Using SAS® to Examine Internal Consistency and to Develop Community Engagement Scores

Renee Gennarelli, MS, Melody S. Goodman, PhD
Division of Public Health Sciences, Department of Surgery, and Alvin J. Siteman Cancer Center
Washington University School of Medicine, St. Louis, MO, USA

ABSTRACT

Comprehensive cancer centers have been mandated to engage communities in their work; thus, measurement of community engagement is a priority area. Siteman Cancer Center’s Program for the Elimination of Cancer Disparities (PECaD) seek to align projects with 11 Engagement Principles (EP) previously developed in the literature. Participants in a PECaD pilot project were administered a survey with questions on community engagement in order to evaluate how well the project aligns with the EPs. Internal consistency is examined using PROC CORR with the ALPHA option to calculate Cronbach’s alpha for questions that relate to the same EP. This allows items that have a lack of internal consistency to be identified and to be edited or removed from the assessment. EP-specific scores are developed on quantity and quality scales. Lack of internal consistency was found for six of the EP’s (alpha<.70). After editing the items, all EP question groups had strong internal consistency (alpha>.85). There was a significant positive correlation between quantity and quality scores (r=.918, P<.001). Average EP-specific scores ranged from 6.87 to 8.06; this suggests researchers adhered to the 11 EPs between sometime and most of the time on the quantity scale and between good and very good on the quality scale. Examining internal consistency is necessary to develop measures that accurately determine how well PECaD projects align with EPs. Using SAS® to determine internal consistency is an integral step in the development of community engagement scores.

INTRODUCTION

Community engagement is a powerful vehicle for bringing about changes that can improve community health; engaging community members in the research process is often the missing link to improving the quality and outcomes of health promotion activities, disease prevention initiatives, and research studies. Community engagement requires a long term process that builds trust and values contributions of all stakeholders through collaboration. Health concerns unique to marginalized communities can be addressed through the help of collaboration between researchers and those communities.

Comprehensive cancer centers have been mandated to engage communities in their work. Thus, the measurement of community engagement in projects supported by cancer centers is a priority area. While the importance of community engagement in research has been established, there are few evidence-based approaches for measuring the level of community engagement in research projects. We developed an instrument to measure community engagement through the examination of internal consistency and the development of informative community engagement scores using SAS®.

Siteman Cancer Center’s Program for the Elimination of Cancer Disparities (PECaD) works with communities to reduce cancer disparities through outreach, education, and training. PECaD projects seek to align with 11 Engagement principles (EP) based on the principles of community-based participatory research (CBPR) that have been previously developed in the literature. The 11 CBPR Engagement Principles (EP) utilized by PECaD are:

1. Focus on local relevance and determinants of health
2. Acknowledge the community
3. Disseminate findings and knowledge gained to all partners
4. Seek and use the input of community partners
5. Involve a cyclical and iterative process in pursuit of objectives
6. Foster co-learning, capacity building, and co-benefit for all partners
7. Build on strengths and resources within the community
8. Facilitate collaborative, equitable partnerships
9. Integrate and achieve a balance of all partners
10. Involve all partners in the dissemination process
11. Plan for a long-term process and commitments

Community-Based Participatory Research (CBPR) engages communities as partners in the research process. Community members share their knowledge and experience to help identify key problems and develop culturally sensitive research questions. The Community Research Fellows Training (CRFT) program is a current PECaD pilot project that focuses on enhancing the infrastructure for CBPR and seeks to promote the role of racial, ethnic,
and other underserved populations in the research enterprise. Fifty community members were selected to participate in this 15-week long training program based on the CARES Fellows training program.

**METHODS**

Forty-seven fellows and 11 members of the CRFT Community Advisory Board (CAB) were administered a survey with 96 questions on community engagement to evaluate how well the project aligns with the principles. These items, each pertaining to a specific EP, with Likert scale response options are used for this analysis; half of the questions measure quality (strength) and the other half measure quantity (frequency) of community engagement. Each Engagement Principle is measured by 3-5 quality items and a corresponding 3-5 quantity items. All 48 quality questions had response options: 1=Very Poorly, 2=Poorly, 3=Adequately, 4=Well, 5=Very Well. The 48 quantity questions had response options: 1=Never, 2=Rarely, 3=Somewhat, 4=Often, 5=Always.

