

The Asian Sarcoma Consortium Sarcoma Preceptorship Program: A Program Evaluation Study Utilizing the Kirkpatrick Model (Level 1 and 2)

Gracieux Y. Fernando, MD, MHPEd,¹ Eileen Poon Yi Ling, MD,² Goh Wei Lin,²
Hong Jiancheng² and Richard Quek Hong Hui, MD³

¹*Division of Medical Oncology, Department of Medicine, College of Medicine and Philippine General Hospital, University of the Philippines Manila*

²*Division of Medical Oncology, National Cancer Center, Singapore*

³*Parkway Cancer Centre, Singapore*

ABSTRACT

Background. Sarcomas, a heterogenous group of malignancies consisting of tumors, are rarely encountered in most general oncology clinics. As a result, most general oncologists have very little experience in dealing with these tumors and resulting in poorer treatment and survival outcomes. Having successful collaborative efforts in research, the Asian Sarcoma Consortium (ASC) pioneered a Sarcoma Preceptorship Program with its main goal of increasing awareness of the multidisciplinary approach in the management of sarcomas.

Objective. The objective of this study is to evaluate the ASC Sarcoma Preceptorship Program using the first 2 levels of the Kirkpatrick Evaluation Model: Level 1) Reaction – through using satisfaction surveys; and Level 2) Knowledge – focusing on the amount of information gained from the activity evaluated by test-retest method.

Methods. The 2-day preceptorship program held in Singapore utilized educational activities such as didactic lectures, multidisciplinary case discussions, and case presentations. The program was evaluated using: 1) pre-test and post-test; 2) satisfaction survey from ASC; and 3) satisfaction survey made by the sponsor, Novartis Pharma.

Results. The preceptorship program enrolled 30 participants, most from the Asia-Pacific region. The overall results from the satisfaction surveys were generally optimistic, with results showing high levels of satisfaction with regard to the time allotment, scope of topics, and how helpful each session was. Test results showed that participants scored lower in post-test as compared to the pre-test. Computation of the Coefficient of Test reliability showed that although the pre-test was highly reliable, this was not true of the post-test.

Conclusions. Based on the Level 1 evaluation, the program was successful in terms of both participant and faculty satisfaction. However, certain areas for improvement were identified and recommendations were made in order to effect improvements for future iterations of this preceptorship.

Keywords: soft tissue sarcoma, clinical preceptorship, Kirkpatrick model, Asian Sarcoma Consortium

INTRODUCTION

Sarcomas are a heterogeneous group of malignancies consisting of tumors that arise from mesenchymal tissues such as bone, muscle, fat and cartilage. Epidemiologically, they represent around 20% of pediatric and 1% of adult solid tumors.¹ Because they are derived from different mesenchymal tissues of origin, they can have very varied presentations, with each variant having its own clinical course, pathologic behavior and prognosis. Because they are not very common malignancies, they are rarely encountered in most general oncology clinics. As a result, most general oncologists have

Poster presented at the European Society of Medical Oncology – Asia, November 22-24, 2019, Suntec Singapore Convention and Exhibition Center.

Corresponding author: Gracieux Y. Fernando, MD, MHPEd
Division of Medical Oncology
Department of Medicine
College of Medicine and Philippine General Hospital
University of the Philippines Manila
Taft Avenue, Ermita, Manila 1000, Philippines
Email: gyfernando@up.edu.ph

very little experience in dealing with these tumors and this diminished expertise has resulted in poor patient outcomes. This is further compounded by the fact that no screening examination exists that allows for the early identification of sarcomas. Consequently, the opportunity for early diagnosis and work-up depends greatly on the physician who initially saw the case being capable of recognizing and diagnosing the condition. This is made difficult because of the factors previously mentioned, and results in a delay in diagnosis. This in turn is associated with the occurrence of events that can negatively impact on sarcoma patient survival, such as large tumors, lost opportunities for limb-sparing surgeries and the occurrence of distant metastases.² A study done in Denmark estimated that the median time to diagnosis was 176 days from the start of the patient's symptoms.³ In a study from the UK, the average time from start of symptoms to diagnosis was 96 weeks and by that time, most sarcomas were at least 10 cm. in size or larger.⁴ It is no surprise then that most treatment recommendations have advised that patients with sarcomas be seen in specialist sarcoma centers^{5,6} where those with adequate experience can institute multidisciplinary management of the cases. Aside from this, primary health care specialists, who often see these cases before the sarcoma oncologist does, should be provided opportunities for education in the recognition, diagnosis and early referral of potential sarcoma cases in order to bridge the gap between initial consult and subsequent referral to a sarcoma specialist center in the shortest time possible.

The Asian Sarcoma Consortium (ASC) represents an international study group of oncologists who specializes in the treatment of sarcomas. The group was first convened in November 2015 under the pioneering efforts of Dr. Richard Quek of the National Cancer Centre, Singapore. Its members come from Asian countries that include Singapore, Japan, Hong Kong, Taiwan, the Philippines, Malaysia, Myanmar and Thailand. The main intentions of the ASC are: (1) to encourage international exchange of the latest up-to-date developments in the field of sarcoma; (2) to encourage international research collaborations; and (3) to educate our trainees and other involved health care personnel in order to widen their exposure and experience in sarcoma management. This collaborative effort has proven to be very successful, with the ASC presenting its first research output on Angiosarcoma at the ESMO Asia Conference in December 2016. This consists of 2 projects on Angiosarcoma^{7,8}, both of which were presented in paper presentation sessions and were well-received.

Having encountered a successful collaborative effort in research, the group has since started to increase sarcoma awareness among trainees and general oncologists. To this end, the group has initiated a Sarcoma Preceptorship Program with its main goal being an increased awareness of the multidisciplinary approach to the management of sarcomas. The initial offering of this program was held on July 14-15, 2017 at the National Cancer Centre, Singapore.

The targeted audience consisted of oncology fellows-in-training and junior consultants. The preceptorship program was conceived and delivered by members of the ASC. It consisted of an intensive 2-day workshop utilizing both didactic lectures and small group break-out sessions, where the participants were expected to produce and present outputs on several theoretical sarcoma cases provided. (Details on the Instructional design of the ASC Sarcoma Preceptorship can be viewed in **Appendix 1: ASC Preceptorship Instructional Design**). It was hoped that the program will increase sarcoma awareness among its participants and that they would then carry and disseminate this awareness upon returning to their respective institutions.

Significance of the Study

This sarcoma-centric preceptorship is a first in Asia and the ASC hopes to conduct such similar sessions in the different ASC member countries. In order to determine the effectiveness and utility of our program, an evaluation was carried out to assess if the intended objectives were met. This would also identify problems and difficulties encountered in the conduct of the program. Such information would be valuable in improving future iterations of this program. The primary objective of this study, therefore, is to conduct an evaluation of the ASC Sarcoma Preceptorship Program. This evaluation was conducted using the 4-tier evaluation model initially proposed by Donald Kirkpatrick.

