PharmaSUG 2017 - PO19

Analyzing Singly Imputed Utility Based Weighted Scores Using the EQ-5D for Determining Patients' Quality of Life in Oncology Studies

Vamsi Krishna Medarametla, Liz Thomas, Gokul Vasist

Seattle Genetics, Inc., Bothell, WA

ABSTRACT:

Quality of Life (QoL) analysis provides complementary information to standard safety and efficacy analysis in clinical trials. This analysis allows regulators to take patient-reported outcomes (PROs) on quality of life into account when making decisions about drug approvals and labeling. A standard technique in QoL analysis is to combine patient-reported outcome data collected on a questionnaire into a single utility score. Missingness in PRO data is common, and must be addressed and accounted for in analysis using an imputation method.

The EQ-5D is a standardized instrument for use as a measure of health outcome. Analysis of the EQ-5D requires computation of a weighted utility score. Country-specific time trade-off (TTO) weights are commonly used. Calculating the index weighted scores of the EQ-5D profiles under several different single imputation schemes will help in understanding the health state of the patient.

This paper demonstrates computation of the US and UK weighted TTO EQ-5D scores using LOCF, WOCF, Time point average, and death only single imputation methods for a variety of time point indexed analyses.

INTRODUCTION:

Quality of life of a patient is an important endpoint in clinical trials along with other traditional efficacy endpoints like survival data, as it takes into consideration the physical, mental and social well-being of a patient. Patient reported outcomes are becoming increasingly important in supporting the claims in approved product labeling. A PRO instrument consists of a questionnaire along with the information and documentation to assist the use of the data. The questionnaire information obtained from the PRO data is combined into a single utility score to carry out the analysis.

WHAT IS EQ-5D:

EQ-5D is a five dimensional questionnaire (mobility, self-care, usual activities, pain/discomfort, anxiety/depression) consisting of EQ-5D-3L descriptive system and EQ Visual Analog Scale (Overall Health Assessment). This questionnaire is filled out by the patient and is used to assess the health state. Each questionnaire has three levels of severity (no problems, moderate problems and extreme problems) and are given the scores 1, 2 or 3. This is illustrated in the sample questionnaire CRF page below.

Mobility:	I have no problems in walking about ○
	I have some problems in walking about ○
	I am confined to bed ○
	Not answered ○

Self-Care	I have no problem with self-care ○
	I have some problems washing or dressing myself ○
	I am unable to wash or dress myself ○
	Not answered ○

Usual Activities	I have no problems with performing my usual activities ○
	I have some problems with performing my usual activities ○
	I am unable to perform my usual activities ○
	Not answered ○

Pain/Discomfort	I have no pain or discomfort ○
	I have moderate pain or discomfort ○
	I have extreme pain or discomfort ○
	Not answered ○

Anxiety/Depression	I am not anxious or depressed ○
	I am moderately anxious or depressed ○
	I am extremely anxious or depressed ○
	Not answered ○

Overall Health Assessment (0-100): _____

Once the data is available, Time Trade-off (TTO) scores can be calculated from the population-based preference weights. TTO scores rank health states according to population-based preferences, with 10 years in a given health state equivalent to TTO score * 10 years in a perfect health state. Death has a natural interpretation of zero, and some states considered worse than death can have negative TTO scores (e.g. extreme problems in every dimension results in US-weighted TTO score of -0.109);

EQ-5D WIGHTED SCORES IMPUTATION:

Missing values in PRO data are common and must be addressed and accounted for in analysis. This can be done by imputing the weighted TTO scores. The data used for the analysis consist of questionnaire information collected during the scheduled treatment visits and also during the follow up visits after the end of treatment. These data need to be converted to visits with an interval of 3-months from the time points of interest for the purpose of summarizing the results. For imputation to be carried out, the study specific dates when visits are expected to occur and windows for each visit must be defined. The expected visits and windows from our example are indexed from three time points, in three-month intervals since the sparsest visit schedule required a visit every three months for all patients.

- Indexed from study day 1: The visits are indexed in 3-month intervals (0 to 24 months) starting from the treatment start date. If treatment start date was missing date of randomization was used.
- Indexed from EOT: The visits are indexed in 3-month intervals (0 to 24 months) starting from the end of treatment.
- Indexed from Investigator progression date: The visits are indexed in 3 month intervals (0 to 24 months) from the date of progression.