**INSTRUMENT EVALUATION**

Cronbach’s Alpha is used to assess reliability by measuring the degree to which different items are correlated and measure a single engagement principle (internal consistency). For this study, it is of interest to calculate Cronbach’s alpha for each set of community engagement questions that are meant to measure the same Engagement Principle. While there is no clear threshold, it is widely accepted that in the early stages of validation research $\alpha$ should exceed 0.70 to show internal consistency.

**WHAT YOU NEED TO KNOW ABOUT PROC CORR**

SAS® can calculate the Cronbach’s Alpha statistic using the ALPHA option of PROC CORR.

```sas
proc corr data=data_name alpha nomiss;
  var var1 var2 var3;
run;
```

The NOMISS option should be used to exclude observations with missing values from the calculation. It is also important to ensure that all items in the analysis have the same scale. PROC CORR assumes that all items are scored in the same direction. Therefore, items may need to be reverse-scored or rescaled before analysis. In our previous work we provide a detailed explanation on how to interpret PROC CORR output to examine internal consistency and further discuss the use of Cronbach’s alpha.

Larger sample sizes will lead to a more precise Cronbach’s Alpha estimate. However, previous studies have shown that sample sizes of $n=30$ are just about as precise as a larger sample of $n=200$ when the number of items used to calculate $\alpha$ is at least 5. This is a strength of Cronbach’s Alpha, especially for early stage validation research where sample sizes tend to be small.

**SCORING**

After ensuring the survey items for each EP are internally consistent it is necessary to develop a way to summarize survey responses. Developing summary scores for the community engagement items will create a way to accurately determine how well cancer centers are engaging communities in their research. Since there are few evidence-based approaches for measuring community engagement, several scoring methods are utilized and compared to develop the most informative community engagement scores. The scores developed assess the level of community engagement with higher scores corresponding to higher quality/frequency of community engagement.

Quality scores are created using the 48 quality items and quantity scores are created using the corresponding 48 quantity items. Five different scoring methods are used to calculate community engagement scores. (1) **Simple sum score** is the sum of all 48 responses for the specified scale (quality/quantity). (2) **Average sum score** is created for each scale by calculating the average score of the items for each individual EP and then summing those scores across all 11 principles. (3) An **overall average sum score** is created by summing the quality and quantity average sum scores. (4) Principle components analysis is used to create **weighted community engagement scores** for each scale. (5) **Overall average sum scores for each principle** are also calculated by summing the quantity and quality EP-specific average scores.

In an effort to account for the loss of observations due to unanswered questions, resulting in a small sample size (only 22 participants completed all community engagement items), observations with some missing data were included in the score development for overall and scale-specific average sum scores. Observations with missing data were included based on the following criteria: for each scale-specific (quality/quantity) average sum score, each
observation had to have at least 3 answered items for EPs that contained 4-5 items and at least 2 answered items for EPs that contained 3 items. Sensitivity analysis was performed using complete case data only (22 observations with responses to all community engagement items) to compare the score distributions with and without missingness and examine the how sensitive scores are to missing data.

**USING PROC FACTOR TO DEVELOP WEIGHTED SCORES**

Weighted scores for each scale-specific (quantity/quality) EP item group were calculated using principal components analysis (PCA). Since different items in the assessment are designed to relate to different constructs (engagement principles), we want to determine the weighting scheme of each construct and therefore determine which items have the greater influence on each principle. For this analysis, only components with eigenvalues ≥1 were used to calculate the weighted scores. As a result, each EP item group was scored using only the first component that accounted for most of the total variance. Once the weighted scores for each specific EP were calculated, these scores were added together to produce one weighted community engagement score per scale (quantity/quality). Only those observations with complete-case data were included in this score calculation.

It is important to note that PCA is based on standardized items, which are calculated for each item by subtracting the mean of the original item and dividing by the standard deviation of that item. The item-specific coefficients produced by PCA are multiplied by these new standardized items to produce the weighted scores for each principle.

PCA weighted scores were developed using the SAS® procedure PROC FACTOR.

```sas
proc factor data=data_name
simple corr
method=prin
priors=one
mineigen=1.0
ev score;
var var1 var2 var3;
run;
```

The SIMPLE option requests simple statistics while the CORR option displays the correlation matrix. The METHOD=PRIN option specifies that the method to be used is the principal axis (principal factor) method, which should be used for a principal components analysis. Using PRIORS=ONE sets all prior communalities to 1.0, which is appropriate for a principal components analysis where the diagonals of the correlation matrix are all 1.0; this is the default setting. It is important to note that if you wish to perform a factor analysis rather than a principal components analysis, estimates of prior communalities need to be supplied. More information on the PRIORS option can be found in the SAS® manual.

