evaluating training effectiveness, offers a comprehensive approach to assessing the impact of educational interventions. This section outlines the fundamental principles of the Kirkpatrick Model and its application in measuring the outcomes of training programs conducted in SAFE-RH project.

1.1. Measuring Training Effectiveness of the Kirkpatrick Model

To ensure the success of a training program, a thorough evaluation is essential. According to the Kirkpatrick, evaluation involves the systematic collection of information and data. This evaluation should be planned alongside the training program, aligning with the organization's objectives and goals. In this project, the purpose of the evaluation is to gather information about the outcomes of the training program. This evaluation will provide valuable feedback on participants' reactions, learning outcomes and overall results (Kirkpatrick, 1998).

Evaluating the effectiveness of medical training programs is crucial for ensuring that medical students and professionals acquire the necessary knowledge, skills, and attitudes to become competent healthcare professionals. One widely recognized framework for training evaluation is Kirkpatrick's Model, which consists of four levels: Reaction, Learning, Behavior, and Results. This report aims to apply Kirkpatrick's Model to evaluate a medical training program for medical students and professionals under SAFE-RH project.

The model comprises four levels.

1. Reaction
2. Learning
3. Behaviour
4. Results

Level 1: Reaction

Reaction measures whether learners find the training engaging, favorable, and relevant to their jobs. It is usually access through post-taining observation and interviewstraining surveys. Some of the areas that the survey might focus on are:

1. Program objectives
2. Course materials
3. Content relevance
4. Facilitator knowledge

Level 2: Learning

Learning measures how well participants have acquired knowledge, skills, and confidence as a result of the training. It is usually assessed through pre-tests and post-tests. Identical pre- and post-tests are essential because the difference between the pre-test and post-test scores indicates the amount of learning that took place.

Level 3: Behaviour

Behavior examines how effectively participants can apply what they learned in training to their jobs. This level of evaluation focuses on the practical application of skills and knowledge gained during the

training. It is typically assessed through direct observation of participants in their work environment and through structured interviews with participants and their supervisors. This helps determine whether the training has led to changes in job performance and workplace behavior.

Level 4: Results

This level evaluates the overall impact of the training program on the organization. It examines how well the training has achieved the predefined goals and objectives. This includes measuring the tangible outcomes that can be attributed to the training, such as increased productivity, improved quality of work, higher sales, reduced costs, and enhanced employee retention. Data for this level is typically gathered through key performance indicators (KPIs), business metrics, and performance records. The aim is to determine the return on investment (ROI) of the training program and to assess its effectiveness in contributing to the organization's strategic objectives.

1.2. Objectives

The objectives of the trainings include:

1. To train healthcare people on the fundamental concepts of ICT and its application in healthcare.
2. To introduce SAFE-RH framework and its importance in remote healthcare delivery
3. To demonstrate the role of IOT in enhancing healthcare services in patient monitoring
4. To guide the healthcare people how data analytics can be performed to healthcare data to make informed decisions.
5. To gain insights into MIS for SAFE-RH.

2. Description of Training

The training program for medical students and professionals within the framework of SAFE-RH focuses on equipping participants with the necessary skills and knowledge to effectively use remote health technologies. This initiative aims to enhance participants' understanding of digital health tools and their practical applications in clinical settings. Based on the training modules, the program starts with an introduction to the SAFE-RH healthcare system, covering its components and transformative potential in remote healthcare delivery. It then delves into the basics of ICT, focusing on digital communication and data management principles crucial for modern healthcare settings. Participants learn about the functionality and application of SAFE-RH diagnostic equipment and receive practical training to ensure proficiency in operating these devices. The program also addresses the management information system (MIS), highlighting its role in supporting decision-making processes, and includes training on e-prescribing to enhance patient care through efficient medication management.

The curriculum further enriches the participants' knowledge with modules on data analytics and telemedicine, areas of growing importance in healthcare. Participants explore big data principles and predictive analytics techniques, learning to harness data for improving patient outcomes and optimizing healthcare services. The telemedicine module covers the technical and regulatory aspects of conducting remote consultations. This comprehensive training provides a robust understanding of digital health tools and their practical applications, empowering healthcare professionals to enhance healthcare delivery in remote areas effectively.