The expected visits data is then merged with the non-missing nominal time points (available for records indexed from study day 1), and these are identified as non-imputed records. Among the remaining records the closest record within the window (+/- 7 days for dosing visits, +/- 30 days for remaining visits) are identified. After identifying the nominal visits and the window matches, the remaining visits are flagged as records requiring imputations. The imputed visits which occur 24 months after the treatment start date as well as the visits occurring after the end of study for reasons other than death are deleted, since those would not be collected according to study schedule and do not have a clear imputation. After categorizing the visits as nominal, window match or imputed visits the imputations for the missing data are performed. For all the visits occurring after the date of death the scores are imputed to '0'. The different types of imputations used are:

- LOCF – Last Observation Carried Forward. The missing TTO score of a visit is imputed with the score from the previous visit.
- WOCF – Worst Observation Carried Forward. The missing TTO score of a visit is imputed with the worst (lowest TTO) score of all the visits for the patient.
- Time point average – The averages for each time point are calculated and the missing scores are imputed with the averages.
- Only death imputed – The scores for visits expected after the date of death are imputed to '0'.

ALGORITHM:

Step 1: Identifying the records when the visits were expected to occur. Visits that matched were identified as nominal while the non-matching ones were categorized into a separate dataset named here as nomatch.

```
data nominal nomatch;
merge uktto(in=a) alldt(in=b);
       by usubjid visitnum;
       if a and b then output nominal;
       if b>a then output nomatch;
run;
```

Step 2: Determining the nominal and imputed time points as per the visit schedule. Certain visit windows had a +/- 7 day range while the other had a +/- 30 day range in it. We used the expected dates of a visit (variable newdt) and created variables newdt1 indicating the upper and lower limit of the windowed range. We also created a visit description (visdes) variable that indicates the start and the end of the windows. The start of the visit window would have VISDES=-1, the end of the window would have VISDES=1, and the original records before imputation would get a VISDES=0.

```
data windows;
set nomatch;
       if month ne . or visitnum le 160001 then do;

       if newdt ne . then newdt1=newdt-7;
       visdes=-1;
       output;
       if newdt ne . then newdt1=newdt+7;
       visdes=1;
       output;
end;
else do;

       if newdt ne . then newdt1=newdt-30;
       visdes=-1;
output;
       if newdt ne . then newdt1=newdt+30;
       visdes=1;
output;
end;
       format newdt1 date9.;
run;

/* Actual records get visdes=0 */
data trans2;
set uktto;
       visdes=0;
       newdt1=input(qsdtc,is8601da.);
       format newdt1 date9.;
run;

data trans3;
set trans2 windows;
run;

proc sort data=trans3;by usubjid newdt1 visdes;run;
```

Step 3: Once the basic dataset was created and sorted, we retained variables over time and over windows to impute both LOCF and WOCF and assign visits in one pass through the data. Assignment of visit variables over the window (mnth, mnth_eot, mnth_pd) and the expected date of the visit (expdt, eot_exp, pd_ex) allow us to create distance variables (dist, dist_eot, dist_pd) to find the closest windowed match in the next step.

```sas
data tto2;
set trans3;
    by usubjid newdt1;
    retain locfus  locfuk  nadir  wocf  mnth  mnth_eot  mnth_pd  expdt  eot_exp
pd_exp;
if first.usubjid then do;
   locfus=.;
   nadir=.;
   locfuk=.;
   mnth=.;
   mnth_eot=.;
   mnth_pd=.;
   expdt=.;
   eot_exp=.;
   pd_exp=.;
end;
if eq5dtto1 ne . then do;
   locfus=eq5dtto1;
   wocf=eq5dtto1;
   locfuk=eq5dtto;
   if nadir<-999 or eq5dtto1<nadir then nadir=eq5dtto1;
end;

if newdt ne . then do;

  if visdes = -1 then do;
    if month ne . then do;
       mnth=month;
         expdt=newdt;
    end;
      if month_eot ne . then do;
       mnth_eot=month_eot;
         eot_exp=newdt;
    end;
      if month_pd ne . then do;
       mnth_pd=month_pd;
         pd_exp=newdt;
    end;
  end;

      if visdes = 1 then do;
    if month ne . then do;
       mnth=.;
         expdt=.;
    end;
      if month_eot ne . then do;
       mnth_eot=.;
         eot_exp=.;
    end;
      if month_pd ne . then do;
       mnth_pd=.;
         pd_expdt=.;
    end;
   end;
  end;
else do;
  if expdt ne . then dist=abs(newdt1-expdt);
```

```sas
    if eot_exp ne . then dist_eot=abs(newdt1-eot_exp);
    if pd_exp ne . then dist_pd=abs(newdt1-pd_exp);
end;
format expdt eot_exp pd_exp date9.;
run;
```