For this PCA, only components with eigenvalues ≥1 are used to develop the weighted scores; to specify this within PROC FACTOR, use the MINEIGEN=1.0 option. If you wish to output the results of the analysis, the NFACT= option should be used in place of the MINEIGEN= option. This option specifies the number of factors (principal components) to be retained in the analysis. For this particular analysis, NFACT=1 would be used since only one factor met the eigenvalue criteria for each group of questions. It is important to note that the OUT= option only works when NFACT= is specified.

Lastly, the EV and SCORE options display the eigenvectors of the reduced correlation matrix and the factor scoring coefficients, respectively. The scoring coefficients are then applied to the standardized variables to create weighted scores. The VAR statement specifies the variables to be included in the principal components analysis.

**INTERPRETING PROC FACTOR RESULTS**

<table>
<thead>
<tr>
<th>Eigenvalues of the Correlation Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Eigenvalue</strong></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

Figure 1.0
SAS® outputs principal components and corresponding eigenvalues as seen in table Figure 1.0. According to the MINEIGEN=1 parameter, only the first component will be retained for analysis. After this preliminary analysis, the NFACT=1 option is used to create a dataset that contains the weighted scores for the first factor. The weighted score, labeled Factor 1 in the dataset, is a linear combination of the standardized variables for each variable (question) multiplied by the corresponding factor coefficients that are output in Figure 1.1.

<table>
<thead>
<tr>
<th>Standardized Scoring Coefficients</th>
<th>Factor1</th>
</tr>
</thead>
<tbody>
<tr>
<td>quest_20a</td>
<td>0.30173</td>
</tr>
<tr>
<td>quest_20b</td>
<td>0.30379</td>
</tr>
<tr>
<td>quest_20c</td>
<td>0.24497</td>
</tr>
<tr>
<td>quest_20d</td>
<td>0.28422</td>
</tr>
</tbody>
</table>

These weighted scores represent the final weighted score for the specific principle that was included in the analysis. For the purpose of this study, each principle will have two weighted scores: one for the quantity scale and one for the quality scale.

RESULTS

INTERNAL CONSISTENCY
During the first stage of instrument evaluation lack of internal consistency was found for six of the EP question groups. One of these groups was the quality scale items for EP 6. The results of the Cronbach’s alpha calculation using SAS® are displayed in Figures 2.0 and 2.1.

EP 6: Foster co-learning, capacity building, and co-benefit for all partners.

<table>
<thead>
<tr>
<th>Cronbach Coefficient Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables</td>
</tr>
<tr>
<td>Raw</td>
</tr>
<tr>
<td>Standardized</td>
</tr>
</tbody>
</table>

Figure 2.0

<table>
<thead>
<tr>
<th>Cronbach Coefficient Alpha with Deleted Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deleted Variable</td>
</tr>
<tr>
<td>Correlation with Total</td>
</tr>
<tr>
<td>quest_47i</td>
</tr>
<tr>
<td>quest_48b</td>
</tr>
<tr>
<td>quest_47j</td>
</tr>
</tbody>
</table>

Figure 2.1

Between 2.0 shows that the standardized Cronbach’s Alpha statistic is drastically lower than the desired α>0.70 level which indicates a lack of consistency among the questions. However, removing question 48b increases α to .82, which would then show internal consistency among the two remaining items (Figure 2.1). The negative item-total correlation for question 48b further justifies its removal. With question 48b removed, the standardized alpha (Figure 3.0) and item-total correlations (Figure 3.1) for the remaining two items increases.

<table>
<thead>
<tr>
<th>Cronbach Coefficient Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables</td>
</tr>
<tr>
<td>Raw</td>
</tr>
<tr>
<td>Standardized</td>
</tr>
</tbody>
</table>

Figure 3.0
Considering these results, new questions pertaining to EP 6 were developed to increase the internal consistency of the question group. This re-development was done for all question groups that showed evidence of a lack of internal consistency. The newly designed questions along with those that showed strong internal consistency during this first stage of evaluation were used for the second round of pilot testing.

