ABSTRACT
The Health-Related Quality of Life (HRQOL) measurements have been frequently applied in clinical trials of chronic diseases (such as cancers). There are many measurement tools of HRQOL questionnaires and diverse statistical analysis methods accordingly, which make the analysis of HRQOL data rather complicated and hard to be standardized. This article provides a SAS® macro application based on the standard method of HRQOL data analysis of the National Cancer Institute – Clinical Trials Group (NCIN-CTG) [1]. The program integrates data manipulation, statistical analyses and reports generation processes. It can calculate HRQOL domains from different measurement tools, determine visit/cycle through predefined time windows and summarize the HRQOL domains' response rates. The analysis results are organized into standard statistical reports, which includes the compliance rates, baseline scores comparison, the analysis of the changes of the domain scores from baseline and the summary responses analysis results. The strategies and the kernel program codes are discussed.

KEY WORDS
Health-related Quality of Life (HRQOL), Statistical analysis, SAS macros

INTRODUCTION
As the usage of HRQOL data analysis in clinical trials becomes more and more frequent, the statistical methodologies and questionnaire tools have been developed rapidly. Some complex analysis methods were criticized for the difficulties in modeling and in results interpreting [2]. A method emphasizing the clinical meaning while avoiding complex statistical modeling is desired by clinicians and other non-statistics researchers.

The National Cancer Institute of Canada – Clinical Trials Group (NCIC-CTG) has been carrying out HRQOL assessments since 1987. From the plentiful experiences, NCIC-CTG has developed an approach that focuses on the clinical interpretation with a summary response (improved, stable and worsened). This approach consists of four standard steps: Calculating the compliance rates, comparing baseline scores between groups, comparing the changes of the scores from baseline, and compare the proportion of responses between the treatment groups [1]. Following this standard procedure, a SAS macro application is developed to perform the analyses.

STATISTICAL ANALYSIS METHODS
The baseline scores comparison uses non-parametric (the Wilcoxon Rank-sum) test. It supposes no statistically significant difference between the treatment groups; otherwise the validity of randomization will be challenged.

The parametric test (comparing mean score-changes) is suggested in comparing score-changes, because it has been found that there is little practical difference between using non-parametric and parametric approaches (unpublished data).

There are many approaches determining the summary response [3]. A 10% cut-off scale is recommended by NCIC-CTG standard approach, but the scale can vary according to the specific trials. If the scores exceed the cut-off scale in the defined study period, the response is classified as ‘Improved’; while scores decline below the scale in the study period, the response is classified as ‘Worsened’. The situation that is neither ‘improved’ nor ‘worsened’ is classified as ‘Stable’. The Chi-square test is used to analyze the difference of the response rates between the treatment groups.

PROGRAMMING STRATEGIES
Although the statistical analyses methods are straightforward, there are several crucial points in the programming. First, there is usually no cycle/visit information in the HRQOL questionnaire and the cycle needs to be determined with the predefined time window. Second, the designed analyses will be conducted for all the domains. This needs looping the analyses to all domains. Third, the logic of classifying the response categories is the key part of the program. The program will detect the QOL score changes and define the response categories. Fourth, if trial specific questions/domains are used, it is hard to automatically include them into the macro application. Moreover, the domains calculation from different questionnaire tools is hard to be automated because so many forms are available and may be used.
To accommodate these crucial points, the following strategies are considered in the SAS programming:

1. Divide the analyses into functional modules. The modules include domain calculation, data preparation, compliance rates calculation and reporting, baseline score comparison, score-changes comparison, and response rate comparison. The cycle determination and response classification are included in the data preparation step.
2. Use a macro parameter entry to control the time window, as it might be study dependent. Specific coding may be necessary if the follow-up time is not designed in an equal period.
3. Keep the names of all the domains with the contents procedure and convert them to a macro variable array. Use do-loop to repeat the analyses on each of the domains.
4. Create the domain calculation program based on the specific questionnaire tools. So far we have programs for several most commonly used QOL tools in cancer clinical trials (EORTC30, EORTC13, SF36 and combined EORTC). More programs can be added if other alternative questionnaire tools are used and the assessment methods are provided.
5. For the extra trial specific domains, a macro string is defined in the macro parameter entries. The string lists extra domain names and will be used as an array in the data preparing step.
6. Use the logic conditions to classify the response. If the score changes reach the upper boundary it is defined as 'Improved'; otherwise, if the score changes reach the lower boundary it is defined as 'Worsened'; for the rest situations the response is defined as 'Stable'.
7. Use standard templates to generate all the reports and save them to output files.