Donald Kirkpatrick developed the Kirkpatrick Evaluation Model in order to determine the impact of training on its target participants.⁹ This widely used evaluation system is both simple to use and adaptable to different training scenarios. In its purest form, it consists of a 4-tiered hierarchy that seeks to determine the depth of training effect on its participants. Briefly, these 4 tiers include the following:

1. **Reaction:** consists of the participants' impressions of the program and their level of satisfaction.
2. **Learning:** the acquisition of knowledge, skills and attitudes from the training program.
3. **Application:** determination of on-the-job performances of the participants after they have completed the training program.
4. **Results:** designates any changes made by the participants to the organization as a result of the training received.

In order to achieve the goals of the Kirkpatrick model of evaluation, several outcomes were selected as targets for evaluation and evaluation schemes were prescribed for each one. (Details of these may be viewed in **Appendix 2: Evaluation Summary and Appendix 3: Study Concept Map**.)

For this study, the following research questions were formulated:

1. What were the general impressions of the participants towards the program?
2. Were the participants satisfied with the content and conduct of the program?

3. Did the participants learn any new information from participating in the program?

Conceptual Framework

The targeted aspects of the program were evaluated in the following manner:

1. Level 1 (Reaction): satisfaction levels of the participants were gauged for topic content, speaker performance, conduct of the Multi-Disciplinary Conference (MDC) and the logistical aspects of the program (venue, comfort level, refreshments) using a developed tool consisting of Likert scales and open-ended questions.
2. Level 2 (Learning): the acquisition of knowledge from participation in this program was gauged using a pre-and post-test consisting of soft tissue sarcoma-based Multiple-Choice Questions (MCQs). The difference in performance between both tests was then used as a gauge for knowledge improvement in sarcoma management. Pre-test was given before the start of the program while the post-test was administered after the final MDC was completed.

Limitations of the Program

In an ideal evaluation scenario, an educational program should be subjected to both short-term outcomes, corresponding to the first 2 levels of the Kirkpatrick model as well as long-term outcomes which would correspond to the next 2 levels of the Kirkpatrick model. Being an international preceptorship, with a heterogeneous population with diverse backgrounds, this was going to be difficult to attain with our program. The biggest issue involved the evaluation of long-term outcomes, since follow-up at periods of 6 months and 1-year post-attendance would pose both problems in follow-up, collection of data and reliability of results. Therefore, we had decided to restrict the present evaluation schema to only the first 2 levels of the Kirkpatrick Evaluation Model. This would therefore involve the following levels and their corresponding outcomes:

1. Level 1: Reaction – focusing mainly on satisfaction surveys of the participants as to the content and conduct of the activity
2. Level 2: Knowledge – focusing on the amount of information gained from the activity and evaluated through a pre-test given prior to the start of the activity proper and a post-test administered at the end of the final session.

The Asian Sarcoma Consortium Sarcoma Preceptorship Program: Activity Mechanics

Program Duration: 2 days (8:00 am to 5:00 pm)

Venue: The Academia, Singapore General Hospital

Faculty: Members of the Asian Sarcoma Consortium

Participants: A total of 30 attendees participated in the first preceptorship program. The criteria for selection were as follows:

1. 2-3 nominated participants from each of the 8 member institutions of the ASC as well as participants invited from other countries and/or institutions with interest in Soft Tissue Sarcoma management
2. Clinical background: Medical Oncology, Orthopedic Oncology, Radiation Oncology, Surgical Oncology
3. Experience level: Fellows-in-training or Junior consultants

The summary of program activities for this 2-day preceptorship may be viewed in **Appendix 4**.

Educational Activities Utilized

All sessions of the Sarcoma Preceptorship were conducted in the English language by the faculty of the ASC. The activities included the following:

1. Didactic Lectures
 - a. Preliminary sessions aimed at setting the knowledge base for each selected topic
 - b. Handled by an assigned faculty from the ASC
 - c. 30–40-minute presentations followed by 5-10 minutes for open forum
 - d. Prior to the start of all activities, all participants completed a 20-item MCQ pre-test.
2. Multidisciplinary Case (MDC) Presentation
 - a. Summative activity occurring at the conclusion of each major topic.
 - b. Triggered by a case (actual or theoretical) prepared by the ASC faculty.
 - c. Participants were divided into smaller groups, with each group assigned to discuss the case and formulate their management plans. All groups were facilitated by a faculty of the ASC, who guided the discussion. Each group then presented their outputs to the group. The activity was moderated by selected faculty of the ASC, who provided constructive comments and suggestions. Further input was obtained from the ASC faculty who had delivered the didactic lectures. Responses of each subgroup was elicited with a final consensus obtained as a summative activity.
3. Summative Multidisciplinary Case Presentation
 - a. This was the concluding activity of the preceptorship.
 - b. A trigger case, based on any of the topics discussed in the preceptorship, albeit more complex than those initially utilized, was assigned to all groups.
 - c. Group outputs were presented in the MDC conference, moderated by the entire faculty of the ASC.

METHODS

A descriptive method was used to analyze outcomes. Data to ascertain the level of participant satisfaction (Level 1) were collected using a tool developed by the study faculty

specifically for this event. The tool was administered after every session of the program and each participant was asked to sign an informed consent prior to actual data collection. Factors included adequacy of time for a topic and/or discussion, topic coverage and degree of helpfulness. Degree of helpfulness was assessed using a 10-point scale, with 10 indicating the highest level of satisfaction. Results were reported as averages for all respondents. Open-ended questions were also included and focused on feedback, difficulties encountered, and future topics participants would want covered.

The Level 2 Evaluation conducted on the ASC Sarcoma Preceptorship program was done under the following assumptions:

- Hypothesis: The Sarcoma Preceptorship will improve knowledge in the general management of sarcomas.
- Null Hypothesis: There will be no difference in knowledge on sarcoma management even after participation in the preceptorship program.

Level 2 evaluation in the Kirkpatrick system seeks to determine if the participation in this particular program improved the level of knowledge of its participants, in this case their knowledge on the diagnosis and management of sarcomas. In order to measure this improvement in learning, the following activities were conducted:

1. Pre-test: A 20-point MCQs was developed as a pre-test for the program participants. All questions were formatted as MCQs with a maximum of 4 distractors per question. All questions were submitted to members of the ASC for review and validation. The pre-test was administered to all participants prior to the beginning of the first session of the program and they were given no prior warning that this would be done. A 30-minute time limit was budgeted for this activity.
2. Post-test: To determine any change in the level of knowledge on sarcoma diagnosis and management as a result of their participation in this preceptorship program, a post-test was also administered after the final session of the program was completed. To be able to get a good gauge of these changes, we administered the same 20-point questions that were given in the pre-test. Again, no prior warning was given that a post-test would be done.