2.1. Training Assessment Models

The purpose of this evaluation is to assess the effectiveness of the medical training program and the improvement in skills among medical students and professionals at medical colleges in Bahawalpur. The evaluation utilizes the Kirkpatrick model to analyze and ensure the quality of the training provided. For the training, ensuring that participants are satisfied with the training and have gained the necessary knowledge and skills can be a solid foundation for effective education, Therefore Kirkpatrick model is employed to evaluate the training program across two levels:

Level 1: Reaction

This level aims to capture participants' immediate reactions to the training modules, focusing on their satisfaction and perceived relevance. The survey assesses several critical areas: the clarity of the introduction to the SAFE-RH healthcare system, evaluating how effectively the background, objectives, and functionalities were communicated; the quality of ICT basics course materials, including their comprehensiveness, organization, and applicability to participants' roles; the relevance of diagnostic equipment training, gauging whether the content aligns with practical needs and daily responsibilities; and the knowledge and effectiveness of facilitators in teaching management information systems (MIS), data analytics, and remote health monitoring, focusing on their ability to convey complex concepts and engage the audience. Collecting feedback on these aspects helps determine if the training is engaging and meets participants' needs in using remote health technologies, providing valuable insights to refine the program and enhance learning outcomes.

Level 2: Learning

This level measures the extent to which participants have acquired knowledge and skills as a result of the training.

The training program is implemented across three medical colleges in Bahawalpur:

1. Shahida Islam Medical College, Bahawalpur
2. Quaid-e-Azam Medical College, Bahawalpur
3. UCON, Bahawalpur

A total of 200 participants were involved across these three institutions.

2.1.1. Reaction Assessment

The Reaction Assessment includes 10 questions aimed at evaluating participants' satisfaction and perceptions regarding the training program. The questionnaire (provided in Table 1) focuses on several key areas: the alignment with training objectives, the relevance of the content, the knowledge of the facilitators, and the overall training environment. Each question uses a Likert scale to gauge the extent of agreement or disagreement with various statements. Participants will choose the response that best represents their views, with the Likert scale offering five options:

- Strongly Disagree
- Disagree
- Neutral

- Agree
- Strongly Agree

This approach helps capture a detailed assessment of participants' reactions to the training.

Table 1: Questionnaire for Reaction Assessment

Sr #	Question	Strongly Disagree (0)	Somewhat Disagree (1)	Neutral (2)	Somewhat Agree (3)	Strongly Agree (4)
1	The objectives of the training were clearly defined.					
2	The topics covered were relevant to my job.					
3	The course content was sufficient.					
4	The trainer was well prepared.					
5	The time allotted for the training was adequate.					
6	This training experience will be useful in my work.					
7	The training was relevant to what I might be expected to do.					
8	The meeting room and facilities were adequate and comfortable.					
9	I am satisfied with the training overall.					
10	I would recommend this training to others.					

2.1.2. Learning Assessment

The learning assessments consists of two identical questionnaires as a pre test and post test, allows direct comparison of responses before and after intervention. This approach helps to assess changes in knowledge, skills, and attitudes over time and evaluate the effectiveness of the intervention. Each questionnaire has six Likert scale questions and two open-ended questions designed to evaluate the knowledge, skills, and attitudes of participants.

A. Pre-Learning

The Pre-test of knowledge skills evaluates the medical knowledge and skills of participants before they participate in our training. There is one assessment that medical students and professionals have to complete before beginning the training.

B. Post-Learning

The Post-test of knowledge, skills evaluates the knowledge and skills of participants after they participate in training.

The questionnaire for pre-learning and post-learning assessments is provided in Table 2.

Table 2: Questionnaire for Pre-Learning and Post-Learning Assessments

Sr #	Question	Strongly Disagree (0)	Somewhat Disagree (1)	Neutral (2)	Somewhat Agree (3)	Strongly Agree (4)
1	I have basic knowledge of Sensing Devices*.					
2	I have basic IT skills to run an online health system.					
3	I feel confident to apply what I will learn to my work.					
4	I feel confident in learning new things during a training.					
5	I have sufficient knowledge on Remote Health Monitoring.					
6	The training program motivated me and made me interested in learning.					
7	Which topics should be covered in the training?					
8	In your opinion, what topics are not relevant to your field?					

3. Results by Venues

This section details the assessment of trainings conducted at Shahida Islam Medical College, Quaid-e-Azam Medical College, and UCON, aimed at evaluating their effectiveness.

3.1. Training 1: Shahida Islam Medical College(SIMC)

At Shahida Islam Medical College, total number of participants were 56. Results in Table 3 shows that the reaction level demonstrated a high Cronbach's Alpha of 0.886, indicating a high level of internal consistency among participant responses. The mean score of 2.99 suggests a generally positive reaction. In terms of learning, there was a notable increase in Cronbach's Alpha from 0.679 in the pre-test to 0.826 in the post-test. This improvement in reliability is accompanied by an increase in the mean scores from 2.30 to 2.68, indicating that participants' understanding and perception improved after the intervention.

