

Dynamic document corresponding to 2021 Stata Conference presentation

Chappell LJ. Use of the bayesmh command in Stata to calculate Excess Relative and Absolute Risk for Radiation-Induced Health Estimates. Presented at the 2021 Stata Conference, Virtual, August 5-6, 2021

Data for example model fitting

Edmondson EF, Gatti DM, Ray FA, Garcia EL, Fallgren CM, Kamstock DA, et al. Genomic mapping in outbred mice reveals overlap in genetic susceptibility for HZE ion- and γ -ray-induced tumors. Science Advances. 2020 Apr 1;6(16):eaax5940.

The data is publicly available. This presentation only uses the phenotype data. The data have previously been imported into Stata and prepared as a .dta file for analyses.

```
. use outbred_all, replace

. describe

Contains data from outbred_all.dta
  obs:      1,804
  vars:      31
  size:    184,008
  1 Apr 2021 09:41
```

variable name	storage type	display format	value label	variable label
AnimalNumber	int	%10.0g		Animal number
group1	float	%12.0g	group1	
group2	float	%12.0g	group2	
agedays	int	%10.0g		age (days)
SolidTumor	byte	%10.0g		Solid Tumor
sex_str	str1	%9s		sex
Weight	double	%10.0g		Weight
coatcolor_str	str17	%17s		coat color
family	byte	%10.0g		family
group_str	str15	%15s		group
neoplasia	byte	%10.0g		neoplasia
nonneoplastic	byte	%10.0g		non-neoplastic
euthanized_str	str12	%12s		euthanized
MyeloidLeukemia	byte	%10.0g		Myeloid Leukemia
Lymphoma	byte	%10.0g		Lymphoma
PulmonaryAden~a	byte	%10.0g		Pulmonary Adenocarcinoma
GranulosaCell~r	byte	%10.0g		Granulosa Cell Tumor
HarderianTumor	byte	%10.0g		Harderian Tumor
HCC	byte	%10.0g		HCC
Mammarycarcin~a	byte	%10.0g		Mammary carcinoma
SoftTissueSar~s	byte	%10.0g		Soft Tissue Sarcomas
Osteosarcoma	byte	%10.0g		Osteosarcoma
PituitaryAden~a	byte	%10.0g		Pituitary Adenoma
ThyroidTumor	byte	%10.0g		Thyroid Tumor
BcellLymphoma	byte	%10.0g		B-cell Lymphoma
PreTLymphoma	byte	%10.0g		PreT Lymphoma
sex	long	%8.0g	sex	sex
coatcolor	long	%11.0g	coatcolor	coat color
group0	long	%15.0g	group0	group
euthanized	long	%12.0g	euthanized	euthanized
exposure_age_~s	float	%9.0g		

```
Sorted by:

. generate exposure_age_months = exposure_age_days/30

. tab group1
```

group1	Freq.	Percent	Cum.
unirradiated	601	33.31	33.31
gamma	596	33.04	66.35
HZE-Fe	307	17.02	83.37
HZE-Si	300	16.63	100.00
Total	1,804	100.00	

```

. * Only use the control and gamma irradiated data
. drop if group1 > 1
(607 observations deleted)

. * Define the radiation information for the experiment
. generate dose_Gy = 3

. replace dose_Gy = 0 if group1 == 0
(601 real changes made)

. label variable dose_Gy "Dose (Gy)"

. stset agedays, failure(SolidTumor) id(Animalnumber) scale(30)

      id: Animalnumber
failure event: SolidTumor != 0 & SolidTumor < .
obs. time interval: (agedays[_n-1], agedays]
exit on or before: failure
t for analysis: time/30

```

```

-----
1,197 total observations
    0 exclusions

```

```

-----
1,197 observations remaining, representing
1,197 subjects
    516 failures in single-failure-per-subject data
23,729.2 total analysis time at risk and under observation
              at risk from t =          0
              earliest observed entry t =          0
              last observed exit t =        27.2

```

Split the age variable into categories

```
. stsplot agegroup, at(14(2)25)
(4,167 observations (episodes) created)

. *calculate rates by age group, sex, and dose
. strate agegroup dose_Gy sex, per(100) graph output(ratesbyageYgroup2, replace)

      failure _d: SolidTumor
analysis time _t: agedays/30
      id: Animalnumber
```

Estimated rates (per 100) and lower/upper bounds of 95% confidence intervals
(5364 records included in the analysis)

	agegroup	dose_Gy	sex	D	Y	Rate	Lower	Upper
301.	0	0	M	6	41.4893	0.144615	0.064970	0.321897
601.	0	0	F	4	41.7107	0.095899	0.035993	0.255513
903.	0	3	M	17	39.7590	0.427576	0.265807	0.687797
1197.	0	3	F	13	38.4580	0.338031	0.196280	0.582153
1476.	14	0	M	4	5.4753	0.730549	0.274188	1.946481
1765.	14	0	F	1	5.6787	0.176098	0.024806	1.250130
1995.	14	3	M	16	4.2207	3.790870	2.322408	6.187843
2216.	14	3	F	12	4.1813	2.869898	1.629843	5.053441
2481.	16	0	M	9	5.0963	1.765976	0.918863	3.394052
2759.	16	0	F	2	5.3753	0.372070	0.093054	1.487699
2948.	16	3	M	25	3.2497	7.693097	5.198295	1.1e+01
3143.	16	3	F	17	3.5433	4.797742	2.982567	7.717624
3385.	18	0	M	23	4.4127	5.212268	3.463689	7.843583
3645.	18	0	F	10	4.9597	2.016265	1.084861	3.747323
3784.	18	3	M	18	2.4410	7.374027	4.645951	1.2e+01
3941.	18	3	F	39	2.5003	1.6e+01	1.1e+01	2.1e+01
4135.	20	0	M	26	3.4750	7.482014	5.094296	1.1e+01
4365.	20	0	F	13	4.2637	3.049019	1.770432	5.250987
4471.	20	3	M	17	1.8830	9.028147	5.612442	1.5e+01
4568.	20	3	F	25	1.5817	1.6e+01	1.1e+01	2.3e+01
4721.	22	0	M	36	2.5090	1.4e+01	1.0e+01	2.0e+01
4913.	22	0	F	20	3.4673	5.768122	3.721345	8.940645
4992.	22	3	M	34	1.1037	3.1e+01	2.2e+01	4.3e+01
5050.	22	3	F	19	0.8573	2.2e+01	1.4e+01	3.5e+01
5151.	24	0	M	41	1.7317	2.4e+01	1.7e+01	3.2e+01
5296.	24	0	F	38	2.7070	1.4e+01	1.0e+01	1.9e+01
5335.	24	3	M	20	0.6540	3.1e+01	2.0e+01	4.7e+01
5364.	24	3	F	11	0.5073	2.2e+01	1.2e+01	3.9e+01