In order to determine if the results of both the pre-test and the post-test were both reliable and homogenous, the Test Reliability (KR21) Coefficient was calculated for both. Test reliability is a quality of a test that seeks to determine if the evaluation tool is capable of yielding consistent results for each student. The particular characteristic of Test Reliability that we were interested in was the test's internal consistency. This pertains to the ability of the different parts of a single assessment tool to make similar conclusions about the examinee's achievements. In order to quantify this, we calculated for the Kuder-Richardson (KR) Coefficient. The

$$KR21 = [n/(n - 1)] \times [1 - (M \times (n - M) / (n \times Var))]$$

KR21 = estimated reliability of the full-length test
 n = number of items
 Var = variance of the whole test (standard deviation squared)
 M = mean score on the test

Figure 1. Formula for the KR21 Coefficient.

KR21 formula was utilized based on the assumption that the scored items were of the same difficulty. A KR21 Coefficient of > 0.70 is widely accepted as the cut-off to declare an assessment tool as being both reliable and homogenous. Figure 1 illustrates the formula for calculating the KR21 Coefficient.¹⁰

Copies of the actual evaluation forms as well as the exam questions utilized for the pre- and post-test may be viewed in **Appendix 5: Evaluation Forms and Appendix 6: Examination Questions**. It should be noted that the same set of questions were given for the pre- and post-test. **Appendix 7** presents the consent form used for all participants in this preceptorship. **Appendix 8** presents the data gathering form for the particulars of each attendee to this preceptorship program.

Sponsor Participation

The first ASC Sarcoma Preceptorship would not have been possible without the generous support from our partners in the Pharmaceutical Industry. Specifically, this first foray of the consortium into an educational activity was funded by an unrestricted grant from Novartis Pharmaceuticals, a long-time supporter of sarcoma education and research in the region. This funding supported the following activities:

1. Travel grants to and from Singapore for all foreign preceptorship faculty and participants.
2. Accommodations in Singapore for all foreign preceptorship faculty and participants.
3. Operating costs for the preceptorship activity, including rental of venue and daily operating costs of the activity.

On the other hand, Novartis Pharmaceuticals did not participate in the following preceptorship-related activities:

1. Design of the academic program.
2. Selection of the preceptorship topics, curricula, program and faculty.
3. Selection of the preceptorship participants, who were selected based on recommendations from the various members of the ASC.
4. Conduct and management of the preceptorship program itself.
5. Preparation and administration of program-related evaluation tools (pre- and post-tests, program evaluation tools).
6. Evaluation of program outcomes.

In addition, Novartis did not provide any honorarium for all preceptorship participants, including the participants themselves as well as the faculty. Also, the activity was rendered free of any commercial advertisements of any of the sponsor's products and there was no session during the entire preceptorship that was dedicated solely to the promotion or discussion of any product or service provided by Novartis. Indeed, most participants highly rated the activity as being fair, balanced and free from commercial bias, with an average score of 3.75 out of 4 (Table 5). However, as part of their funding activities, Novartis was required to submit a report on the conduct of this activity and this included collection of feedback data from the participants of the activity. This yielded additional feedback information which was used to cross-check the responses given to our own evaluation forms. As this was a useful source of additional information, the results of the evaluation scheme conducted by Novartis will also be presented, in parallel with the results obtained from our own evaluation forms. A sample of the form used by Novartis for their data collection may be viewed in **Appendix 9: Novartis Evaluation Form**.

RESULTS

Demographic Data of Preceptorship Participants

The first ASC Sarcoma Preceptorship program enrolled a total of 30 participants coming from 16 countries. There were 18 males and 12 females. The average age was 38.9 years, ranging from 29 to 62 years. The demographics of the preceptorship participants are summarized in Table 1.

Level 1 Evaluation Results: Satisfaction Surveys and Feedback

Session 1 of the ASC Sarcoma Preceptorship consisted of topics on Localized Soft Tissue Sarcomas. The activities in Session 1 consisted of 4 didactic lectures followed by 1 MDC discussion. The results of the evaluation of this session utilizing the ASC feedback form were as follows:

- As to the questions, "Was the time allocated for each topic enough?" and "If not, what would be the sufficient length of time in minutes?" 29 of 30 respondents (97%) felt that the time allocated was sufficient. 1 of 30 (3%) respondent felt that the time was insufficient and suggested a time of 30-40 minutes per topic.
- As to the question, "Was the time allocated for the multidisciplinary conference enough?" 27 of 30 respondents (90%) felt that the time was sufficient but 3 of 30 (10%) respondents replied that the time allocated was not sufficient and suggested an average time of 22 minutes per case (Range: 8 – 30 minutes per case).
- As to the question, "Was the scope covered adequate for each of the topics covered?" and "Were the number of topics covered adequate?" all 30 respondents felt that the scope of the topics and the number of topics covered were adequate.

Table 1. Demographics of Preceptorship Participants

Characteristic	Number	Percentage
Nationalities		
Thailand	7	23
Kazakhstan	3	10
Myanmar	3	10
Taiwan	3	10
Egypt	2	7
United Kingdom	2	7
Algeria	1	3
Bangladesh	1	3
Hong Kong	1	3
India	1	3
Indonesia	1	3
Malaysia	1	3
Pakistan	1	3
Philippines	1	3
Sudan	1	3
Ukraine	1	3
Designation		
Doctor	25	84
Pharmaceutical Industry	3	10
Doctor/Nurse	1	3
Pharmacist	1	3
Place of Work		
University/Government Hospital	24	80
Pharmaceutical Industry	4	13
Private Hospital	2	7
Area of Specialization		
Medical Oncology	13	43
Radiation Oncology	3	10
Clinical Oncology	3	10
Orthopaedic Oncology	3	10
Surgical Oncology	3	10
Others	2	7
Medical/Surgical Oncology	1	3
Medical/Radiation Oncology	1	3
Pharmacy	1	3
Experience Level		
Consultant	15	50
Specialist in Training	15	50

- When asked to recommend other topics for this session, the following were suggested by the respondents:
 - Types of flaps and /or prosthesis available
 - More discussion on the use of adjuvant radiotherapy
 - The role of neoadjuvant therapy
 - The use of chemotherapy
 - Chondrosarcomas
- The respondents reported the following as positive comments for Session 1:
 - Very good
 - Good educational program
 - Very informational, to the point, excellent
 - More case discussions

6. On the other hand, difficulties encountered in Session 1 are as follows:
 - a. Difficulties with language barrier
 - b. Not enough scope

With regard to the question, “On a scale of 0-10, how helpful did you find the session?” the response of the participants revealed that the mode of the scores was 8 (Range 6 – 10) and no respondent scored the session lower than 5. The evaluation form utilized by Novartis looked at the perception of the participants with regard to the Speakers’ expertise in handling the session. 18 Of 24 (75%) respondents strongly agreed that the speakers were experts in the session while 6 of 24 (25%) agreed. The average score given by the 24 respondents was 3.75 out of a possible 4. This level of satisfaction with session 1 remains consistent with those reported using the ASC evaluation form.

