The Stata Journal (2010)
10, Number 3, pp. 386–394

Projection of power and events in clinical trials with a time-to-event outcome

Patrick Royston
Hub for Trials Methodology Research
MRC Clinical Trials Unit and University College London
London, UK
pr@ctu.mrc.ac.uk

Friederike M.-S. Barthel
Oncology Research & Development
GlaxoSmithKline
Uxbridge, UK
FriederikeB@ctu.mrc.ac.uk

Abstract. In 2005, Barthel, Royston, and Babiker presented a menu-driven Stata program under the generic name of ART (assessment of resources for trials) to calculate sample size and power for complex clinical trial designs with a time-to-event or binary outcome. In this article, we describe a Stata tool called ARTPEP, which is intended to project the power and events of a trial with a time-to-event outcome into the future given patient accrual figures so far and assumptions about event rates and other defining parameters. ARTPEP has been designed to work closely with the ART program and has an associated dialog box. We illustrate the use of ARTPEP with data from a phase III trial in esophageal cancer.

Keywords: st0013_2, artpep, artbin, artsurv, artmenu, randomized controlled trial, time-to-event outcome, power, number of events, projection, ARTPEP, ART

1 Introduction

Barthel, Royston, and Babiker (2005) presented a menu-driven Stata program under the generic name of ART (assessment of resources for trials) to calculate sample size and power for complex clinical trial designs with a time-to-event or binary outcome. Briefly, the features of ART include multiarm trials, dose–response trends, arbitrary failure-time distributions, nonproportional hazards, nonuniform rates of patient entry, loss to follow-up, and possible changes from allocated treatment. A full report on the methodology and its performance—in particular, regarding loss to follow-up, nonproportional hazards, and treatment crossover—is given by Barthel et al. (2006).

In this article, we concentrate on a new tool that addresses a practical issue in trials with a time-to-event outcome. Because of staggered entry of patients and the gradual maturing of the data, the accumulation of events from the date the trial opens is a process that occurs over a relatively long period of time and with a variable course. Trials are planned and their resources are assigned under certain critical assumptions.

If those assumptions are unrealistic, timely completion of the trial may be threatened. Because the cumulative number of events is the key indicator of trial maturity and is the parameter targeted in the sample-size calculation, it is of considerable interest and relevance to monitor and project this number at particular points during the trial.

The new tool is called ARTPEP (ART projection of events and power). ARTPEP comprises an ado-file (`artpep`) and an associated dialog box. It works in conjunction with the ART system, of which the latest update is included with this article.

2 Example: A trial in advanced esophageal cancer

2.1 Sample-size calculation using ART

As an example, we describe sample-size calculation and ARTPEP analysis of a "typical" cancer trial. The OE05 trial in advanced esophageal carcinoma is coordinated by the MRC Clinical Trials Unit. The protocol is available online at http://www.ctu.mrc.ac.uk/plugins/StudyDisplay/protocols/OE05%20Protocol%20Version%205%2031st%20July%202008.pdf. The design, which comprises two randomized groups of patients with equal allocation, aims to test the hypothesis that a new chemotherapy regimen, in conjunction with surgery, improves overall survival at 3 years.

According to the protocol, the probability of 3-year survival in this patient group is 30%, and the trial has 82% power at the 5% two-sided significance level to detect an improvement in overall survival to 38%. The overall sample size is stated to be 842 patients, and the required number of events is 673. The plan is to recruit patients over 6 years and to follow up with them for a further 2 years before performing the definitive analysis of the outcome (overall survival).

The description in the protocol provides nearly all the ingredients for an ART sample-size and power calculation. The only missing item is the target hazard ratio, which is $\ln(0.38) / \ln(0.30) = 0.80$ under proportional hazards of the treatment effect (a standard assumption). We first use the `artsurv` command (Barthel, Royston, and Babiker 2005) to verify the sample-size calculation and to set up some of the parameter values needed by ARTPEP. We supply the other design features, and then we run the `artsurv` command to compute the power and events:

(*Continued on next page*)

```
. artsurv, method(l) nperiod(8) ngroups(2) edf0(0.3, 3) hratio(1, 0.80) n(842)
> alpha(0.05) recrt(6)
ART - ANALYSIS OF RESOURCES FOR TRIALS        (version 1.0.7, 19 October 2009)
```

```
A sample size program by Abdel Babiker, Patrick Royston & Friederike Barthel,
MRC Clinical Trials Unit, London NW1 2DA, UK.
```