Fit a restricted cubic spline in log-attained age for the background hazard function

A continuous function can be fit to the background hazard using restricted cubic splines in log-attained age similar to the quadratic splines used to model the Life Span Study of atomic bomb survivors background data (Grant et al. 2017):

Grant EJ, Brenner A, Sugiyama H, Sakata R, Sadakane A, Utada M, et al. Solid Cancer Incidence among the Life Span Study of Atomic Bomb Survivors: 1958–2009. Radiation Research. 2017 Mar 20.

```
. generate time_exposed = _t - _t0

. gen agedmed = agegroup + 1

. replace agedmed = 7 if agegroup == 0
(1,197 real changes made)

. gen lnage = ln(agedmed/20)

. * create a cubic spline with knots at ages 15 20 and 23
. mkspline lnagesp = lnage, cubic knot(-.28768207 0 .13976194)

.
. poisson _d lnagesp* sex if group2 == 0, exposure(time_exposed)

Iteration 0:  log likelihood = -866.7205
Iteration 1:  log likelihood = -862.99851
Iteration 2:  log likelihood = -862.99625
Iteration 3:  log likelihood = -862.99625

Poisson regression              Number of obs   =    3,229
                               LR chi2(3)         =    626.38
                               Prob > chi2        =    0.0000
Log likelihood = -862.99625     Pseudo R2      =    0.2663

-----+-----
      _d |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
 lnagesp1 |  3.077664   .3958627    7.77  0.000   2.301788   3.853541
 lnagesp2 |  2.513981   .4878475    5.15  0.000   1.557817   3.470144
      sex | -.7850166   .1354861   -5.79  0.000  -1.050565  -.5194687
      _cons | -3.37016   .1632716  -20.64  0.000  -3.690166  -3.050154
ln(time_e~d) |           1 (exposure)
-----+-----
```

Bayesian Poisson regression model

```
. set seed 45

. bayesmh (_d, likelihood(dpoisson(exp(ln(time_exposed) + lnagesp1*{theta_1=10} + lnagesp2*{theta_2=-3} + {theta_0=
> -2.5} + sex*{theta_s})*(1 + exp({delta_s}*sex + {delta_e}*(exposure_age_months-2) + {delta_a}*lnage) * {beta_gamm
> a=0.5}*dose_Gy))), prior({theta_0} {theta_1} {theta_2} {theta_s}, normal(0,10000)) prior({beta_gamma}, lognormal
> (0.33,2)) prior({delta_s}, normal(0.59,0.5)) prior({delta_e}, normal(-0.21,0.5)) prior({delta_a}, normal(-1.66,0.
> 5)) block({theta_0} {theta_1} {theta_2} {delta_a}) block({theta_s} {delta_s}) block({beta_gamma}) block({delta_e}
> ) saving(sim_Solid_ERR, replace) thinning(20) burnin(50000)
note: discarding every 19 sample observations; using observations 1,21,41,...
```

```
Burn-in ...
note: invalid initial state
Simulation ...
```

Model summary

```
-----
Likelihood:
_d ~ poisson(<expr1>)
```

```
Priors:
{theta_s} ~ normal(0,10000)
{delta_s} ~ normal(0.59,0.5)
{delta_e} ~ normal(-0.21,0.5)
{delta_a} ~ normal(-1.66,0.5)
```

```
Hyperpriors:
{theta_0 theta_1 theta_2} ~ normal(0,10000)
{beta_gamma} ~ lognormal(0.33,2)
```

```
Expression:
expr1 : exp(ln(time_exposed) + lnagesp1*{theta_1=10} + lnagesp2*{theta_2=-3} + {theta_0=-2.5} +
sex*{theta_s})*(1 + exp({delta_s}*sex + {delta_e}*(exposure_age_months-2) + {delta_a}*lnage) *
{beta_gamma=0.5}*dose_Gy)
```

```
-----
Bayesian Poisson model                MCMC iterations = 249,981
Random-walk Metropolis-Hastings sampling  Burn-in = 50,000
                                           MCMC sample size = 10,000
                                           Number of obs = 5,364
                                           Acceptance rate = .3486
                                           Efficiency: min = .3581
                                           avg = .5394
                                           max = .968
Log marginal likelihood = -1869.5654
```

```
-----
|               |               |               |               |               |               |
|               | Mean   Std. Dev.   MCSE   Median   Equal-tailed
|               |       |             |       |         [95% Cred. Interval]
|-----+-----|-----+-----|-----+-----|-----+-----|
beta_gamma | .3523971 | .0740951 | .001238 | .347643 | .2218528 | .5130183
delta_a   | -1.184415 | .3838861 | .005625 | -1.179685 | -1.943432 | -.4291951
delta_e   | -.125632 | .211929 | .002154 | -.1246816 | -.5465446 | .2803918
delta_s   | 1.048715 | .2600569 | .004006 | 1.045351 | .5465176 | 1.571688
theta_0   | -3.07478 | .119328 | .001752 | -3.073381 | -3.310085 | -2.842623
theta_1   | 3.863657 | .356535 | .005316 | 3.851435 | 3.199128 | 4.599245
theta_2   | 1.311634 | .3317564 | .004253 | 1.311573 | .654233 | 1.957713
theta_s   | -.7127724 | .129681 | .001704 | -.7116233 | -.9656503 | -.4610413
|-----+-----|-----+-----|-----+-----|
```

