

It is Time to Get Answers: Handling Questionnaire Data in Respiratory Studies



Still waiting for the perfect diary data...

Valeriia Oreshko, Intego Group LLC, Kyiv, Ukraine

Ihor Savostianov, Intego Group LLC, Kharkiv, Ukraine

AGENDA

- Introduction
- Structure of TTE dataset
- Common pitfalls
- Dealing with missing data
- ADaMs arrangement for TTE Calculation
- Conclusion

INTRODUCTION

Questionnaire data in respiratory studies

Patients are asked to estimate their symptoms using the scale:

0 – None

1 – Mild

2 – Moderate

3 – Severe

№	Symptom	Day 1 Date: _____		Day 2 Date: _____	
		Morning Time: _____	Evening Time: _____	Morning Time: _____	Evening Time: _____
1	Headache				
2	Fainting/dizziness				
3	Fatigue				
4	Sore throat				
5	Shortness of breath				
6	Muscle Aches				
7	Vomiting				
8	Runny nose				
9	Chills/night sweats				
10	Cough				
11	Unable to care for self				
12	Little or no urination				
13	Loss of smell				
14	Loss of taste				

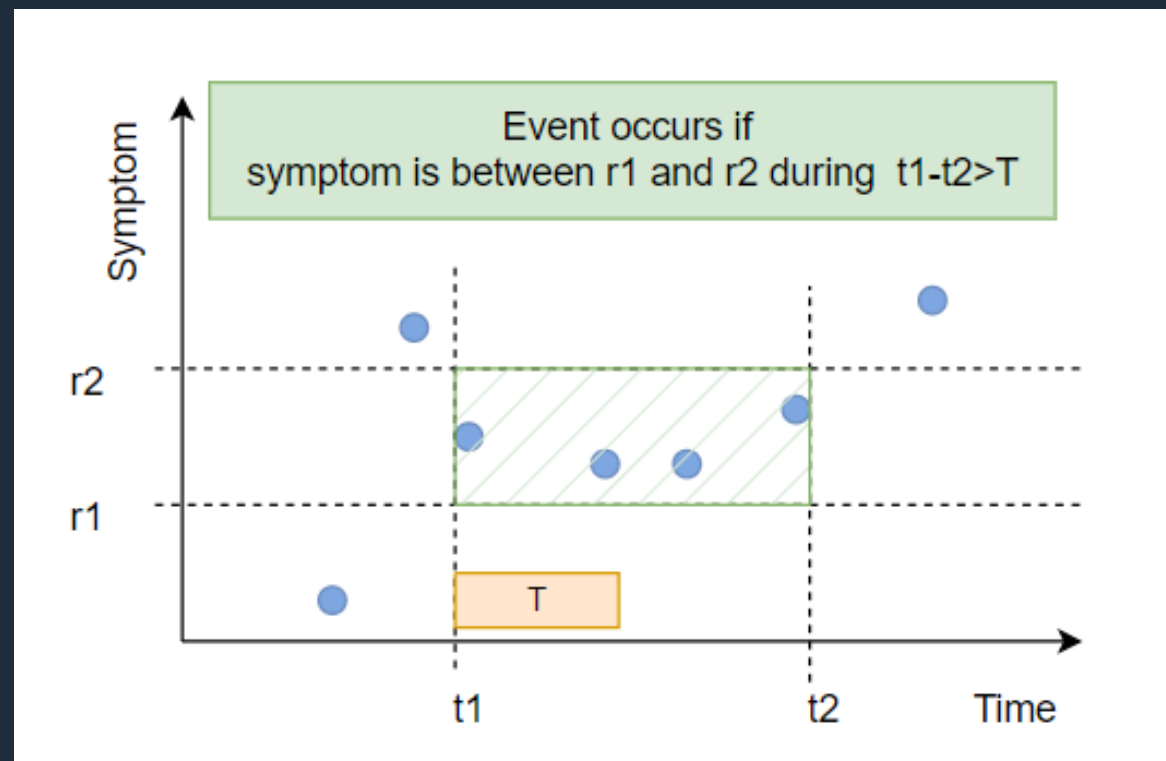
TTE endpoints based on patient's diaries - 1

For TTE endpoints an appropriate criterion should be maintained for a certain period

In the further examples:

Event: $AVAL \leq 1$ for all questions

T = 43 hours



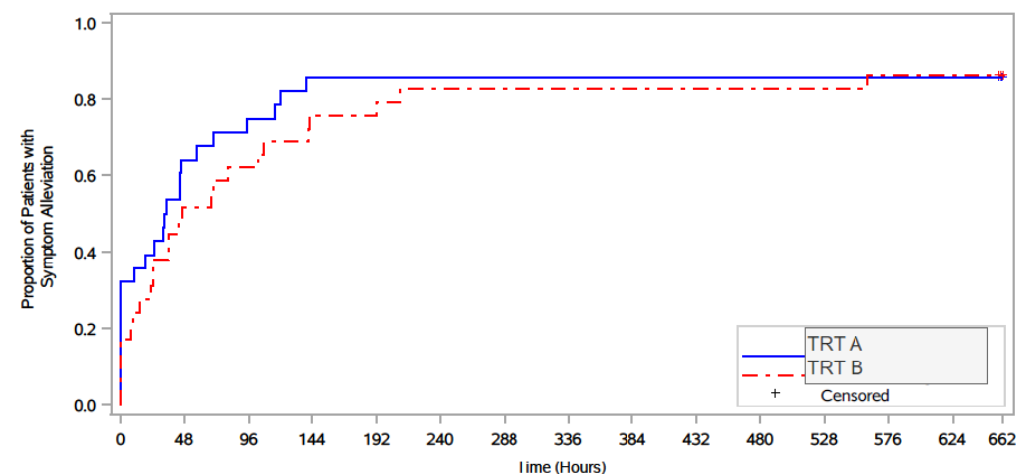
TTE endpoints based on patient's diaries - 2

Time to Event Summary for Time to Alleviation of COVID-19 Signs and Symptoms for 43 hours

	TRT A	TRT B
Patients included in analysis	28	29
Patients with event (%)	24 (85.7%)	25 (86.2%)
Patients without event (%)	4 (14.3%)	4 (13.8%)
Time to event (hours)		
Median	33.6	46.3
80% CI	(18.7, 45.9)	(35.3, 80.4)
25% and 75%-ile	0.0 - 105.2	14.1 - 142.1

Time to alleviation of COVID-19 symptoms is defined as the length of time taken from start of treatment to the point at which the following criterion is met and maintained for at least 43 hours: Score of 0 or 1 on Items 1-12 of the COVID-19 Symptom Diary, regardless of if the symptom is pre-existing or new.

Kaplan-Meier Plot of Time to Alleviation of COVID-19 Signs and Symptoms for 43 hours.



Time to alleviation of COVID-19 symptoms is defined as the length of time taken from start of treatment to the point at which the following criterion is met and maintained for at least 43 hours: Score of 0 or 1 on Items 1-12 of the COVID-19 Symptom Diary, regardless of if the symptom is pre-existing or new.

STRUCTURE OF TTE DATASET

TTE specific variables

Variable Name	Variable Label	Core	Description
ADTM	Analysis Date/Time	Perm	Analysis of date/time of event or censoring associated with AVAL in numeric format.
STARTDTM	Time to Event Origin Date/Time	Perm	This is generally the time at which a subject is first at risk for the event of interest evaluation (as defined in the Protocol or Statistical Analysis Plan). For example, this may be the randomization date or the date of the first study therapy exposure.
AVAL	Analysis Value	Req	AVAL is the elapsed time to the event of interest from the origin based on ADTM and STARTDTM.
CNSR	Censor	Cond	CNSR displays whether the event has occurred or not. Usually, CNSR = 0 for the event and CNSR = 1 or > 0 otherwise.
EVNTDESC	Event or Censoring Description	Perm	Describes the event of interest or an event that warrants censoring.
CNSDTDSC	Censor Date Description	Perm	Describes the <u>date</u> that was used for censoring

COMMON PITFALLS

Origin of the event and baseline value

PARAMCD	AVISIT	ATPT	AVAL
QS00101	Baseline	Morning	1
QS00101	Day 1	Evening	1
QS00101	Day 2	Morning	1
QS00101	Day 2	Evening	1
QS00101	Day 3	Morning	1

Event starts at baseline before the treatment

PARAMCD	AVISIT	ATPT	AVAL
QS00101	Baseline	Morning	2
QS00101	Day 1	Evening	1
QS00101	Day 2	Morning	1
QS00101	Day 2	Evening	1
QS00101	Day 3	Morning	1

