

```

> library(nlme)
> library(gof)
>
> fev <- read.table(file="fev1a.dat",header=T)
> head(fev)
  id ht  age baseht baseage logfev1
1  1 1.20 9.3415  1.2 9.3415 0.21511
2  1 1.28 10.3929  1.2 9.3415 0.37156
3  1 1.33 11.4524  1.2 9.3415 0.48858
4  1 1.42 12.4600  1.2 9.3415 0.75142
5  1 1.48 13.4182  1.2 9.3415 0.83291
6  1 1.50 15.4743  1.2 9.3415 0.89200
> dim(fev)
[1] 1994  6
> fev$ID <- factor(fev$id)
> fev$loght <- log(fev$ht)
> fev$logbaseht <- log(fev$baseht)
>
> fev1 <- groupedData(logfev1~age|ID,data=fev[fev$id != 197,], order.groups=F)
>
> # Part 10.1.1
> m1 <- lme(logfev1~age+loght+baseage+logbaseht,data=fev1,random=~1+age|ID)
> summary(m1)
Linear mixed-effects model fit by REML
Data: fev1
      AIC      BIC logLik
-4549.882 -4499.528 2283.941

Random effects:
Formula: ~1 + age | ID
Structure: General positive-definite, Log-Cholesky parametrization
      StdDev  Corr
(Intrcpt) 0.110485537 (Intr)
age        0.007078381 -0.553
Residual  0.060237879

Fixed effects: logfev1 ~ age + loght + baseage + logbaseht
      Value Std.Error DF t-value p-value
(Intrcpt) -0.2883233 0.03871675 1692 -7.44699 0.0000
age        0.0235286 0.00139534 1692 16.86231 0.0000
loght      2.2371985 0.04353724 1692 51.38586 0.0000
baseage    -0.0165088 0.00745785 296 -2.21362 0.0276
logbaseht  0.2182147 0.14552086 296  1.49954 0.1348
Correlation:
      (Intr) age  loght  baseag
age      0.023
loght   -0.077 -0.875
baseage -0.822 -0.184 0.180

```

logbaseht 0.370 0.239 -0.275 -0.815

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-6.45672772	-0.52535014	0.05352018	0.60114673	2.76671589

Number of Observations: 1993

Number of Groups: 299

```
> m1.res0 <- resid(m1,type="r", level=0)
```

```
> m1.res1 <- resid(m1,type="r", level=1)
```

```
>
```

```
>
```

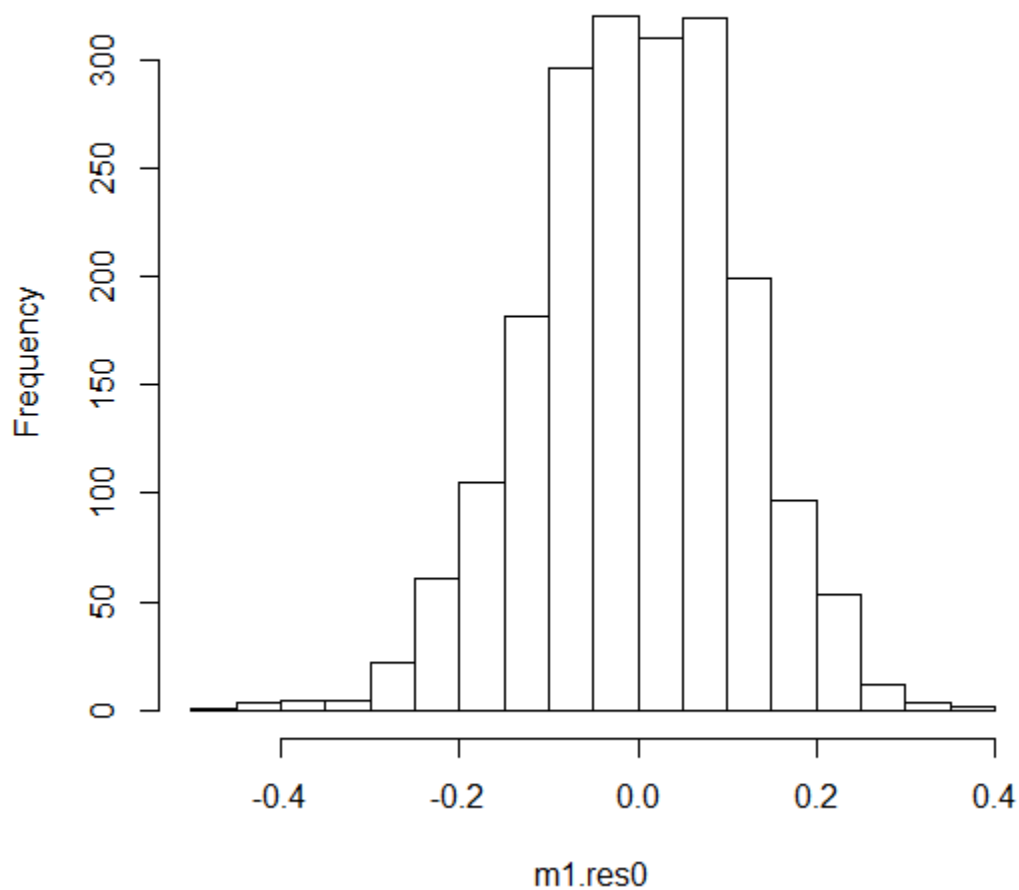
```
> # Part 10.1.2
```

```
>
```

```
> hist(m1.res0,main="Histogram of pop. level raw resids, model M1")
```

```
> # Histogram looks fine (approximately symmetric, normal)
```

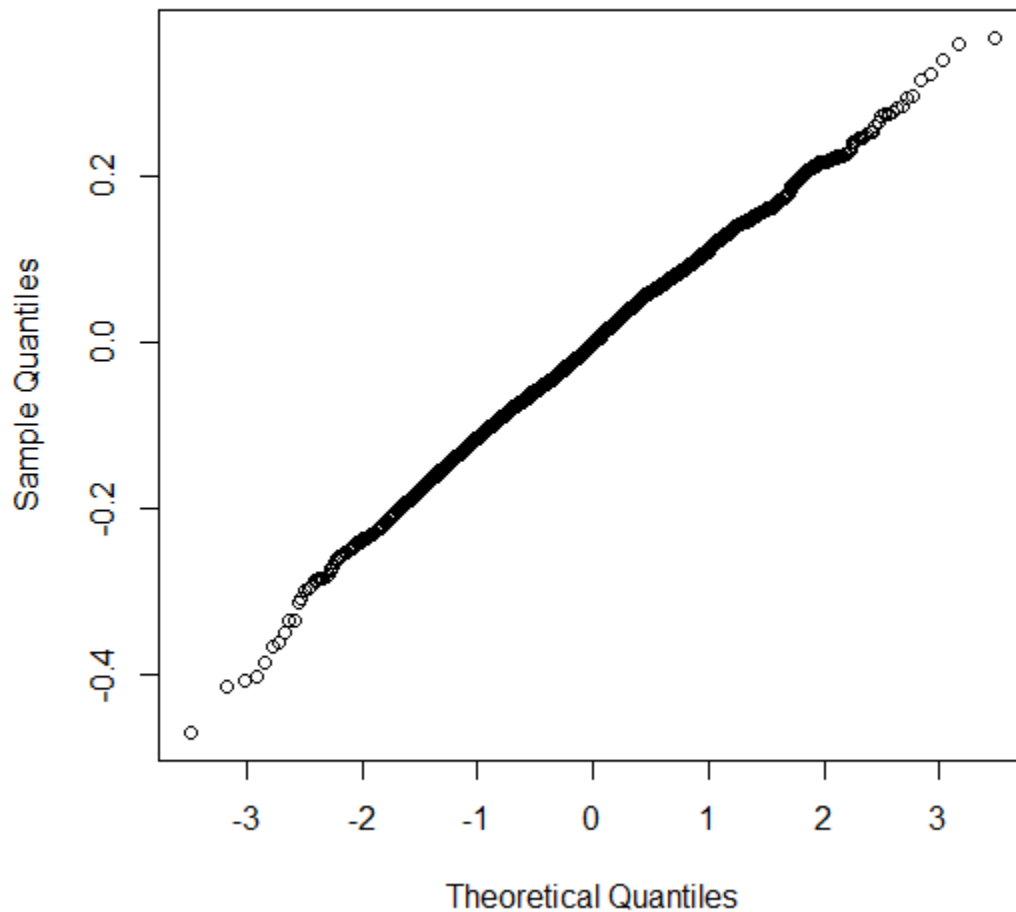
Histogram of pop. level raw resids, model M1



```
>
```

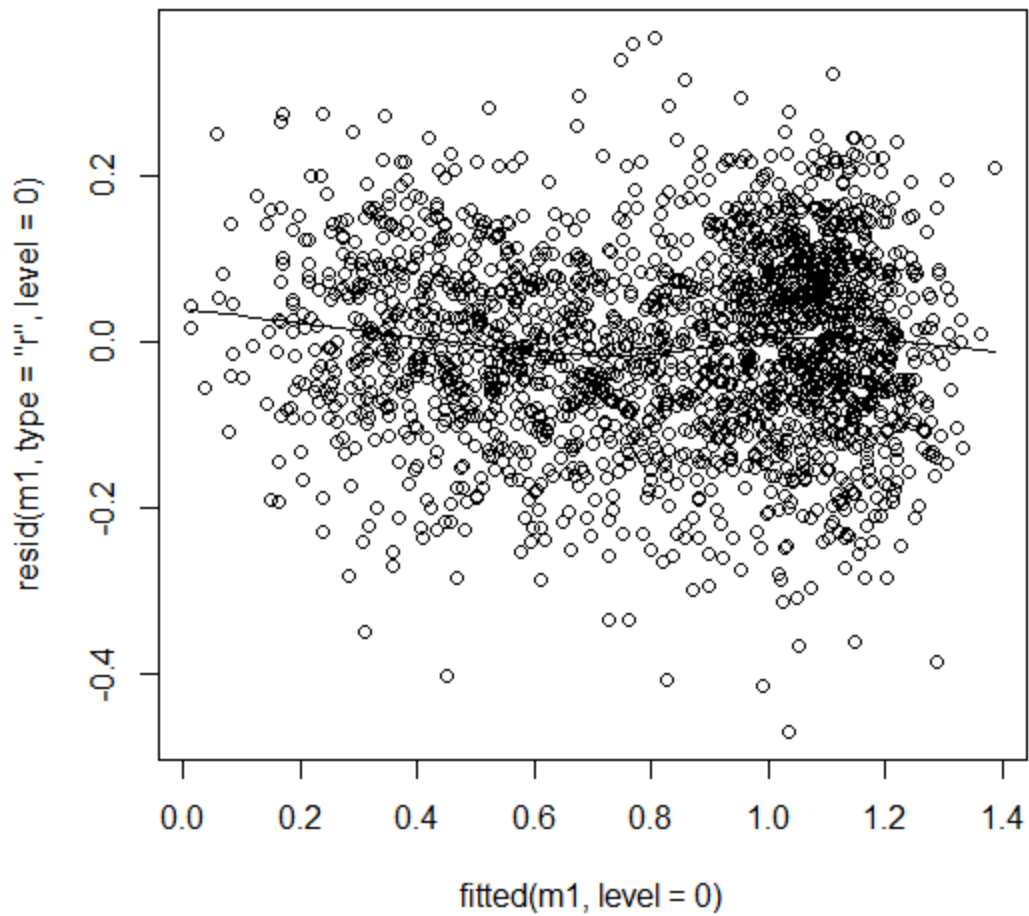
```
> # Part 10.1.3
>
> qqnorm(m1.res0,main="Normal qq plot of pop. level raw resids, model M1")
> # QQ plot also looks fine. No important departures from normality. No outliers.
```