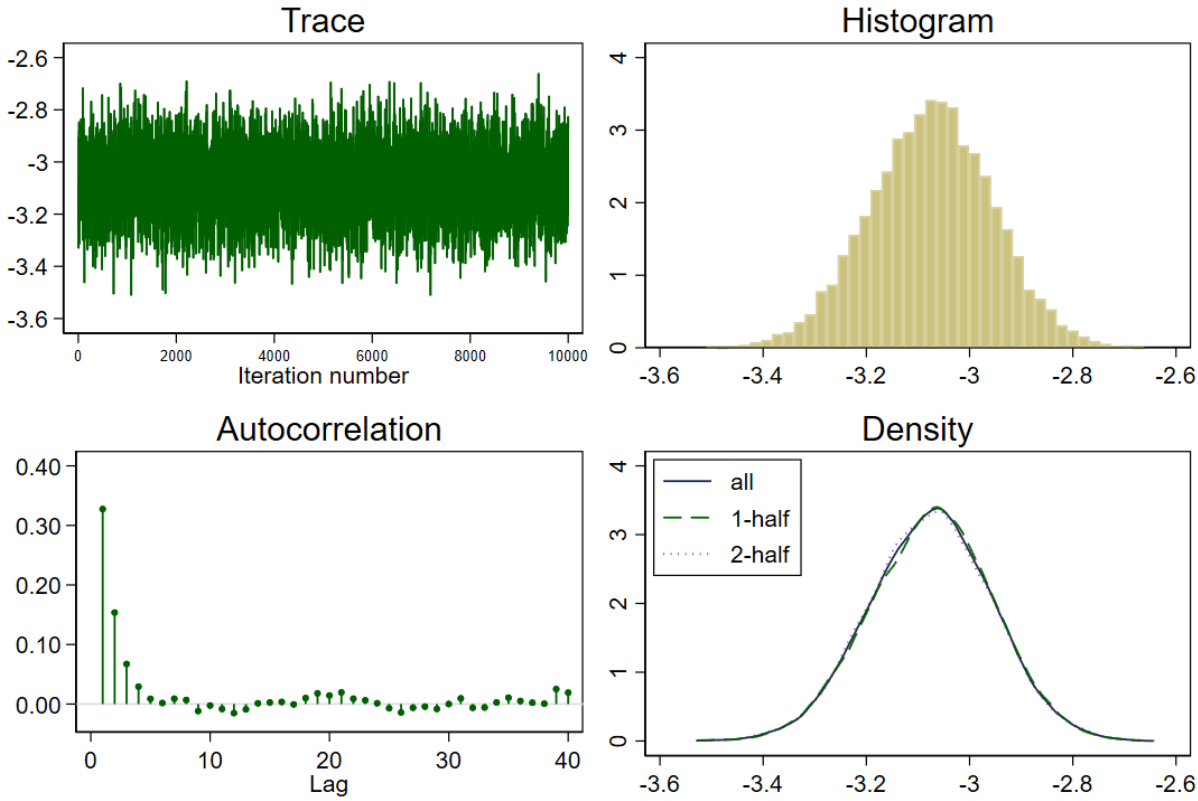
file sim_Solid_ERR.dta saved

```
.
. estimates store Solid_ERR
. matrix Solid_ERR_means = e(mean)
```

Verify the convergence of the background theta_0 term

```
. bayesgraph diagnostic {theta_0}, name(Solid_ERR_theta_0_d, replace)
```

theta_0

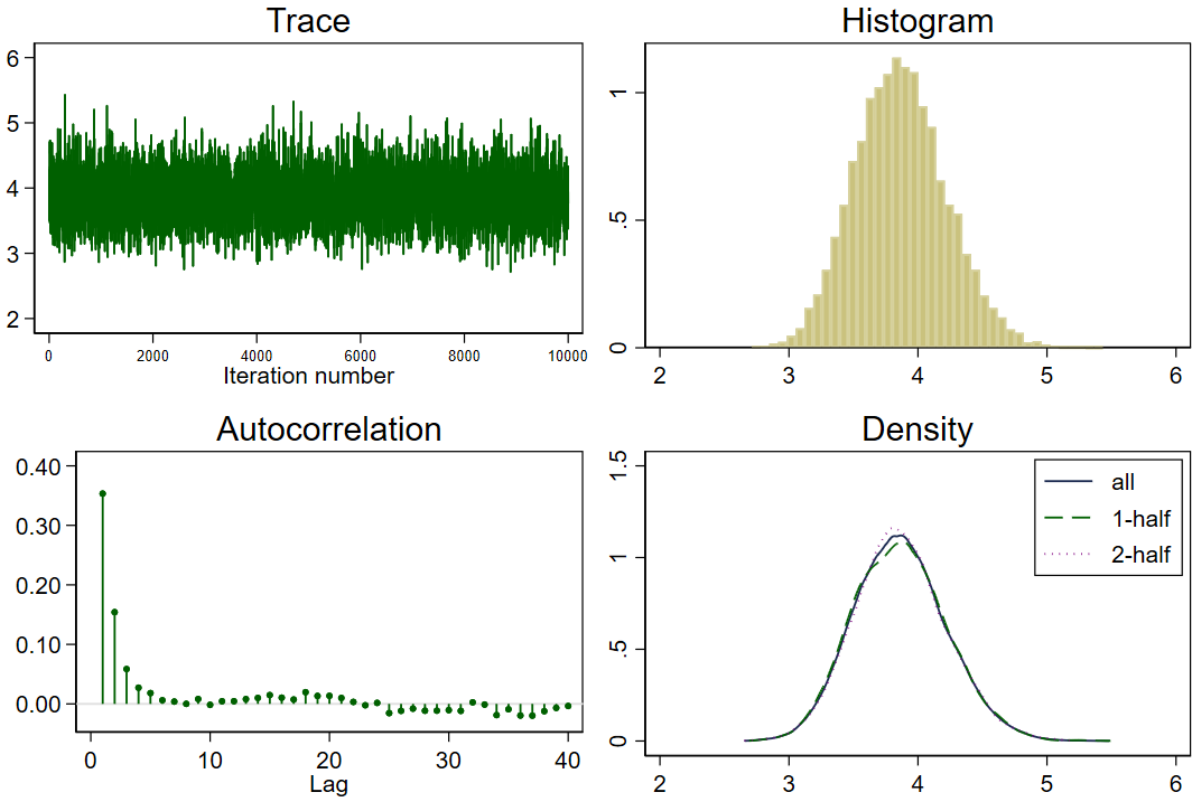


- Good convergence

Verify the convergence of the background theta_1 spline term

```
. bayesgraph diagnostic {theta_1}, name(Solid_ERR_theta_1_d, replace)
```