Session 2 of the ASC Sarcoma Preceptorship consisted of topics dealing with Metastatic Soft Tissue Sarcomas. The activities in Session 2 consisted of 4 didactic lectures followed by 1 MDC discussion. The results of the evaluation of this session utilizing the ASC feedback form were as follows:

1. As to the question, “Was the time allocated to the session sufficient?” 29 of 30 (97%) respondents reported that the time allocated was sufficient while 1 of 30 respondents (3%) reported that it was not and suggested 40 minutes per topic as being sufficient.
2. As to the question, “Was the time allocated for the multidisciplinary conference enough?” 29 of 30 respondents (94%) felt that the time was sufficient but 1 of 30 (3%) respondents replied that the time allocated was not sufficient and suggested a time of 30 – 45 minutes per case as being sufficient.
3. As to the question, “Were the number of topics covered adequate?” all 30 respondents felt that the scope of the topics and the number of topics covered were adequate. All 30 respondents also agreed that the topics were helpful.
4. When asked to recommend other topics for this session, a respondent replied that he would like more cases to illustrate clinical points discussed.
5. Difficulties reported by respondents with regard to Session 2 included the following:
 - a. Some language barriers
 - b. Some topics are not familiar to surgeons
6. Additional topics recommended for Session 2 by the respondents include the following:
 - a. Radiologic images for sarcomas (PET-CT, MRI, CT scans)
 - b. Management of metastatic Gastrointestinal Stromal Tumors
 - c. Role of the interventional radiologist

With regard to the question, “On a scale of 0-10, how helpful did you find the session?” the response of the participants revealed that the mode of the scores was 8 (Range 6 – 10) and no respondent scored the session lower than 5. The evaluation form utilized by Novartis looked at the perception of the participants with regard to the Speakers’ expertise in handling the session. The response of the participants showed 18 Of 24 (75%) respondents strongly agreed that the speakers were experts in the session while 6 of 24 (25%) agreed. The average score given by the 24 respondents was 3.75 out of a possible 4. This level of satisfaction with session 2 remains consistent with those reported using the ASC evaluation form.

Session 3 of the ASC Sarcoma Preceptorship consisted of topics dealing with Bone Sarcomas. The activities in Session 3 consisted of 4 didactic lectures followed by 1 MDC discussion. The results of the evaluation of this session utilizing the ASC feedback form were as follows:

1. As to the question, “Was the time allocated to the session sufficient?” 29 of 30 (97%) respondents reported that the time allocated was sufficient while 1 of 30 respondents (3%) reported that it was not sufficient but did not suggest what was felt to be a sufficient time.
2. As to the question, “Was the time allocated for the multidisciplinary conference enough?” 29 of 30 respondents (94%) felt that the time was sufficient but 1 of 30 (3%) respondents replied that the time allocated was not sufficient but gave no suggestion as to what they felt was a sufficient time.
3. As to the question, “Were the number of topics covered adequate?” all 30 respondents felt that the scope of the topics and the number of topics covered were adequate. 29 of 30 (97%) respondents also agreed that the topics were helpful but 1 of 30 (3%) reported that the topic, “Surgical Overviews in Bone Sarcomas” was not helpful.
4. When asked to recommend other topics for this session, 2 respondents replied the following:
 - a. Palliative care after radiotherapy
 - b. Other choices of treatment such as use of carbon ion particles
5. Difficulties reported by respondents with regard to Session 2 included the following:
 - a. Language barriers
 - b. Too many case discussions
 - c. Too complicated
6. Additional topics recommended for Session 3 by the respondents include the following:
 - a. Palliative care after treatment and Psychiatric support
 - b. Metastatic Synovial Sarcoma
 - c. More case sessions

With regard to the question, “On a scale of 0-10, how helpful did you find the session?” the response of the

participants revealed that the mode of the scores was 8 (Range 5 – 10) and no respondent scored the session lower than 5. The evaluation form utilized by Novartis looked at the perception of the participants with regard to the Facilitator for MDC Case 1, Speakers’ expertise in handling the session and Facilitator for MDC Case 2. 19 Of 24 (79%) respondents strongly agreed that the Facilitator for MDC 1 performed well while 5 of 24 (21%) agreed. The average score given by the 24 respondents was 3.79 out of a possible 4. 17 of 24 respondents (71%) strongly agreed that the speakers were experts in the case while 7 of 24 (29%) respondents agreed. The average score given was 3.71 out of a possible 4. 14 of 24 respondents (58%) strongly agreed that the Facilitator for MDC 2 performed well while 10 of 24 (42%) agreed. The average score given was 3.58 out of a possible 4. This level of satisfaction with session 3 remains consistent with those reported using the ASC evaluation form.

Session 4 of the ASC Sarcoma Preceptorship consisted of topics related to Ewing’s Sarcoma and Rhabdomyosarcoma. The activities in Session 4 consisted of 3 didactic lectures followed by 1 MDC discussion. The results of the evaluation of this session utilizing the ASC feedback form were as follows:

1. As to the question, “Was the time allocated to the session sufficient?” all of 30 (100%) respondents reported that the time allocated was sufficient.
2. As to the question, “Was the time allocated for the multidisciplinary conference enough?” all of 30 (100%) respondents replied that the time allocated was sufficient.
3. As to the question, “Were the number of topics covered adequate?” all 30 (100%) respondents felt that the scope of the topics and the number of topics covered were adequate. 30 (100%) respondents also agreed that the topics were helpful.
4. When asked to recommend other topics for this session, a respondent asked for continuing workshop for radiation oncologists.
5. Difficulties reported by respondents with regard to Session 4 included the following:
 - a. Language barriers
 - b. Not practical for day-to-day use
 - c. Too complicated
 - d. May not be applicable for surgeons if too much focus on chemotherapy
6. In addition, the following comments were reported by the participants with regard to all four sessions:
 - a. Pens do not work
 - b. Some groups did not conduct group discussions
 - c. Suggestions for complete history and relevant reports to be given for case discussions
 - d. More coffee
 - e. Suggestion for including Psychiatrist for multi-disciplinary case discussion

With regard to the question, “On a scale of 0-10, how helpful did you find the session?” the response of the participants revealed that the mode of the scores was 8 (Range 6 – 10) and no respondent scored the session lower than 5. The evaluation form utilized by Novartis looked at the perception of the participants with regard to the Speakers’ expertise in handling the session and the handling of the Team Multidisciplinary Conference. 17 of 24 (71%) respondents strongly agreed that the speakers were experts in the session while 7 of 24 (29%) agreed. The average score given by the 24 respondents was 3.71 out of a possible 4. 16 of 23 (70%) respondents strongly agreed that the Team Multidisciplinary Conference was well conducted while 7 of 23 (30%) agreed. This level of satisfaction with session 4 remains consistent with those reported using the ASC evaluation form.