Normal qq plot of pop. level raw resids, model M1



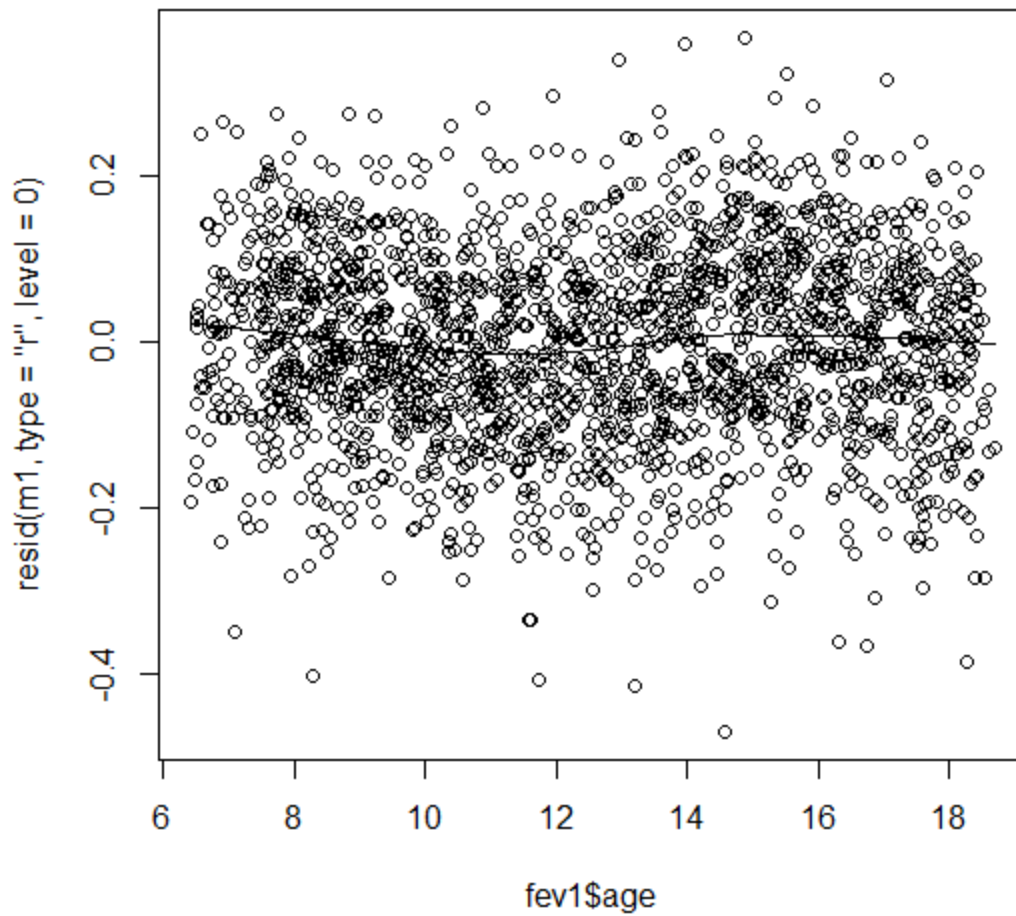
```
>
> # Part 10.1.4
>
> #plot(m1, resid(.,type="r",level=0)~fitted(.,level=0),id=0.001, lowess=T)
> plot(fitted(m1,level=0),resid(m1,type="r",level=0),main="Level 0 raw resids vs fitteds, m1")
> lines(lowess(fitted(m1,level=0),resid(m1,type="r",level=0)))
```

Level 0 raw resid vs fitteds, m1



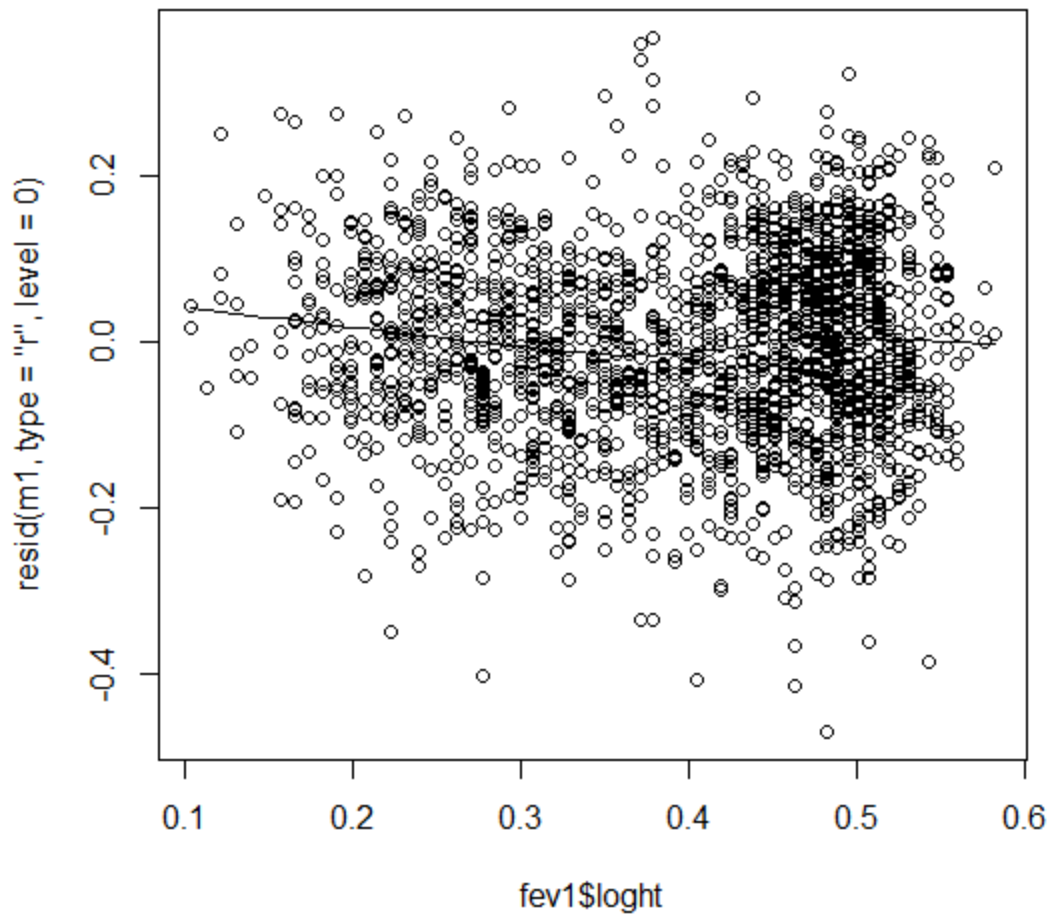
```
>  
> # Part 10.1.5  
>  
> #plot(m1, resid(.,type="r",level=0)~age,id=0.001, lowess=T)  
> plot(fev1$age,resid(m1,type="r",level=0),main="Level 0 raw resid vs age, m1")  
> lines(lowess(fev1$age,resid(m1,type="r",level=0)))
```

Level 0 raw resid vs age, m1

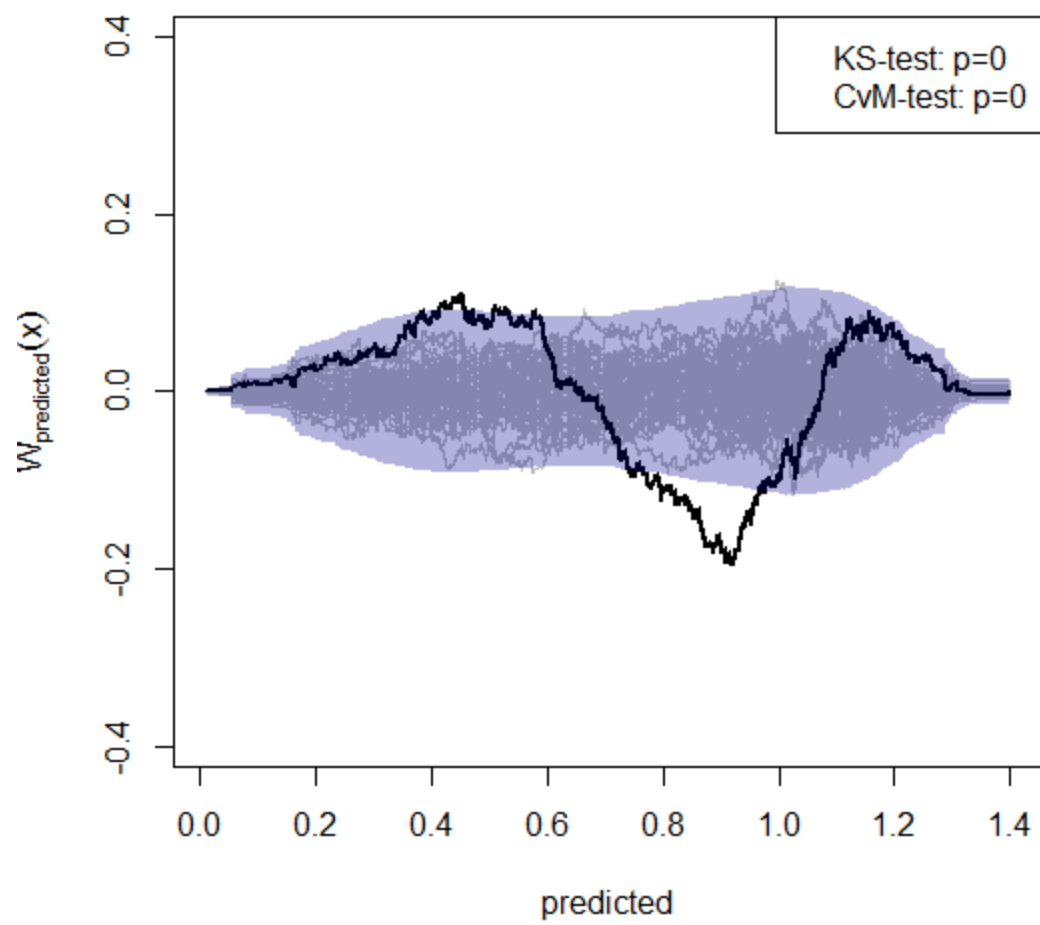


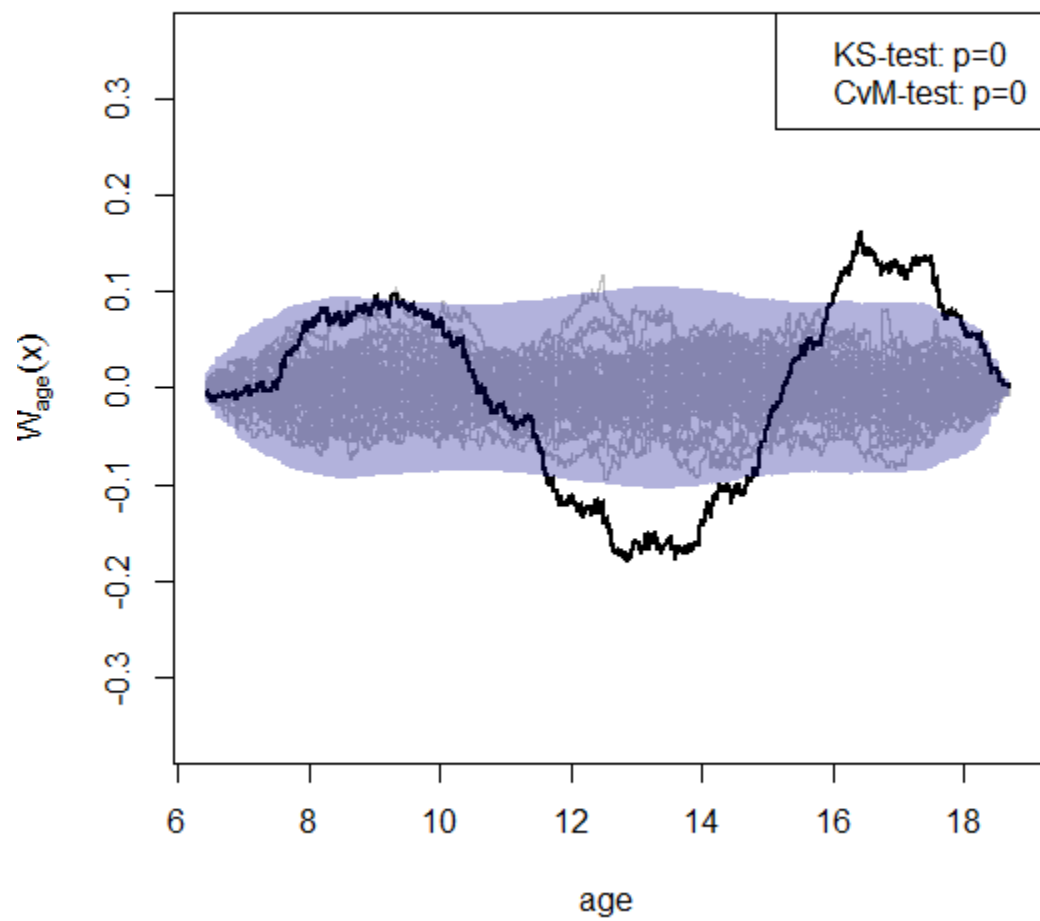
```
>  
> # Part 10.1.6  
>  
> #plot(m1, resid(.,type="r",level=0)~loght,id=0.001, lowess=T)  
> plot(fev1$loght,resid(m1,type="r",level=0),main="Level 0 raw resid vs loght, m1")  
> lines(lowess(fev1$loght,resid(m1,type="r",level=0)))
```