theta_1

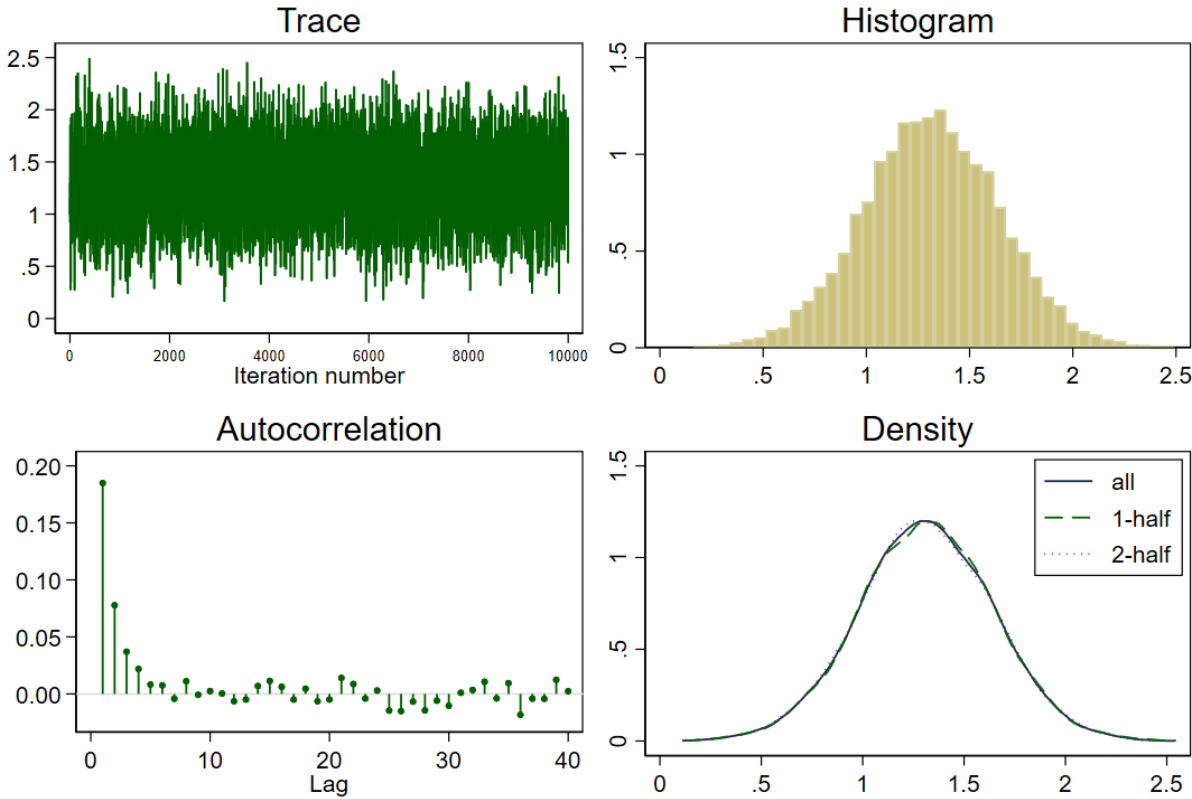


- Good convergence

Verify the convergence of the background theta_2 spline term

```
. bayesgraph diagnostic {theta_2}, name(Solid_ERR_theta_2_d, replace)
```

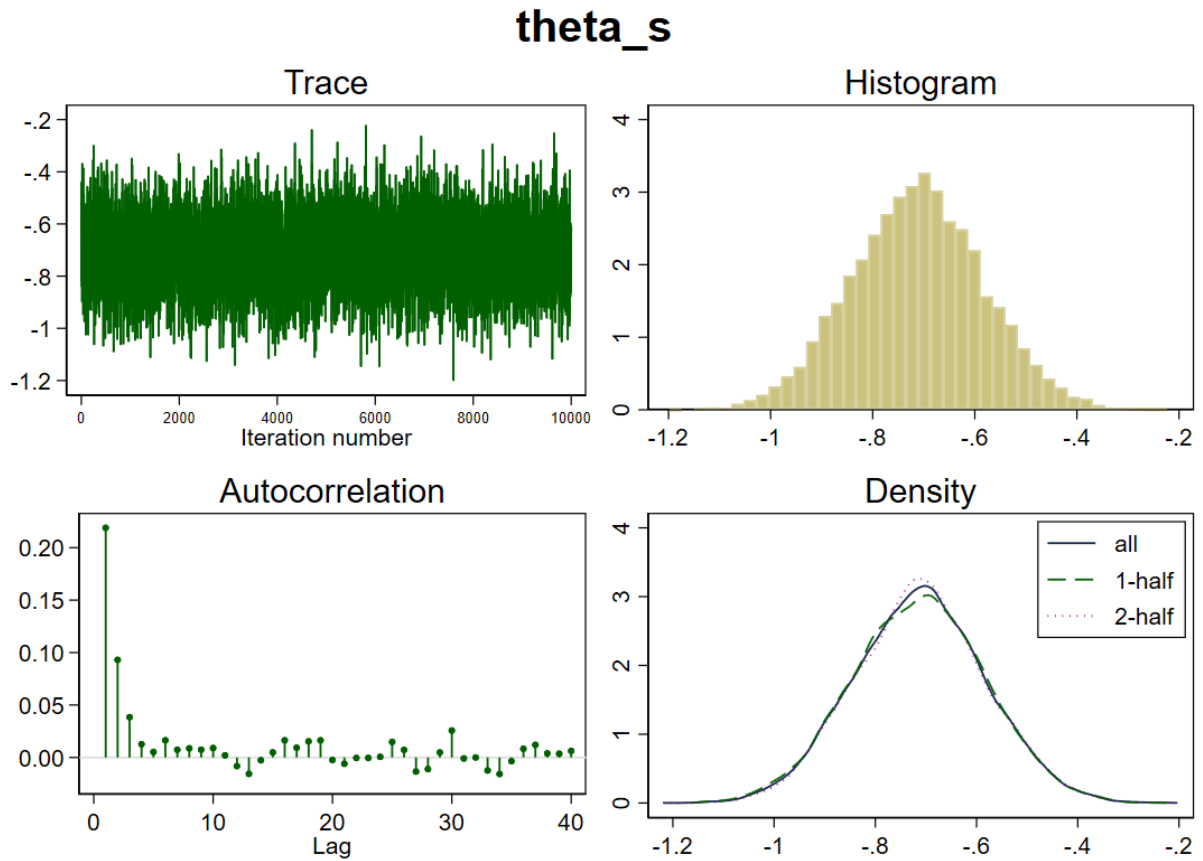
theta_2



- Good convergence

Verify the convergence of the background θ_s sex modification term

```
. bayesgraph diagnostic {theta_s}, name(Solid_ERR_theta_s_d, replace)
```

