Bayesian meta-analysis of time to benefit

W. John Boscardin
Departments of Medicine and Epidemiology & Biostatistics
University of California, San Francisco

July 20, 2023

Co-authors and Acknowledgements

- Co-authors
  - Irena Cenzer
  - Sei Lee
  - Matthew Growdon
  - James Deardorff
- Acknowledgements
  - Jasmine Kang
  - UCSF Statistical Laboratory for Aging Research
  - UCSF Pepper Center

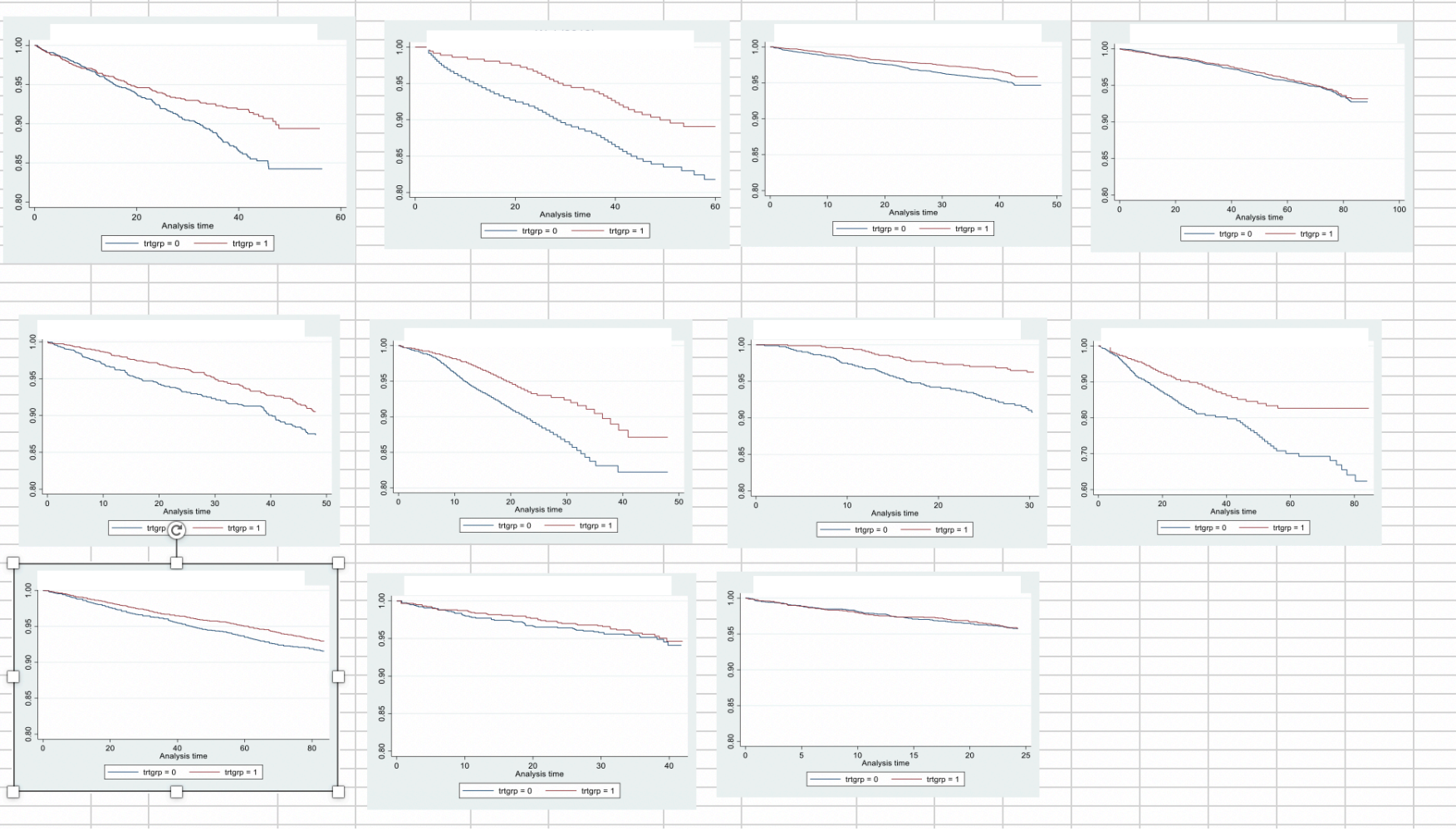
How our work fits in to Stata conference

- I direct UCSF Statistical Laboratory for Aging Research (10 full time statisticians; based out of Pepper Center and Division of Geriatrics)
- Team-science framework with emphasis on deep, longitudinal collaboration with clinical researchers
- We are not currently Stata programming experts at level of others in this conference
- Historically, users of Stata, and users/programmers in SAS/R
- Stata-specific tools have become incredibly useful for our research in general (e.g. `margins`, `svy`, `mi`) and specifically for today's topic (e.g. `ipdfc`, `meta`, `bayes:streg`, `bayesmh`)
- Have end-to-end Stata script for this project; hope to create proper Stata command in near future
- Stata potentially better fit at UCSF (e.g. training in clinical research is Stata-centric from Vittinghoff et al. textbook)

Topics for today

- Reconstruction of individual patient survival data from Kaplan-Meier figures in publications of clinical trials
- Alternatives to hazard ratios (time to benefit; difference in restricted mean survival)
- Estimation of these with Bayesian parametric survival
- Combining across multiple studies (meta-analysis)

Meta-analysis worksheet



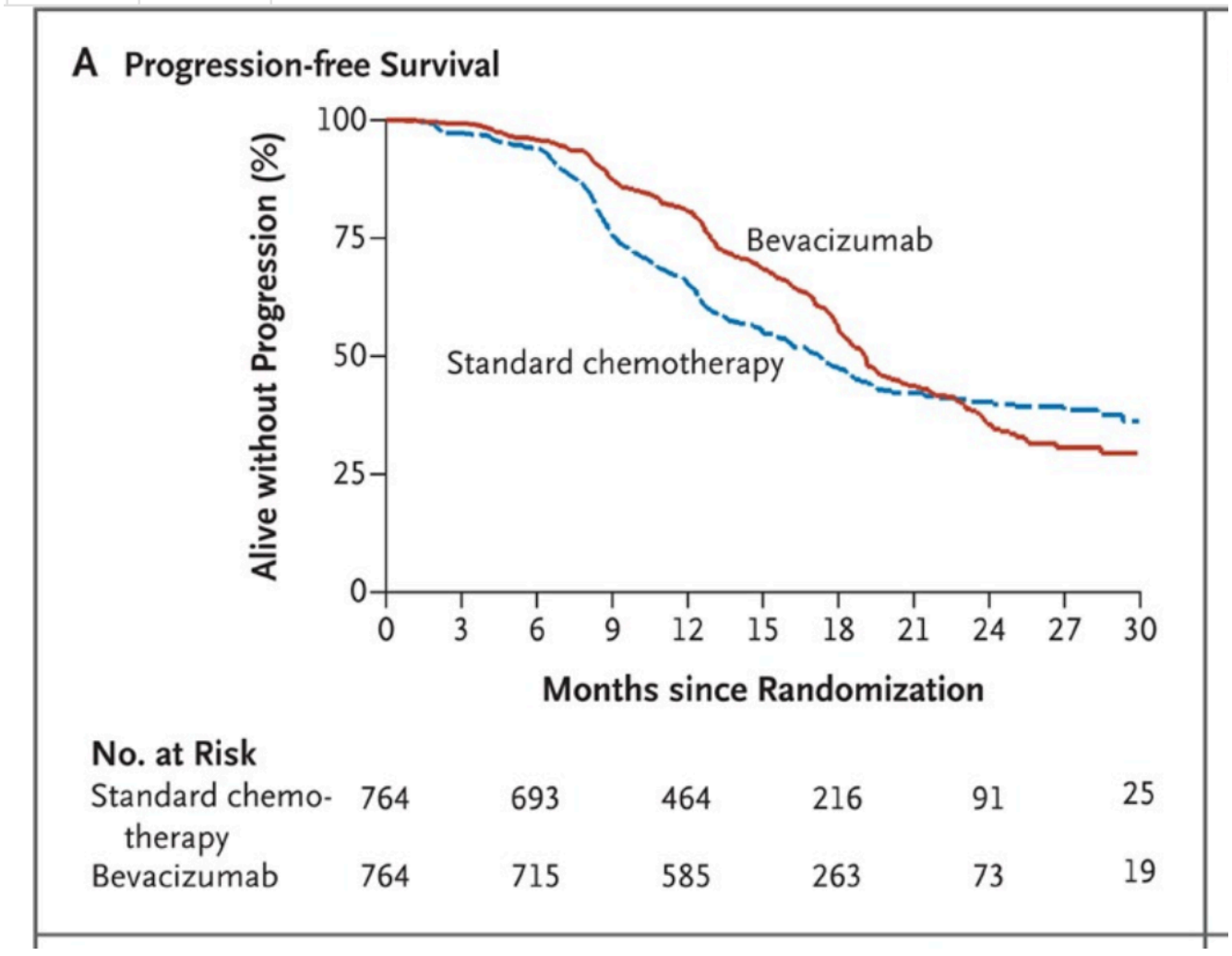
Reconstruction of individual patient data