Event starts after the treatment start

Event: AVAL $\leq$ 1

Censoring catches

PARAMCD	AVISIT	ATPT	AVAL
QS00101	Day 28	Morning	2
QS00101	Day 28	Evening	2
QS00101	UV 1	Morning	2
QS00101	UV 1	Evening	2
QS00101	UV 2	Morning	2

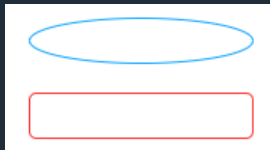
Case1 - Event doesn't occur

PARAMCD	AVISIT	ATPT	AVAL
QS00101	Day 28	Morning	2
QS00101	Day 28	Evening	2
QS00101	UV 1	Morning	1
QS00101	UV 1	Evening	0
QS00101	UV 2	Morning	0

Case 2 - Event starts after the planned data collection period

PARAMCD	AVISIT	ATPT	AVAL
QS00101	Day 28	Morning	2
QS00101	Day 28	Evening	1
QS00101	UV 1	Morning	0
QS00101	UV 1	Evening	0

Case 3 - Maintenance of needed criteria is after the planned data collection period



Source line for censoring date

Event

DEALING WITH MISSING DATA

Fully and Partially missing data

№	Symptom	Day 1 Date: 24-02-2022		Day 2 Date: 25-02-2022	
		Morning Time: 09:30	Evening Time: 18:30	Morning Time:	Evening Time: 19:00
1	Headache	1	1		1
2	Fainting/dizziness	1	1		1
3	Fatigue	1	1		1
4	Sore throat	2			2
5	Shortness of breath	0	1		0
6	Muscle Aches	1	2		1
7	Vomiting	1	0		1
8	Runny nose	1	1		1
9	Chills/night sweats	1	2		1
10	Cough	1			1
11	Unable to care for self	0	1		0
12	Little or no urination	0	1		0
13	Loss of smell	1	2		1
14	Loss of taste	1	3		1

Partially missing data: only symptoms 4 and 10 are absent

Fully missing data

Handling of partially missing data

Total rows: 13 Total columns: 6

	USUBJID	QSTESTCD	QSSTRESN	VISIT	QSDTC
1	XXX-0001	QUEST1	1	DAY 5	2021-05-06T08:56:08
2	XXX-0001	QUEST2	2	DAY 5	2021-05-06T08:56:08
3	XXX-0001	QUEST3	1	DAY 5	2021-05-06T08:56:08
4	XXX-0001	QUEST4	2	DAY 5	2021-05-06T08:56:08
5	XXX-0001	QUEST5	2	DAY 5	2021-05-06T08:56:08
6	XXX-0001	QUEST6	2	DAY 5	2021-05-06T08:56:08
7	XXX-0001	QUEST7	1	DAY 5	2021-05-06T08:56:08
8	XXX-0001	QUEST9	2	DAY 5	2021-05-06T08:56:08
9	XXX-0001	QUEST10	1	DAY 5	2021-05-06T08:56:08
10	XXX-0001	QUEST11	1	DAY 5	2021-05-06T08:56:08
11	XXX-0001	QUEST12	1	DAY 5	2021-05-06T08:56:08
12	XXX-0001	QUEST13	0	DAY 5	2021-05-06T08:56:08
13	XXX-0001	QUEST14	0	DAY 5	2021-05-06T08:56:08

Total rows: 14 Total columns: 7

	PARAMCD	AVAL	ADTM	ATPT	AVISIT	DTYPE
1	Q500101	1	06MAY2021:08:56:08	MORNING	DAY 5	
2	Q500102	2	06MAY2021:08:56:08	MORNING	DAY 5	
3	Q500103	1	06MAY2021:08:56:08	MORNING	DAY 5	
4	Q500104	2	06MAY2021:08:56:08	MORNING	DAY 5	
5	Q500105	2	06MAY2021:08:56:08	MORNING	DAY 5	
6	Q500106	2	06MAY2021:08:56:08	MORNING	DAY 5	
7	Q500107	1	06MAY2021:08:56:08	MORNING	DAY 5	
8	Q500108	3	06MAY2021:08:56:08	MORNING	DAY 5	WC
9	Q500109	2	06MAY2021:08:56:08	MORNING	DAY 5	
10	Q500110	1	06MAY2021:08:56:08	MORNING	DAY 5	
11	Q500111	1	06MAY2021:08:56:08	MORNING	DAY 5	
12	Q500112	1	06MAY2021:08:56:08	MORNING	DAY 5	
13	Q500113	0	06MAY2021:08:56:08	MORNING	DAY 5	
14	Q500114	0	06MAY2021:08:56:08	MORNING	DAY 5	