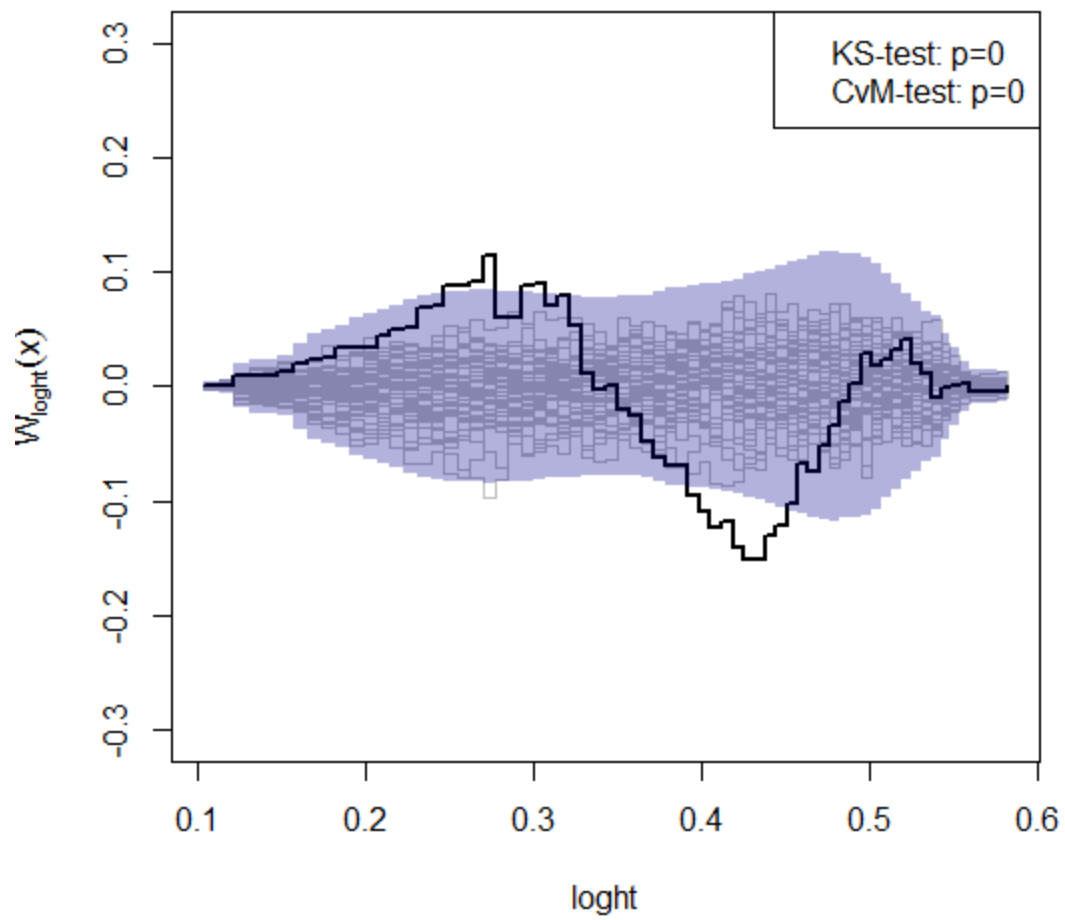
Level 0 raw resid vs loght, m1

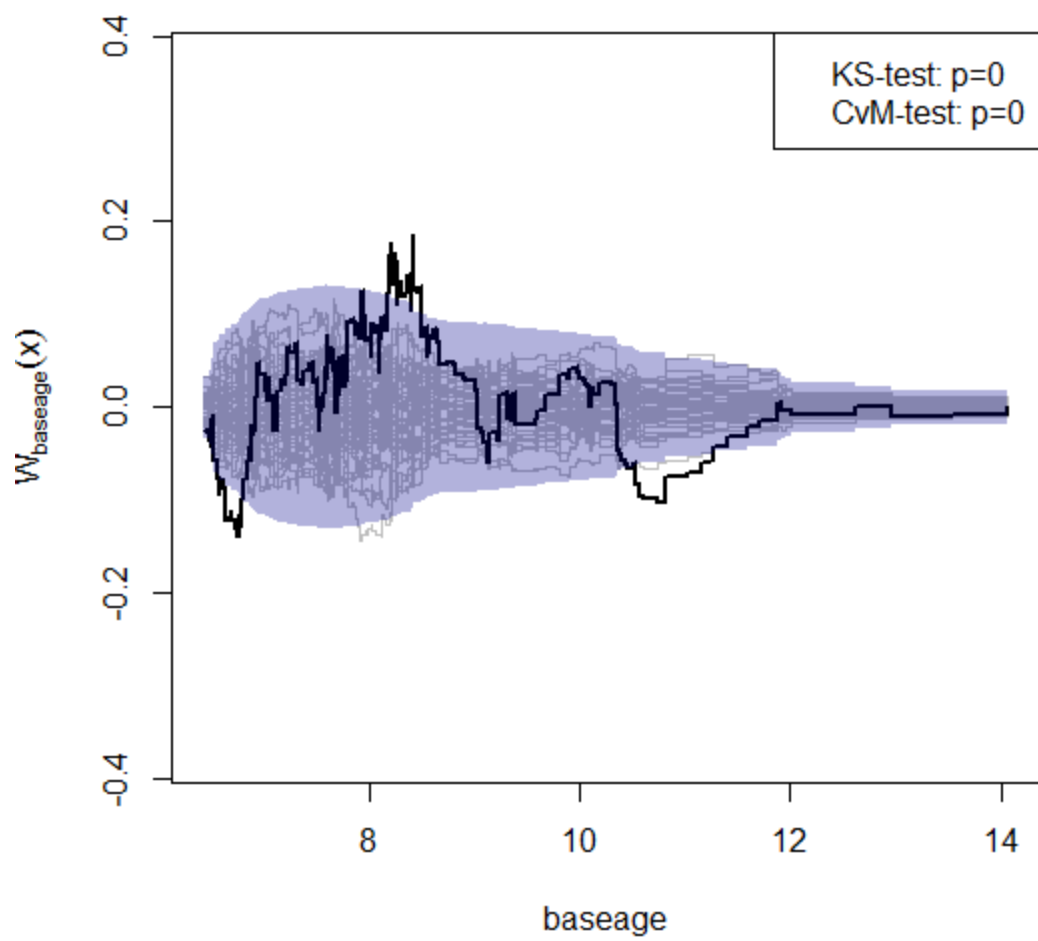


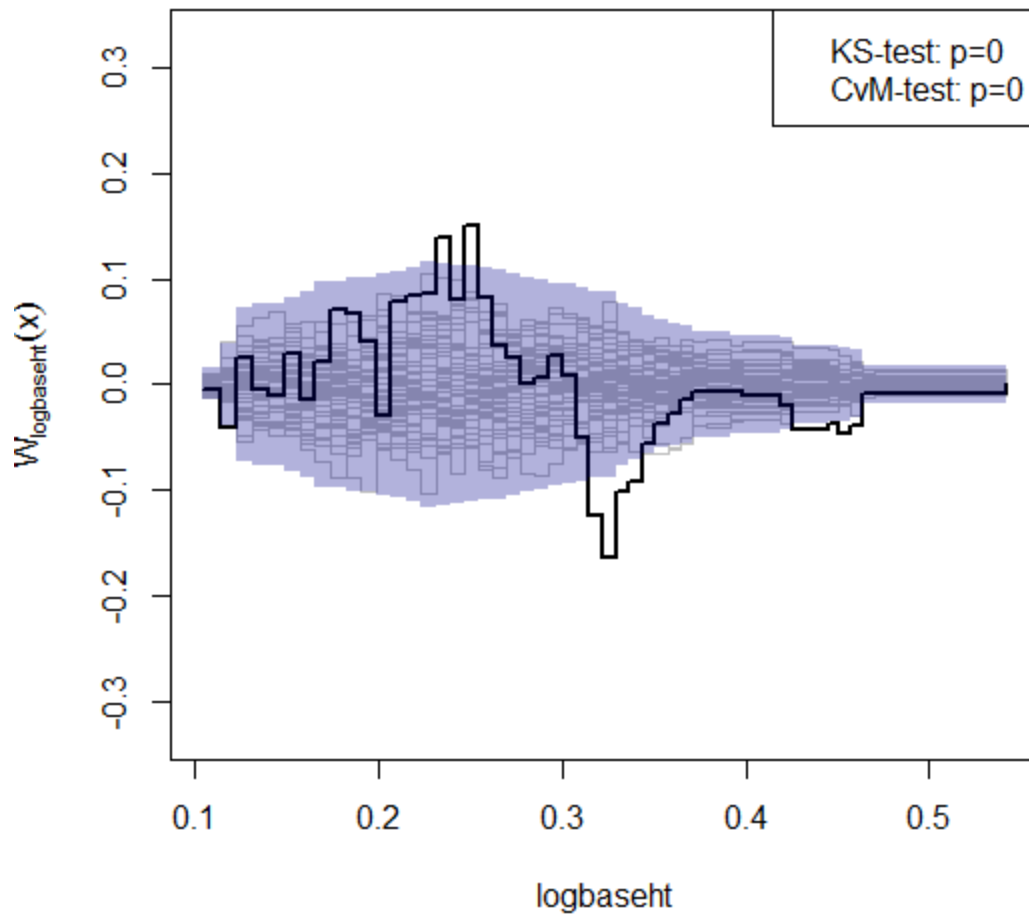
```
>
> # Part 10.1.7
> # The plots obtained above all have a wave in them that can be seen in the
> # lowest curves. This indicates a systematic misspecification of the mean.
> # The mean response does not seem linearly related to age or to log ht
> # As you will recall, in FEVExample1.sas I tried models with higher order
> # terms in age and had to include a quartic term in age in order to get a
> # good fit. Alternatively, we can try spline models as in FEVExample2.sas
>
> # Part 10.1.8
>
> m1.lm <- lm(logfev1~age+loght+baseage+logbaseht,data=fev1)
> cres1 <- cumres(m1.lm,R=500)
> plot(cres1)
> # These plots do not look good. The wave in the residuals versus fitteds,
> # resid vs age, and resid vs loght plots shows up as an unusual path in the
> # cumulative resid plots
```











```

>
> # Part 10.1.9
>
> fev1$age14plus <- (fev1$age-14)*as.numeric(fev1$age>14)
>
> m3 <- lme(logfev1~age+age14plus+loght+baseage+logbaseht,data=fev1,
+ random=~1+age+age14plus|ID)
> summary(m3)
Linear mixed-effects model fit by REML
Data: fev1
      AIC   BIC logLik
-4589.328 -4516.601 2307.664

```

Random effects:
Formula: ~1 + age + age14plus | ID

Structure: General positive-definite, Log-Cholesky parametrization

```
StdDev  Corr
(Intercept) 0.13812653 (Intr) age
age      0.01217911 -0.758
age14plus 0.02235398 0.587 -0.844
Residual  0.05748941
```

Fixed effects: logfev1 ~ age + age14plus + loght + baseage + logbaseht

```
Value Std.Error DF t-value p-value
(Intercept) -0.2931855 0.03862689 1691 -7.590192 0.0000
age          0.0412973 0.00422625 1691 9.771602 0.0000
age14plus   -0.0195652 0.00468252 1691 -4.178348 0.0000
loght       1.8310344 0.09987144 1691 18.333915 0.0000
baseage     -0.0311311 0.00823223 296 -3.781617 0.0002
logbaseht   0.5350851 0.16534351 296 3.236203 0.0013
```

Correlation:

```
(Intr) age ag14pl loght baseag
age      -0.023
age14plus 0.041 -0.935
loght    -0.014 -0.964 0.852
baseage  -0.723 -0.474 0.436 0.468
logbaseht 0.304 0.536 -0.477 -0.556 -0.849
```

Standardized Within-Group Residuals:

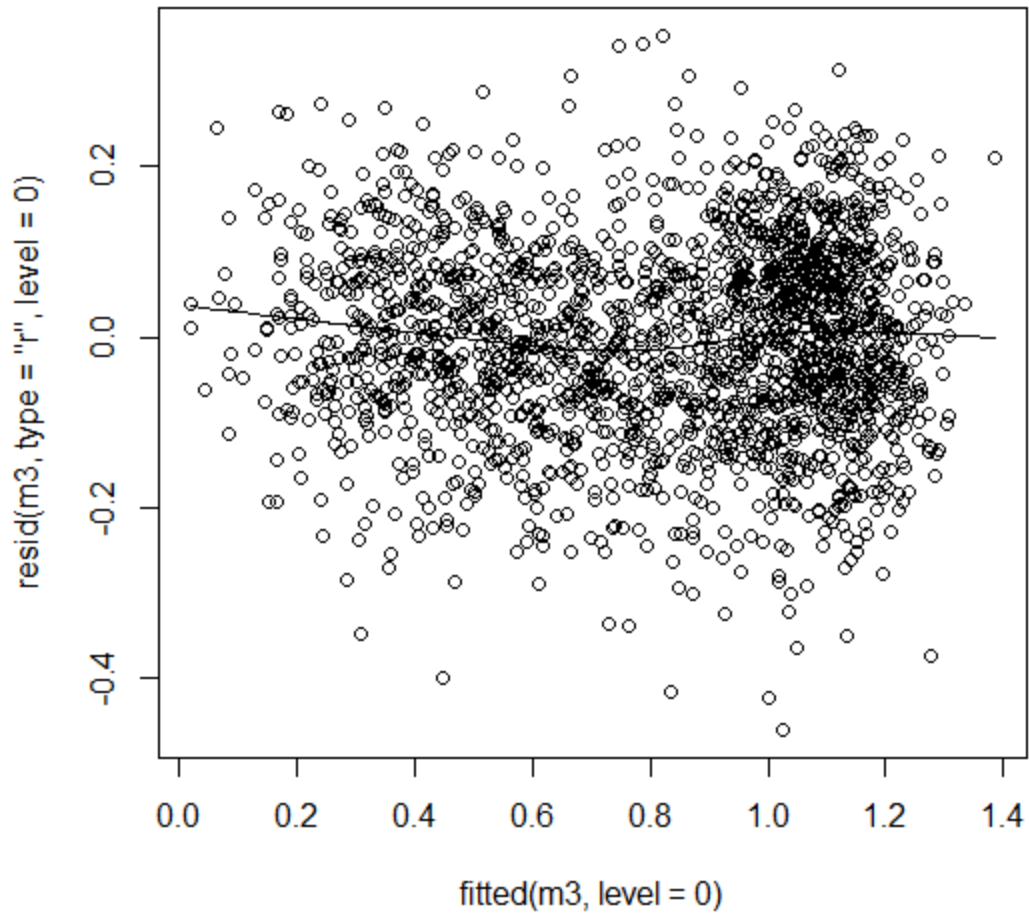
```
Min      Q1      Med      Q3      Max
-6.71096860 -0.49406071 0.08395634 0.54493873 2.62345627
```

Number of Observations: 1993

Number of Groups: 299

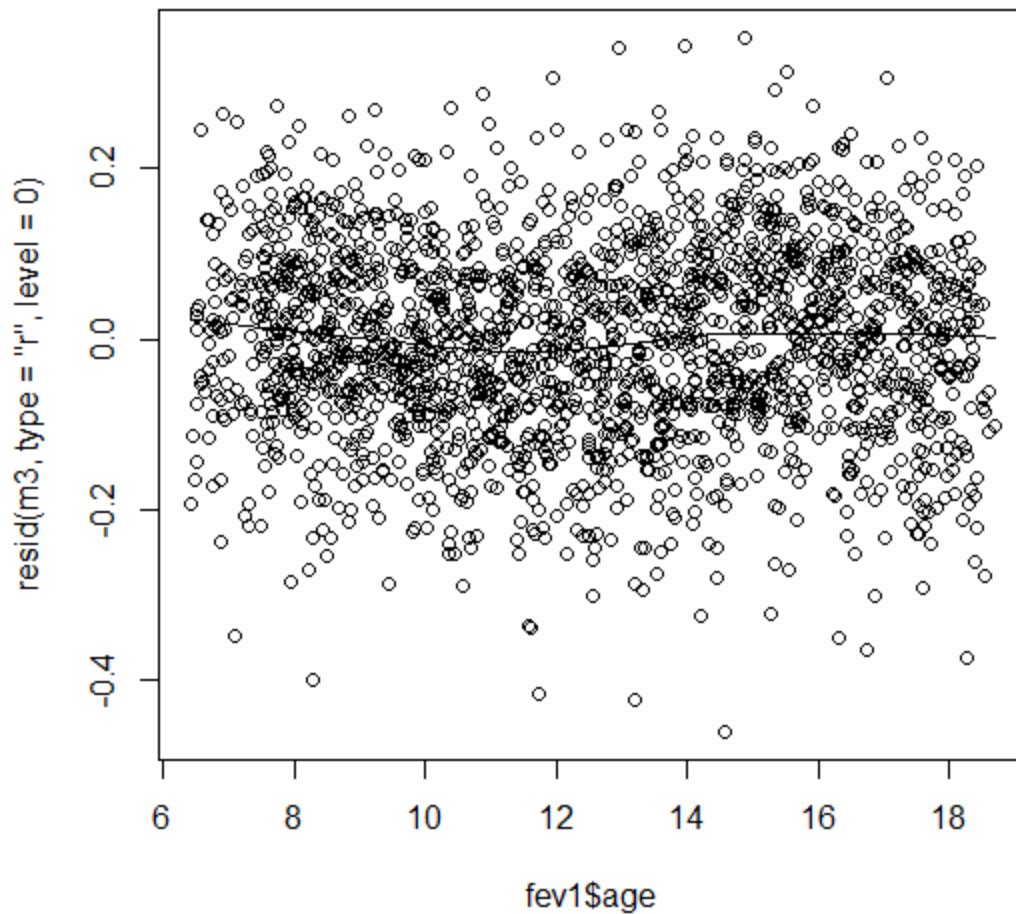
```
>
> plot(fitted(m3,level=0),resid(m3,type="r",level=0),
+ main="Level 0 raw resid vs fitteds, m3")
> lines(lowess(fitted(m3,level=0),resid(m3,type="r",level=0)))
```

Level 0 raw resid vs fitteds, m3

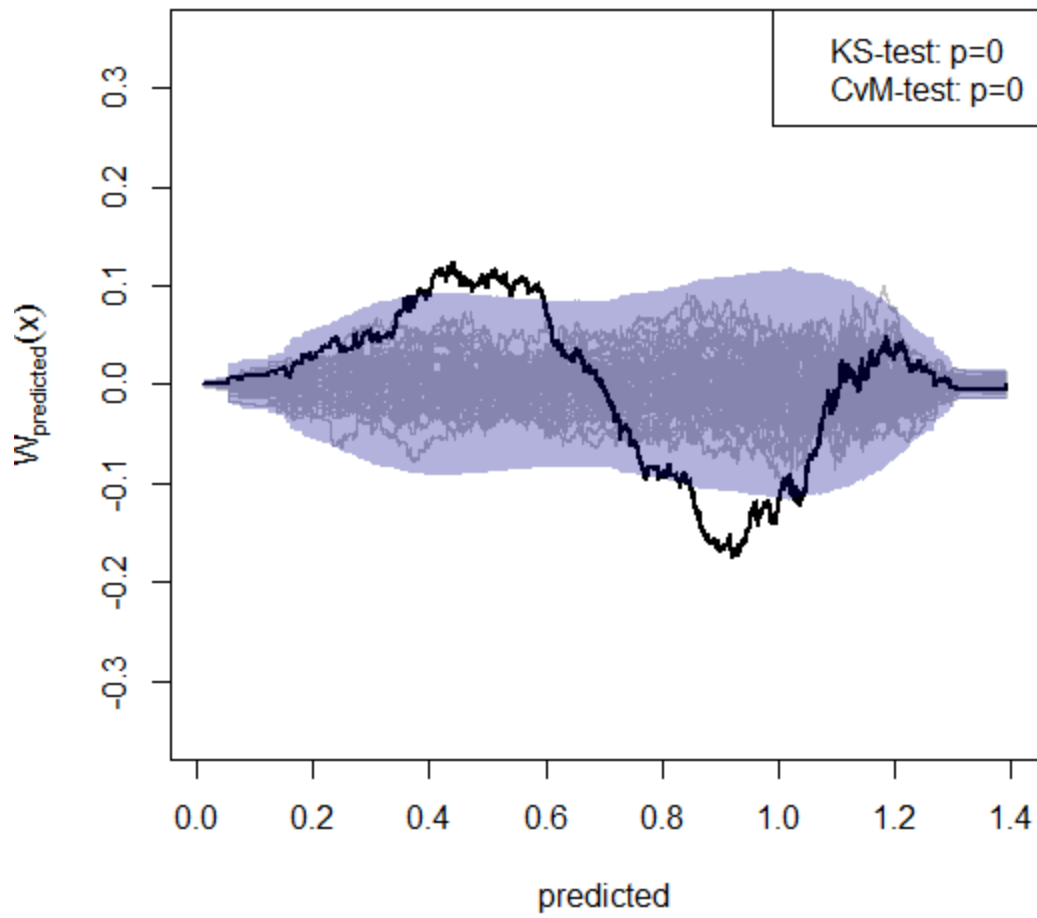


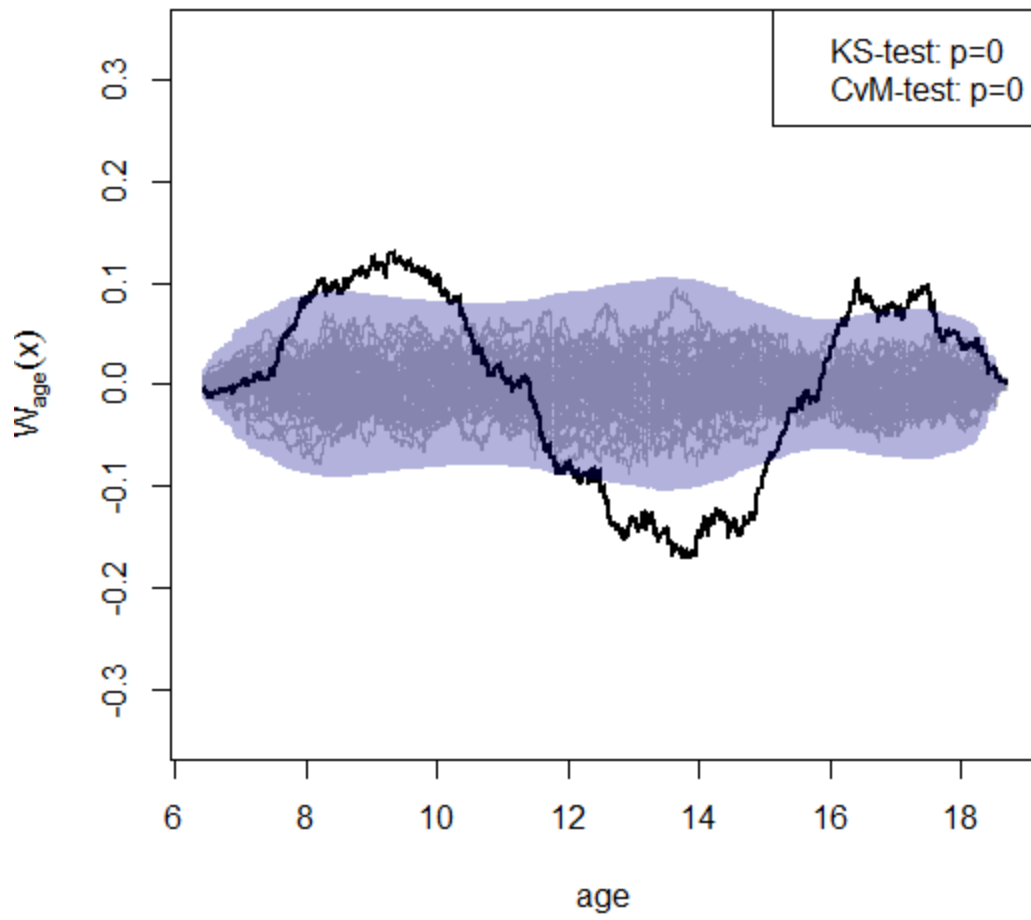
```
>  
> plot(fev1$age, resid(m3, type="r", level=0),  
+ main="Level 0 raw resid vs age, m3")  
> lines(lowess(fev1$age, resid(m3, type="r", level=0)))
```