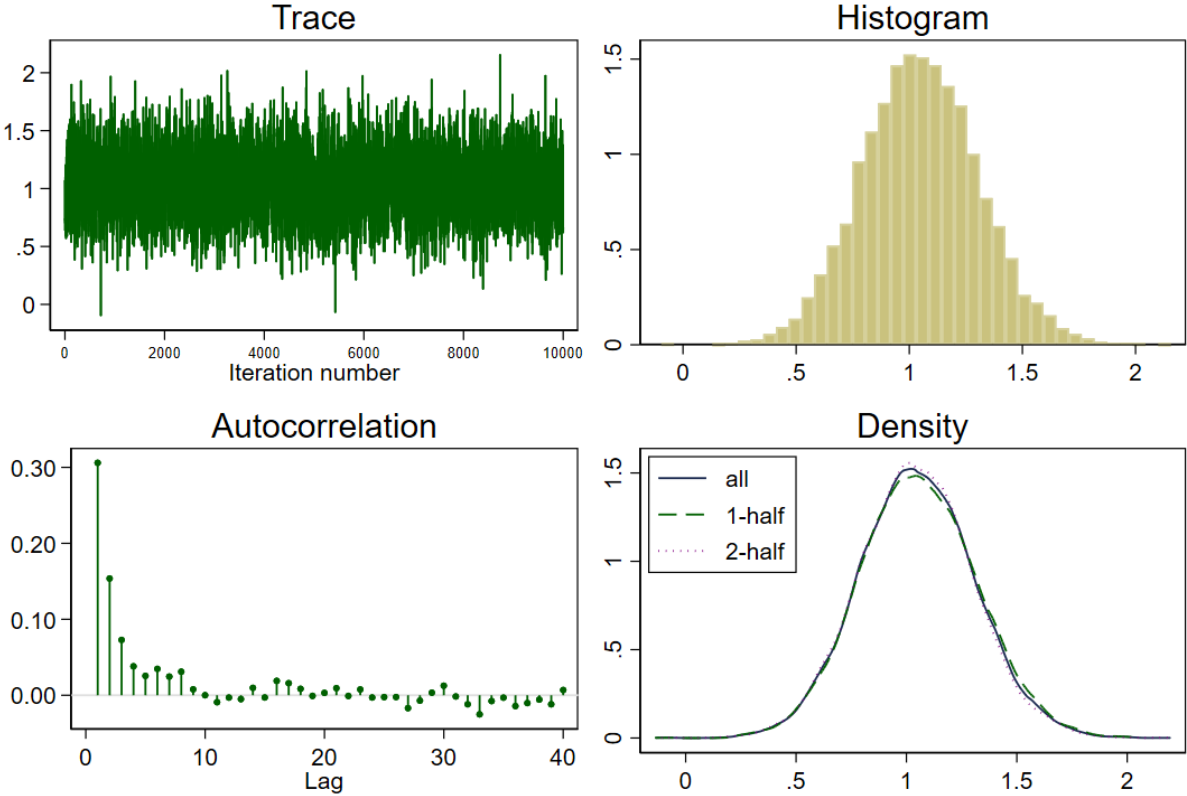


- Good convergence

Verify the convergence of the radiation delta_s sex modification term

```
. bayesgraph diagnostic {delta_s}, name(Solid_ERR_delta_s_d, replace)
```

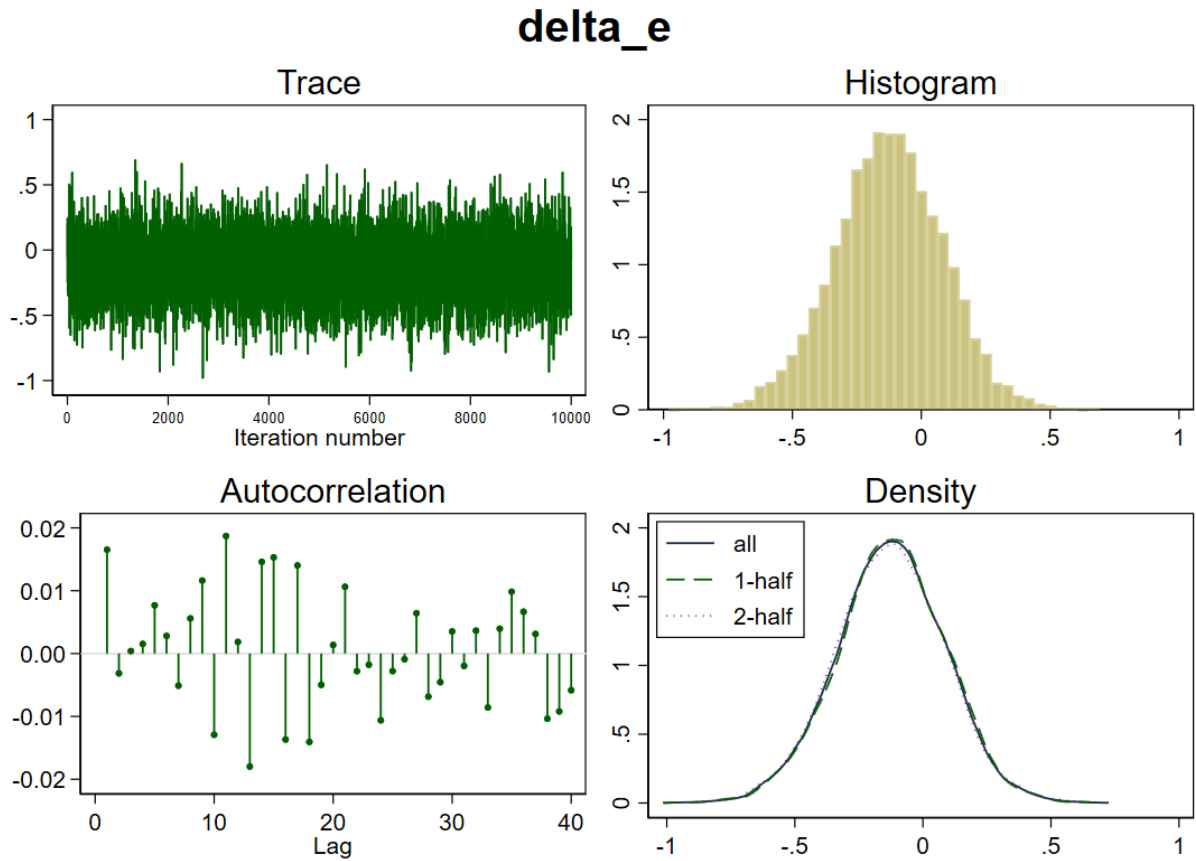
delta_s



- Good convergence

Verify the convergence of the radiation delta_e exposure age modification term

```
. bayesgraph diagnostic {delta_e}, name(Solid_ERR_delta_e_d, replace)
```