- Clinical trials often publish Kaplan-Meier curves for each arm and hazard ratio with 95% CI
- If want to look at other metrics, would be great if had the individual patient data
- Turns out this can be reconstructed from the Kaplan-Meier curves with high fidelity (Guyot 2012; Parmar 1998; Earle 2002)
- First step: extract the coordinates of the steps on the figure and number at risk information
- Second step: use this info to figure out number of events and censored at each jump in curve; this allows creation of a standard individual patient dataset

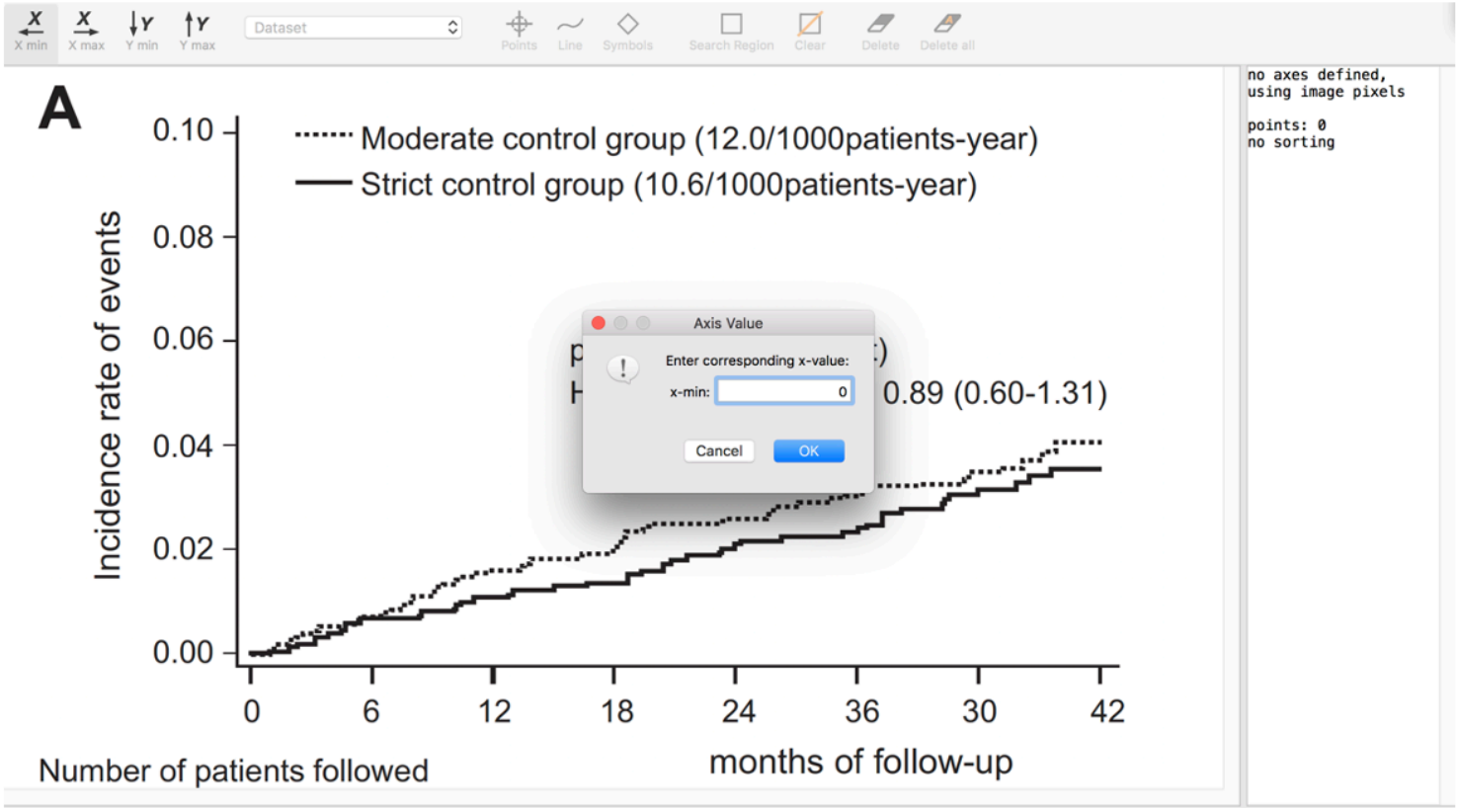
First step: a picture worth a thousand numbers

- Numerous packages and methods to turn a figure from a published paper back into the underlying numbers
- Raster figures: ycasd (Gross 2013), g3plot, WebPlotDigitizer, Engauge, DigitizeIt
- Vector figures: exact numbers are computable from file and can be extracted (Liu 2015)

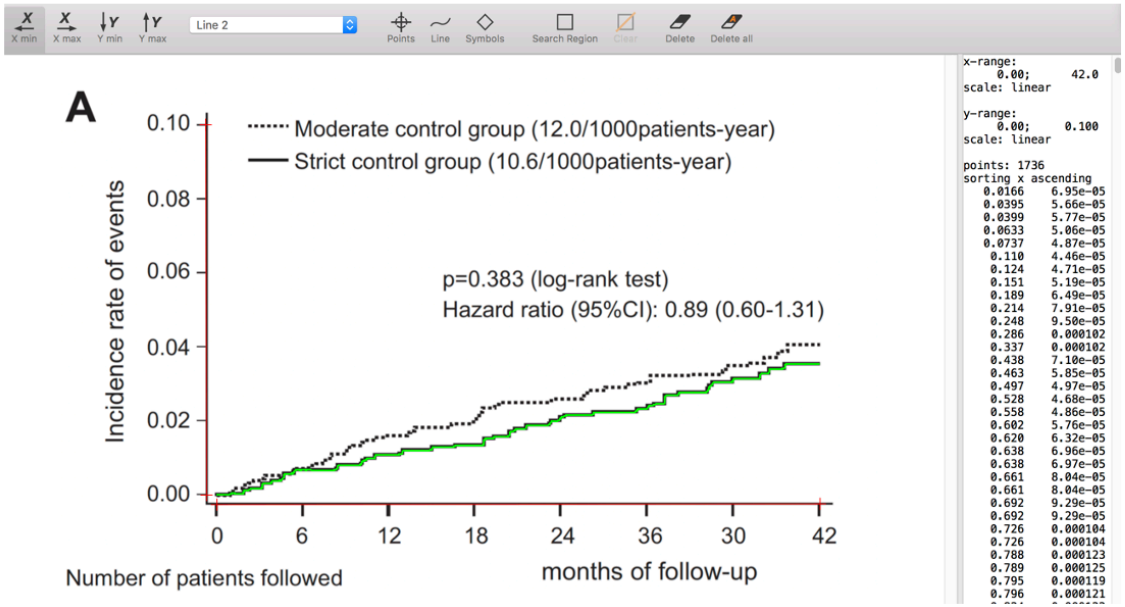
Example (Perren et al. 2011, NEJM)