EXAMPLES OF THE KERNEL SAS CODES
The example codes for some of the key issues are shown and discussed as below.

Domains calculation:
QOL domains calculation is measurement tools specific. Most of the QOL questions use four or seven categories recording the patients’ answer on the QOL questions. Generally, the categorical scores are converted into a 100-based scale number. The calculations for functional domains and symptom domains are different. For example, in EORTC30:

```sas
*** domain parameter calculation***;
*** physical (functional domain) ***;
   nm=nmiss(of q_1-q_5);
   if nm>2 then phys=.;
   else do;
       sum=sum(q_1,q_2,q_3,q_4,q_5);
       phys=100-((sum/(5-nm)-1)*100/3);
   end;
*** Fatigue (symptom domain) ***;
   nm=nmiss(q_10,q_12,q_18);
   if nm>1 then fati=.;
   else do
       sum=sum(q_10,q_12,q_18);
       fati=((sum/(3-nm)-1)*100/3);
   end;
```

For example if the answers for question 10, 12 and 18 are 2, 3, and 3, the converted domain score for fatigue will be 55.56.

Determining cycles with time window
The time window is defined as a macro variable (for example, 7 days for treatment cycles). There may be different time windows for follow-up (for example, 30 days). The difference in days between the date the questionnaire is filled-out and the treatment date for the cycle is calculated. If it falls within the time window, the HRQOL data belongs to the same treatment cycle/follow-up. The total number of cycles / follow-ups can be determined in the data step.

The code below is the example of combining QOL and treatment data:
Where macro variable &cno is the total number of cycles and &cwd is the value of time window. The new variable ‘time’ is the cycle variable assigned to HRQOL data. Now all QOL records that fall into the time window find their cycles. Follow-up QOL data can be handled in the same way.

Classifying the QOL response

As described above, the response is classified as ‘Improved’, ‘Worsened’ and ‘Stable’. The program code firstly checks if the scores can be defined as ‘Improved’ based on the score changing criteria. If so, the response is ‘Improved’ even if the scores are somewhere below the lower-boundary. If it is not ‘improved’, the program checks if the response is ‘Worsened’. Finally, if neither ‘Improved’ nor ‘Worsened’ the response is classified as ‘Stable’.

In the program code (see textbox at the right side), &varno is the number of domain variables. The macro variable &fdm is a predetermined string of all functional domain names. The program merges the cycle / follow-up QOL data and the baseline QOL data, and calculates the score changes. The determination of responses is different for functional and symptom domains.

The response classification is performed to each domain by looping. The program gives the ‘improved’ cases the value of 3 and the ‘worsened’ and ‘Stable’ cases 2 and 1, respectively. The largest value is kept to ensure the priority of improved cases. After this step, the responses are changed back to 1 (improved), 0 (stable) and –1 (worsened) for standard analysis.
SELECTED REPORT EXAMPLES

Compliance table:

<table>
<thead>
<tr>
<th></th>
<th>TREATMENT A</th>
<th></th>
<th>TREATMENT B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eligible</td>
<td>Received (%)</td>
<td>Eligible</td>
<td>Received (%)</td>
</tr>
<tr>
<td>Baseline</td>
<td>153</td>
<td>139(90.8)</td>
<td>152</td>
<td>125(82.2)</td>
</tr>
<tr>
<td>Cycle 2</td>
<td>153</td>
<td>130(85.0)</td>
<td>152</td>
<td>123(80.9)</td>
</tr>
<tr>
<td>Cycle 3</td>
<td>153</td>
<td>92(60.1)</td>
<td>152</td>
<td>103(67.8)</td>
</tr>
<tr>
<td>Cycle 4</td>
<td>153</td>
<td>81(52.9)</td>
<td>152</td>
<td>84(55.3)</td>
</tr>
<tr>
<td>Cycle 5</td>
<td>153</td>
<td>67(43.8)</td>
<td>152</td>
<td>69(45.4)</td>
</tr>
<tr>
<td>Cycle 6</td>
<td>153</td>
<td>55(35.9)</td>
<td>152</td>
<td>59(38.8)</td>
</tr>
<tr>
<td>Cycle 7</td>
<td>153</td>
<td>47(30.7)</td>
<td>152</td>
<td>48(31.6)</td>
</tr>
<tr>
<td>Cycle 8</td>
<td>153</td>
<td>1(0.7)</td>
<td>152</td>
<td>2(1.3)</td>
</tr>
<tr>
<td>Month 1</td>
<td>153</td>
<td>60(39.2)</td>
<td>152</td>
<td>62(40.8)</td>
</tr>
<tr>
<td>Month 3</td>
<td>153</td>
<td>31(20.3)</td>
<td>152</td>
<td>24(15.8)</td>
</tr>
<tr>
<td>Month 6</td>
<td>153</td>
<td>14(9.2)</td>
<td>152</td>
<td>9(5.9)</td>
</tr>
<tr>
<td>Month 9</td>
<td>153</td>
<td>9(5.9)</td>
<td>152</td>
<td>5(3.3)</td>
</tr>
<tr>
<td>Month 12</td>
<td>153</td>
<td>10(6.5)</td>
<td>152</td>
<td>4(2.6)</td>
</tr>
<tr>
<td>Month 15</td>
<td>153</td>
<td>4(2.6)</td>
<td>152</td>
<td>3(2.0)</td>
</tr>
<tr>
<td>Month 18</td>
<td>153</td>
<td>2(1.3)</td>
<td>152</td>
<td>3(2.0)</td>
</tr>
<tr>
<td>Month 21</td>
<td>153</td>
<td>1(0.7)</td>
<td>152</td>
<td>2(1.3)</td>
</tr>
</tbody>
</table>

Only part of the compliance and response analyses tables are shown as examples. The reports of cross-sectional analysis by cycle/follow-up and of baseline score test are not shown here. The baseline score test displays the mean values of all QOL domains by treatment group with the test p-values. The cross-sectional comparison shows the similar results (mean domain values by treatment with test p-values) for each of the treatment cycle and follow-up.

DISCUSSION

HRQOL data analysis is a common practice in cancer clinical trials. Standard analysis approaches and corresponding analysis applications are required. The NCIC-CTG’s standard analysis approach provides the basis for building up a robust analysis system of analyzing the HRQOL data.

The NCIC-CTG’s approach is simple and straightforward in statistical methodology. The data management procedures are standard. These facts make it possible to used SAS macro based programs to automate the analysis steps. The difficulties in defining cycles, determining response and repeating the analyses for all the domain variables can be concurred with the SAS programming techniques.

The SAS application introduced in this article is still in its primary stage. The most important pre-requisite is that the structure and contents of the clinical datasets should be relatively standard. For example, the QOL variables should be named in the same manner (Q_1-Q_## for question number, qol_dt for the date when the HRQOL questionnaire is filled out, and so on); the dataset name and contents should also be standard (F1data for baseline, F3data for treatment, F5data for follow-up etc.). If the data structure and contents are not in the standard format, an extra step is needed to modify the datasets into the pattern that meets the requirements of the SAS application. A specific format section may be needed too, to apply the format for domain names, response levels and cycle / follow-up descriptions.

There might be concerns that the NCIC-CTG standard QOL data analysis approach summarizes the response only based on one point and it may result in losing information of all the time period. Our ongoing study revealed that there is well consistency of the result of this method with the results of more complicated statistical modeling. This article is not purposed to compare statistical methodologies.
CONCLUSION
This article introduces the NICI-CTG standard HRQOL data analysis approach and a SAS macro oriented analysis application based on the approach. The programming strategies are discussed. For the key technical issues in data managing and analyzing, solutions with SAS programming techniques are provided. The example codes of some substantial steps are included, too. This application can be a model of automating the data analysis processes with a standard data analysis approach.

REFERENCES

CONTACT INFORMATION
Don (Dongguang) Li, PhD, MPH, Senior Biostatistician
National Cancer Institute of Canada - Clinical Trials Group
10 Stuart Street, Queen's University, Kingston, Canada, K7L 3N6
613 533 6000 Ext. 78337 dli@ctg.queensu.ca