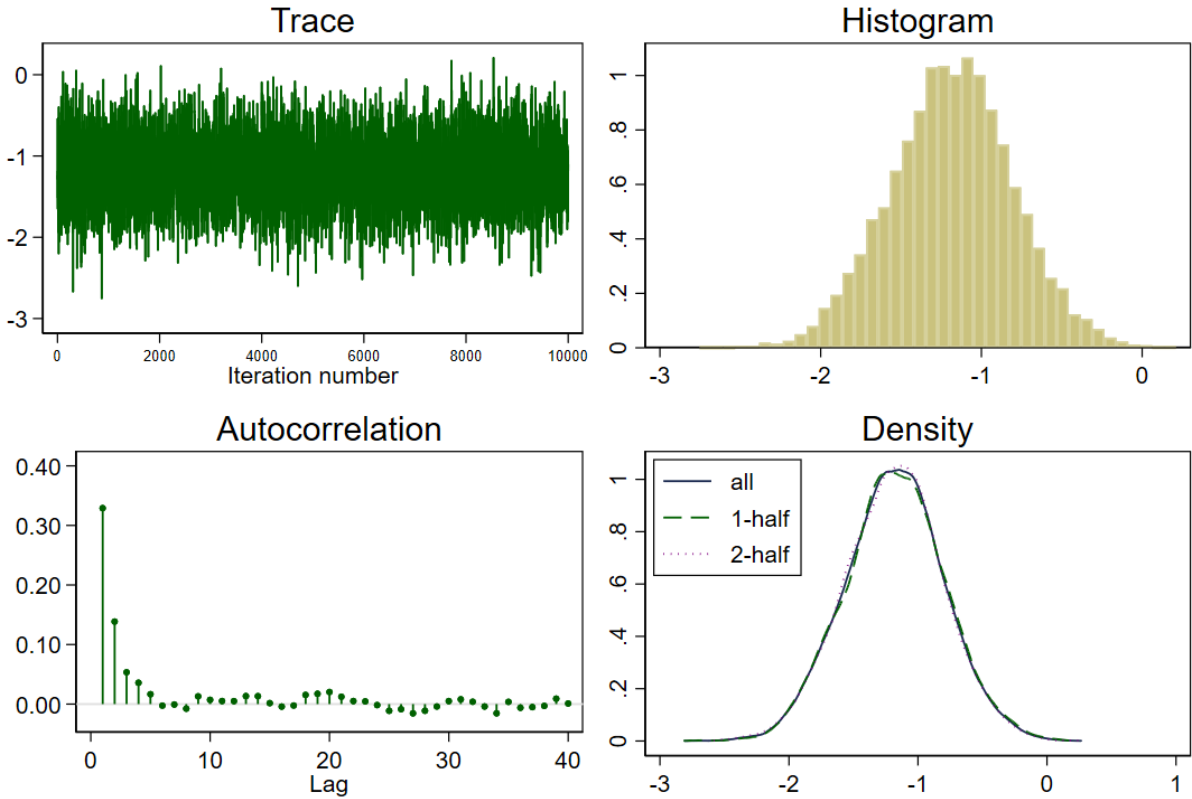


- Good convergence

Verify the convergence of the radiation delta_a age modification term

```
. bayesgraph diagnostic {delta_a}, name(Solid_ERR_delta_a_d, replace)
```

delta_a

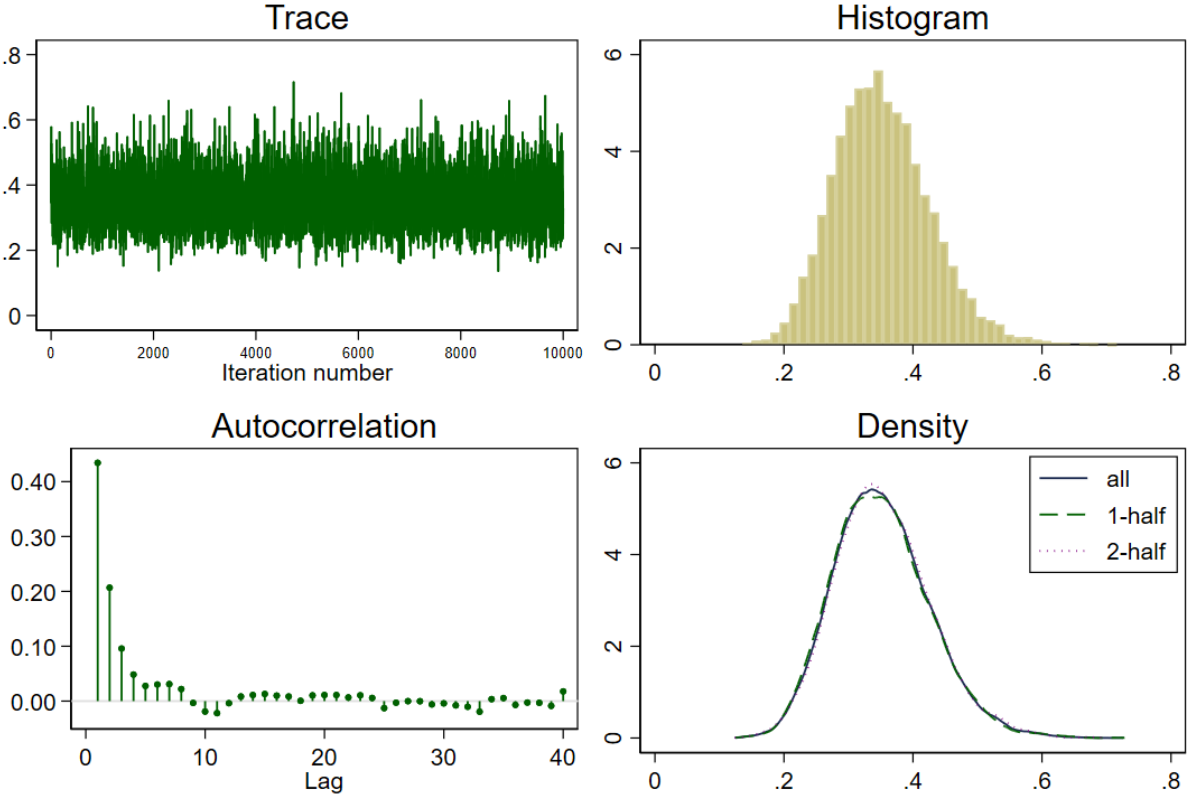


- Good convergence

Verify the convergence of the ERR beta_gamma term for gamma irradiation

```
. bayesgraph diagnostic {beta_gamma}, name(Solid_ERR_beta_gamma_d, replace)
```