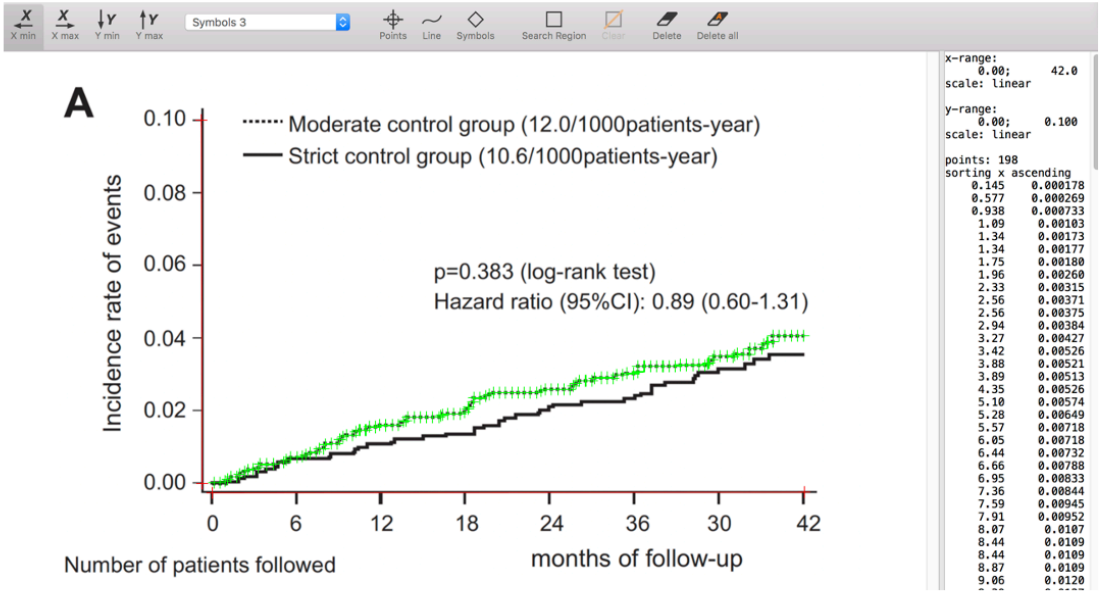


Example extraction process (1)

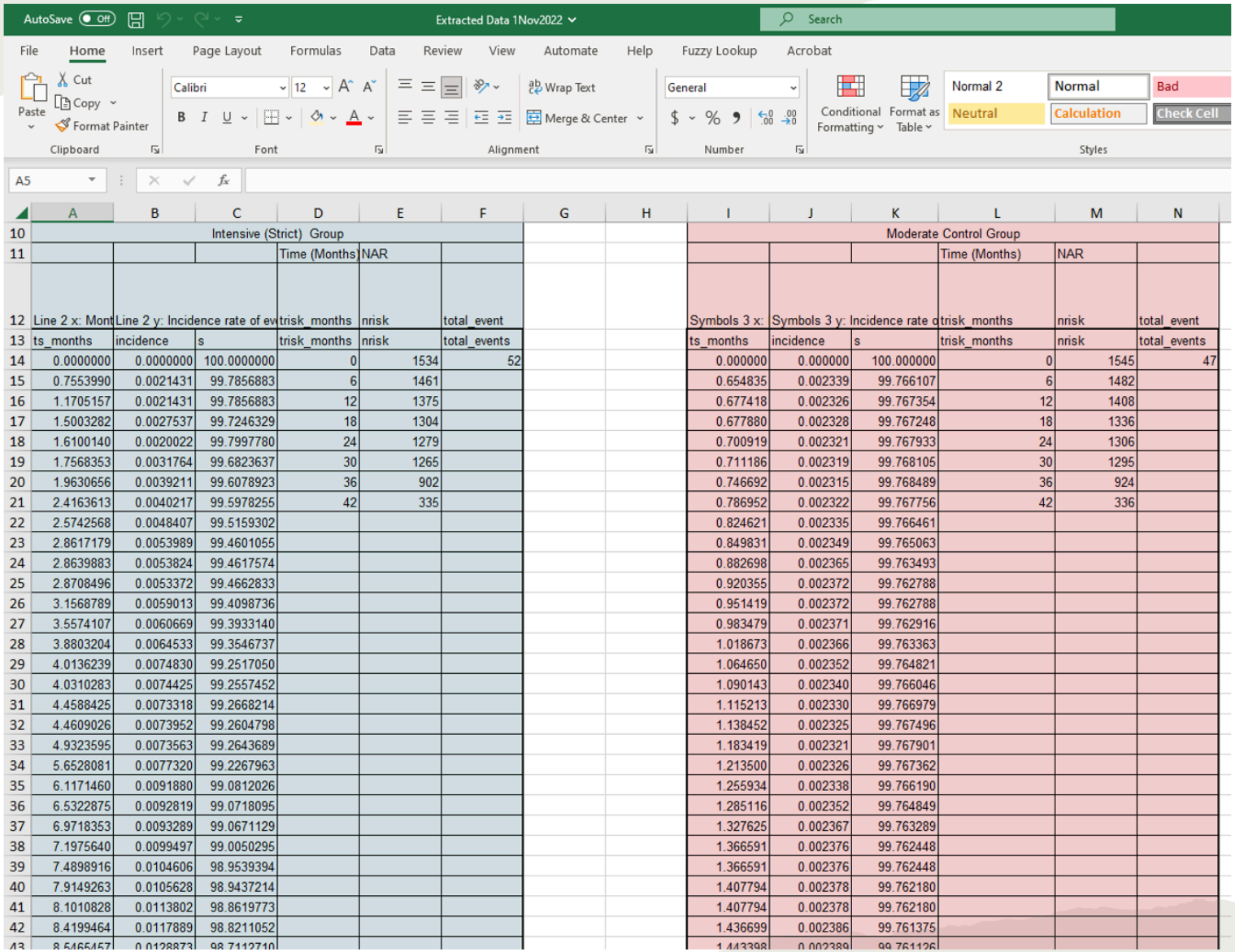


Extraction in process (2)





Result of extraction (3)



AutoSave Off | Extracted Data 1Nov2022 | Search

File Home Insert Page Layout Formulas Data Review View Automate Help Fuzzy Lookup Acrobat

Cut Copy Format Painter | Paste | Clipboard | Calibri 12 | Font | Wrap Text Merge & Center | Alignment | General | Number | Conditional Formatting | Format as Table | Normal 2 Normal Bad Neutral Calculation Check Cell | Styles

A5

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
10	Intensive (Strict) Group								Moderate Control Group					
11				Time (Months)	NAR							Time (Months)	NAR	
12	Line 2 x: Mont	Line 2 y: Incidence rate of ev		trisk_months	nrisk	total_event			Symbols 3 x:	Symbols 3 y: Incidence rate o		trisk_months	nrisk	total_event
13	ts_months	incidence	s	trisk_months	nrisk	total_events			ts_months	incidence	s	trisk_months	nrisk	total_events
14	0.0000000	0.0000000	100.0000000	0	1534	52			0.000000	0.000000	100.000000	0	1545	47
15	0.7553990	0.0021431	99.7856883	6	1461				0.654835	0.002339	99.766107	6	1482	
16	1.1705157	0.0021431	99.7856883	12	1375				0.677418	0.002326	99.767354	12	1408	
17	1.5003282	0.0027537	99.7246329	18	1304				0.677880	0.002328	99.767248	18	1336	
18	1.6100140	0.0020022	99.7997780	24	1279				0.700919	0.002321	99.767933	24	1306	
19	1.7568353	0.0031764	99.6823637	30	1265				0.711186	0.002319	99.768105	30	1295	
20	1.9630656	0.0039211	99.6078923	36	902				0.746692	0.002315	99.768489	36	924	
21	2.4163613	0.0040217	99.5978255	42	335				0.786952	0.002322	99.767756	42	336	
22	2.5742568	0.0048407	99.5159302						0.824621	0.002335	99.766461			
23	2.8617179	0.0053989	99.4601055						0.849831	0.002349	99.765063			
24	2.8639883	0.0053824	99.4617574						0.882698	0.002365	99.763493			
25	2.8708496	0.0053372	99.4662833						0.920355	0.002372	99.762788			
26	3.1568789	0.0059013	99.4098736						0.951419	0.002372	99.762788			
27	3.5574107	0.0060669	99.3933140						0.983479	0.002371	99.762916			
28	3.8803204	0.0064533	99.3546737						1.018673	0.002366	99.763363			
29	4.0136239	0.0074830	99.2517050						1.064650	0.002352	99.764821			
30	4.0310283	0.0074425	99.2557452						1.090143	0.002340	99.766046			
31	4.4588425	0.0073318	99.2668214						1.115213	0.002330	99.766979			
32	4.4609026	0.0073952	99.2604798						1.138452	0.002325	99.767496			
33	4.9323595	0.0073563	99.2643689						1.183419	0.002321	99.767901			
34	5.6528081	0.0077320	99.2267963						1.213500	0.002326	99.767362			
35	6.1171460	0.0091880	99.0812026						1.255934	0.002338	99.766190			
36	6.5322875	0.0092819	99.0718095						1.285116	0.002352	99.764849			
37	6.9718353	0.0093289	99.0671129						1.327625	0.002367	99.763289			
38	7.1975640	0.0099497	99.0050295						1.366591	0.002376	99.762448			
39	7.4898916	0.0104606	98.9539394						1.366591	0.002376	99.762448			
40	7.9149263	0.0105628	98.9437214						1.407794	0.002378	99.762180			
41	8.1010828	0.0113802	98.8619773						1.407794	0.002378	99.762180			
42	8.4199464	0.0117889	98.8211052						1.436699	0.002386	99.761375			
43	8.5465457	0.0128873	98.7112710						1.443398	0.002389	99.761126			