Table 3: Shahida Islam Medical College Results

Level	Number of Questions	Cronbach's Alpha	Mean	Responses	Venue
Reaction	10	0.936	3.00	56	SIMC
Pre-Learning	6	0.784	2.31	56	SIMC
Post-Learning	6	0.905	2.64	56	SIMC

3.2. Training 2: Quaid-e-Azam Medical College(QMC)

At Quaid-e-Azam Medical College (QMC) total number of participants were 74, Results in Table 4 shows that the reaction level had a slightly lower Cronbach's Alpha of 0.856 compared to SIMC, but it still signifies good internal consistency. The mean score of 2.86 shows that participants generally reacted positively, though slightly less consistently than those at SIMC. In the learning phase, Cronbach's Alpha increased from 0.642 in the pre-test to 0.797 in the post-test. The mean scores also improved from 2.18 to 2.53, reflecting an enhanced understanding and perception post-intervention.

Table 4: Quaid-e-Azam Medical College Results

Level	Number of Questions	Cronbach's Alpha	Mean	Responses	Venue
Reaction	10	0.856	2.86	74	QMC
Pre-Learning	6	0.642	2.18	74	QMC
Post-Learning	6	0.797	2.53	74	QMC

3.3. Training 3: UCON

At UCON, total number of participants were 64, Results in Table 5 shows that the reaction level exhibited a Cronbach's Alpha of 0.839, indicating good internal consistency similar to QMC. The mean score of 3.12 suggests a positive reaction overall, though with a bit less consistency. For the learning aspect, the Cronbach's Alpha values ranged from low to moderate, increasing from 0.630 in the pre-test to 0.755 in the post-test. The mean scores showed an improvement from 2.45 to 2.91, indicating a reasonable level of internal consistency and better participant understanding post-intervention.

Table 5: UCON Results

Level	Number of Questions	Cronbach's Alpha	Mean	Responses	Venue
Reaction	10	0.839	3.12	64	UCON
Pre-Learning	6	0.630	2.45	64	UCON
Post-Learning	6	0.755	2.91	64	UCON

3.4. Results of all three trainings

Table 6 shows the results from all three trainings at SIMC, QMC, and UCON, here we can see that the **reaction** level yields a strong Cronbach's Alpha, signifying that overall, participant reactions are positive and consistently measured across different venues. In terms of **learning**, the combined data shows low to moderate Cronbach's Alpha values. However, there is a significant improvement in the mean scores from pre-test to post-test across all venues, indicating enhanced understanding and perception after the intervention.

Table 6: Results of three trainings

Level	Number of Questions	Cronbach's Alpha	Mean	Responses	Venue
Reaction	10	0.886	2.99	200	SIMC-UCON-QMC
Pre-Learning	6	0.679	2.30	200	SIMC-UCON-QMC
Post-Learning	6	0.826	2.68	200	SIMC-UCON-QMC

Table 7: Agreement Results from Participants Across Three Trainings

	SIMC-UCON-QMC	Pre-test			Post-test		
Sr #	**Question**	**Somewhat Agree (3)**	**Strongly Agree (4)**	**Result**	**Somewhat Agree (3)**	**Strongly Agree (4)**	**Result**
1	I have basic knowledge of Sensing Devices*.	n=51	n=21	36%	n=63	n=52	57%
2	I have basic IT skills to run an online health system.	n=41	n=19	30%	n=60	n=40	50%
3	I feel confident to apply what Iwill learn to my work.	n=64	n=72	68%	n=63	n=83	73%
4	I feel confident in learning newthings during a training.	n=46	n=116	81%	n=54	n=96	75%
5	I have sufficient knowledge on Remote Health Monitoring.	n=37	n=11	24%	n=63	n=44	53%
6	The training program motivated me and made me interested in learning.	n=58	n=89	73%	n=51	n=96	73%

The Table 7 illustrates the pre-test and post-test results for participants' agreement across three venues. There is a general improvement in participants knowledge and skills across various aspects of the training program. The largest improvement is seen in the understanding of Remote Health

Monitoring. The participants generally felt motivated and interested in the training, which is a positive sign for the effectiveness of the program.