beta_gamma



- Good convergence

Solid tumor hazard rates figures

```
. save outbred_edit1, replace
file outbred_edit1.dta saved

. use ratesbyageYgroup2

. gen agemed = agegroup + 1

. replace agemed = 7 if agegroup == 0
(4 real changes made)

.

. gen age = _n + 4 in 1/21
(7 missing values generated)

. gen lnage = ln(age/20)
(7 missing values generated)

. * create a cubic spline with knots at ages 15 20 and 23
. mkspline lnagesp = lnage, cubic knot(-.28768207 0 .13976194)

.

. gen ERR_male_fit_bg_haz = exp(Solid_ERR_means[1,5] + Solid_ERR_means[1,6] * lnagesp1 + Solid_ERR_means[1,7] * ln
> agesp2)*100
(7 missing values generated)

. gen ERR_female_fit_bg_haz = exp(Solid_ERR_means[1,5] + Solid_ERR_means[1,6] * lnagesp1 + Solid_ERR_means[1,7] *
> lnagesp2 + Solid_ERR_means[1,8])*100
(7 missing values generated)

. gen ERR_male_fit_gamma_haz = ERR_male_fit_bg_haz * (1 + Solid_ERR_means[1,1] * 3 * exp(Solid_ERR_means[1,2] * ln
> age))
(7 missing values generated)

. gen ERR_female_fit_gamma_haz = ERR_female_fit_bg_haz * (1 + Solid_ERR_means[1,1] * 3 * exp(Solid_ERR_means[1,2]
> * lnage + Solid_ERR_means[1,4]))
(7 missing values generated)

.

. twoway scatter _Rate agemed if dose_Gy == 0 & sex == 0, color(gs12) || line ERR_male_fit_bg_haz age, color(navy)
> ylabel(0(5)40,angle(horizontal) format(%2.0f) grid) legend(off order(2 1 3) label(1 "Rate estimates for age ca
> tegories") label(2 "ERR model fits") col(1) ring(0) pos(11)) xtitle("") ytitle("") title("Baseline, Male", box b
> expand) saving(ERR_bg_solid_m, replace) name(ERR_bg_solid_m, replace)
(file ERR_bg_solid_m.gph saved)

.

. twoway scatter _Rate agemed if dose_Gy == 3 & sex == 0, color(gs12) || line ERR_male_fit_gamma_haz age, color(na
> vy) ylabel(0(5)40,angle(horizontal) format(%2.0f) grid) legend(off order(2 1 3) label(1 "Rate estimates for age
> categories") label(2 "ERR model fits") col(1) ring(0) pos(11)) xtitle("") ytitle("") title("Gamma Irradiated, M
> ale", box bexpand) saving(ERR_gamma_solid_m, replace) name(ERR_gamma_solid_m, replace)
(file ERR_gamma_solid_m.gph saved)

.

. twoway scatter _Rate agemed if dose_Gy == 0 & sex == 1, color(gs12) || line ERR_female_fit_bg_haz age, color(nav
> y) ylabel(0(5)40,angle(horizontal) format(%2.0f) grid) legend(off order(2 1 3) label(1 "Rate estimates for age
> categories") label(2 "ERR model fits") col(1) ring(0) pos(11)) xtitle("") ytitle("") title("Baseline, Female", b
> ox bexpand) saving(ERR_bg_solid_f, replace) name(ERR_bg_solid_f, replace)
(file ERR_bg_solid_f.gph saved)

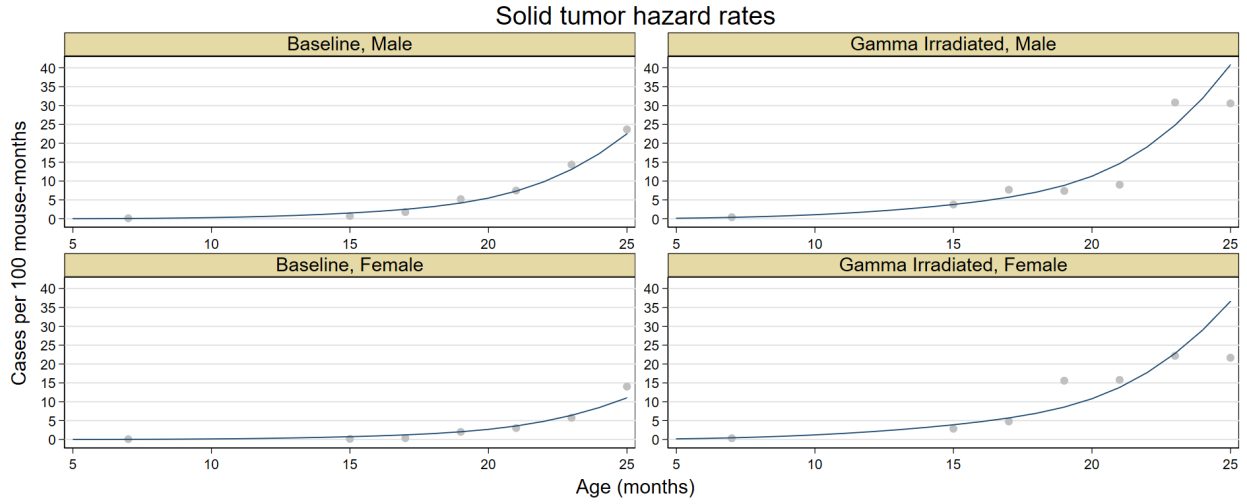
.

. twoway scatter _Rate agemed if dose_Gy == 3 & sex == 1, color(gs12) || line ERR_female_fit_gamma_haz age, color(
> navy) ylabel(0(5)40,angle(horizontal) format(%2.0f) grid) legend(off order(2 1 3) label(1 "Rate estimates for a
> ge categories") label(2 "ERR model fits") col(1) ring(0) pos(11)) xtitle("") ytitle("") title("Gamma Irradiated,
> Female", box bexpand) saving(ERR_gamma_solid_f, replace) name(ERR_gamma_solid_f, replace)
(file ERR_gamma_solid_f.gph saved)
```

```

.
.
. graph combine ERR_bg_solid_m ERR_gamma_solid_m ERR_bg_solid_f ERR_gamma_solid_f, rows(2) ycommon b1title("Age (m
> onths)") l1title("Cases per 100 mouse-months") title("Solid tumor hazard rates") imargin(0 0 0 0) ysize(5) xsize
> (12) saving(ERR_solid, replace) name(ERR_solid, replace)
(file ERR_solid.gph saved)