The overall results of the participant satisfaction survey were in general optimistic, with results from the ASC evaluation forms showing a high level of satisfaction among all the participants with regard to the time allotment and scope of topics for each session. Also, the level of satisfaction was high when the participants considered how helpful the topics of each session were. For all sessions, the participants’ most common score was an 8 out of a possible 10 and no session received a rating of less than 5. These results are summarized in Tables 2 and 3.

It should be noted that, when compared to the responses provided by the participants, the responses of the faculty showed similarly high rates for overall satisfaction in terms of time allotment, scope of the topics and helpfulness (See: RESULTS - Evaluation Data for Program Faculty). These further bolsters the high satisfaction ratings garnered by this initial program.

The high satisfaction ratings provided by the participants remained consistent when the results from the Novartis evaluation form were considered. These results are summarized in Table 4. The form used the same responses as those used in the ASC form (strongly agree, agree, disagree

Table 2. Satisfaction Ratings for Each Session

Session Number	Sufficiency of Time for Each Topic	Sufficiency of Time for Multidisciplinary Case Conference	Adequacy of Scope of Topics
1	97%	87%	100%
2	97%	94%	100%
3	97%	94%	100%
4	100%	100%	100%

Table 3. Mode Scores for Helpfulness of Session

Session Number	Helpfulness of Session (MODE)
1	8
2	8
3	8
4	8

Table 4. Mean Scores for Evaluation of Speakers' and Facilitators' Expertise from Novartis Evaluation Forms

Session Number	Mean Score
Session 1	3.75
Session 2	3.75
Session 3	
Multidisciplinary Case 1	3.79
Speakers	3.71
Multidisciplinary Case 2	3.58
Session 4	
Speakers	3.71
Team Multidisciplinary Conference	3.70
Average Score for All Sessions	3.71

Table 5. Mean Scores for Factors Specific to the Novartis Evaluation Form

Factor	Mean Score
Program met its educational objectives	3.67
Opportunities for interactive learning from faculty and participants	3.63
Content is relevant to clinical practice	3.57
Addressed issues that could improve professional competence	3.58
Change in clinical practice within the next 6 months	3.29
Fair, balanced, free from commercial bias	3.75
Average Score	3.58

and strongly disagree) but added a numerical score for each so that an average score for each evaluated factor could be generated (strongly agree = 4, agree = 3, disagree = 2, strongly disagree = 1). As can be seen with the results, there was a high satisfaction rating given to the conduct of the program, with the mean average score given being 3.71 out of a possible 4 and no factor scored lower than 3.58.

In addition, the Novartis evaluation form included several factors that were not seen in the ASC's form. We felt that the results from these were also important as they could give a broader perspective on the degree of satisfaction of the participants in this program. A summary of these results can be seen in Table 5.

Overall, these results reflected the high satisfaction rating from the participants of the program. Specifically, most participants agreed that the program did meet its objectives, provided opportunities for interactive learning, had content that was relevant to clinical practice, addressed issues that improved professional competence, and allowed participants the opportunity to actually change their practice within the next 6 months. In addition, despite the fact that this was a program that was assisted in its delivery by a pharmaceutical company, most participants agreed that it was presented in a fair and balanced manner, free from any commercial bias.

In summary, participants felt that the strongest points of the program included the speakers, the multidisciplinary

case presentation format, the small group dynamics, interactivity between faculty and participants, and the conduct of discussions that were well organized, educational and friendly. These are very important points to consider and must be re-incorporated and re-emphasized in the design of any future sarcoma preceptorship.

Additional Information from the Novartis Evaluation Forms

The Novartis Evaluation forms assessed the preceptorship sessions, speakers and facilitators. In addition, they also explored several other areas of inquiry that could also serve as satisfaction measures and are included here as additional information. Reactions to specific statement and/or questions posed by this survey tool are summarized as follows:

1. To the statement, "The program met the stated educational objectives" 16 of 24 (67%) strongly agreed while 8 of 24 (33%) agreed. The average score given was 3.67 out of a possible 4.
2. To the statement, "This program included opportunities to learn interactively from faculty and participants" 15 of 24 (63%) strongly agreed while 9 of 24 (37%) agreed. The average score given was 3.63 out of a possible 4.
3. To the statement, "The content of this program is of significant relevance to my practice" 12 out of 21 (57%) strongly agreed while 9 out of 21 (43%) agreed. The average score given was 3.57 out of a possible 4.
4. To the statement, "This program addressed issues that will help me to improve my professional competence" 15 of 24 (63%) strongly agreed, 8 of 24 (33%) agreed and 1 of 24 (4%) disagreed. The average score given was 3.58 out of a possible 4.
5. To the statement, "I intend to make changes in my clinical practice in the next 6 months" 9 of 24 (37%) strongly agreed, 13 of 24 (54%) agreed and 2 of 24 (9%) disagreed. The average score given was 3.29 out of a possible 4.
6. To the question, "Do you believe this activity was fair, balanced and free of commercial bias?" 18 of 24 (75%) strongly agreed and 6 of 24 (25%) agreed. The average score given was 3.75 out of a possible 4.

In addition to the survey questions presented above, the Novartis Evaluation Form also included three open-ended questions for additional feedback from the participants. Their responses to these questions are summarized in **Appendix 10: Participant Responses to Open-Ended Questions**.

In summary, the additional information provided by the Novartis Evaluation served to bolster the results of the evaluation conducted by the ASC. A very high satisfaction rating was given by all participants to the content and conduct of the program as well as to the performance of its faculty, the activities conducted and the logistical resources used. This was consistent with what the ASC evaluation forms had revealed.

Evaluation Data from Program Faculty

The faculty participating in the Sarcoma Preceptorship were also asked to evaluate the program. This would allow a perspective from the faculty, that would help to improve future iterations of the program. Of the 14 members of the ASC who served as faculty of the program, 7 were able to submit completed evaluation forms. The results from these are as follows:

1. 2/7 felt that time allocated was insufficient.
2. All felt that the time allocated for the multi-disciplinary case was enough
3. All felt the range was adequate
4. All felt the number of topics covered were sufficient
5. All felt the number of sessions were adequate

With regard to the question of helpfulness of this activity to participants, the responses of the faculty showed that mode of the scores was 9 with no faculty evaluating the helpfulness of the activity with a score less than 8. This supported the feedback of the participants who also gave similarly high rates for this particular factor.