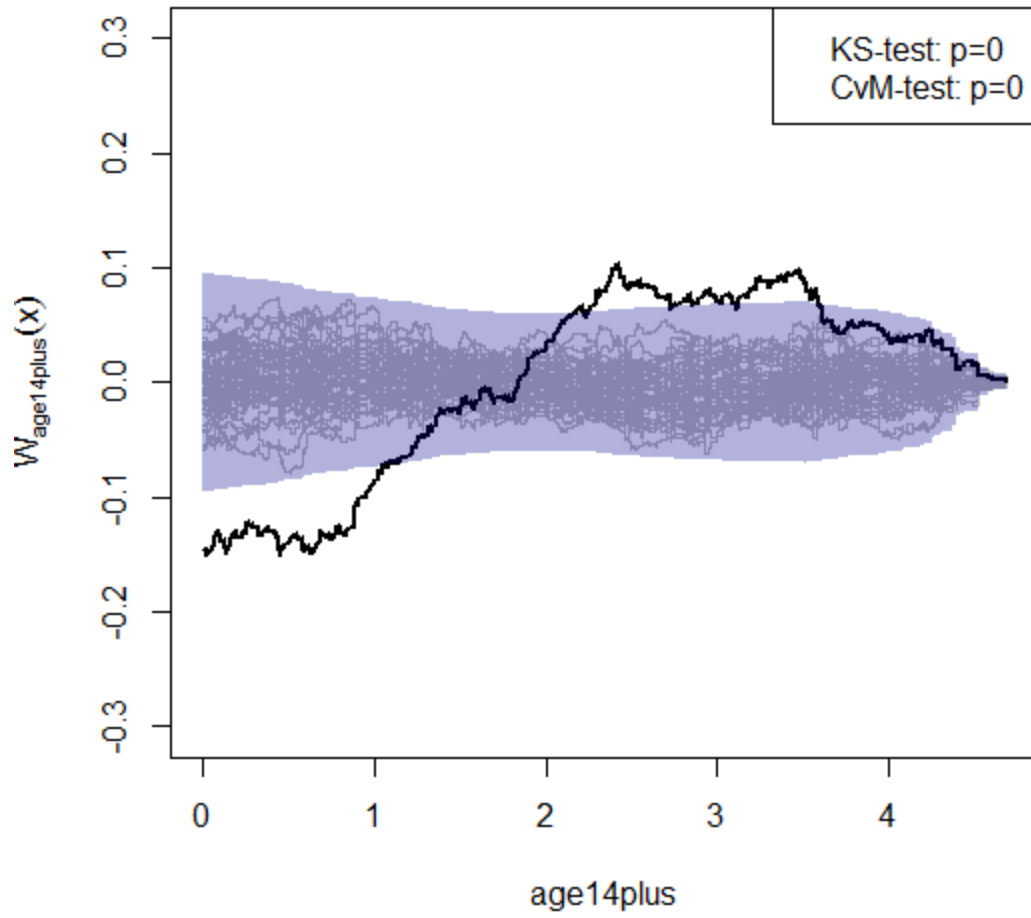
Level 0 raw resid vs age, m3

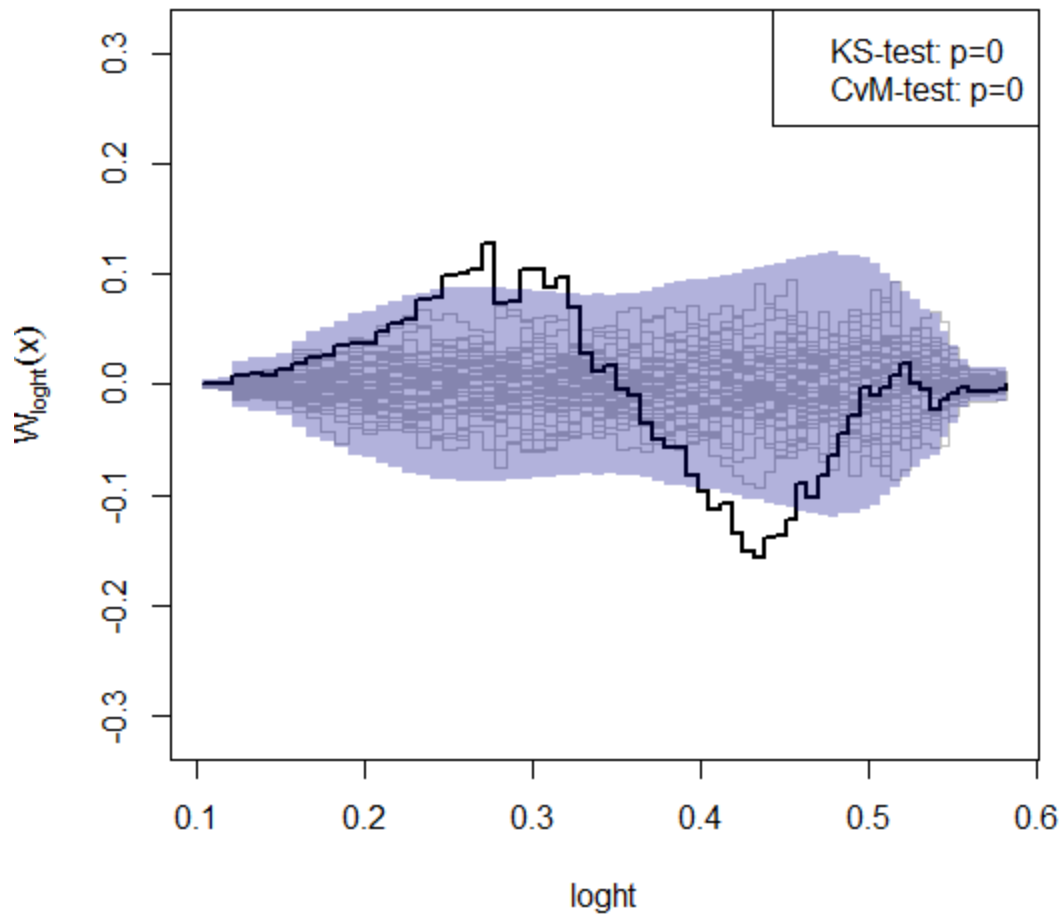


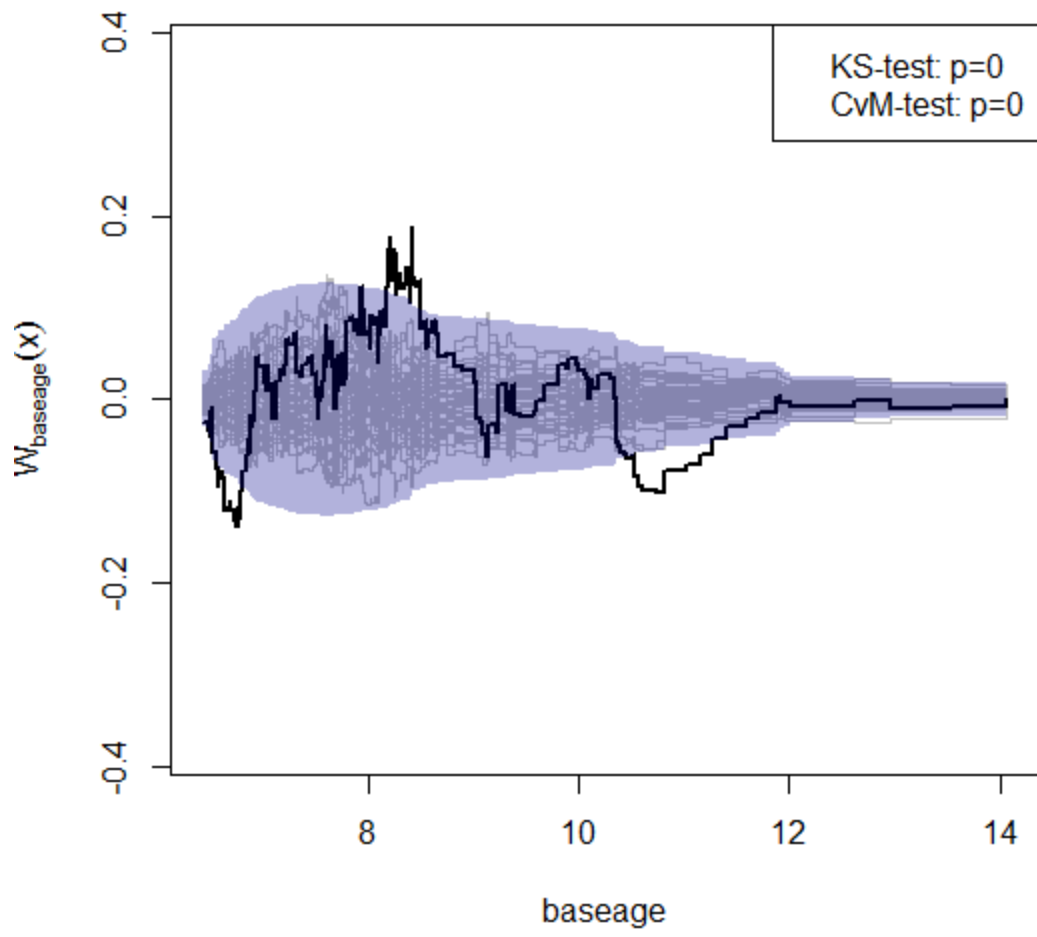
```
>  
> # These plots don't look much better!! There is still a wave in the plots.  
>  
> m3.lm <- lm(logfev1~age+age14plus+loght+baseage+logbaseht,data=fev1)  
> cres3 <- cumres(m3.lm,R=500)  
> plot(cres3)  
> # The wave in the resid vs fitteds and resid vs age plots for model  
> # m3 shows up in the cumulative residuals plots too. Model's mean is still  
> # not specified very well
```

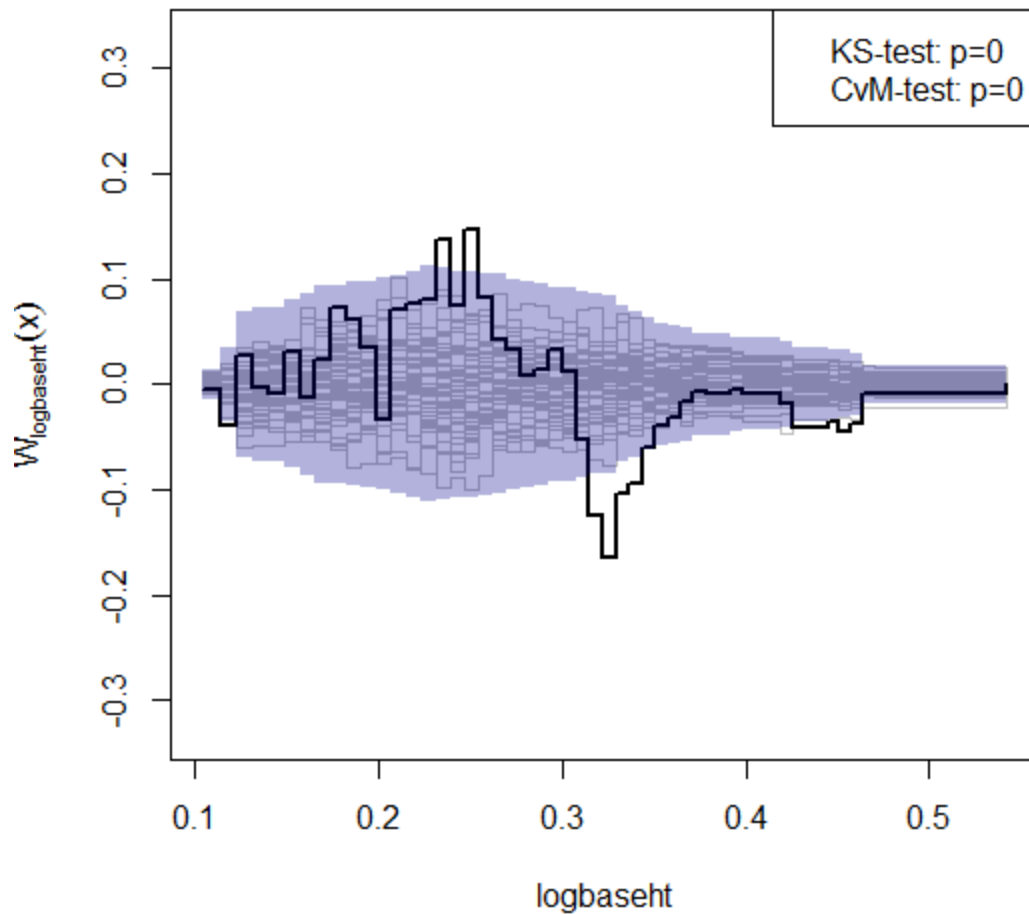












```

>
> # Now try to fix it by changing to a quadratic spline instead of a linear
> # spline. Call this model m3a:
>
> fev1$age2 <- fev1$age^2
> fev1$age14plus2 <- fev1$age14plus^2
> m3a <- lme(logfev1~age+age2+age14plus2+loght+baseage+logbaseht,data=fev1,
+ random=~1+age+age14plus2|ID)
> summary(m3a)
Linear mixed-effects model fit by REML
Data: fev1
      AIC      BIC logLik
-4659.266 -4580.951 2343.633