Type of trial	Superiority - time-to-event outcome
Statistical test assumed	Unweighted logrank test (local)
Number of groups	2
Allocation ratio	Equal group sizes
Total number of periods	8
Length of each period	One year
Survival probs per period (group 1)	0.669 0.448 0.300 0.201 0.134 0.090 0.060 0.040
Survival probs per period (group 2)	0.725 0.526 0.382 0.277 0.201 0.146 0.106 0.077
Number of recruitment periods	6
Number of follow-up periods	2
Method of accrual	Uniform
Recruitment period-weights	1 1 1 1 1 1 0 0
Hazard ratios as entered (groups 1,2)	1, 0.80
Alpha	0.050 (two-sided)
Power (calculated)	0.824
Total sample size (designed)	842
Expected total number of events	673

Apart from small, unimportant differences, the protocol power (0.82) and the number of events (673) are consistent with ART's results.

2.2 Analysis with ARTPEP

To run ARTPEP successfully, three preliminary steps are required:

1. You must activate the ART and ARTPEP items on the **User** menu by typing the command `artmenu on`.
2. You must compute the relevant sample size for the trial using either the ART dialog box or the `artsurv` command. This automatically sets up a global macro called `$S_ARTPEP` whose contents are used by the `artpep` command. (A slightly more convenient alternative with the same result is to use the **ART Settings...** button on the ARTPEP dialog box to set up the necessary quantities for ART without having to run ART or `artsurv` separately.)
3. To set up additional parameters that ARTPEP needs, you must use the ARTPEP dialog box, either by typing `db artpep` or by selecting **User** > **ART** > **Artpep** from the menu.

As a worked example, we now imagine that the OE05 trial has been running for 1 year and has accrued 100 patients so far. Assuming the survival distribution to be

correct, when may we expect to complete the trial (that is, obtain the required number of events)? To answer this question, we complete the three steps described above. The resulting empty dialog box is shown in figure 1.

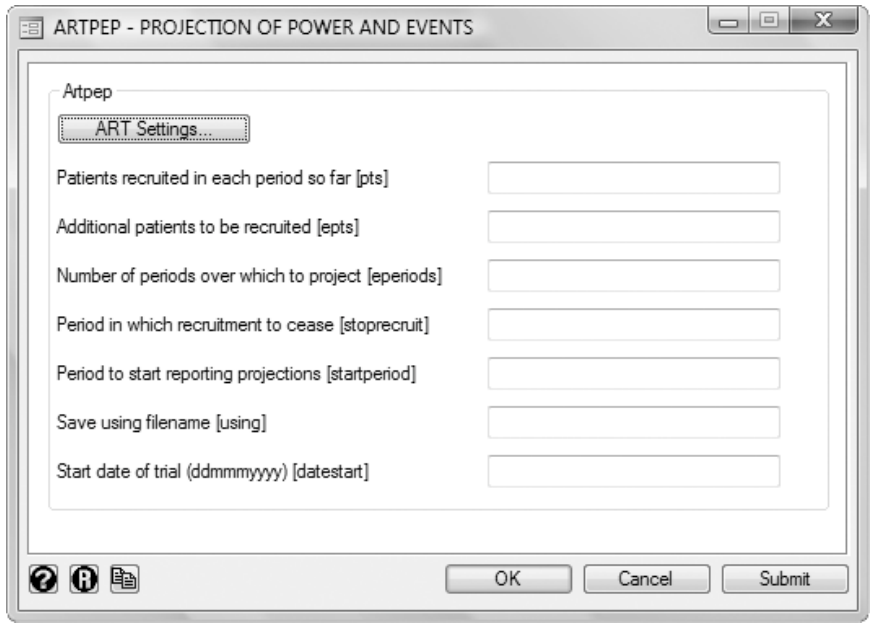

Figure 1. Incomplete ARTPEP dialog box

We now explain the various items that the dialog box needs. The name of the corresponding option for the `artpep` command is given in square brackets:

- **ART Settings...**: As already mentioned, this button may be used to set up the parameters of an ART run if that has not been done already. It accesses the ART dialog box.
- *Patients recruited in each period so far [pts]*: A "period" here is 1 year, and we have recruited 100 patients in the first period. We therefore enter 100 for this item.
- *Additional patients to be recruited [epts]*: To get to the 842 patients (we will use 850), we hope to recruit at about 150 patients per year for the next 5 years, making a total of 6 years' planned recruitment. We enter `150`. The program knows the period in which recruitment is to cease and, by default, repeats the number 150 over the next 5 periods. If we had expected a differing recruitment rate (say, accelerating toward the end of the trial), we could have entered a different number of patients to be recruited in each period.

- *Number of periods over which to project [eperiods]*: Let us say we wish to project events and power over the next 10 years. We enter `10`.
- *Period in which recruitment cease [stoprecruit]*: Here enter the number of periods after which recruitment is to cease. The number must be no smaller than the number of periods implied by *Patients recruited in each period so far [pts]*. If the option is left blank, it is assumed that recruitment continues indefinitely. As already noted, we wish to stop recruitment at 850 patients, which we will achieve by the end of period 6. We therefore enter `6` for this item.
- *Period to start reporting projections [startperiod]*: Usually, we want to enter `1` here, signifying the start of the trial. By default, if the item is left blank, the program assumes that the current period is intended. We enter `1`.
- *Save using filename [using]*: The numerical results of the `artpep` run can be saved to a `.dta` file for a permanent record or for plotting. We leave the item blank.
- *Start date of trial (ddmmmyyyy) [datestart]*: If we enter the start date, the output from `artpep` is conveniently labeled with the calendar date of the end of each period. We recommend using this option. We enter `01jan2009`.

The completed ARTPEP dialog box is shown in figure 2.

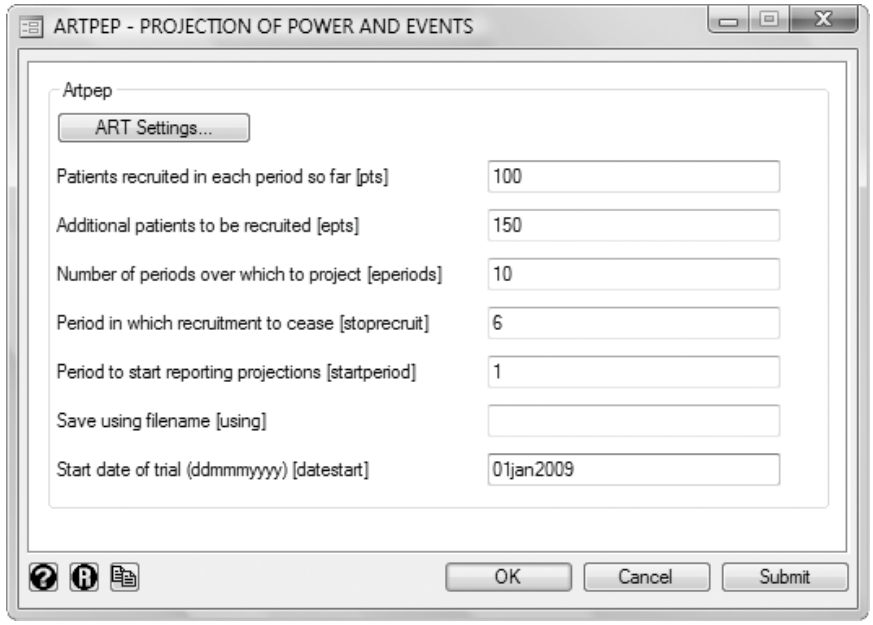

Figure 2. Completed ARTPEP dialog box for the OE05 trial

After submitting the above setup to Stata (version 10 or later), we get the following result:

```
. artpep, pts(100) $S_ARTPEP epts(150) eperiods(10) startperiod(1)
> stoprecruit(6) datestart(01jan2009)
```

Date	year	#pats	#C-events	#events	Power
31dec2009	1	100	9	17	0.06498
31dec2010	2	250	36	66	0.14480
31dec2011	3	400	79	146	0.26850
31dec2012	4	550	132	247	0.41622
31dec2013	5	700	193	362	0.56360
31dec2014	6	850	258	488	0.69209
31dec2015	7	850	314	597	0.77737
31dec2016	8	850	351	673	0.82423
31dec2017	9	850	375	726	0.85155
31dec2018	10	850	392	763	0.86825
31dec2019	11	850	403	789	0.87882