Second step: infer the individual patient data

- `ipdfc` package (Wei and Royston, 2017)
- Start with the sheet created by digital extraction (one line per step in the KM curve)
- Convert to one line per patient data with a time variable and event indicator (event vs. censored)
- Many options to improve fidelity of reconstruction

Working with ipdfc in Stata

Example 2: ICON7 trial

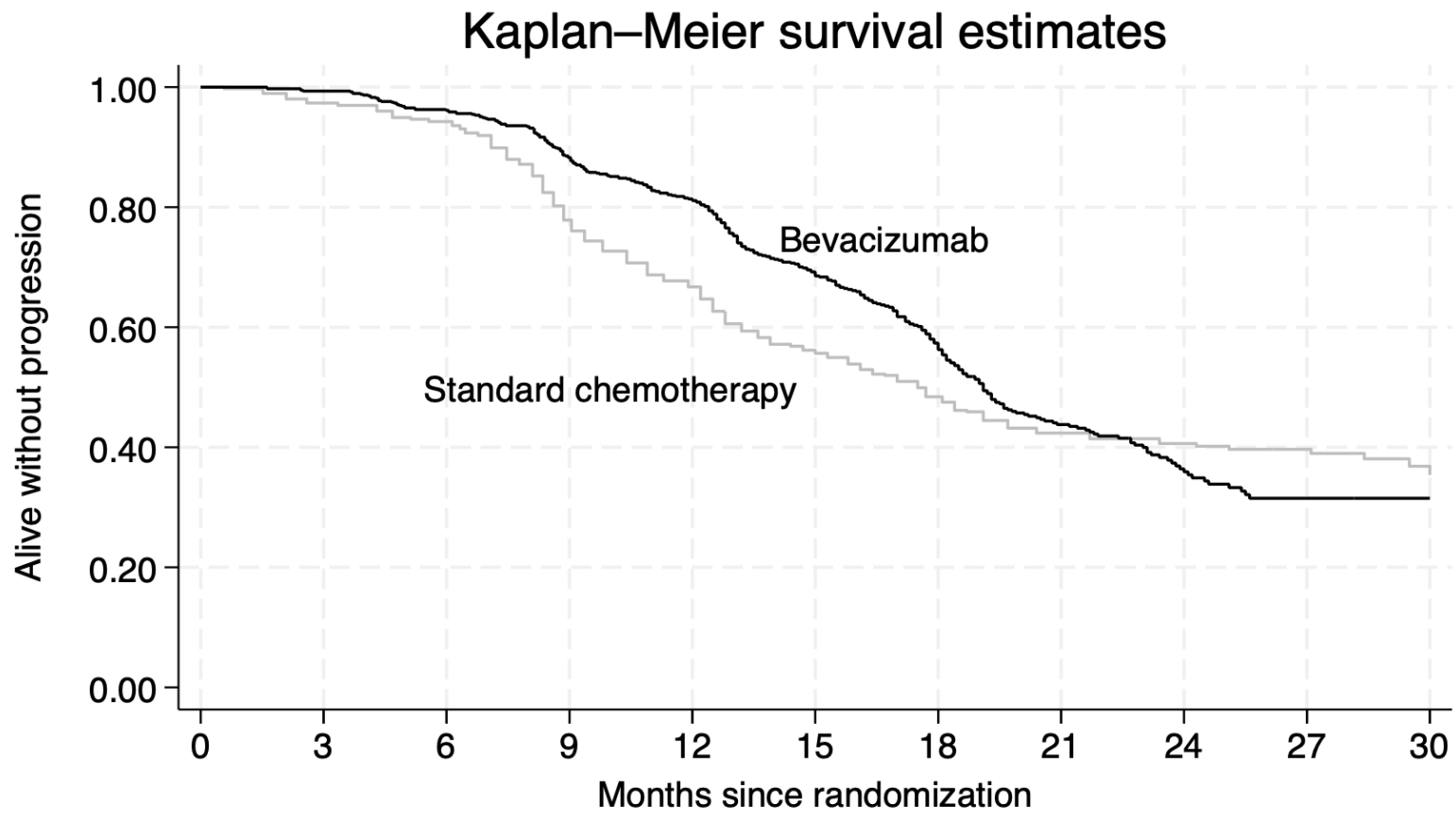
This example is from ICON7, a two-arm randomized controlled trial of bevacizumab in advanced ovarian cancer (Perren et a
probabilities instead of percentages were extracted across 30 months of follow up. The following code shows how to use
extracted survival probabilities to time-to-event data.

```
. local tot0=464
. local tot1=470
. import delimited using "ICON7_data_arm0.txt", clear
. ipdfc, surv(s) tstart(ts) trisk(trisk) nrisk(nrisk) generate(t_ipd event_ipd) saving(temp0, replace) probability i
    totevents(`tot0')
. import delimited using "ICON7_data_arm1.txt", clear
. ipdfc, surv(s) tstart(ts) trisk(trisk) nrisk(nrisk) generate(t_ipd event_ipd) saving(temp1, replace) probability i
    totevents(`tot1')
```

The following code amalgamates the data from both arms and then conducts survival analysis.

```
. use temp0, clear
. generate byte arm = 0
. append using temp1
. replace arm = 1 if missing(arm)
. stset t_ipd, failure(event_ipd)
. stcox arm
. sts graph, by(arm) xlabel(0(3)30) ylabel(0(0.2)1) risktable(0(6)30, order(1 "Bevacizumab" 2 "Standard chemo-")) le
    xtitle("Months since randomization") l2title("Alive without progression") plot1opts(lpattern(solid) lcolor(gs12)
    plot2opts(lpattern(solid) lcolor(black)) text(-0.38 -3.2 "therapy") text(0.75 14 "Bevacizumab", place(e)) text(0
    chemotherapy")
```

The reconstructed KM curves for the inferred data



Number at risk						
Bevacizumab	764	694	465	216	94	25
Standard chemotherapy	764	715	585	263	73	19

Summary of reconstructing IPD

- Clinical trials often publish Kaplan-Meier curves for each arm and hazard ratio with 95% CI
- Use specialized software to digitally extract the underlying coordinates of the Kaplan-Meier curves
- Run `ipdfc` to create a one line per participant version of the original survival data
- Why go to this trouble? Lots of things we can do with these data (e.g fit our own survival models, calculate other metrics besides hazard ratio)

Metrics of interest

- Hazard ratio is useful for comparing survival curves, but there are other quantities of interest
- Difference in Restricted Mean Survival Time (RMST; Royston & Parmar, 2013) is popular with statisticians and is clinically interpretable
- Time to Benefit (TTB) less well known but extremely appealing to clinical researchers to weigh risks and benefits (Lee, 2013)

Restricted Mean Survival Time (RMST)

- For one arm, $RMST(t) = \int_0^t S(u)du$ is the area under the survival curve out to some given time t
- Difference in RMST, $dRMST(t) = RMST^1(t) - RMST^0(t)$, is the area between the survival curves out to that time
- Interpreted as average gain in life from intervention over a t-year period
- Can compute this using non-parametric Kaplan Meier curves or by fitting a parametric model
- We use parametric models (e.g. Weibull or Gompertz) for simplicity/stability of estimation
- Bayesian estimation of the parametric survival curves makes computation of both estimate and CI straightforward