```

Random effects:

Formula: ~1 + age + age14plus2 | ID
 Structure: General positive-definite, Log-Cholesky parametrization

	StdDev	Corr
(Intercept)	0.123227678	(Intr) age
age	0.009721090	-0.683
age14plus2	0.003274632	0.434 -0.802
Residual	0.056927352	

Fixed effects: logfev1 ~ age + age2 + age14plus2 + loght + baseage + logbaseht

	Value	Std.Error	DF	t-value	p-value
(Intercept)	-0.0227229	0.05355166	1690	-0.424317	0.6714
age	-0.0295080	0.01092614	1690	-2.700681	0.0070
age2	0.0026083	0.00037162	1690	7.018779	0.0000
age14plus2	-0.0097583	0.00088517	1690	-11.024150	0.0000
loght	2.1314187	0.09244917	1690	23.055032	0.0000
baseage	-0.0152658	0.00815926	296	-1.870975	0.0623
logbaseht	0.2761781	0.16093615	296	1.716073	0.0872

Correlation:

	(Intr) age	age2	ag14p2	loght	baseag
age	-0.679				
age2	0.703	-0.968			
age14plus2	-0.355	0.300	-0.508		
loght	0.417	-0.795	0.630	0.200	
baseage	-0.275	-0.421	0.351	0.064	0.467
logbaseht	-0.004	0.420	-0.332	-0.113	-0.530 -0.849

Standardized Within-Group Residuals:

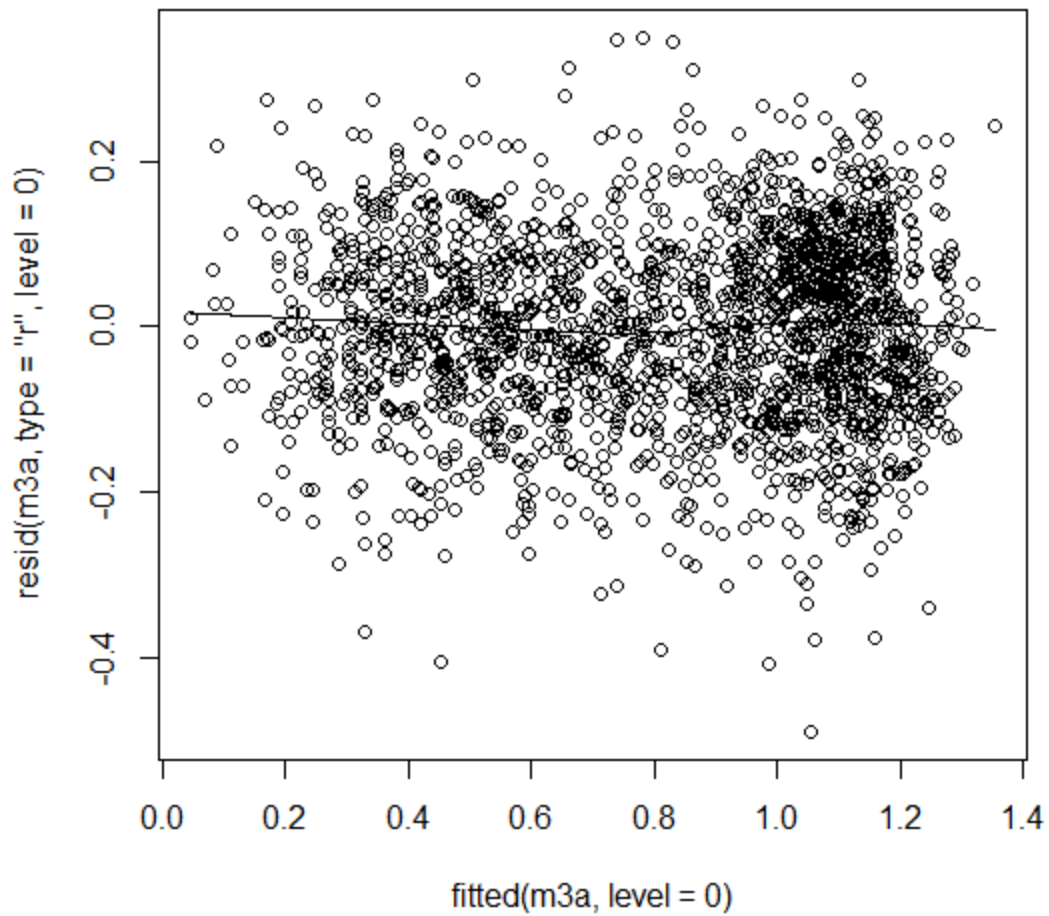
Min	Q1	Med	Q3	Max
-6.57102393	-0.47703330	0.07090773	0.60705236	2.76374181

Number of Observations: 1993

Number of Groups: 299

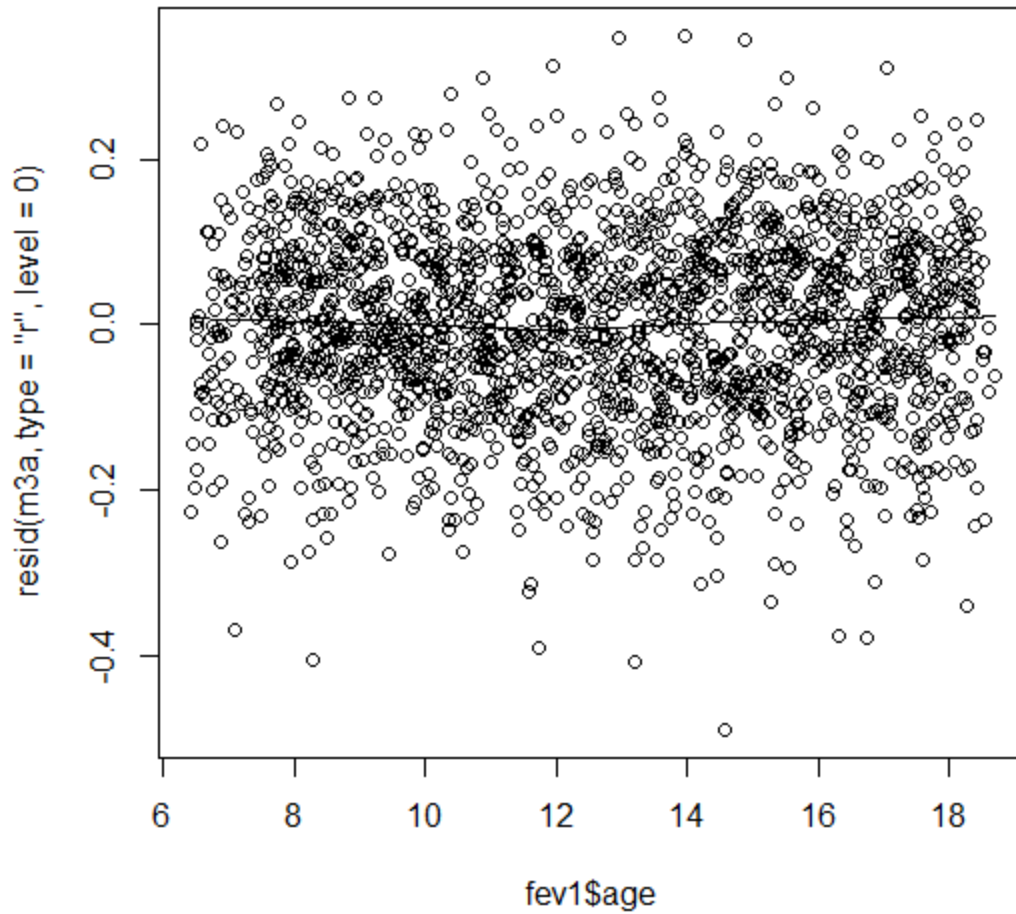
```
>
> plot(fitted(m3a,level=0),resid(m3a,type="r",level=0),
+ main="Level 0 raw resids vs fitteds, m3a")
> lines(lowess(fitted(m3a,level=0),resid(m3a,type="r",level=0)))
```

Level 0 raw resid vs fitteds, m3a

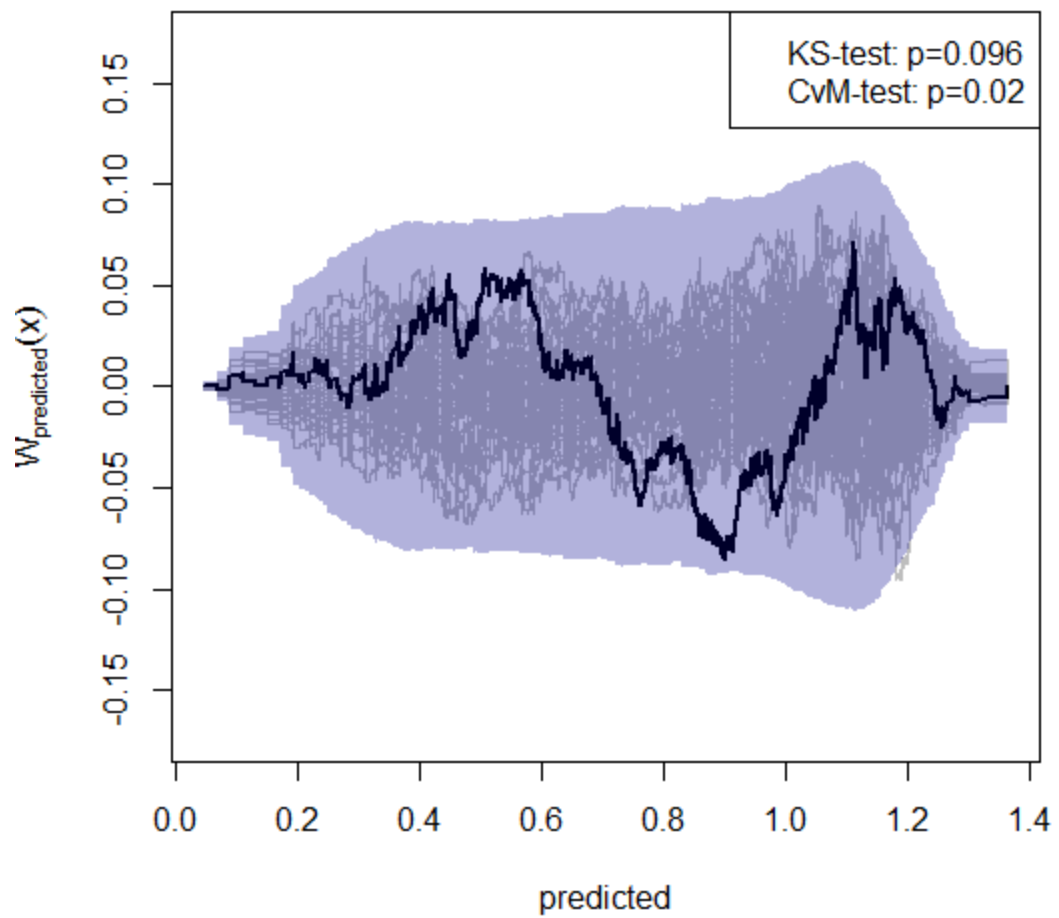


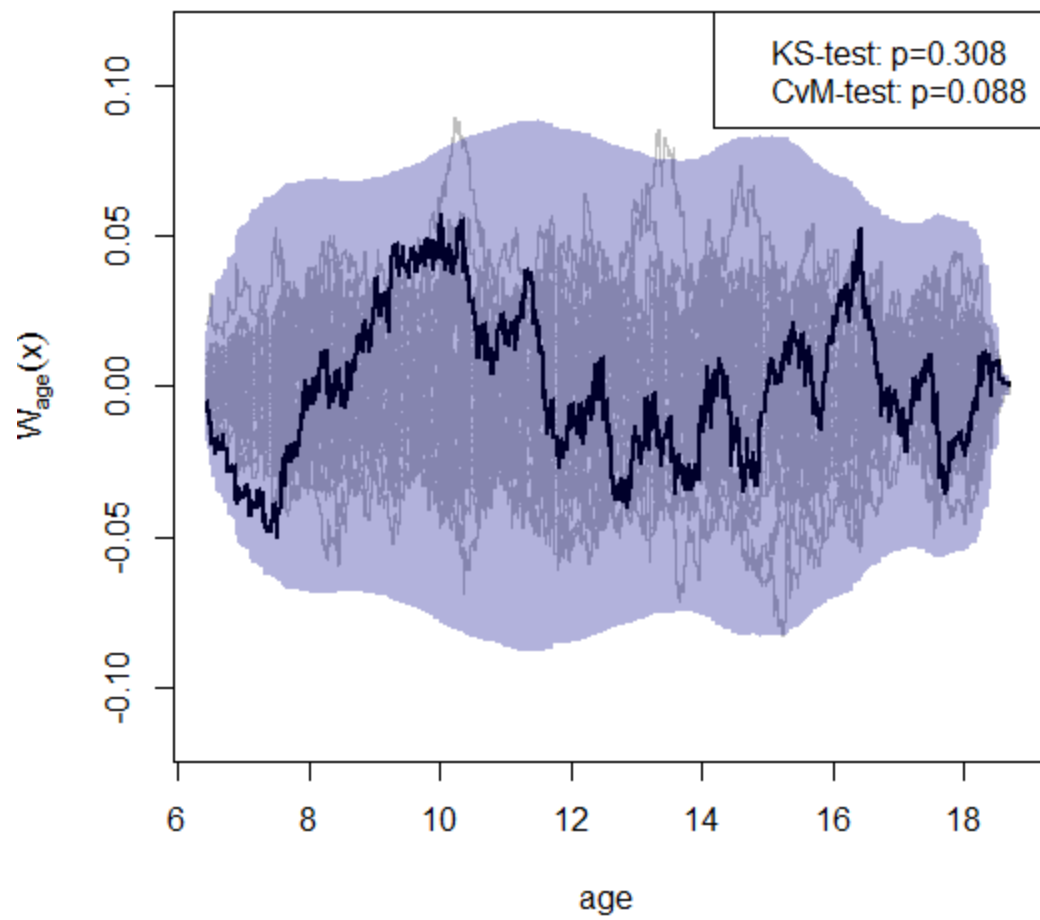
```
>  
> plot(fev1$age, resid(m3a, type="r", level=0),  
+ main="Level 0 raw resid vs age, m3a")  
> lines(lowess(fev1$age, resid(m3a, type="r", level=0)))
```