```



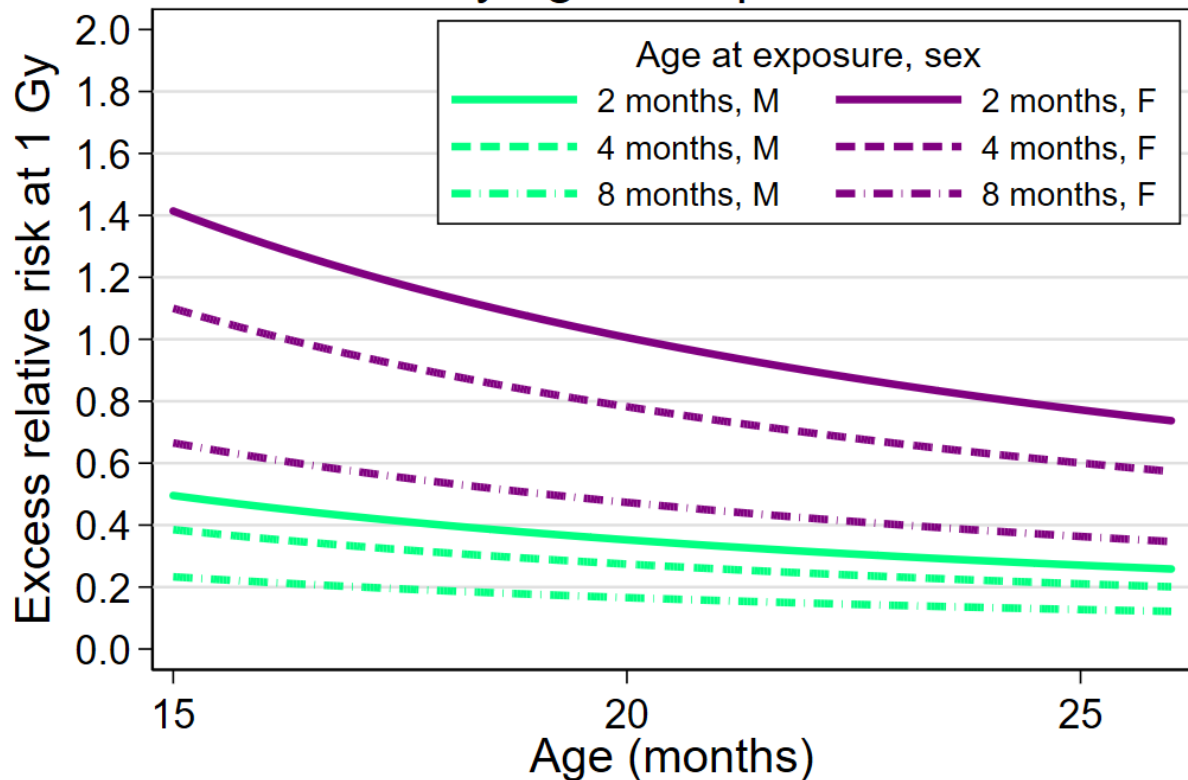
ERR temporal patterns by sex and age by age at exposure

```

. twoway function exp(Solid_ERR_means[1,2]*ln(x/20))*Solid_ERR_means[1,1], range(15 26) color(mint) lwidth(thick)
> || function exp(Solid_ERR_means[1,2]*ln(x/20) + Solid_ERR_means[1,4])*Solid_ERR_means[1,1], range(15 26) color(
> purple) lwidth(thick) || function exp(Solid_ERR_means[1,2]*ln(x/20) + Solid_ERR_means[1,3]*2)*Solid_ERR_means[1
> ,1], range(15 26) color(mint) lpattern(dash) lwidth(thick) || function exp(Solid_ERR_means[1,2]*ln(x/20) + Soli
> d_ERR_means[1,3]*2 + Solid_ERR_means[1,4])*Solid_ERR_means[1,1], range(15 26) color(purple) lpattern(dash) lwidth
> h(thick) || function exp(Solid_ERR_means[1,2]*ln(x/20) + Solid_ERR_means[1,3]*6)*Solid_ERR_means[1,1], range(15
> 26) color(mint) lpattern(dash_dot) lwidth(thick) || function exp(Solid_ERR_means[1,2]*ln(x/20) + Solid_ERR_mea
> ns[1,3]*6 + Solid_ERR_means[1,4])*Solid_ERR_means[1,1], range(15 26) color(purple) lpattern(dash_dot) lwidth(thi
> ck) xtitle("Age (months)", size(large)) ytitle("Excess relative risk at 1 Gy", size(large)) title("ERR temporal
> patterns by sex and age" "by age at exposure",size(vlarge)) legend(order(1 2 3 4 5 6) label(1 "2 months, M") lab
> el(2 "2 months, F") label(3 "4 months, M") label(4 "4 months, F") label(5 "8 months, M") label(6 "8 months, F")
> subtitle("Age at exposure, sex") ring(0) pos(1) ylabel(0(0.2)2,angle(horizontal) format(%2.1f) grid labsz
> med) xlabel(,labsize(medlarge)) ysize(5) xsize(6.5) saving(ERR_temporal, replace) name(ERR_temporal, replace)
(file ERR_temporal.gph saved)

```

ERR temporal patterns by sex and age by age at exposure



Solid cancer dose-response functions for males and females

```
. twoway function x*Solid_ERR_means[1,1], range(0 2) color(mint) lwidth(thick) || function x*exp(Solid_ERR_means[
> 1,4])*Solid_ERR_means[1,1], range(0 2) color(purple) lwidth(thick) lpattern(dash) xtitle("Dose (Gy)", size(large
> )) ytitle("Excess relative risk", size(large)) legend(order(1 2) label(1 "Male") label(2 "Female") ring(0) pos(1
> 1)) ylabel(0(0.2)2,angle(horizontal) format(%2.1f) grid labsize(medlarge)) xlabel(,labsize(medlarge)) title("Sol
> id cancer dose-response functions" "for males and females", size(vlarge)) ysize(5) xsize(6.5) saving(ERR_dose, r
> eplace) name(ERR_dose, replace)
(file ERR_dose.gph saved)
```