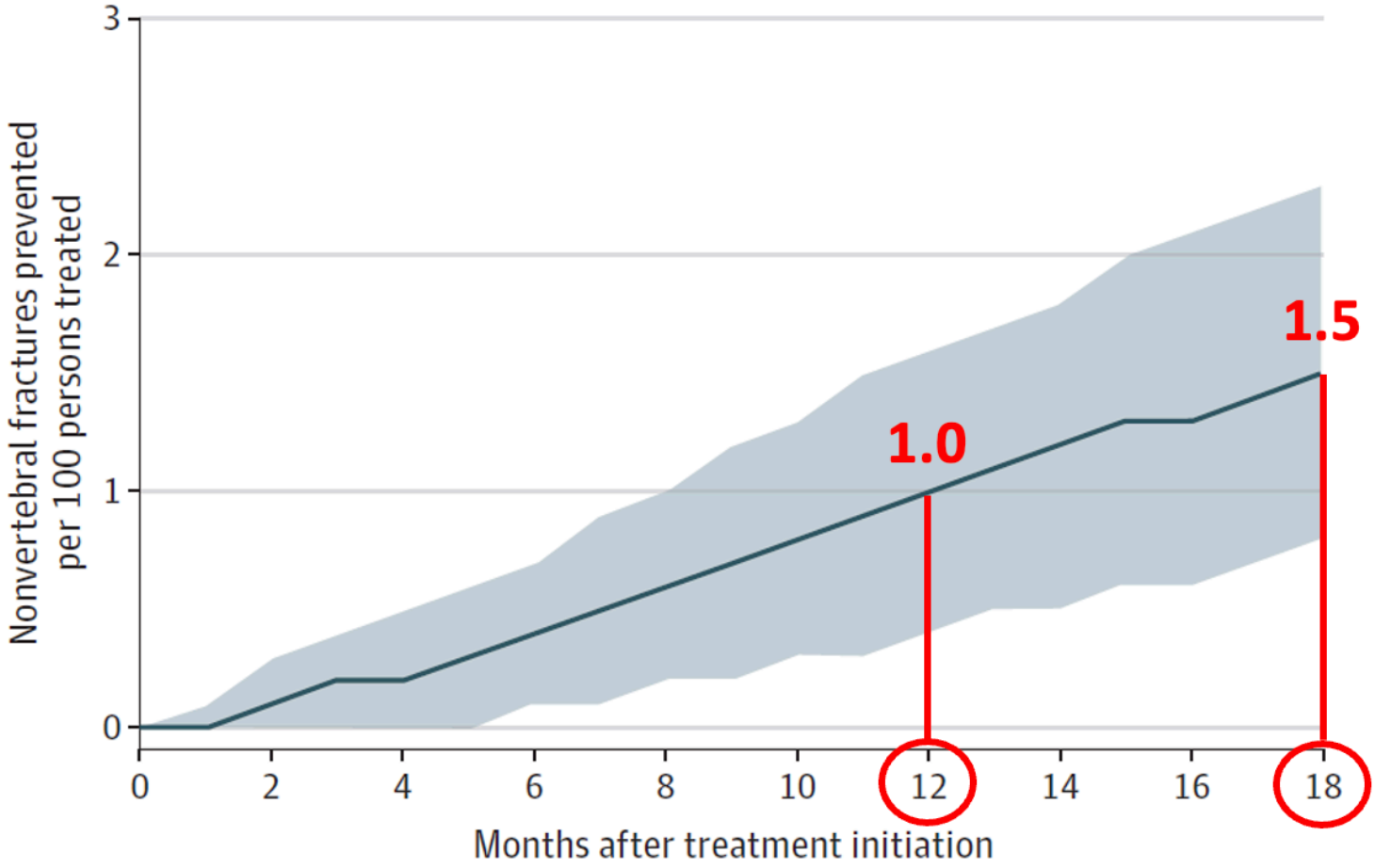
Time to Benefit

- $TTB(r)$ is the amount of time until the survival curves are separated by an absolute amount of risk r; $TTB(r)$ = smallest t such that $S^1(t) - S^0(t) \geq r$
- Suppose survival curves are separated by $r = 0.01$ at 3 years
- Then the number needed to treat (NNT) to save one life with the intervention is 100 patients after 3 years
- Can compare this to life expectancy of patient to aid in decision-making
- And/or can contrast with the expected number out of 100 that will be harmed over 3 years (NNH)
- This framework is very natural for clinicians
- We again use parametric models and Bayesian estimation to make computation straightforward (so can do both $TTB(r)$ and $dRMST(t)$ for same price!)

Time to Benefit examples in literature

- Statins for primary prevention of ASCVD (Yourman 2021): 30 months needed to avoid 1 MACE for 100 patients ($r = 0.01$)
- Intensive blood pressure treatment (Chen 2022): 19.1 months needed to avoid 1 MACE per 200 patients ($r = 0.005$)
- Mammography for breast cancer (Lee 2013): 10.7 years needed to avoid 1 breast cancer death per 1000 women screened ($r = 0.001$)
- Bisphosphonates in osteoporosis (Deardorff 2020): 12.4 months to prevent 1 fracture in 100 women treated ($r = 0.01$)

TTB figure (Deardorff, 2020)



Bayesian analysis of TTB, RMST, etc. in a single study

- Weibull or Gompertz provide excellent fit in our settings
- We allow both parameters (shape and scale for Weibull) to be different for two arms of study (4 parameters)
- We start with the data set reconstructed from `ipdfc` and then use `bayes:streg` to generate large number of MCMC realizations from the posterior distribution of the four parameters
- For each realization, we can create the Weibull survival curve given those 4 parameters. $RMST(t)$ and $TTB(r)$ then numerically evaluated
- Use posterior quantiles across the set of realizations to get estimate and CI for $RMST(t)$ and $TTB(r)$

Bayesian $TTB(r)$ in more detail

- Have simulations $\theta_1, \ldots, \theta_M$ from the posterior distribution of the survival curve parameters
- For each simulated parameter vector θ_m, create the survival curves and solve for the first time they are more than r apart ($TTB(r)_m$)
- Take $\widehat{TTB(r)}$ as median of $TTB(r)_1, \ldots, TTB(r)_M$
- Take the 2.5th and 97.5th percentiles as a 95% credible interval

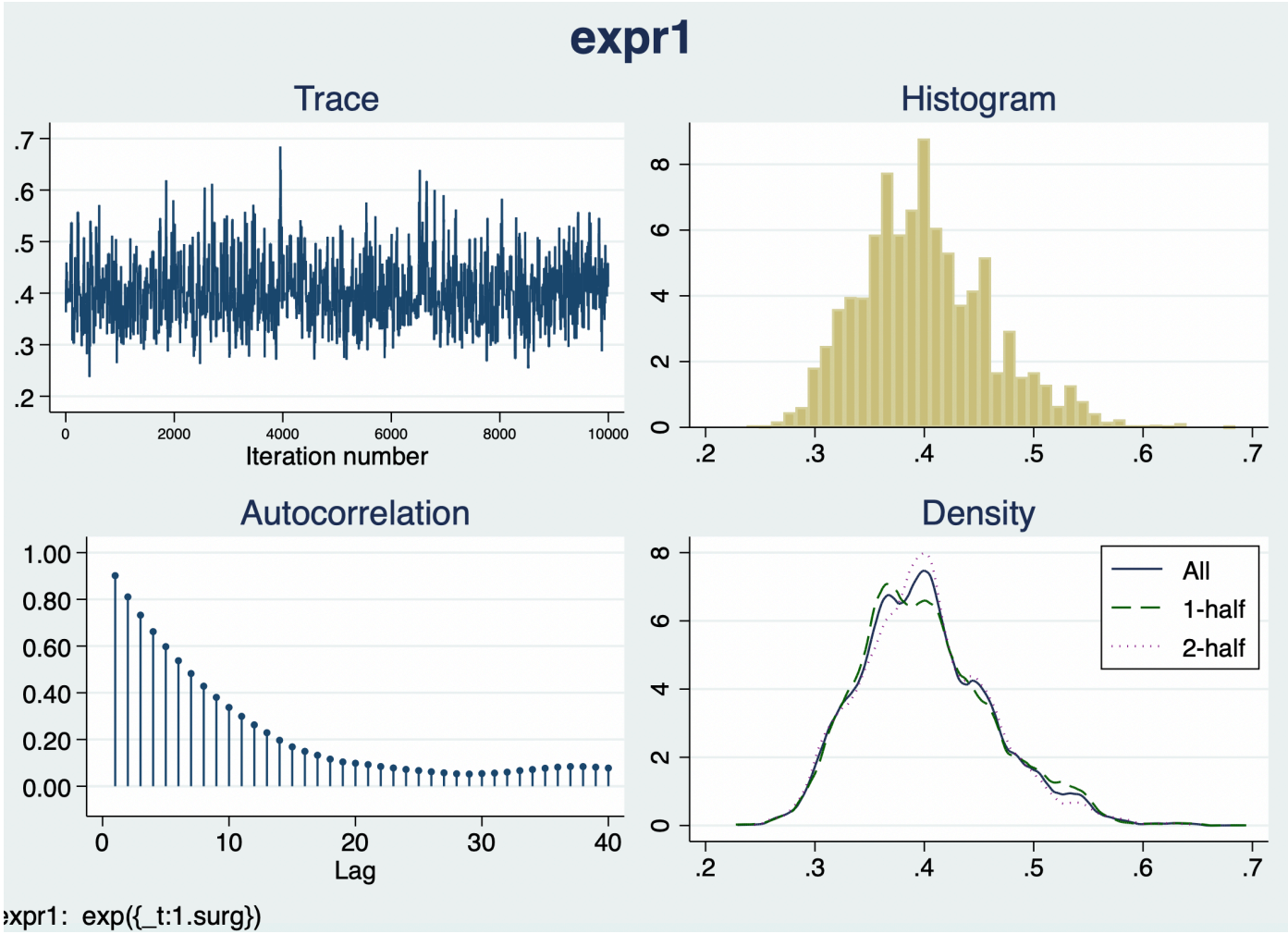
Bayesian $dRMST(t)$ in more detail

- Have simulations $\theta_1, \ldots, \theta_M$ from the posterior distribution of the survival curve parameters
- For each simulated parameter vector θ_m, create survival curves and take difference when numerically integrate them from 0 to t ($dRMSTt_m$)
- Take $\widehat{dRMST}(t)$ as median of the M values
- Take the 2.5th and 97.5th percentiles as a 95% credible interval (or other methods)