Handling of fully missing data - template

1. Define the list of subjects for TTE analysis - SUBJ
2. Form the list of questions - QUEST
3. Generate the full list of planned time points – TPOINT
4. Produce diary template – cartesian product of $\text{SUBJ} \times \text{TPOINT} \times \text{QUEST}$
5. Limit the list of timepoints for each subject according to the latest available visit in SV

Handling of fully missing data - rules

USUBJID	AVISIT	ATPT	QS00101	QS00102	QS00103	QS00104	QS00105
XXX-0001	DAY 10	MORNING	2	1	1	2	2
XXX-0001	DAY 10	EVENING	2	2	2	3	2
XXX-0001	DAY 11	MORNING	.	.	.	.	.
XXX-0001	DAY 11	EVENING	2	1	1	2	2
XXX-0001	DAY 12	MORNING	.	.	.	.	.
XXX-0001	DAY 12	EVENING	.	.	.	.	.
XXX-0001	DAY 13	MORNING	.	.	.	.	.
XXX-0001	DAY 13	EVENING	.	.	.	.	.
XXX-0001	DAY 14	MORNING	2	1	1	2	2
XXX-0001	DAY 14	EVENING	2	2	1	2	3
XXX-0001	DAY 15	EVENING	2	0	1	2	3
XXX-0001	DAY 16	EVENING	1	1	1	3	3
XXX-0001	DAY 17	EVENING	2	1	0	3	3
XXX-0001	DAY 18	EVENING	.	.	.	.	.
XXX-0001	DAY 19	EVENING	.	.	.	.	.
XXX-0001	DAY 20	EVENING	1	2	1	3	3
XXX-0001	DAY 21	EVENING	2	1	2	2	2

Case 1

Case 2

Case 3

- one or two consecutive entire symptom scores are missing -> AVAL takes the result of the next non-missing assessment: Case 1, Case 3
- more than two consecutive entire symptom scores are missing -> AVAL considers that visit as having failed the endpoint: Case 2

Handling of fully missing data – result of imputation

USUBJID	AVISIT	ATPT	QS00101	QS00102	QS00103	QS00104	QS00105	
XXX-0001	DAY 10	MORNING	2	1	1	2	2	
XXX-0001	DAY 10	EVENING	2	2	2	3	2	
XXX-0001	DAY 11	MORNING	2	1	1	2	2	Case 1
XXX-0001	DAY 11	EVENING	2	1	1	2	2	
XXX-0001	DAY 12	MORNING	99	99	99	99	99	
XXX-0001	DAY 12	EVENING	99	99	99	99	99	Case 2
XXX-0001	DAY 13	MORNING	99	99	99	99	99	
XXX-0001	DAY 13	EVENING	99	99	99	99	99	
XXX-0001	DAY 14	MORNING	2	1	1	2	2	
XXX-0001	DAY 14	EVENING	2	2	1	2	3	
XXX-0001	DAY 15	EVENING	2	0	1	2	3	
XXX-0001	DAY 16	EVENING	1	1	1	3	3	
XXX-0001	DAY 17	EVENING	2	1	0	3	3	
XXX-0001	DAY 18	EVENING	1	2	1	3	3	Case 3
XXX-0001	DAY 19	EVENING	1	2	1	3	3	
XXX-0001	DAY 20	EVENING	1	2	1	3	3	
XXX-0001	DAY 21	EVENING	2	1	2	2	2	

□ one or two consecutive entire symptom scores are missing -> AVAL takes the result of the next non-missing assessment: Case 1, Case 3

□ more than two consecutive entire symptom scores are missing -> AVAL considers that visit as having failed the endpoint: Case 2

