ISPOR EUROPE 2026 · VIENNA · 8 TO 11 NOVEMBER

MSR88

The merlin ecosystem: an integrated open-source platform for survival analysis, multi-state modeling, and health economic evaluation in R and Stata

Michael J. Crowther · Red Door Analytics, Stockholm, Sweden

Red Door Analytics

OBJECTIVES

Health technology assessment increasingly demands complex statistical models: joint longitudinal-survival models, competing risks, multi-state models and Bayesian uncertainty. Yet the software is fragmented, so analysts stitch tools together across estimation, prediction and cost-effectiveness. The **merlin** ecosystem is four integrated open-source commands, for R and for Stata, that cover the whole pipeline with one general syntax: the model fitted by **merlin** is the model sampled by **morgana**, costed by **pendragon** and checked by **avalon**.

The case study: a cancer progression model, an illness-death model of relapse, death without relapse and death after relapse, for 2,982 women with breast cancer from the Rotterdam tumour bank, followed for up to 19 years. The comparison of hormonal therapy is **illustrative**: the data are observational, and we adjust for age, number of positive nodes, tumour size, grade and progesterone receptor status.

01 · FIT THE MODEL

merlin

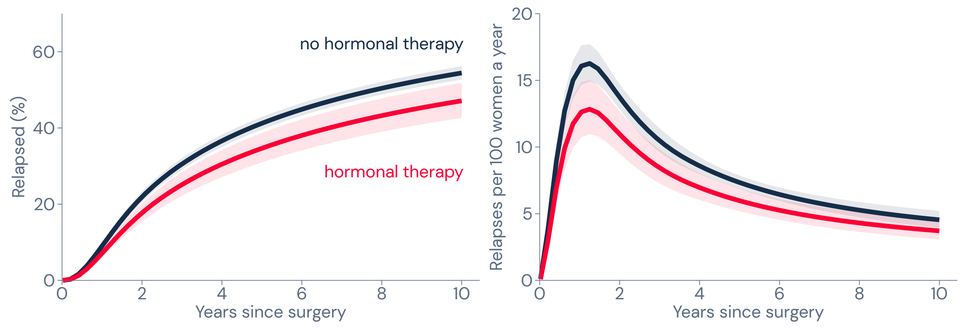
A long-established maximum likelihood engine: parametric survival, relative survival, cure, competing risks, multi-state and joint models, with multilevel random effects.

STATA · THE RELAPSE TRANSITION

```
merlin (_stop hormon age enodes pr_1 size2 size3 grade3, family(rp, df(3) failure(_status))  ///
       timevar(_stop)) if _trans1==1
```

R · THE SAME MODEL

```
merlin(Surv(stop, status) ~ hormon + age + enodes + pr_1 + size2 + size3 + grade3,
       data = d1, family = "rp", df = 3)
```



Left: women relapsed (%), death without relapse competing. **Right:** relapses per 100 women a year. Standardised over the 2,982 women; 95% confidence bands. **By 10 years:** 54.4% (52.7 to 56.2) without hormonal therapy, 47.1% (42.6 to 51.8) with it: 7.3 points fewer (2.5 to 12.1). STATA MERLIN 3.0.0 · R MERLIN 2.0.0 · DEVELOPMENT VERSIONS

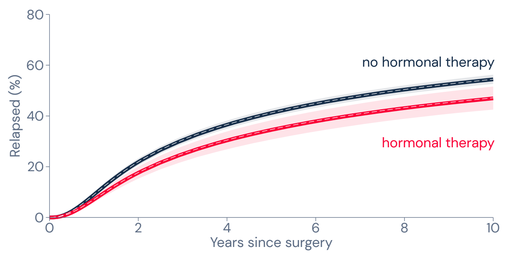
02 · TAKE THE BAYESIAN APPROACH

morgana

Extends merlin with a Bayesian sampler, Metropolis-Hastings in Stata (through Stata's built-in bayesmh) and Hamiltonian Monte Carlo in R (through tmbstan): priors go in, and the posterior uncertainty is carried into every prediction.

