Evaluation of Clinical Preceptor Training and Its Impact on Inter-Rater Reliability

Matthew J. Mendoza, MSc, RCP, RRT-ACCS, RRT-NPS, AE-C – Department of Respiratory Therapy, San Joaquin Valley College, Visalia, California
Thomas Barnes, PhD, RRT, FAARC – College of Professional Studies, Northeastern University, Boston, Massachusetts

Introduction

The consistency of clinical evaluations of respiratory therapy (RT) students by clinical preceptors is an area where research is needed. The purpose of this study is to evaluate if standardized clinical preceptor training will improve inter-rater reliability (IRR) scores for faculty involved in evaluating the clinical skills of respiratory therapy students. We hypothesized that standardized clinical preceptor training will not significantly affect IRR scores for faculty members.

Methods

Study Design: A one-way repeated measures design evaluating mean IRR scores by clinical preceptors before and after training will be used. The repeated measures design is one of the most efficient methods for controlling inter-subject differences. Participants will score a student completing a clinical final in SJCVC’s clinical simulation lab using the institution’s RT32 Clinical Final Rubric. Scores will be recorded on a scorecard designed for this research study. Then, participants will complete preceptor training before completing a second evaluation of the student completing another clinical simulation using the RT32 Clinical Final Rubric. The participants will record their second scores on their scorecard as well.

Results

All faculty members met inclusion criteria and results showed that standardized training improved IRR scores for faculty members. Mean pre-training IRR scores for participants resulted in a score of 63% which was a "Needs Improvement" score. Mean post-training IRR scores for participants resulted in a score of 78% which was a "Good" score. Standardized preceptor training improved IRR scores by 15% and had a P value <0.05 showing that scores were significantly different.

Data Analysis

Data will be collected on the study scorecard. Results will be analyzed by comparing percent agreement and the mean IRR score difference between groups by using a paired-sample t-test using SPSS software. Percent agreement is examined due to its ease of calculation and the paired-samples t-test was selected for this study’s variables when analyzing the statistical data recorded. Inter-rater reliability will be measured by percentage agreement. Agreement rate (A) is the observed agreement (O) divided by the total possible agreement (P). P = A/O. Scores will be categorized as ≤70% = Needs improvement; 70-79% = good; 80-89% = great; and ≥90% = excellent. Significance can be seen with a p value of <0.05 using the paired-sample t-test.

Discussion

The results demonstrate that preceptor training significantly improves IRR scores. Focused training sessions for clinical preceptors resulted in higher IRR between preceptors involved with student evaluations. Standardized preceptor training improved percent agreement IRR scores by 15%. A paired-samples t-test was used to determine whether there was a statistically significant mean difference between the IRR scores before and after preceptor training. Data are mean ± standard deviation, unless otherwise stated. There were no outliers in the data, as assessed by inspection of a boxplot. The assumption of normality was not violated, as assessed by the Shapiro-Wilk’s test (p>0.05). Participants had higher IRR scores (76.17 ± 4.26) after preceptor training compared to no training (62.67 ± 11.71), a statistically significant increase of 15.5 (95% CI, 6.88-24.12, t=4.621, p<0.05, d=1.89). This is relevant to all educators because standardized preceptor training has shown in our study to improve inter-rater reliability and IRR scores required for accreditation. Alternative explanations for our results could be that the group did better on the second evaluation because they had already gone through the testing procedure itself or that the second clinical evaluation was easier to score.

These results are clinically significant because the data shows standardized preceptor training improves inter-rater reliability among preceptors. This will promote consistent and fair evaluations of respiratory students allowing for a more effective clinical experience. More efficient evaluations of students will allow more time to be dedicated to training in focused areas of individual need for the students.

Our data reflects the earlier works of Dunlevy (2016) and Rye (2009) demonstrating that inter-rater reliability can be improved by developing standardized training for clinical preceptors. The AARC did a fantastic job with their Clinical PEP training addressing many professional behavior issues that arise with clinical precepting, but failed to address the need to make available strategies or tools to help respiratory therapy programs improve IRR efficiency between clinical preceptors. Our research hopes to add to the existing evidence showing that standardized preceptor training improves IRR scores and that this should be the next major update to the AARC’s Clinical PEP training program.

One potential limitation with this study is how to evaluate the lasting effects of the completed preceptor training by San Joaquin Valley College-Visalia’s clinical educators. Since this training will be completed and evaluated over 1 day, measuring the retention of training by staff will need further evaluation and follow-up training. Another limitation is the sample size. With the small sample size in this research study, the estimated power of the study is also low leading to a more likely chance of a Type II error. Significant effects may be missed because of inadequate sample size or failure to control extraneous sources of variation. Unanswered questions after this study are if there any correlations between IRR scores and any secondary participant characteristics i.e. age, education, gender, experience, earned specialty credentials?

Conclusions

Standardized preceptor training significantly affects IRR scores. Our research shows that preceptor training significantly improves mean IRR scores recorded by clinical preceptors. The mean difference was statistically significant from zero (p<0.006). Therefore, we can reject the null hypothesis and accept the alternative hypothesis. Clinical preceptor training does affect inter-rater reliability. With our small sample size, we speculate that the results of standardized training in larger samples would be similar and this is an area of future research that is needed. Another question that would require future research would be establishing a standard preceptor training interval that would allow minimum decline in IRR scores between clinical preceptors.

Disclosures

Sources of Financial Support: None
Conflict of Interest: None

Table 1: Participant Characteristics

<table>
<thead>
<tr>
<th>Participant</th>
<th>Age (Years)</th>
<th>Educational Level</th>
<th>Gender</th>
<th>Experience (Years)</th>
<th>Specialty Credentials</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>60-69</td>
<td>AS</td>
<td>Male</td>
<td>&gt;35</td>
<td>RRT</td>
</tr>
<tr>
<td>#2</td>
<td>40-49</td>
<td>MA</td>
<td>Female</td>
<td>11-15</td>
<td>None</td>
</tr>
<tr>
<td>#3</td>
<td>50-59</td>
<td>BS</td>
<td>Female</td>
<td>&gt;35</td>
<td>None</td>
</tr>
<tr>
<td>#4</td>
<td>30-39</td>
<td>BS</td>
<td>Female</td>
<td>2-5</td>
<td>RRT-NPS</td>
</tr>
<tr>
<td>#5</td>
<td>30-39</td>
<td>BS</td>
<td>Male</td>
<td>2-5</td>
<td>RRT-ACCS</td>
</tr>
<tr>
<td>#6</td>
<td>30-39</td>
<td>AS</td>
<td>Male</td>
<td>11-15</td>
<td>AE-C</td>
</tr>
</tbody>
</table>