TTB: fitting Weibull model

```
bayes: streg if trtgrp==0 , dist(weibull) nohr

Model summary
------------------------------------------------------------------------------
Likelihood:
  _t ~ streg_weibull({_t:_cons},{ln_p})

Priors:
  {_t:_cons} ~ normal(0,10000)
      {ln_p} ~ normal(0,10000)
------------------------------------------------------------------------------

Bayesian Weibull PH regression                  MCMC iterations  =     12,500
Random-walk Metropolis-Hastings sampling        Burn-in          =      2,500
                                                MCMC sample size =     10,000
No. of subjects =       4243                    Number of obs    =      4,243
No. of failures =        146
Time at risk    =165789.8329238
                                                Acceptance rate  =      .4285
                                                Efficiency:  min =    .006646
                                                             avg =    .006841
Log marginal-likelihood = -788.94239                         max =    .007036

------------------------------------------------------------------------------
             |                                                 Equal-tailed
             |      Mean   Std. dev.     MCSE     Median  [95% cred. interval]
-------------+----------------------------------------------------------------
_t           |
       _cons | -7.129357    .333975   .040966  -7.112837  -7.851203  -6.508836
-------------+----------------------------------------------------------------
        ln_p |  .0208444   .0847516   .010104   .0202927  -.1411043   .1896877
------------------------------------------------------------------------------
Note: Default priors are used for model parameters.
```

TTB: plotting results from basic Weibull



Calculate TTB from posterior simulations

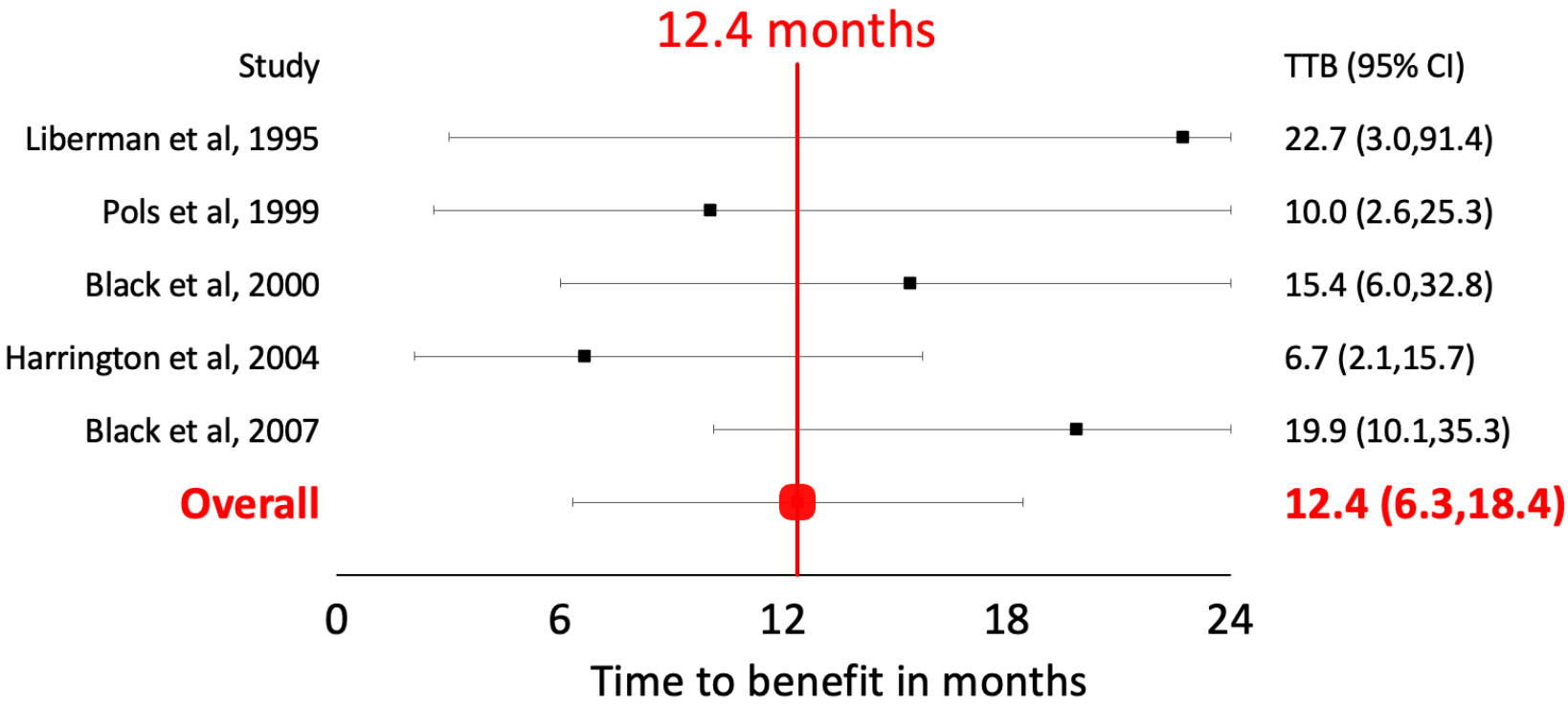
```
/* GENERATE SURVIVAL USING THE RANDOM SAMPLES. */
forv t=1/120 {
    generate surv_c`t' = exp(-(exp(b0_control)) * `t'^(exp(lnp_control)))
}
forv t=1/120 {
    generate surv_t`t' = exp(-(exp(b0_treatment)) * `t'^(exp(lnp_treatment)))
}
forv t=1/120 {
    generate surv_d`t' = surv_t`t'-surv_c`t'
}

/* ESTIMATE LTTBs. */

/* Find the first time difference is bigger than 0.005*/
generate lttb_005=84
forvalues t = 84(-1)1 {
    replace lttb_005=`t' if surv_d`t'> 0.005
}
count if lttb_005==84
_pctile lttb_005, percentiles(2.5 25 50 75 97.5)
return list
```

Two ideas for meta-analyses of TTB, RMST, etc. across multiple studies

1. Calculate estimate and CI of TTB (or dRMST) for each study, then meta-analyze with usual random-effects `meta` command
   - Pros: straightforward to explain given similarity to how one would typically do meta-analysis for hazard ratios
   - Cons: does not easily handle curves that do not separate out in time range of data
2. Use hierarchical model for the underlying Weibull parameters (Ouwens 2010). This implies a (meta-analyzed) survival curve in each group. Can calculate estimate and CI for TTB for this pair of meta-analyzed survival curves in same way as was done for single curve
   - Pros: Easily accomodates "null" studies that have arbitrarily long TTB
   - Cons: Need to program using `bayesmh` so not quite so easy to implement

TTB (1): forest plot (Deardorff et al. 2020)

TTB for non-vertebral fracture prevention (ARR=0.01)



TTB (1): summary of results (Deardorff et al. 2020)

Table 2. Time to Benefit of Bisphosphonate Therapy for the Prevention of Nonvertebral Fractures Among Postmenopausal Women With Osteoporosis