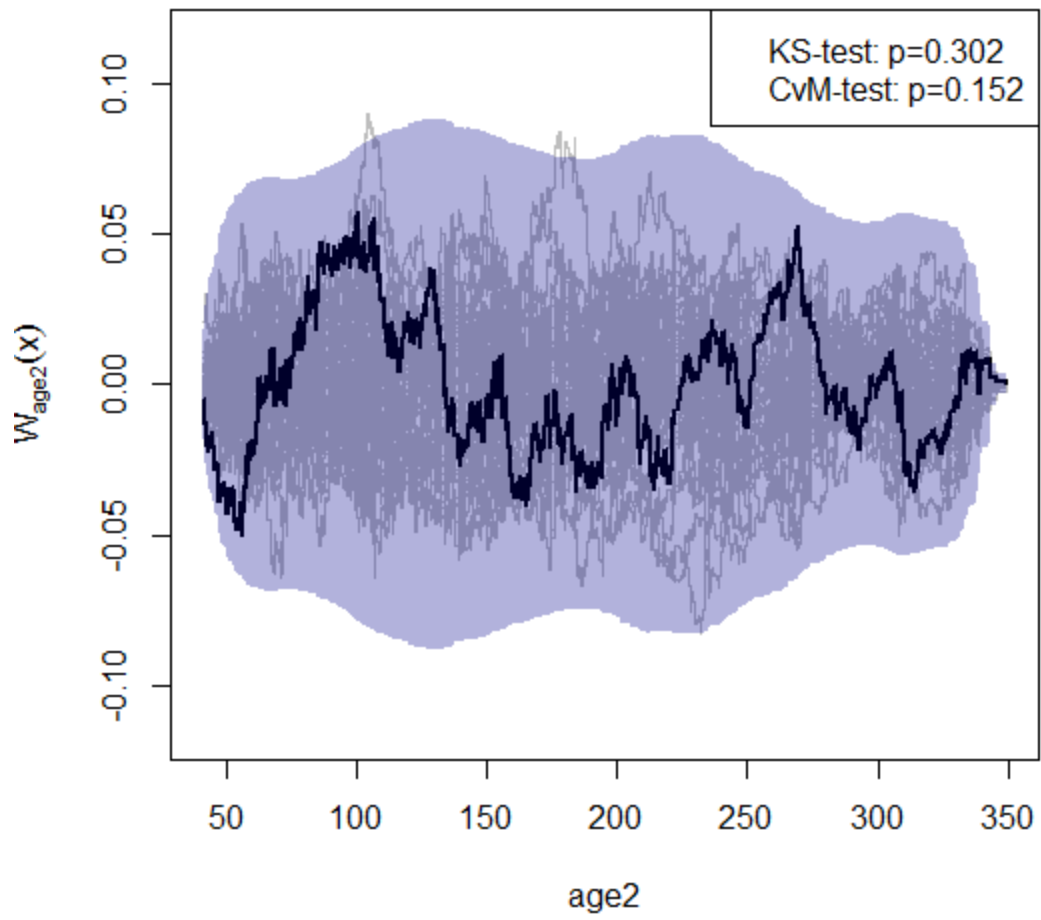
Level 0 raw resid vs age, m3a

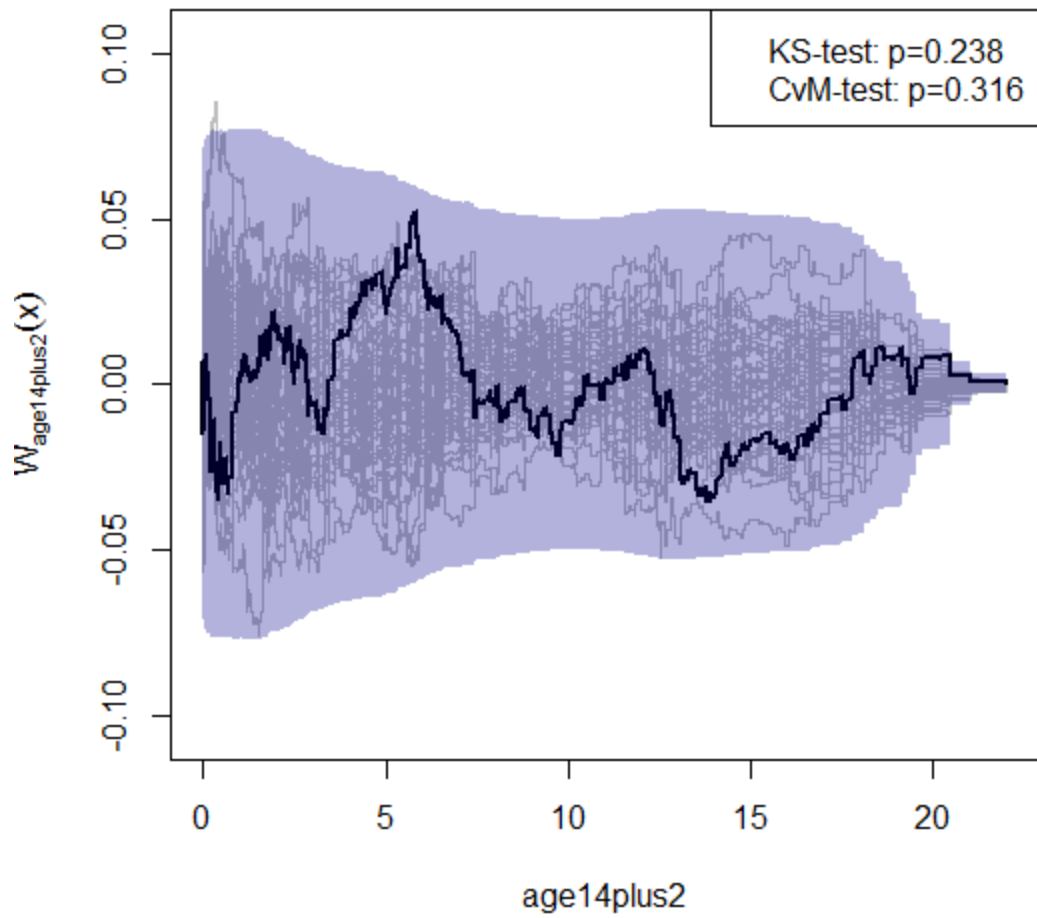


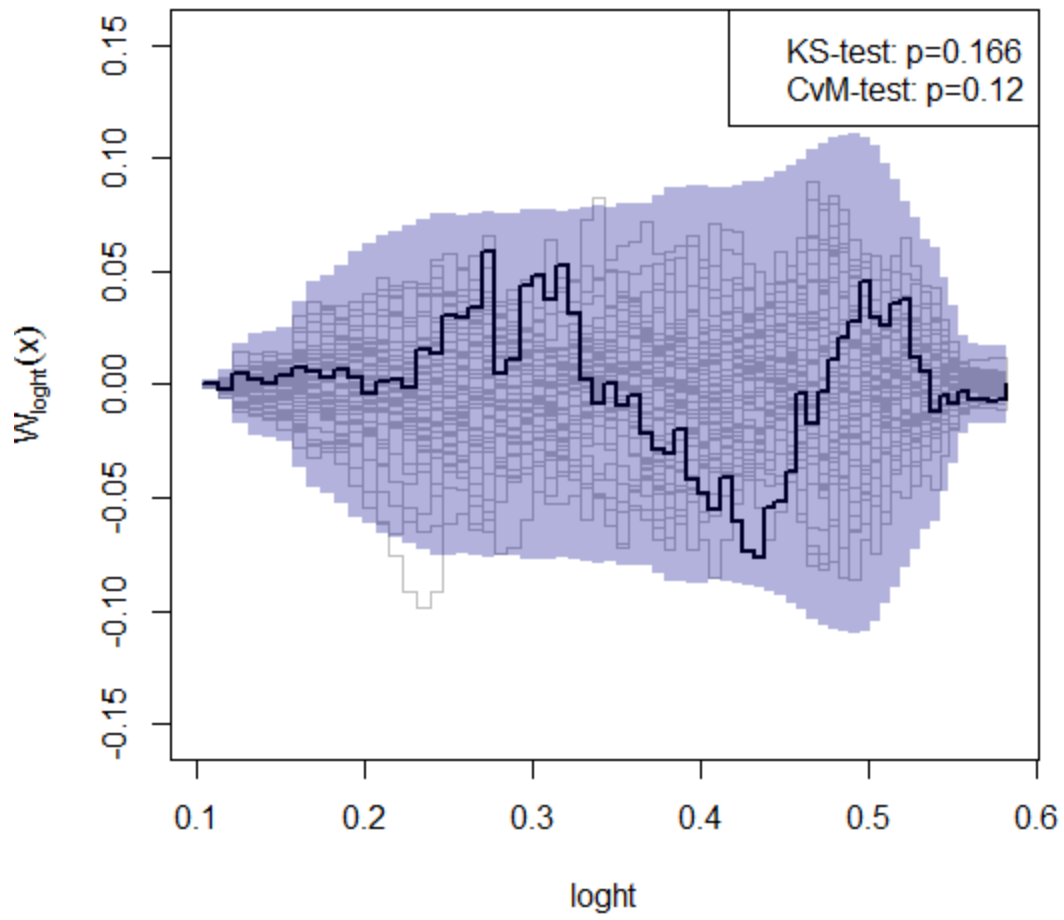
```
>  
>  
> m3a.lm <- lm(logfev1~age+age2+age14plus2+loght+baseage+logbaseht,data=fev1)  
> cres3a <- cumres(m3a.lm,R=500)  
> plot(cres3a)  
> # These now look pretty good
```