The program reports the total number of events (`#events`) and the number of events in the control arm (`#C-events`), which are often of interest. The required total number of events (that is, both arms combined) of 673 is projected to be reached on 31 December 2016, the end of period 8. We expect 351 events in the control arm by that time. The projection is not surprising because the accrual figures that have been entered more or less agree with the trial plan. Nevertheless, the output shows us the expected progress of the number of events and the power over time. The trial may be monitored (and the ARTPEP analysis updated) to follow its progress.

The dialog box has, as usual, created and run the necessary `artpep` command line. The second item in the command is `$S_ARTPEP`. As already mentioned, it contains additional information needed by `artpep`. On displaying its contents, we find

```
. display "$S_ARTPEP"
alpha(.05) aratios() hratio(1, 0.80) ngroups(2) ni(0) onesided(0) trend(0)
> tunit(1) edf0(0.3, 3) median(0) method(1)
```

The key pieces of information here are `hratio(1, 0.80)` and `edf0(0.3, 3)`, which specify the hazard ratios in groups 1 and 2, and the survival function in group 1, respectively. All the other items are default values and could be omitted in the present example. The present example could have been run directly from the command line as follows:

```
. artpep, pts(100) edf0(0.3, 3) epts(150) eperiods(10) startperiod(1)
> stoprecruit(6) datestart(01jan2009) hratio(1, 0.8)
```

2.3 Sensitivity analysis of the event rate

We have assumed a 30% survival probability 3 years after recruitment. Suppose, in fact, that the patients do better than that—their 3-year survival is 40% instead. What effect would that have on the power and events timeline?

We need only change the `edf0()` option to `edf0(0.4, 3)`:

```
. artpep, pts(100) epts(150) edf0(0.4, 3) eperiods(10) startperiod(1)
> stoprecruit(6) datestart(01jan2009) hratio(1, 0.8)
```

Date	year	#pats	#C-events	#events	Power
31dec2009	1	100	7	13	0.05869
31dec2010	2	250	29	53	0.12410
31dec2011	3	400	65	119	0.22732
31dec2012	4	550	111	205	0.35714
31dec2013	5	700	165	306	0.49586
31dec2014	6	850	224	419	0.62612
31dec2015	7	850	277	522	0.72135
31dec2016	8	850	316	600	0.77974
31dec2017	9	850	345	660	0.81685
31dec2018	10	850	366	705	0.84128
31dec2019	11	850	382	739	0.85784

The time to observe the required number of events has advanced by more than 1 year, to period 9 (`31dec2017`).

3 Syntax

Once you have gained a little experience with using the ARTPEP dialog box, you will find it more natural and efficient to use the command line. The syntax of `artpep` is as follows:

`artpep` [`using` *filename*], `pts(`*numlist*`)` `edf0(`*slist0*`)` [`epts(`*numlist*`)`
`eperiods(`#`)` `stoprecruit(`#`)` `startperiod(`#`)` `datestart(`*ddmmmyyyy*`)`
`replace` *artsurv_options*]

4 Options

`pts(`*numlist*`)` is required. *numlist* specifies the number of patients recruited in each period since the start of the trial, that is, since randomization. See help on `artsurv` for the definition of a "period". The number of items in *numlist* defines the number of periods of recruitment so far. For example, `pts(23 12 25)` specifies three initial periods of recruitment, with recruitment of 23 patients in period 1, 12 in period 2, and 25 in period 3. The "current" period would be period 3 and would be demarcated by parallel lines in the output.