STATA · THE SAME THREE TRANSITIONS, SAMPLED IN ONE CALL

```
morgana, rseed(20261012) mcmcsize(2000) burnin(5000) thinning(100) :        ///
    merlin (_stop hormon age enodes pr_1 size2 size3 grade3 if _trans1==1,  ///
              family(rp, df(3) failure(_status)) timevar(_stop))            ///
          (_stop hormon age enodes pr_1 size2 size3 grade3 if _trans2==1,  ///
              family(rp, df(3) failure(_status)) timevar(_stop))            ///
          (_stop hormon age enodes pr_1 size2 size3 grade3 if _trans3==1,  ///
              family(rp, df(3) failure(_status) ltruncated(_start)) timevar(_stop))
```



The same curve, sampled: posterior median and 95% credible band; dashed white line, merlin's maximum likelihood. **By 10 years:** 54.4% (52.6 to 56.2) without therapy, 46.9% (42.5 to 51.6) with it. Metropolis-Hastings, 2,000 draws, normal priors (mean 0, SD 100); smallest effective sample size 520. MORGANA 2.0.0 · DEVELOPMENT VERSION

03 · ESTIMATE COST-EFFECTIVENESS

pendragon

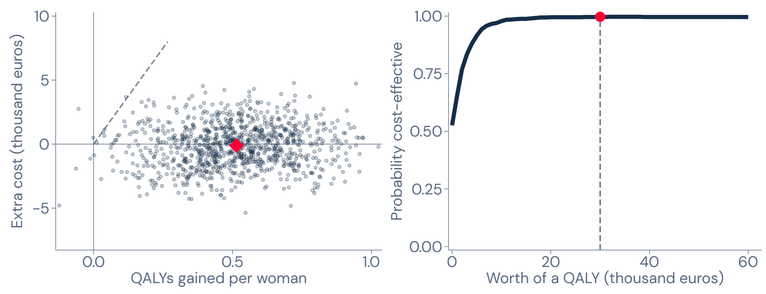
Reads merlin or morgana fits and returns multi-state predictions, life-years, QALYs, costs, ICERs, cost-effectiveness planes and acceptability curves, all with probabilistic sensitivity analysis.

STATA · FROM THE SAME THREE FITS

```
pendragon_cea, transmatrix(tmat) models(m1 m2 m3) timevar(tt) at1(hormon 1) at2(hormon 0)  ///
    standardise standif(_trans==1) utility(0.8 0.6 0) cost(83.33333 1000 0)                ///
    discount(0.002873774) discmethod(annual) wtp(0(2500)60000) reps(1000) seed(20261016) los
```

R · THE SAME STEP, ON THE THREE R MERLIN FITS (WOMEN: THE 2,982 WOMEN, ONE ROW EACH)

```
msm <- pendragon_msm(tmat, list(m1, m2, m3), clock = "forward")
ic  <- icer(msm, newdata    = transform(women, hormon = 1),
                 comparator = transform(women, hormon = 0),
            times = 0:228, utilities = c(0.8, 0.6, 0), state_costs = c(83.33333, 1000, 0),
            discount = 0.002873774, disc_method = "annual",
            uncertainty = "auto", B = 1000L, seed = 20261016)
```



Therapy against none, per woman, over 19 years; 1,000 probabilistic draws. **Left:** the cost-effectiveness plane (diamond: estimate; dashed: 30,000 euros per QALY). **Right:** probability cost-effective. **Cost-saving, with 0.514 QALYs gained** (0.116 to 0.882); extra cost -117 euros (-3,381 to 3,186); probability cost-effective at 30,000 euros 0.997. Illustrative inputs: 1,000 and 12,000 euros a year relapse-free and relapsed (entered per month), the therapy 2,500 euros once, 3.5% discounting. ILLUSTRATIVE INPUTS · PENDRAGON 0.5.0 · PENDRAGON FOR R · DEVELOPMENT VERSIONS

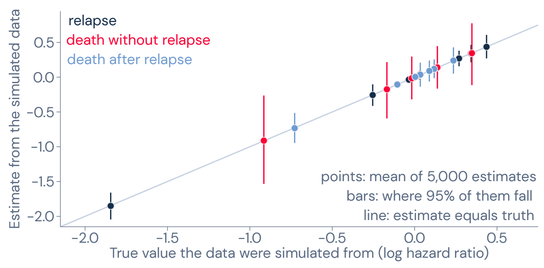
04 · CHECK BY SIMULATION

avalon

Simulates complex survival data, including time-varying effects, recurrent events and competing risks, so a model and its assumptions can be tested against a known truth.