Table 1.0

<table>
<thead>
<tr>
<th>Cronbach’s Alpha</th>
<th>Quantity EPs</th>
<th>N</th>
<th>Percent</th>
<th>Quality EPs</th>
<th>N</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.85 &lt; \alpha \leq 0.90)</td>
<td>5</td>
<td>45%</td>
<td>0</td>
<td>0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.90 &lt; \alpha \leq 0.95)</td>
<td>6</td>
<td>55%</td>
<td>6</td>
<td>55%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\alpha &gt; 0.95)</td>
<td>0</td>
<td>0%</td>
<td>5</td>
<td>45%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Using standardized Cronbach’s alpha >0.70 as an acceptable level, all quantity and quality EP question groups were found to be internally consistent in the second round of pilot testing. This means that all questions designed to pertain to a specific EP were sufficiently correlated with the remaining questions pertaining to that EP. In fact, every EP question group had a Cronbach’s alpha > .85 (Table 1.0), which indicates strong internal consistency for all question groups in the survey. On the quantity scale, alpha coefficients ranged from 0.87 to 0.95 with a mean alpha of 0.91 and standard deviation of 0.03. On the quality scale, coefficients ranged from 0.91 to 0.97, with a mean alpha of 0.95 and a standard deviation of 0.02. The standardized Cronbach’s alpha for all EP-question groups and scales can be seen in Table 1.1.
COMMUNITY ENGAGEMENT SCORES

Simple Sum Scores

Simple sum scores were created using only those observations with complete survey responses. 32 observations were used to calculate the quantity community engagement scores and 25 observations were used to calculate the quality community engagement scores. The lowest possible simple sum score for each scale (quantity/quality) was 48, while the highest possible score was 240. Summary statistics of the quantity and quality simple sum scores are shown in Table 2.0.

<table>
<thead>
<tr>
<th>Simple Sum Score</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Min</th>
<th>Max</th>
<th>Median</th>
<th>5%</th>
<th>25%</th>
<th>75%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity Simple Sum Score</td>
<td>32</td>
<td>186.31</td>
<td>30.89</td>
<td>133.00</td>
<td>240.00</td>
<td>189.00</td>
<td>134.00</td>
<td>164.00</td>
<td>209.50</td>
<td>235.00</td>
</tr>
<tr>
<td>Quality Simple Sum Score</td>
<td>25</td>
<td>176.12</td>
<td>41.59</td>
<td>96.00</td>
<td>240.00</td>
<td>181.00</td>
<td>108.00</td>
<td>140.00</td>
<td>214.00</td>
<td>234.00</td>
</tr>
</tbody>
</table>

Table 2.0

The quantity and quality simple sum scores were highly positively correlated with each other (r=.92, P<.001). On average, quantity simple sum scores were significantly higher than quality simple sum scores (P=0.012).

Average Sum Scores

Scale-specific average sum scores were calculated using observations that met the missingness criteria; quantity average sum scores were calculated for 41 observations and quality average sum scores were calculated for 33 observations. Of these, 31 participants completed enough questions to have average sum scores for both scales. Average sum scores were created for each scale by calculating the average score of the items for each individual EP (EP-specific average scores) and then summing those scores across all 11 principles.

Overall EP-specific average scores were created for each of the 11 principals by adding the quantity EP-specific average scores and the quality EP-specific average scores. These scores measure how well the project adhered to each specific engagement principle. This is especially useful to evaluate the program and pinpoint specific areas of community engagement that need to be improved. Summary statistics for the EP-specific scores are shown in Table 2.1.