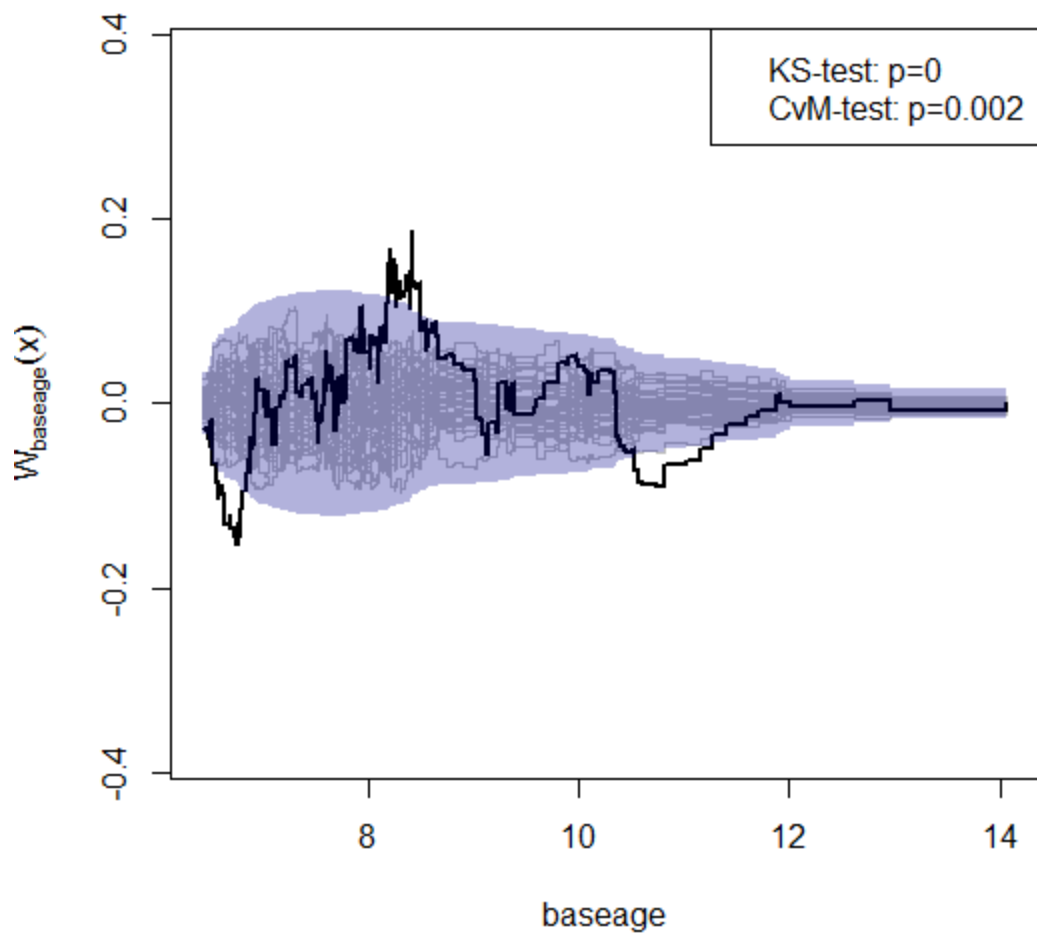


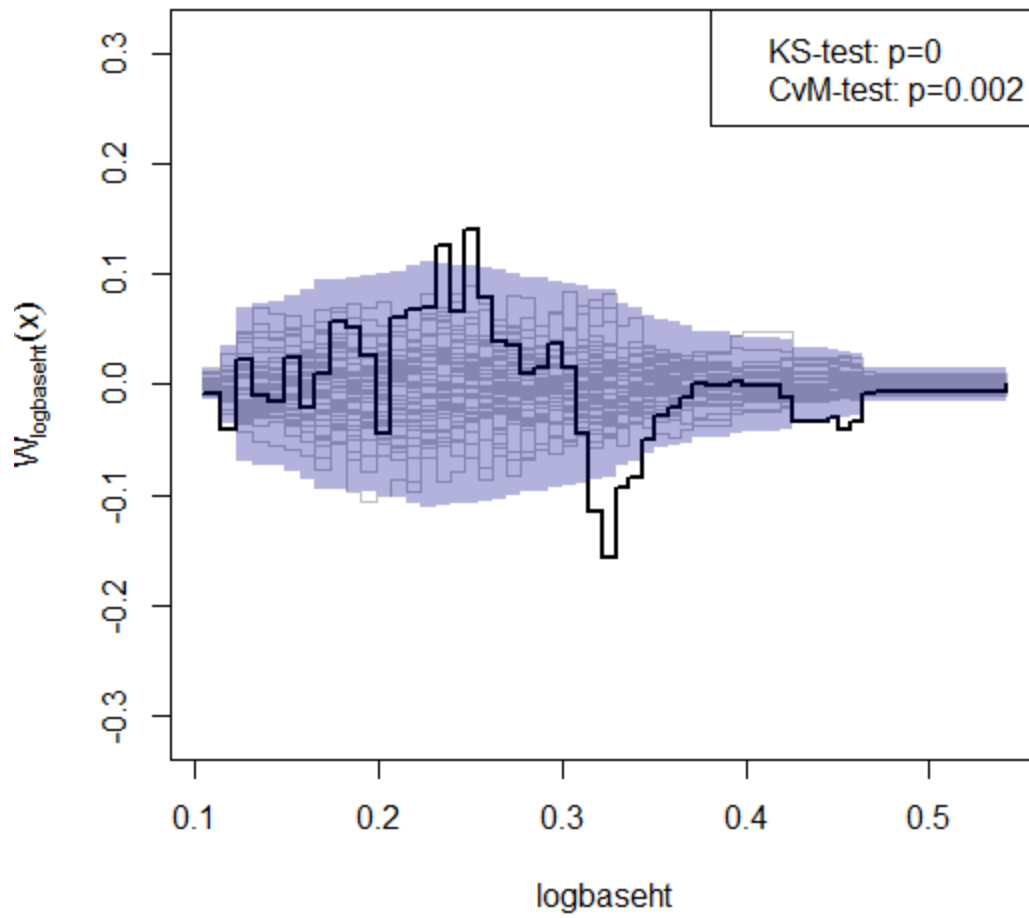




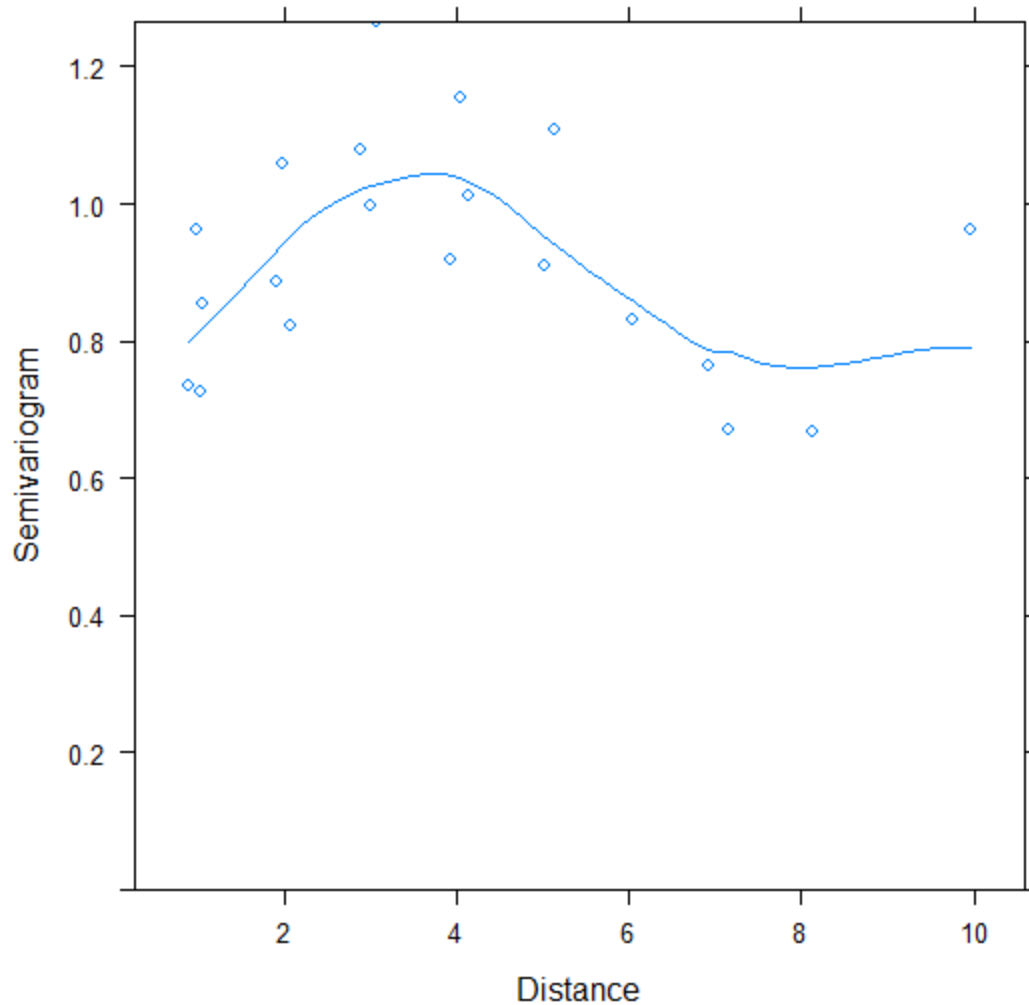








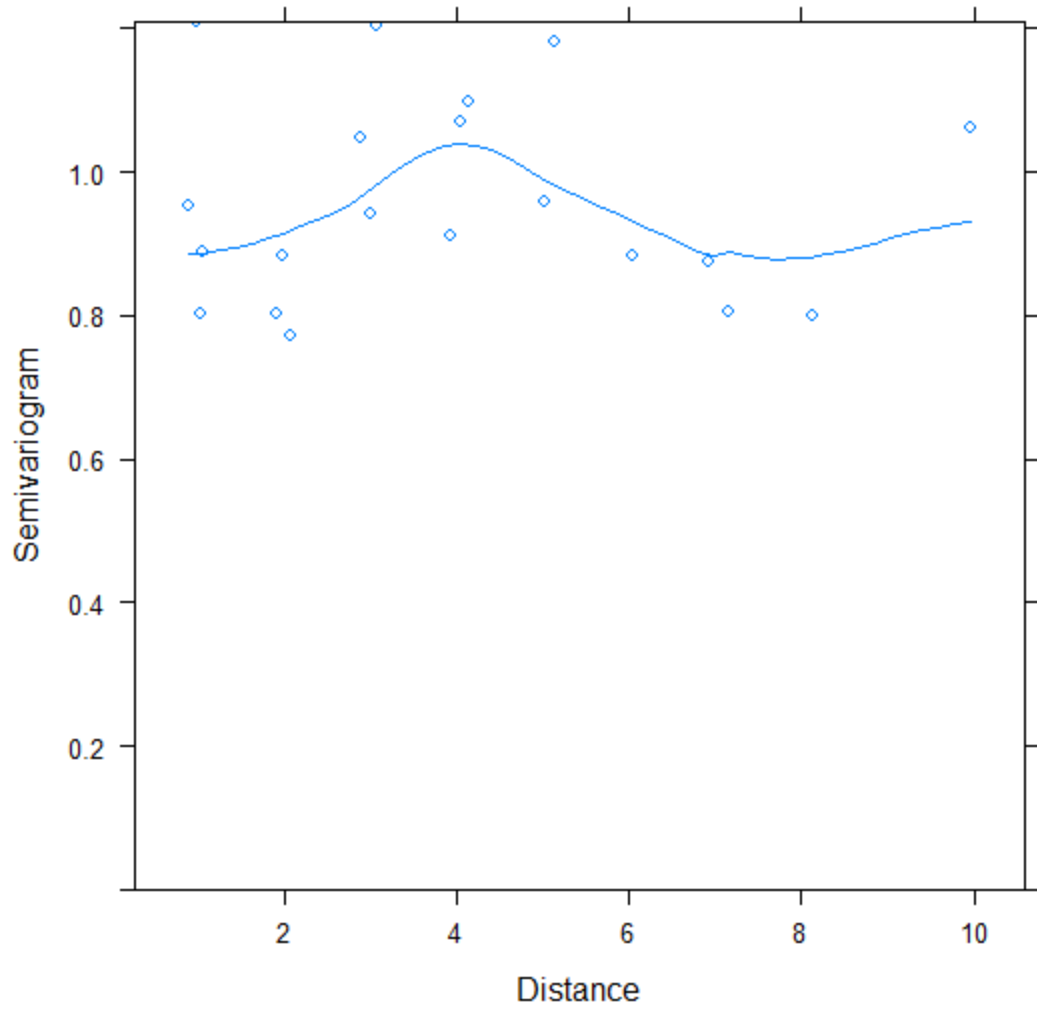
```
>  
> # Part 10.1.10  
>  
> plot(Variogram(m3a,form=~age,resType="n"))  
> # doesn't look great. Try adding a residual correlation structure.
```



```

>
> m4a <- update(m3a,corr=corGaus(form=~age|ID,nugget=F))
Error in lme.formula(fixed = logfev1 ~ age + age2 + age14plus2 + loght + :
  nlminb problem, convergence error code = 1
  message = iteration limit reached without convergence (10)
> # did not converge. Try reducing the random effects structure and
> # including residual spatial corr structure
>
> m4 <- update(m3a,random=~1+age|ID,corr=corGaus(form=~age|ID,nugget=F))
> AIC(m3a,m4) #m4 has smaller AIC
  df   AIC
m3a 14 -4659.266
m4   12 -4705.137
>
> plot(Variogram(m4,form=~age,resType="n"))
> # and better semivariogram

```



>
>
>