STATA · ONE DATA SET SIMULATED FROM THE THREE FITTED MODELS

```
avalon msm time state event, hazard1(model(m1)) hazard2(model(m2)) hazard3(model(m3))  ///
    transmatrix(tmat) maxtime(120) seed(20261013) long
```



5,000 data sets simulated with this line, each refitted with merlin. One point per covariate effect (21): the 95% intervals covered the truth 94.3% to 95.3% of the time; 14,998 of 15,000 fits converged.
AVALON 0.1.0 · MERLIN 3.0.0 · DEVELOPMENT VERSIONS

BENCHMARKS · THE ENGINES' TEST RECORDS, OCTOBER 2026

Every command checked four ways

AGAINST	WHAT EACH COMMAND IS CHECKED AGAINST
Stata	**merlin:** streg, stcox, stintreg, mixed, mestreg, meintreg, gsem, glm, poisson, nbreg, betareg, stpm2, strcs. **morgana:** Bayesian streg (five families) and Bayesian mestreg. **pendragon:** sts, stcox, streg. **avalon:** streg recovers five families and a covariate from its simulated data, each within 3%
R	**merlin:** the survival (survreg, coxph), flexsurv and rstpm2 packages. **pendragon:** survival, flexsurv, rstpm2, hesim. **avalon:** survival, flexsurv
Oracle code	likelihoods and predictions written out by hand, apart from the commands; closed forms and analytic identities; analytic derivatives against numerical ones; simulated data with a known truth
Each other	morgana's posteriors against merlin's maximum likelihood; pendragon reproduces merlin's own predictions in nine families; avalon against merlin's predictions; **Stata against R:** 441 of 441 quantities agree (merlin 20 cases, pendragon 5, avalon 1 across eight families)

Test scripts run, none failed: merlin 229 (for R 127), morgana 38, pendragon 58, avalon 24. Records of 7 October 2026 (Stata merlin, morgana, pendragon; Stata against R) and 21 October 2026 (merlin for R, avalon), development versions; pendragon and avalon for R at development grade. On this poster's example, R and Stata give the same three log likelihoods to four decimals.

SOFTWARE, NOVEMBER 2026

What you can install

COMMAND	STATA	R
merlin	v2.5.0 released · install from reddooranalytics.se/software/merlin/stata · v3.0.0, used for this poster, coming at the end of November	v2.0 development version · built and tested · install from reddooranalytics.se/software/merlin/r
morgana	v1.0.2 released · install from reddooranalytics.se/software/morgana · v2.0.0, used for this poster, is not yet public	in development, not yet public
pendragon	in development, not yet public	in development, not yet public
avalon	in development, not yet public	in development, not yet public

Every result on this poster was run on a development version. **Try them on your own data:** merlin and morgana for Stata are ready to install, and so is the development version of merlin for R. Install lines and what is coming next: reddooranalytics.se/software

CONCLUSIONS

- One general syntax from the fit to the decision, in R and in Stata.
- Each step reads the model the step before it fitted: no manual transfer between steps, so the whole pipeline can run from one analysis script.

References. Crowther MJ. merlin: a unified modeling framework for data analysis and methods development in Stata. Stata Journal 2020;20(4):763–784. Crowther MJ, Lambert PC. Parametric multistate survival models: flexible modelling allowing transition-specific distributions with application to estimating clinically useful measures of effect differences. Statistics in Medicine 2017;36(29):4719–4742. **Data.** The Rotterdam tumour bank, 2,982 patients. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. BMC Medical Research Methodology 2013;13:33.



PRESENTING AUTHOR
Michael J. Crowther, PhD
CEO & Director of Statistical Methodology
michael.crowther@reddooranalytics.se

Scan for the poster PDF