<table>
<thead>
<tr>
<th>Overall EP-Specific Average Scores</th>
<th>Engagement Principle N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Median</th>
<th>5%</th>
<th>25%</th>
<th>75%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>EP 1</td>
<td>51</td>
<td>7.93</td>
<td>1.33</td>
<td>5.00</td>
<td>10.00</td>
<td>8.00</td>
<td>5.75</td>
<td>7.00</td>
<td>9.00</td>
<td>10.00</td>
</tr>
<tr>
<td>EP 2</td>
<td>50</td>
<td>8.06</td>
<td>1.48</td>
<td>4.50</td>
<td>10.00</td>
<td>8.25</td>
<td>5.00</td>
<td>7.25</td>
<td>9.00</td>
<td>10.00</td>
</tr>
<tr>
<td>EP 3</td>
<td>47</td>
<td>7.30</td>
<td>1.54</td>
<td>3.80</td>
<td>10.00</td>
<td>7.60</td>
<td>5.00</td>
<td>6.00</td>
<td>8.40</td>
<td>10.00</td>
</tr>
<tr>
<td>EP 4</td>
<td>43</td>
<td>7.30</td>
<td>1.67</td>
<td>4.60</td>
<td>10.00</td>
<td>7.40</td>
<td>4.75</td>
<td>6.00</td>
<td>8.60</td>
<td>10.00</td>
</tr>
<tr>
<td>EP 5</td>
<td>43</td>
<td>7.08</td>
<td>1.86</td>
<td>2.80</td>
<td>10.00</td>
<td>7.20</td>
<td>4.40</td>
<td>5.50</td>
<td>8.60</td>
<td>9.80</td>
</tr>
<tr>
<td>EP 6</td>
<td>47</td>
<td>7.63</td>
<td>1.63</td>
<td>3.73</td>
<td>10.00</td>
<td>8.00</td>
<td>5.00</td>
<td>6.25</td>
<td>9.00</td>
<td>10.00</td>
</tr>
<tr>
<td>EP 7</td>
<td>44</td>
<td>7.17</td>
<td>1.84</td>
<td>3.00</td>
<td>10.00</td>
<td>7.00</td>
<td>5.00</td>
<td>5.75</td>
<td>8.75</td>
<td>10.00</td>
</tr>
<tr>
<td>EP 8</td>
<td>48</td>
<td>7.77</td>
<td>1.53</td>
<td>5.00</td>
<td>10.00</td>
<td>7.87</td>
<td>5.00</td>
<td>6.90</td>
<td>9.20</td>
<td>10.00</td>
</tr>
<tr>
<td>EP 9</td>
<td>48</td>
<td>7.88</td>
<td>1.48</td>
<td>4.75</td>
<td>10.00</td>
<td>8.00</td>
<td>5.00</td>
<td>6.88</td>
<td>9.00</td>
<td>10.00</td>
</tr>
<tr>
<td>EP 10</td>
<td>42</td>
<td>7.12</td>
<td>1.77</td>
<td>3.50</td>
<td>10.00</td>
<td>7.50</td>
<td>4.25</td>
<td>5.50</td>
<td>8.00</td>
<td>9.50</td>
</tr>
<tr>
<td>EP 11</td>
<td>40</td>
<td>6.87</td>
<td>1.99</td>
<td>3.00</td>
<td>10.00</td>
<td>7.17</td>
<td>3.17</td>
<td>5.50</td>
<td>8.17</td>
<td>10.00</td>
</tr>
</tbody>
</table>

Table 2.1
The lowest possible overall EP-specific score is 2, while the highest possible score is 10. Engagement principle 11, plan for a long-term process and commitments, had the lowest mean score (6.87). Conversely, engagement principle 2, acknowledge the community, had the highest mean score (8.06). This indicates that researchers succeeded in recognizing the community during the pilot project and highlights one of the desired outcomes of community-based participatory research.

Average sum scores ranged from 22 to 55 with 11 being the lowest possible score and 55 being the highest; summary statistics can be seen in Table 2.2. As with the simple sum scores, there was a significant positive correlation between quantity and quality average sum scores ($r=0.91$, $P<.001$). An overall average sum score was created by adding the quantity and quality average sum scores. This allows for a single measurement of community engagement that accounts for the measurement of both the frequency (quantity) and strength (quality) of community engagement.

<table>
<thead>
<tr>
<th>Average Sum Score</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Min</th>
<th>Max</th>
<th>Median</th>
<th>5%</th>
<th>25%</th>
<th>75%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity Average Sum Score</td>
<td>41</td>
<td>42.25</td>
<td>6.91</td>
<td>30.10</td>
<td>55.00</td>
<td>42.65</td>
<td>32.22</td>
<td>37.55</td>
<td>47.38</td>
<td>53.75</td>
</tr>
<tr>
<td>Quality Average Sum Score</td>
<td>33</td>
<td>40.07</td>
<td>8.90</td>
<td>22.00</td>
<td>55.00</td>
<td>41.08</td>
<td>24.60</td>
<td>35.17</td>
<td>46.45</td>
<td>53.75</td>
</tr>
<tr>
<td>Overall Average Sum Score</td>
<td>31</td>
<td>81.87</td>
<td>15.72</td>
<td>55.00</td>
<td>110.00</td>
<td>84.48</td>
<td>57.15</td>
<td>70.30</td>
<td>93.90</td>
<td>107.50</td>
</tr>
</tbody>
</table>

Table 2.2

Simple sum scores and average sum scores were all found to be approximately normally distributed and significantly correlated with each other. This suggests that both scores would be appropriate for measuring quality/quantity of community engagement. Average sum scores are better suited for the community engagement survey since it allows for missing data within observations.