If at any point there are duplicate visits we had to select the record that was closest to the windowed match. This is guaranteed by sorting on the distance variables described above.

```sas
/* these two sort steps keep only the closest windowed match, if more than
one exists */
proc sort data=tto2 out=tto3;
by usubjid mnth dist;
      where mnth ne . and visdes=0;
run;

proc sort data=tto3 out=mnthwin nodupkey;by usubjid mnth;run;

data mnthwin;
set mnthwin;
      month=mnth;
      newdt=expdt;
      visitnum=1000000+mnth;
      drop mnth: dist_pd;
run;


proc sort data=nominal(keep=usubjid visitnum);by usubjid visitnum;run;

proc sort data=tto2;by usubjid visitnum;run;

data all;
merge nominal(in=a) tto2(where=(visdes=0));
      by usubjid visitnum;
if a;
run;

/*winall has all the windows data*/

proc sort data=winall ;by usubjid visitnum;run;

/*Classifying Nominal match and window match into a descriptor vistyp for the
timepoints*/

data all1;
merge all(drop= visdes in=a) winall(drop=visdes in=b);
      by usubjid visitnum;
      length vistyp $20;
      if a then vistyp='Nominal match';
      else if b then vistyp='Window match';
run;

data imp;
merge all1(in=a keep=usubjid visitnum) alldt(in=b);
      by usubjid visitnum;
      if b>a;
run;
```

```sas
data imp1;
merge imp(in=a) sl;
      by usubjid;
      if a;
run;
```

Step 4: When setting up the expected visits, we projected out from study start date. In the case when a subject discontinues for a reason other than death or our expected visit is beyond the planned 24-month schedule for collecting EQ-5D, the records should be deleted. These records are not expected and are not intended to be imputed.

```sas
data imp2;
set imp1;
      if newdt>=eosdt>. and dthdt<0 then delete;
      if newdt>intnx('month',max(trtstdt,randdt),24,'same') then delete;
run;

/*Identifying the imputed visits*/;

data allimp;
merge tto2 imp2(in=a);
      by usubjid visitnum;
      length vistyp $20;
      vistyp='Imputed';
      if visdes=1 and a=1;
run;

data final;
set all1 allimp;
      if newdt>dthdt>. then do;
      locfus=0;
      wocf=0;
      locfuk=0;
      vistyp='Dead';
end;
else do;
      if eq5dtto1=. then wocf=nadir;
end;
run;
```

Step 5: In addition to individual subject-level imputation methods (LOCF, WOCF) that can be computed using a retain step, we did a timepoint-average imputation, requiring computation of timepoint averages and merging these averages back into the dataset. At each visit, the mean of exact or windowed matches at that timepoint was used, since this imputation is intended to only represent the mean of patients alive at that timepoint.

```sas
data tptavg;
set final;
      if vistyp in ('Nominal match' 'Window match') and (eq5dtto1 ne . or
locfus ne .);
      if eq5dtto1 ne . then tptscr=eq5dtto1;
run;

data trt;
set a.adsl;
```

```
       if ittfl='Y';
       keep usubjid trt1pn;
run;

proc sort data=tptavg;by usubjid;run;

data tptavg;
merge tptavg(in=a) trt;
       by usubjid;
       if a;
run;

proc sort data=tptavg;by visitnum trt1pn;run;

proc univariate data=tptavg noprint;
       by visitnum trt1pn;
       var tptscr;
       output out=avg
       n=n
       mean=mean
;
run;

data avg;
set avg;
       if visitnum in (666666 1000099) then delete;
       avg=mean;
       keep visitnum avg trt1pn;
run;

proc sort data=final;
       by usubjid;
run;

data final;
merge final(in=a) trt;
       by usubjid;
       if a;
run;

proc sort data=final;by visitnum trt1pn;run;
```

Step 6: Finally, records past the time of death are imputed as zero, since death is a meaningful state for EQ5D-TTO with an assigned value of zero. This applies to all forms of imputation.

```
data eq5d1;
merge final avg;
       by visitnum trt1pn;
       if vistyp='Dead' then tptavg=0;
       else if eq5dtto1>. then tptavg=eq5dtto1;
       else tptavg=avg;
       if vistyp in ('Nominal match' 'Window match') then raw_dth=eq5dtto1;
       if vistyp='Dead' then raw_dth=0;
       drop avg eosdt month: pd: visdes dsdt mnth: dist: expdt eot_exp;
run;
```

CONCLUSIONS:

Since the PRO data is collected separately for each dimension, collapsing the multi-dimensional data into a single index score will help in carrying out the analysis. Careful consideration of the patterns of data collection and also missing data values is crucial for analyzing the data. The imputation model described in the paper is based on single imputation method. The extension of this method to accommodate multiple imputations is beyond the scope of this paper, and is a future direction of research.

REFERENCES:

Shaw, Johnson, & Coons. (2005). *US valuation of the EQ-5D health states. Development and testing of the D1 valuation model.* Medical Care, 203-220.

ACKNOWLEDGEMENTS:

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

Vamsi Krishna Medarametla   Liz Thomas
Seattle Genetics, Inc.   Seattle Genetics, Inc.
21823 - 30th Drive S.E.   21823 - 30th Drive S.E.
Bothell, WA 98021   Bothell, WA 98021
kmedarametla@seagen.com   lthomas@seagen.com

Gokul Vasist
Seattle Genetics, Inc.
21823 - 30th Drive S.E.
Bothell, WA 98021
gvasist@seagen.com