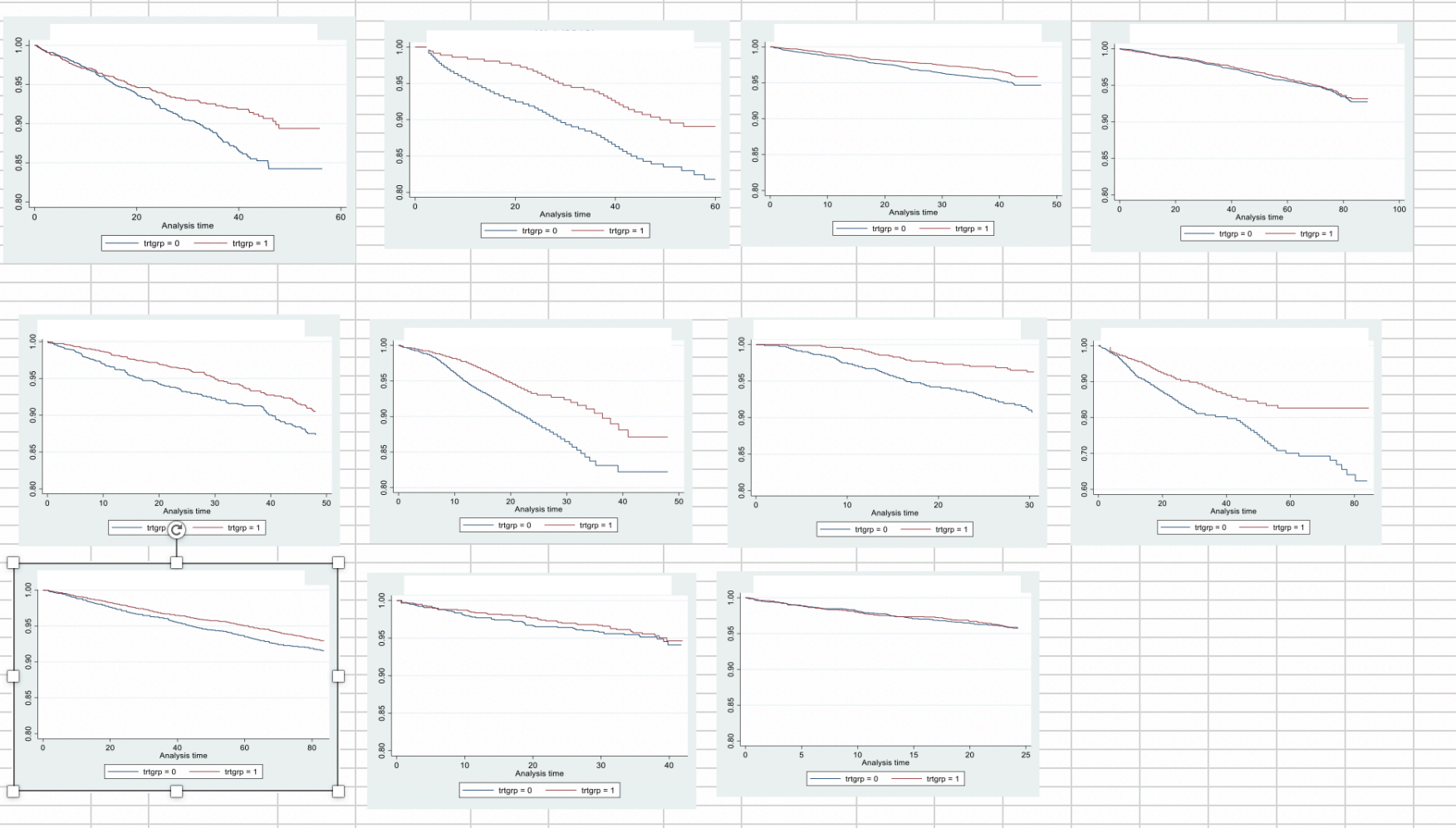
Source	Bisphosphonate type	Time to benefit (95% CI), mo		
		ARR = 0.002[a]	ARR = 0.005[b]	ARR = 0.010[c]
Liberman et al,[33] 1995	Alendronate	12.5 (0.4-77.6)	16.6 (1.1-88.3)	22.7 (3.0-91.4)
Pols et al,[34] 1999	Alendronate	3.4 (0.6-10.6)	5.9 (1.3-16.0)	10.0 (2.6-25.3)
Black et al,[43] 2000	Alendronate	6.9 (1.1-24.0)	10.3 (2.9-26.9)	15.4 (6.0-32.8)
Harrington et al,[44] 2004	Risedronate	1.9 (0.5-4.5)	3.5 (1.0-9.0)	6.7 (2.1-15.7)
Black et al,[42] 2007	Zoledronic acid	7.6 (2.0-20.6)	12.5 (5.0-26.3)	19.9 (10.1-35.3)
Summary time to benefit	NA	3.3 (0.2-6.5)	6.5 (2.2-10.9)	12.4 (6.3-18.4)
Test of heterogeneity				
I^2, %	NA	0	0	0
P value	NA	.70	.56	.49

TTB (2): fully Bayesian hierarchical model

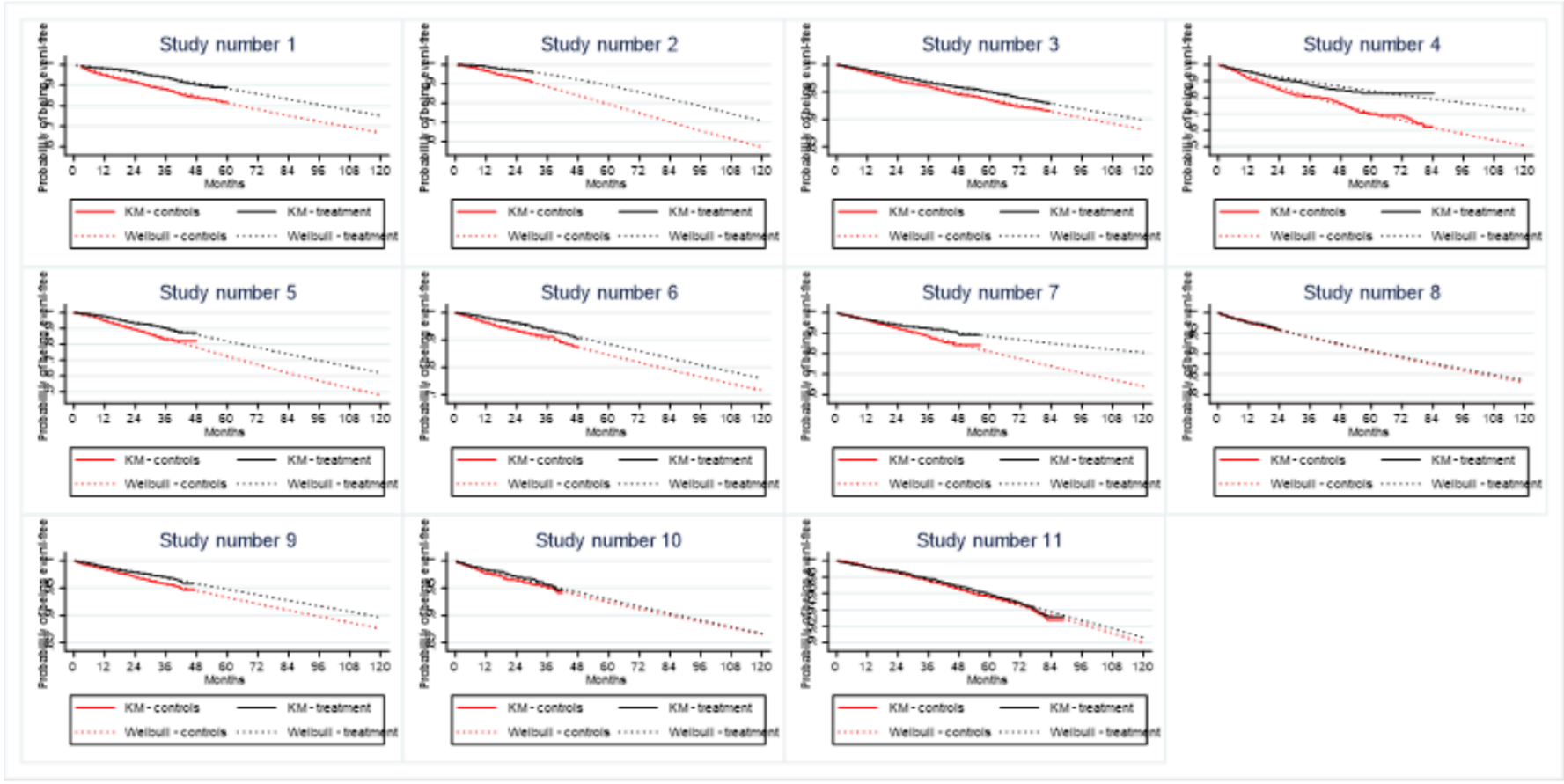
$$
\begin{aligned}
\text{Data for study i, arm k} &\sim \text{Weibull regression}(\beta_i^{(k)}, p_i^{(k)}) \\
(\beta_i^{(0)}, \beta_i^{(1)}, \log p_i^{(0)}, \log p_i^{(1)}) &\sim \text{N}((\beta^{(0)}, \beta^{(1)}, \log p^{(0)}, \log p^{(1)}), \Sigma) \\
p(\beta^{(0)}, \beta^{(1)}, \log p^{(0)}, \log p^{(1)}, \Sigma) &\propto \text{InverseWishart}(\Sigma|\Lambda, \nu)
\end{aligned}
$$