Weighted Scores

Weighted community engagement scores were created using PROC FACTOR to perform principal components analysis and to calculate score coefficients. For each EP-specific question group, a weighted quantity score and a weighted quality score were calculated, similar to the EP-specific average sum scores. These weighted scores were then summed across all 11 principles to create scale-specific (quality/quantity) weighted scores.

Within each EP-specific item group, there was very little variation among the item-specific coefficients used to create the principal component. The weighted scores are essentially a standardized version of the average sum scores and are highly correlated with those scores (Quality: $r=.9998$, $P<.001$; Quantity: $r=.999$, $P<.001$). This is to be expected since the items have previously been shown to be internally consistent and therefore, the structure of the first principal component will be similar to an average of the items within the group.

SENSITIVITY ANALYSIS

The small sample size of the pilot project led to the inclusion of observations with missingness in score calculations. However, it is necessary to perform a sensitivity analysis to explore the effect of including data with missingness. For this analysis, only observations with complete-case data were included in score calculations. The resulting quantity average sum scores ($n=32$) and quality average sum scores ($n=25$) do not differ greatly from the previous average sum scores. Simple statistics for the complete-case scores are shown below in Table 3.0.
Table 3.0

The distribution of quantity average sum scores does not differ much between those calculated with observations with missing values (Figure 4.1) and those calculated with only complete-case data (Figure 4.2). This also holds true for the quality-scale; scores calculated with observations with missing values are shown in Figure 5.1 and with only complete-case data in Figure 5.2.
DISCUSSION

The strong internal consistency for all EP question groups indicates that the community engagement questions in the survey were well designed to measure the principles they pertain to. Given these positive results, the survey will now be disseminated to other PECaD projects to further validate the instrument in a larger sample size. Once the instrument is fully validated, this survey will be disseminated as a standardized tool to all PECaD projects and other CBPR-based projects as a way of quantitatively measuring community engagement. The next step in this instrument development is to perform construct validity on these measures.

The calculation and analysis of the community engagement scores through different scoring mechanisms allowed for the development of informative measurement of community engagement across 11 principles. The simple sum scores and average sum scores were both normally distributed and easily calculated. However, the average sum scores proved to be more informative due to their ability to handle missing data. The development of EP-specific scores was necessary to fully understand the gaps in community engagement and to identify which principles need improved adherence. Since evaluating community engagement in research has been primarily done using qualitative approaches, it is important to develop evidence-based quantitative measures to understand and evaluate the level of community engagement in research quickly and easily, particularly in large scale science where qualitative approaches can be cumbersome. As an increasing number of researchers begin to engage communities in their work, these measurements will go a long way in helping to evaluate and improve the quality of CBPR and understand its impact on the science developed and disseminated.

CONCLUSIONS

Using SAS® to determine internal consistency is an integral step in the development of community engagement scores. The PROC CORR procedure allows for easy and quick analysis of the internal consistency of survey questions while PROC FACTOR can be used to develop weighted scores through principal components analysis. The methods described in this paper can also be applied to many different fields beyond the scope of community engagement to develop validated internally consistent survey items and scoring systems to summarize these measures.

REFERENCES

ACKNOWLEDGMENTS

The CRFT project is funded by National Institutes of Health, National Cancer Institute grant U54CA153460. The work of Dr. Goodman is supported by National Institutes of Health, National Cancer Institute grant U54CA153460 and Washington University School of Medicine Faculty Diversity Scholars Program. The work of Renee Gennarelli is supported by the Siteman Cancer Center and Barnes Jewish Hospital Foundation.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

**Renee Gennarelli, MS**
Division of Public Health Sciences
Department of Surgery
Washington University in St. Louis
School of Medicine
660 S. Euclid Avenue, Campus Box 8100
St. Louis, MO 63110
Phone: 314-747-2385
E-mail: gennarellir@wudosis.wustl.edu

**Melody S. Goodman, MS, PhD**
Division of Public Health Sciences
Department of Surgery
Washington University in St. Louis
School of Medicine
660 S. Euclid Avenue, Campus Box 8100
St. Louis, MO 63110
Phone: 314-362-1183
E-mail: goodmanm@wudosis.wustl.edu

* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.