Handling fully missing data – PHANTOM record

USUBJID	PARAMCD	AVAL	ADTM	ADY	ATPT	AVISIT	AVISITN	DTYPE
XXX-0001	QS00101	2	11MAY2021:09:00:53	10	MORNING	DAY 10	10	
XXX-0001	QS00101	2	11MAY2021:20:01:04	10	EVENING	DAY 10	10	
XXX-0001	QS00101	2	12MAY2021:00:00:00	11	MORNING	DAY 11	11	PHANTOM
XXX-0001	QS00101	2	12MAY2021:21:02:42	11	EVENING	DAY 11	11	

- ✓ ADT is taken from SV
- ✓ Rules to impute time should be defined additionally

Mingle with Baseline and Day 1 data

USUBJID	PARAMCD	AVAL	ADTM	ADY	ATPT	AVISIT	AVISITN	DTYPE
XXX-0001	QS00101	2	02MAY2021:10:03:52	1	MORNING	Baseline	0	
XXX-0001	QS00101	1	02MAY2021:20:13:52	1	EVENING	DAY 1	1	
XXX-0001	QS00101	1	03MAY2021:10:21:16	2	MORNING	DAY 2	2	
XXX-0001	QS00101	1	03MAY2021:20:01:23	2	EVENING	DAY 2	2	

- ✓ Day 1 Morning -> Baseline
- ✓ No need to create additional record for Day 1 Morning

ADaMs ARRANGEMENT FOR TTE CALCULATION

ADQS vs ADDIAR

ADQS						ADDIAR		
usubjid	paramcd	avisit	atpt	adtm	aval	adtm	aval	dtype
1	XXX-0007	QS00101	DAY 8	MORNING	11AUG2021:08:32:39	0	11AUG2021:08:32:39	0
2	XXX-0007	QS00101	DAY 8	EVENING	11AUG2021:19:41:39	1	11AUG2021:19:41:39	1
3	XXX-0007	QS00101	DAY 9	MORNING	12AUG2021:15:00:00	1	12AUG2021:00:00:00	2 PHANTOM
4	XXX-0007	QS00101	DAY 9	EVENING	12AUG2021:19:14:00	2	12AUG2021:19:14:00	2
5	XXX-0007	QS00101	DAY 10	MORNING	13AUG2021:09:17:21	1	13AUG2021:09:17:21	1
6	XXX-0007	QS00101	DAY 10	EVENING	13AUG2021:19:36:13	1	13AUG2021:19:36:13	1

ADDIAR – Analysis Diary Dataset helps:

- to apply additional restrictions to data for TTE analysis
- to facilitate QC process

Symptom data storage in ADaM

DIARY	ADaM
Symptom name	Param/Paramcd
Date, time of a symptom assessment	ADT/ADTM
Time point of a symptom assessment	ATPT/ATPTN
Visit of an assessment	AVISIT/AVISIT
Symptom score	AVAL/AVALC

TEE Calculation

Determine the intervals during which the event requires maintenance

```
data adqs1;
  set adqs;
  by usubjid paramcd adtm;

  length avaln prev_avaln prev_adtm 8 prev_usubjid $50 prev_paramcd $8;
  format prev_adtm datetime20.;

  avaln = ifn(aval in (0 1), 0, 99);

  prev_usubjid = lag(usubjid);
  prev_paramcd = lag(paramcd);
  prev_avaln = lag(avaln);

  gr + 1;

  if prev_usubjid = usubjid
  and prev_paramcd = paramcd
  and prev_avaln = avaln
  and avaln = 0
  then gr = gr - 1;
run;
```

Determine the duration of the maintenance that falls within these intervals and select only those intervals that meet our requirements

```
proc sql;
  create table tte1 as
  select *
  from adqs1
  group by usubjid, paramcd, gr
  having (max(adtm) - min(adtm))/3600 >= 21.5
  order by usubjid, paramcd, gr, adtm;
quit;
```

Select the start time of each qualifying interval

```
data tte;
  set tte1;
  by usubjid paramcd gr adtm;
  if first.paramcd;
run;
```

CONCLUSION

Points to consider during preparing Diary data for TTE analysis:

- define which diary data will be used for TTE analysis;
- define strategy for dealing with missing timepoints;
- establish censoring rules while considering the period of interest;
- pay attention to defining baseline and its possible overlap with scheduled assessments;
- prepare diary data in separate dataset.

THANK YOU FOR YOUR ATTENTION!

QUESTIONS?