4. Combined Results

The results of this study provide a detailed analysis of participants' feedback and learning outcomes across different venues. To ensure a comprehensive understanding of the effectiveness of the training program, various statistical measures I.e. Mean, Standard Deviation, Cronbach's Alpha and dependent T-test were employed. These measures help assess both the central tendencies and the reliability of the participants' responses. By examining these metrics, we gain insights into the overall perception, satisfaction, and consistency of the participants' learning experiences.

To evaluate the data collected from participants, we used key metrics such as the mean, standard deviation, and Cronbach's Alpha. The mean represents the average value of a set of scores, serving as an indicator of the central tendency of participant responses. This measurement is crucial for monitoring the overall perception and satisfaction of participants. In contrast, the standard deviation measures the dispersion or spread of these scores, providing insight into the variability in participant responses. A higher standard deviation signifies greater variability, indicating that participants' responses are more spread out from the mean. Additionally, Cronbach's Alpha is a measure of internal consistency or reliability, used to assess the reliability of survey instruments or assessments. It ensures that the responses of participants are consistent. Cronbach's Alpha values range from 0 to 1, with higher values indicating greater reliability, signifying that the survey or assessment tool is dependable and yields consistent results.

4.1. Level 1: Results

The questionnaires are based on Likert scale items, providing valuable insights into the participants' feedback on various aspects of the training program. The assessment provided in Table 2 gives insight into the participants' feedback on various aspects of the training program. The Mean and Standard Deviations for each question are shown in Table 8.

Table 8:Mean and Standard Deviation of Each Question

Sr #	Questions	Mean	Std.
1	The objectives of the training were clearly defined.	3.03	1.15
2	The topics covered were relevant to my job.	3.12	1.02
3	The course content was sufficient.	2.81	1.08
4	The trainer was well prepared.	3.07	1.09
5	The time allotted for the training was adequate.	2.85	1.10
6	This training experience will be useful in my work.	3.04	1.02
7	The workshop was relevant to what I might be expected to do to.	2.85	1.01
8	The meeting room and facilities were adequate and comfortable.	2.99	1.23
9	I was satisfied with the workshop overall.	3.01	1.04

10	I would recommend this course to others.	3.11	1.04
	Overall	**2.99**	**1.08**

Table 8 shows the results of level-1(Reaction), The highest mean score is for the relevance of the topics to the participants' jobs (3.12). This indicates that participants found the training content relevant and applicable to their work. Recommendation to others also scored highly (3.11), suggesting overall positive feedback and a likelihood of participants advocating for the course. The objectives of the training (3.03), trainer preparedness (3.07), and usefulness of the training in participants' work (3.04) all received moderate scores, indicating general satisfaction but also room for improvement. Overall satisfaction with the workshop (3.01) aligns closely with the average mean score, showing a consistent level of moderate satisfaction across different aspects of the training. The course content sufficiency (2.81) and the relevance of the workshop to expected tasks (2.85) received the lowest scores, suggesting that these areas might need improvement. The time allotted for the training (2.85) also received lower scores, indicating potential issues with the duration of the training sessions.

The standard deviations range from 1.01 to 1.23, indicating a moderate level of variability in the responses. The highest variability is seen in the adequacy and comfort of the meeting room and facilities (1.23), which suggests differing opinions on this aspect. The relatively lower standard deviations for the relevance of the workshop to expected tasks (1.01) and the relevance of topics to the job (1.02) indicate more consistent opinions in these areas.

The overall mean score of 2.99 suggests that participants are moderately satisfied with the training program. However, specific areas such as the sufficiency of course content and the adequacy of time allotted for training need attention. The variability in responses highlights the need for a more tailored approach to meet the diverse expectations and needs of the participants. Improving these areas could lead to higher overall satisfaction and more effective training outcomes. High levels of satisfaction and engagement indicate that the training program is well-received by participants, suggesting a positive learning environment.

4.2. Level 2: Results

The Table 9 shows the results of the participants' feedback on the training program. The feedback is categorized into four main areas: Knowledge, Confidence, Skills, and Attitude. Each category has specific questions associated with it, and the responses are measured using a Likert scale. The table includes the mean scores and standard deviations (SD) for both pre-test and post-test responses.