`edf0(`*slist0*`)` is required and gives the survival function in the control group (group 1). This need not be one of the survival distributions to be compared in the trial, unless `hratio()` = 1 for at least one of the groups. The format of *slist0* is *#1* [*#2* ... *#r*, *#1 #2* ... *#r*]. Thus `edf0(`$p_1\ p_2 \ldots\ p_r$`,`$\ t_1\ t_2 \ldots t_r$`)` gives the value p_i for the survival function for the event time at the end of time period t_i, $i = 1, \ldots, r$. Instantaneous event rates (that is, hazards) are assumed constant within time periods; that is, the

distribution of time-to-event is assumed to be piecewise exponential. When used in a given calculation up to period T, t_r may validly be less than, equal to, or greater than T. If $t_r \leq T$, the rules described in the `edf0()` option of `artsurv` are applied to compute the survival function at all periods $\leq T$. If $t_r > T$, the same calculation is used but estimated survival probabilities for periods $> T$ are not used in the calculation at T, although they may of course be used in calculations (for example, projections of sample size and events) for periods later than T. Be aware that use of the `median()` option (an alternative to `edf0()`) and the `fp()` option of `artsurv` may modify the effects and interpretation of `edf0()`.

`epts(`*numlist*`)` specifies in *numlist* the number of additional patients to be recruited in each period following the recruitment phase defined by the `pts()` option. For example, `pts(23 12 25) epts(30 30)` would specify three initial periods of recruitment followed by two further periods. A projection of events and power is required over the two further periods. The initial recruitment is of 23 patients in period 1, 12 in period 2, and 25 in period 3; in each of periods 4 and 5, we expect to recruit an additional 30 patients. If the number of items in (or implied by expanding) *numlist* is less than that specified by `pts()`, the final value in *numlist* is replicated as necessary to all subsequent periods. If `epts()` is not given, the default is that the mean of the numbers of patients specified in `pts()` is used for all projections.

`eperiods(`#`)` specifies the number of future periods over which projection of power and number of events is to be calculated. The default is `eperiods(1)`.

`stoprecruit(`#`)` specifies the number of periods after which recruitment is to cease. # must be no smaller than the number of periods of recruitment implied by `pts()`. The default is `stoprecruit(0)`, meaning to continue recruiting indefinitely (no follow-up phase).

`startperiod(`#`)` specifies # as the period in which to start reporting the projections of events and power. To report from the beginning of the trial, specify `startperiod(1)`. Note that `startperiod()` does not affect the period at which the calculations are started, only how the results are reported. The default # is the last period defined by `pts()`.

`datestart(`*ddmmmyyyy*`)` signifies the opening date of the trial (that is, when recruitment started), for example, `datestart(14oct2009)`. The date of the end of each period is used to label the output and is stored in *filename* if `using` is specified.

`replace` allows *filename* to be replaced if it already exists.

artsurv_options are any of the options of `artsurv` except `recrt()`, `nperiod()`, `power()`, and `n()`.

(*Continued on next page*)

5 Final comments

We have illustrated ARTPEP with a basic example. However, ARTPEP understands the more complex options of `artsurv`. Therefore, complex features, including loss to follow up, treatment crossover, and nonproportional hazards, can be allowed for in the projection of power and events.

Sometimes it is desirable to make projections on a finer time scale than 1 year, for example, in 3- or 6-month periods. This is easily done by adjusting the period parameters used in ART and ARTPEP.

6 References

Barthel, F. M.-S., A. Babiker, P. Royston, and M. K. B. Parmar. 2006. Evaluation of sample size and power for multi-arm survival trials allowing for non-uniform accrual, non-proportional hazards, loss to follow-up and cross-over. *Statistics in Medicine* 25: 2521–2542.

Barthel, F. M.-S., P. Royston, and A. Babiker. 2005. A menu-driven facility for complex sample size calculation in randomized controlled trials with a survival or a binary outcome: Update. *Stata Journal* 5: 123–129.

About the authors

Patrick Royston is a medical statistician with 30 years of experience, with a strong interest in biostatistical methods and in statistical computing and algorithms. He now works in cancer clinical trials and related research issues. Currently, he is focusing on problems of model building and validation with survival data, including prognostic factor studies; on parametric modeling of survival data; on multiple imputation of missing values; and on novel clinical trial designs.

Friederike Barthel is a senior statistician in Oncology Research & Development at GlaxoSmithKline. Previously, she worked at the MRC Clinical Trials Unit and the Institute of Psychiatry. Her current research interests include sample-size issues, particularly concerning multistage, multiarm trials, microarray study analyses, and competing risks. Friederike has taught undergraduate courses in statistics at the University of Westminster and at Kingston University.