When the faculty was asked for suggestions for additional topics, these were their responses:

- Imaging in STS
- Radiology input in differential diagnosis
- Challenges and pitfalls in interpreting responses to treatment with radiological imaging
- Addition of Pathology and Radiology discussions
- Potential targets for systemic therapy
- Fertility preservation
- Adolescent – Young Adult issues
- In addition, the faculty identified language barriers as the biggest obstacle in conducting the program. Other comments provided included the following: excellent first workshop, appreciative that time was given for questions and answers, afternoon sessions too long.

Level 2 Evaluation Results: Pre-test and Post-test

A total of 28 of the 30 participants were able to complete both the pre-test and the post-test. Two other participants were able to take the pre-test but left the program early due to travel schedules and were thus unable to complete the post-test. The pre-test scores of these 2 were excluded from the analysis. Post-test was administered upon the conclusion of the final MDC. However, because of time constraints and departure schedules for almost all remaining participants, no set time limit was given but it was observed that most participants completed the test in a far shorter time than they did in the pre-test. This was most likely due to the pressure of making it to the airport in order to catch their respective flights back home. Also, because participants had different departure schedules, post-test was not accomplished at the same time, with those leaving at earlier times completing the post-test earlier than those with later departure times. Table 6 shows the comparison between the average scores

Table 6. Comparison of Average Test Scores of the Pre-Test and Post-Test and the Calculated Results of the KR21 Coefficient of Reliability

	Average Score	Range	Test Reliability Coefficient
Pre-test	13.4	5-18	1.033*
Post-test	11.4	2-19	-95.57**
	-2.0		

* Indicates that the test is both homogenous and reliable

** Indicates an unreliable test - may be the effect of external events.

garnered in both the pre-test and the post-test as well as the calculated results of the KR21 Coefficient of Test Reliability.

The results of the Level 2 evaluation of the program showed that there was no improvement in outcomes between the pre-test and post-test results of participants. Instead, comparison of the average scores garnered between both tests showed that the average score of the post-test was 2 points lower than that of the pre-test. However, the discrepancy in the results may be explained by the results of the KR21 Coefficient of Test Reliability. The pre-test KR21 results were very favorable, resulting in a calculated result of 1.033. This was way above the accepted threshold of 0.70 and indicates that the pre-test was both reliable and homogenous and that the results were internally consistent in evaluating the participant's individual achievements in the test. However, the calculated KR21 of the post-test was a disappointing -95.57. This clearly indicates that the test was both unreliable and inhomogeneous and was not consistent in evaluating the individual achievements of the examinees. It must be made clear that Test Reliability may be affected by many factors both internal and external to the test. Since the same examination was administered in both situations, it is believed that several external factors may have contributed to the unreliability of the post-test.

DISCUSSION

This preceptorship in soft tissue sarcomas was a pilot program and it was important to determine its efficacy in obtaining its objectives, particularly because there were plans to offer this on a regular basis. Also, it was important to identify problem areas encountered so that these may be addressed in subsequent preceptorship programs. To this end, this study was designed in order to evaluate the efficacy of the program in terms of participant satisfaction and improvement in sarcoma knowledge base, by using the first two levels of the Kirkpatrick model for program evaluation.

Preceptorship refers to a period of practical training for a student or a novice under the guidance of a specified instructor, termed the preceptor. In a clinical setting, it aims to integrate senior trainees and new graduates into the clinical and practical aspects of medical care. A preceptorship program may play an important role in the improvement of care for rare and unusual cases such as soft tissue sarcomas.

In her review, Poon noted the existence of an unmet need for sarcoma-specific forums in Asia as well as a lack of preceptorships from which medical trainees may learn the basics of sarcoma management.¹¹ This scenario is further compounded by the rarity of this group of tumors thus lowering the experience level of health care professionals who are not specifically engaged in sarcoma care. The Sarcoma Policy Checklist, an infographic produced to identify gaps in and propose solutions to the problem of sarcoma care in Europe, identified that this lack of experience may result in inaccurate diagnosis which in turn leads to mismanagement in 70% of patients.¹² Such a scenario argues strongly for the need to develop educational forums, such as preceptorships, to share knowledge and experience in sarcoma management with other health care personnel in the region.

Participant satisfaction is an important aspect of this evaluation process. Papastavrou in his study on clinical settings in nursing education in Cyprus investigated the student's satisfaction of the clinical settings as a learning environment and considered student satisfaction as an important aspect of such an evaluation. His results showed that nursing students showed high satisfaction ratings for the clinical learning environment, which included pedagogical atmosphere, Ward Manager's style, the premises of Nursing in the ward, the supervisory relationship, and the role of the Nurse teacher. Supervisory relationship was identified as the most influential factor in student satisfaction. They concluded that student satisfaction could be used as an important contributing factor in the development of the clinical learning environment.¹³ Our own satisfaction surveys revealed similar tendencies, with participants giving high satisfaction ratings for sufficiency of time for lecture topics and MDC and the adequacy of topics. Most agreed that sessions were very helpful to them. These high ratings were echoed when the preceptorship faculty were surveyed. Further, the independent survey conducted by the sponsor also produced high satisfaction levels, particularly for attainment of program objectives and the opportunity for interactive learning. Faculty expertise was similarly given high satisfaction ratings, which is comparable with that seen in the nursing experience. A comparable experience was reported by Hsing-Yuan, who evaluated the satisfaction levels of faculty and students of a nursing practicum project workshop in Northern Taiwan. This involved a 1-day workshop that included 14 participants. Similar to our experience and that of the Cyprus study, high satisfaction levels for the activity were reported by the participants in terms of overall performance of speakers, course content, class environment, quality of service and self-evaluation. They concluded that such activities were effective educational tools based on these high levels of satisfaction.¹⁴ Based on the experiences from these two events, we can similarly conclude that the high level of satisfaction we garnered may be indicative of the effectivity of our conducted preceptorship. However, high levels of participant satisfaction do not necessarily mean that there would be no revisions needed for future iterations of

the activity. In the Northern Taiwan experience, which was also a pilot program, they still identified areas that needed refinement in spite of the high satisfaction levels reported. These included strengthening some workshop topics and providing more time for the activity.¹⁴ In our experience, there were several weaknesses identified by our participants in spite of their high levels of satisfaction. These were as follows: time management, manner by which cases were presented in the multidisciplinary case conferences, inconsistent conduct of the MDCs, sessions were too busy, and not enough topics for some sub-specialties (example: radiation oncology). But the factor that was consistently reported as a potential weakness was the language barrier, given that not all participants come from countries with a similar level of English proficiency. These weak points should serve as guides as to how to better improve future sarcoma preceptorships. Specifically, the following improvements were proposed: 1) For time management, dedicated time must be allocated for the pre-test and post-test as well as other evaluation-related activities such as completion of program evaluation forms as well as dedicated "relaxation" time; 2) The content of the program was deemed topic-heavy for a two-day schedule so future iterations should consider a more compact program with a limited list of topics but each approached in a more comprehensive manner; also, more exposure to other subspecialties that manage sarcomas and a reduction in the number of case management discussions with emphasis on audience participation be encouraged; 3) Language issues may have to be addressed and may include the selection of participants with fluency in the English language or the provision for interpreters for participants with language difficulties; 4) There is a need to address the inconsistencies in the faculty's handling of the MDC by developing a standard discussion format, the training of MDC facilitators and the establishment of "rules of engagements" for every MDC case by presenting the objectives and desired outcomes of each MDC case to be discussed.