Table 9: Results of Learning pre and Learning post

	Knowledge		Confidenece		Skills	Attitude	
	Pre-Test						
	Q1	Q5	Q3	Q4	Q2	Q6	Average
Mean	1.77	1.49	2.86	3.27	1.47	3.00	2.31
Std.	1.30	1.23	1.11	1.04	1.40	1.17	1.21
	Post-Test						
Mean	2.44	2.37	2.98	3.07	2.23	3.05	2.69
Std.	1.33	1.30	1.10	1.12	1.34	1.17	1.23

In the pre-test phase, participants' knowledge was measured through questions Q1 and Q5, with mean scores of 1.77 (neutral) and 1.49 (disagree) respectively. This indicates that initially, participants had a mixed perception of their knowledge, leaning slightly towards disagreement. Confidence was assessed through Q3 and Q4, with scores of 2.86 (agree) and 3.27 (strongly agree), showing high confidence levels even before the training. Skills, evaluated by Q2, had a mean score of 1.47 (disagree), suggesting participants felt their skills were lacking. Attitude, measured by Q6, scored 3.00 (agree), indicating a positive attitude towards the training. The overall average pre-test score was 2.31 (neutral), reflecting a balanced yet cautious perspective on the training program's initial impact.

The post-test results show a significant improvement across all areas. Knowledge scores increased to 2.44 (agree) for Q1 and 2.37 (neutral) for Q5, suggesting that the training effectively enhanced participants' knowledge. Confidence remained high, with scores of 2.98 (agree) for Q3 and 3.07 (agree) for Q4, indicating a sustained positive impact on participants' confidence levels. Skills improved notably, with Q2 scoring 2.23 (neutral), showing progress from the initial disagreement. Attitude remained consistently positive, with Q6 scoring 3.05 (agree). The overall average post-test score rose to 2.69 (agree), signifying a positive shift in participants' perceptions following the training.

The standard deviations provide insight into the variability of participants' responses. For the pre-test, standard deviations ranged from 1.04 to 1.40, with an average of 1.21, indicating moderate variability. In the post-test, standard deviations ranged from 1.10 to 1.34, with an average of 1.23, showing a slight increase in variability. Despite this, the overall trend demonstrates that the training program had a beneficial effect on participants' knowledge, confidence, skills, and attitude. The positive changes in mean scores from pre-test to post-test highlight the training program's effectiveness in enhancing these key areas.

Table 10: T-test results of level-2

	Questions					
	T-test					
	Q1	Q2	Q3	Q4	Q5	Q6
T-Statistic	-6.6354	-6.8782	-1.2778	2.2167	-8.2001	-0.3886
P-Value	3.03×10^{-10}	7.76×10^{-11}	0.2028	0.02778	2.98×10^{-14}	0.6980

The Table 10 shows t-test results of Level-2. For Questions 1, 2, 4, and 5, the p-values indicate significant improvements in scores post-training, leading to the rejection of the null hypothesis. The training program significantly improved participants' knowledge of sensing devices, IT skills for running

an online health system, and Remote Health Monitoring, with p-values well below the 0.05 threshold. However, there was no significant improvement in participants' confidence in applying what they learned to their work or in their motivation to learn. The lack of significant improvement in participants' confidence in applying what they learned and their motivation to learn may be due to several factors unrelated to the quality of the training itself. First, participants may have already possessed sufficient foundational knowledge in these areas, resulting in minimal perceived gains. Second, the training's content could have been too theoretical, and the participants might not have had immediate opportunities to apply it in their work, leading to limited impact on their practical confidence.

5. Conclusion

The evaluation of the training program, guided by the Kirkpatrick Model, identified both strengths and areas for improvement. Participants expressed the highest satisfaction with the relevance of training topics to their jobs (mean score of 3.12) and the preparedness of the trainers (mean score of 3.07). High scores for recommending the program to others (mean score of 3.11) indicate overall positive feedback. However, lower scores for the sufficiency of course content (mean score of 2.81) and the time allocated for training (mean score of 2.85) highlight areas needing attention. Despite these concerns, the overall mean score of 2.99 suggests moderate satisfaction, indicating that the training was generally well-received.

Post-test results revealed significant improvements in participants' knowledge, confidence, skills, and attitudes. Knowledge scores increased notably, demonstrating effective knowledge enhancement. Confidence levels remained high, and skills showed substantial improvement from pre-test to post-test. The increase in Cronbach's Alpha for post-test responses across venues indicated higher internal consistency and reliability, reflecting the learning effect, clearer concept understanding, and reduced response bias. These positive changes affirm the program's effectiveness in enhancing professional competencies. Notably, the results of the dependent t-test further support these findings, with significant differences observed in the scores for several questions (Questions 1, 2, 4, and 5) where p-values were well below the 0.05 threshold. This statistical evidence underscores the program's impact on participants' development.

Overall, the training program has successfully improved participants' knowledge and skills, highlighting its success in fostering professional development and readiness in remote health monitoring.

SAFE-RH