Fit with `bayesmh` random effects formulation. The two Weibull survival curves with parameters $(\beta^{(0)}, p^{(0)})$ and $(\beta^{(1)}, p^{(1)})$ are thought of as the underlying survival curves for the control and treatment arms

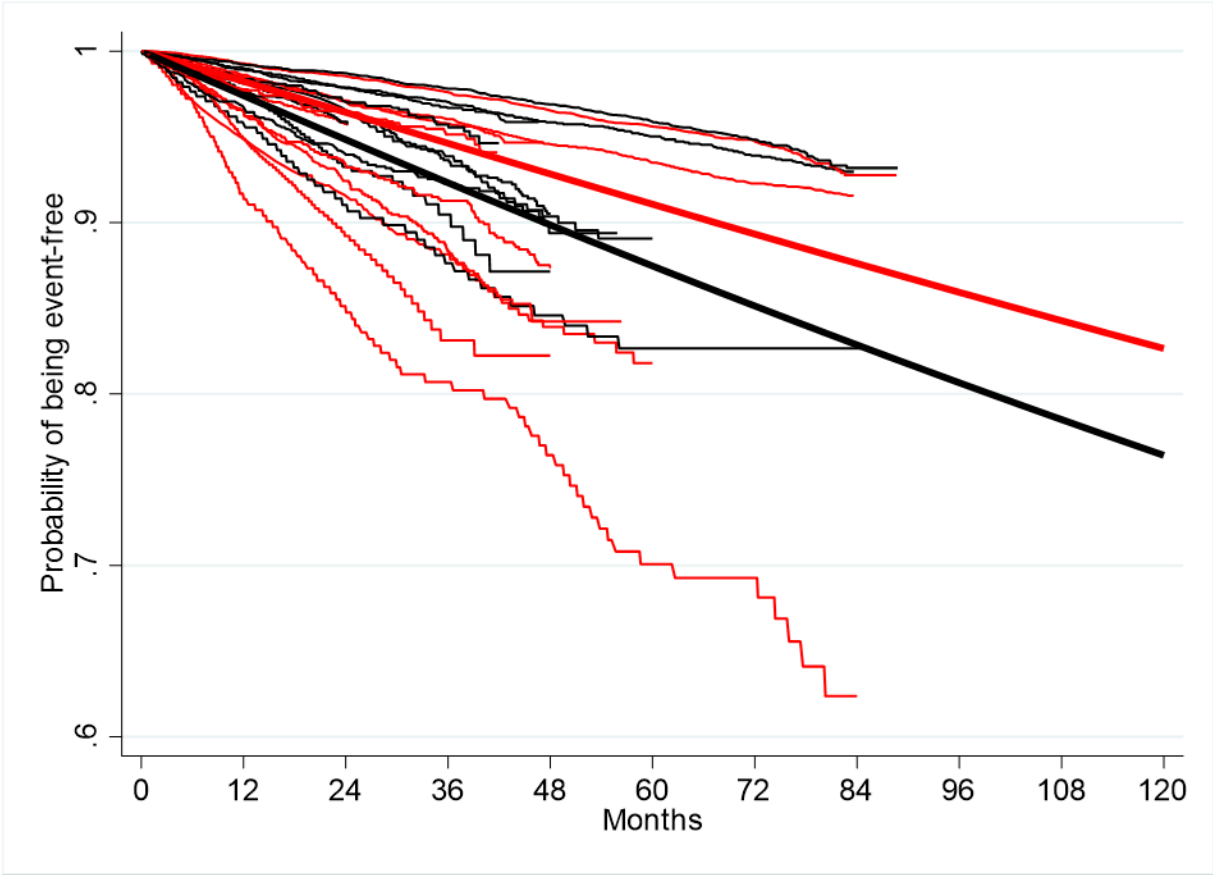
TTB (2): raw figures (Growdon et al., 2023)



TTB (2): curve fitting (Growdon et al., 2023)



TTB (2): meta-analyzed curve (Growdon et al., 2023)



TTB (2): results (Growdon et al., 2023)

TTB: Intensive Antihypertensive Therapy for Prevention of MACE

		Median time to benefit (IQR[a]), mo		
Study	**Trial length, mo**	**ARR = 0.002[b]**	**ARR = 0.005[c]**	**ARR=0.01[d]**
1	84	2 (1,3)	3 (2,5)	5 (3,7)
2	48	2 (2,3)	4 (3,5)	7 (5,9)
3	84	6 (4,11)	18 (11,31)	46 (30,>84)
4	48	2 (1,2)	4 (3,5)	8 (6,10)
5	24	24 (8,>24)	>24 (>24,>24)	>24 (>24,>24)
6	42	4 (2,22)	21 (8,>42)	>42 (29,>42)
7	60	1 (1,2)	2 (2,3)	4 (3,6)
8	84	58 (33,>84)	>84 (68,>84)	>84 (>84,>84)
9	54	11 (7,14)	13 (10,17)	18 (14,22)
10	48	4 (3,6)	12 (9,17)	32 (24,>48)
Summary time to benefit		3 (2,10)	8 (5,37)	16 (9,>84)

Summary

- Use external digitization software and `ipdfc` to turn published Kaplan-Meier curves from two arm trials into Stata datasets
- Analyze these data in Bayesian framework using the `bayes` commands (i.e. create MCMC realizations of underlying parameters)
- Use the simulated parameter realizations for inference on less traditional metrics such as $TTB(r)$ and $dRMST(t)$
- Can do this for single studies or in meta-analysis of multiple studies
- Work in progress but let us know if you are interested!

References

- Chen T, Shao F, Chen K, Wang Y, Wu Z, Wang Y, Gao Y, Cornelius V, Li C, Jiang Z (2022). Time to clinical benefit of intensive blood pressure lowering in patients 60 years and older with hypertension. *JAMA Int Med*
- Deardorff WJ, Cenzer I, Nguyen B, Lee SJ (2020). Time to benefit of bisphosphonate therapy for the prevention of fractures among postmenopausal women with osteoporosis: a meta-analysis. *JAMA Int Med*
- Earle CC, Wells GA (2002). An assessment of methods to combine published survival curves. *Med Dec Making*.
- Gross A, Schirm S, Scholz M (2013). Ycasd – a tool for capturing and scaling data from graphical representations. *BMC Bioinf*
- Growden ME, Cenzer IS, Xie L, Kang J, Boscardin WJ, Lee SJ (2023). Time to benefit for prevention of MACE after blood pressure treatment. *In preparation*
- Guyot P, Ades AE, Ouwens MJ, Welton NJ (2012). Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol*
- Lee SJ, Boscardin WJ, Cenzer IS, Conell-Price J, O'Brien S, Walter LC (2013). Time lag to benefit after screening for breast and colorectal cancer: meta-analysis of survival data. *BMJ*
- Liu Z, Rich B, Hanley JA (2015). Recovering the raw data behind a non-parametric survival curve. *Systematic Reviews*

References (continued)

- Ouwens MJ, Philips Z, Jansen JP (2010). Network meta-analysis of parametric survival curves. *Res synth meth*
- Parmar M, Torri V, Stewart L (1998). Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Stat Med*
- Perren TJ, Swart AM, Pfisterer J, et al. (2011). A phase 3 trial of bevacizumab in ovarian cancer. *N Engl J Med*
- Royston, Parmar (2013). Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC Med Res Methodol*
- Tang V, Boscardin WJ, Stijacic-Cenzer I, Lee SJ (2015). Time to benefit for colorectal cancer screening: survival meta-analysis of flexible sigmoidoscopy trials. *BMJ*
- Vittinghoff E, Glidden DV, Shiboski SC, McCulloch CE (2012). Regression methods in biostatistics: linear, logistic, survival, and repeated measures models. *Springer*
- Wei Y, Royston P (2017). Reconstructing time-to-event data from published Kaplan-Meier curves. *Stata J*
- Yourman LC, Cenzer IS, Boscardin WJ, Nguyen BT, Smith AK, Schonberg MA, Schoenborn NL, Widera EW, Orkaby A, Rodriguez A, Lee SJ (2021). Evaluation of time to benefit of statins. *JAMA Int Med*