The results of the Level 2 evaluation of learning for the sarcoma preceptorship program are another matter. The results were previously summarized in Table 6. What is immediately striking is the fact that the participants actually scored lower in the post-test than in the pre-test. This outcome seems to run counter to the high level of participant satisfaction. However, lower post-test results do not necessarily reflect a lack of learned knowledge from the experience, given the high satisfaction ratings reported for content, speakers' and facilitators' performances and conduct of the different activities. What then could account for this unexpected result? Tests are evaluation tools designed to measure knowledge, skills, attitudes and qualifications of students. Effective tests should always have three basic qualities: validity, reliability and practicality.¹⁵ Validity is a test quality that refers to the extent by which a test measures what it intends to measure. Valid tests have a congruency between the student's attainment of the learning objectives and their performance

as reflected in their test scores. Therefore, the content of the test must reflect the content of the learning objectives. Our pre-/post-test was constructed by the faculty of the program, who were all sarcoma experts in each of their locals. To ensure conformity to the program's learning objectives, a table of specifications was prepared in order to determine the number of questions assigned for each learning objective. Once the questions were constructed, these were circulated among the rest of the program faculty in order to determine fitness for inclusion, clarity and question construction before the final tool was prepared. This process helped ensure that the content of our test was valid as far as the objectives of the program were concerned. Practicality refers to usefulness of the testing procedure and a practical test should be cost-efficient in terms of construction, administration and scoring. Our test was short, with all items presented in a multiple-choice format and was easy to score, making it a practical test. Finally, reliability refers to the consistency by which a test measures what it intends to measure. A reliable test must be capable of producing consistent and reproducible results. A test with no or low reliability cannot reflect the true ability of the examinees. Zhu and Han identified the factors that affect test reliability.¹⁵ These include the following:

1. Change of candidates and the process of testing: refers to personal factors that could affect the consistency of examinee performance and include events such as illness, exhaustion, emotional state and psychological factors; process of testing refers to the testing environment and include such factors as time, presence of distractions such as noise and comfort levels of the testing areas. Lack of time adversely affects test reliability because the examinee may not be able to complete the test in the specified amount of time, which could also introduce emotional factors such as anxiety which ultimately affects consistency of performance. Uncomfortable testing areas, test given too early or too late in the day and areas with a high level of extraneous noise could all result in inconsistencies in test performance.
2. Testing features: this refers to the length of the test and the level of difficulty of the test items. Very long tests may adversely affect reliability by magnifying the time factor while very difficult tests may result in low test scores but do not reflect actual ability since the difficulty level may be beyond their level of competence.
3. Going over test results: this refers to test items that introduce a level of subjectivity to their assessment. It is a particular issue with essay-type tests and may be addressed by using a purely objective examination.

Similarly, Crisostomo listed the following as potential threats to test reliability: unclear directions, insufficient time allotment, lengthy examinations, presence of distractions, disturbances during test administration, ambiguous test items and lack of objectivity in scoring leading to wide disagreement among peers.¹⁶

We are confident that the pre-/post-test prepared for the Level 2 evaluation of this program were both valid and practical. Could the difference in pre- and post-test results be a problem of test reliability? Reliability can be estimated by calculating the reliability coefficient. In this scenario, we chose to use KR21 coefficient based on the assumption that the test items were homogenous and had similar levels of difficulty. For the pre-test, the reliability coefficient was high at 1.033, indicating a test that is highly reliable in measuring the performance of examinees on the content examined. Based on this as well as the analysis of test validity and practicality, we could surmise that we had a valid, reliable and practical pre-test. However, the post-test results were a different matter. The reliability coefficient was very low at -0.9557, indicating a very unreliable test. However, the test was exactly the same as the pre-test, which we had judged as a competent tool. Therefore, the problems of test reliability were not due to factors that were inherent to the test, such as directions, test length, item ambiguity or issues with objectivity in scoring. In other words, the reliability issues with the post-test were most likely from extraneous sources such as time and presence of distractions during the testing period. A review of the conduct of the post-test revealed that the test was administered immediately upon completion of the last session. Furthermore, the administration of the test was conducted at the time when most participants were getting ready to leave for the airport to catch their individual flights back home. Some participants had to leave early and thus took the test at an earlier time compared to those with departure times after the activity had concluded. In two instances, flight schedules were so tight that they failed to accomplish the post-test. This time element factor could have caused the participants undue stress in completing the examination, given the time pressures present and the need to make it to the airport in time. This could have then resulted in a rushed completion of the exam, with the focus more on the need to finish rather than the need to demonstrate what they had actually learned. In fact, a curious observation by many of the preceptorship faculty was that the participants took less time in completing the post-test compared to the pre-test. These extraneous factors are a strong influence on test performance and could have greatly contributed to the resultant unreliability of the post-test results. Therefore, adequate time management, with enough time allocated for the conduct of evaluation exercises such as the completion of satisfaction survey forms and the accomplishment of all items in the pre-test and post-test should be considered when designing the program for future iterations of this activity. Specifically, the program schedule of future preceptorships must also include a time frame for the completion of program evaluation activities such as these. This would avoid the effects of external factors that could introduce undue stress levels on the test takers, allowing them the opportunity to actually demonstrate the level of knowledge they had amassed as a result of the program.

Based on the evaluation exercise of the Sarcoma Preceptorship program pioneered by the ASC, we can consider the initial program a success in terms of participant and faculty satisfaction indices. High marks were consistently given for session times, adequacy of topics, performance of speakers and facilitators and helpfulness in future practice. Specifically, the participants were quite satisfied with the following: Preceptorship as a whole; Performance of the faculty (Faculty was graded as experts); Delivery of topic content (Conduct of the session and Multidisciplinary case discussions); Emphasis on team approach (Insights into how different subspecialties “think”); Preceptorship logistics. The fact that similarly high satisfaction ratings were given in the Novartis evaluation form results further bolsters the fact that both participants and faculty perceived the first sarcoma preceptorship program as a highly satisfying activity to attend and participate in.

CONCLUSIONS AND RECOMMENDATIONS

The evaluation of this first sarcoma preceptorship program has shown that continuing such an activity, if feasible, may help fill a void in sarcoma knowledge and information, allowing less experienced oncologist and other allied health care professionals the opportunity to improve and broaden their abilities in sarcoma management. In doing so, the ASC could well be on its way in achieving the goal of contributing to the improved care of sarcoma patients in Asia.

Nevertheless, any new program is not without its own faults and weaknesses and these are part of the natural growing pains in the development of new educational programs. The first sarcoma preceptorship program is no exception. Indeed, despite the high level of satisfaction provided by participants and faculty, there were also areas identified where weaknesses were noted. These should be viewed as potential opportunities by which this program may improve and grow, thereby benefiting future participants. In the light of these identified weaknesses and areas of improvement, we recommend the following for future programs of a similar nature:

1. Preceptorship must be structured with content to fit preceptorship time frame
 - a. Speakers must concentrate on “must know” areas of each topic
 - b. Multidisciplinary case discussion
 - i. Rules of the activity must be clear: limit number of cases and these must illustrate basic and problematic cases; emphasize interaction between participants within each group and between participants and faculty; create heterogenous small groups in order to maximize interaction between participants, allowing them to experience another specialty’s point of view and overcome the language problem
 - ii. Budget time for evaluations such as pre- and post-tests, accomplishment of evaluation forms and consider preparing online evaluation forms

2. Program evaluation
 - a. Revise tools to ones capable of capturing relevant data, including preceptorship content, preceptorship mechanisms, feedback from participants and faculty
 - b. Develop an on-line evaluation tool for paperless data collection that may be easier to analyze, collate data and generate statistics
 - c. In line with a paperless data collection tool, develop an on-line answer sheet for both pre- and post-test
 - d. Factor in time for evaluation activities, including pre- and post-test, completion of survey forms and faculty de-brief at the end of the preceptorship
3. Utilize data from this evaluation as a “pilot study” to identify areas for further improvement, identify pertinent information needed and apply lessons learned to the next preceptorship
4. Publish study results as a pilot study focusing on lessons learned and as a development exercise, focusing on how improvements can be applied to the next preceptorship program

Statement of Authorship

GYF contributed in the conceptualization, acquisition and analysis of data, drafting and revision, and final approval of the version to be published; EPYL contributed in the acquisition and analysis of data, revision of text, and final approval of the version to be published; GWL and HJ contributed in the acquisition and analysis of data; RQHH contributed in the conceptualization and final approval of the version to be published.

Author Disclosure

All authors declared no conflicts of interest.

Funding Source

None.

Supplemental Data

Appendix Tables and Figures are available upon request to the corresponding author.

REFERENCES

1. Burningham Z, Hashilbe M, Spector L, Schiffman J. The epidemiology of sarcoma. *Clin Sarcoma Res.* 2012 Oct; 2(1):14 doi: 10.1186/2045-3329-2-14.
2. Clark MA, Thomas JM. Delay in referral to a specialist soft-tissue sarcoma unit. *Eur J Surg Oncol.* 2005 May;31(4):443–8. doi: 10.1016/j.ejso.2004.11.016.
3. Dyrop H, Vedsted P, Raedkjaer M, Safwat A, Keller J. Routes to diagnosis for suspected sarcoma: the impact of symptoms and clinical findings on the diagnostic process. *Sarcoma* [Internet]. 2016 [cited 2022 Aug 10]. Available from: <https://www.hindawi.com/journals/sarcoma/2016/8639272/>.
4. Sarcoma UK. Soft tissue sarcoma [Internet]. [cited 2022 Aug 10]. Available from: <https://sarcoma.org.uk/sarcoma-types/soft-tissue-sarcoma>.

5. National Comprehensive Cancer Network. Soft tissue Sarcoma (Version 2.2019). [Internet]. 2019 [cited 2019 May]. Available from: https://www.nccn.org/professionals/physician_gls/pdf/sarcoma.pdf
6. Casali PG, Abecassis N, Aro HT, Bauer S, Biagini R, Bielack S, et al. Soft tissue and Visceral Sarcomas: ESMO-EURACAN Clinical Practice Guidelines for Diagnosis, Treatment and Follow-up. *Ann Oncol.* 2018 Oct;(29 Suppl 4): iv51–iv67. doi: 10.1093/annonc/mdy096.
7. Quek R, Loong HH, Sriuranpong V, Farid M, Tan SH, Goh WL, et al. Epidemiology, real world treatment and outcomes of 423 patients (pts) with angiosarcoma (AS) in Asia: A report from the Asian Sarcoma Consortium (ASC). *Ann Oncol.* 2016; 27(Suppl 9):ix163. doi:10.1093/annonc/mdw597
8. Chen TWW, Pang A, Puhaindran ME, Maw MM, Loong HH, Sriuranpong V, et al. Optimal first line systemic therapy in patients (pts) with metastatic angiosarcoma: A report from the Asian Sarcoma Consortium. *Ann Oncol.* 2016; 27(Suppl 9): ix163.
9. Kirkpatrick DL, Kirkpatrick JL. *Evaluating Training Programs*, 3rd edition. San Francisco, CA: Berrett-Koehler Publishers, Inc.; 2006. pp. 3–26.
10. Statistics How To: Statistics for the Rest of Us! [Internet]. Stephanie Glen. "Kuder-Richardson 20 (KR-20) & 21 (KR-21)" [cited 2019 Nov 15]. Available from: <https://www.statisticshowto.com/kuder-richardson/>.
11. Poon E, Quek R. Soft tissue sarcoma in Asia. *Chin Clin Oncol.* 2018 Aug;7(4):46. doi: 10.21037/cco.2018.08.06.
12. Sarcoma Patients EuroNet. Casali P, Drove N, Dumont S, Eriksson M, Gronchi A, Kasper B, et al. The Sarcoma Policy Checklist [Internet]. [cited 2022 Aug 15]. Available from: <http://www.sarcoma-patients.eu/en/sarcoma-reports2/policy-checklist>.
13. Papastavrou E, Dimitriadou M, Tsangari H, Andreou C. Nursing students' satisfaction of the clinical learning environment: a research study. *BMC Nurs.* 2016 Jul;15:44. doi: 10.1186/s12912-016-0164-4.
14. Liu HY, Kuo CL, Shen MH, Lee PY, Pi-Yu Lee PY, Ku YL. Evaluating Faculties and Students Satisfaction of a Nursing Practicum Project Workshop in Northern Taiwan. *Pulsus Journals* [Internet]. 2018 April 12 [cited 2022 Aug 10]. Available from: <https://www.pulsus.com/scholarly-articles/evaluating-faculties-and-students-satisfaction-of-a-nursing-practicumproject-workshop-in-northern-taiwan-4426.html>.
15. Zhu J, Han L. Analysis on the Main Factors Affecting the Reliability of Test Papers. *J Lang Teach Res.* 2011 Jan;2(1):236–238. doi:10.4304/jltr.2.1.236-238
16. Crisostomo AC. Test Construction and Analysis. In: Erlyn Sana, editor. *Teaching and Learning in the Health Sciences*. Diliman, Quezon City: The University of the Philippines Press; 2010. pp. 203–234.