UPPSALA UNIVERSITET

A strategy for residual error modeling incorporating both scedasticity of variance and distribution shape

Anne-Gaëlle Dosne,

Ron J. Keizer, Martin Bergstrand, Mats O. Karlsson

Pharmacometrics Research Group

Department of Pharmaceutical Biosciences

Uppsala University

Sweden

Traditional error modeling
Models and assumptions

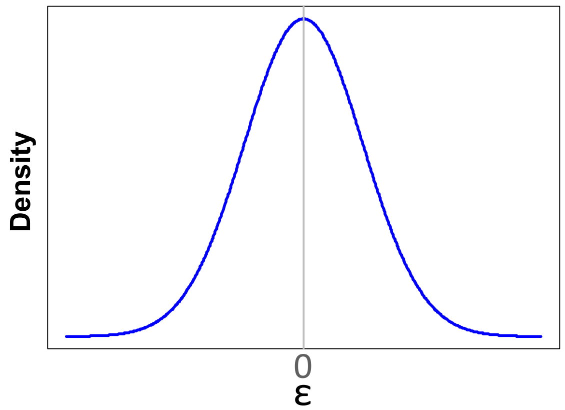
- **Models**

$$y = f(\theta, x, Z) + \varepsilon$$

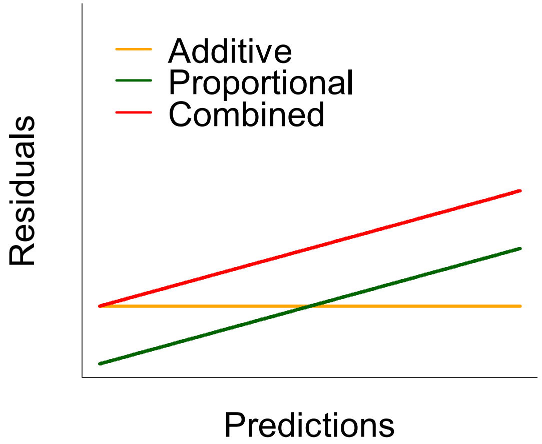
$$\varepsilon \sim N(0, var(\varepsilon))$$

$var(\varepsilon)$
σ^2_{add}
$\sigma^2_{prop} \times f(\theta, x, Z)^2$
$\sigma^2_{add} + \sigma^2_{prop} \times f(\theta, x, Z)^2$

- **Assumptions**



Normal distribution, mean 0



Fixed residual – prediction relationship

UPPSALA UNIVERSITET

What to do if assumptions do not hold?

Common answers, limitations, needs

Untransformed data
best case scenario

Fixed transformations
OFV not comparable
Time-consuming

Estimated transformations, different distribution
OFV comparable

Validity of assumptions on residuals

↓

Validity of likelihood calculation

↓

Validity of inference

UPPSALA UNIVERSITET

Alternative strategies

dTBS and Student's t-distribution

1 **dTBS**

Transform data and prediction so ε has a symmetric density

- Box-Cox (skewness) + Power term (scedasticity)

2 **t-distribution**

Change **distribution assumption**

- Student's t-distribution (heavy-tailed)

UPPSALA
UNIVERSITET

Residual error modeling with dynamic Transform Both Sides (dTBS)

UPPSALA UNIVERSITET

Dynamic Transform Both Sides

Box-Cox transformation in NLMEM

$$h(y) \begin{cases} \lambda \neq 0 & \dfrac{y^{\lambda} - 1}{\lambda} \\ \lambda = 0 & \log(y) \end{cases}$$

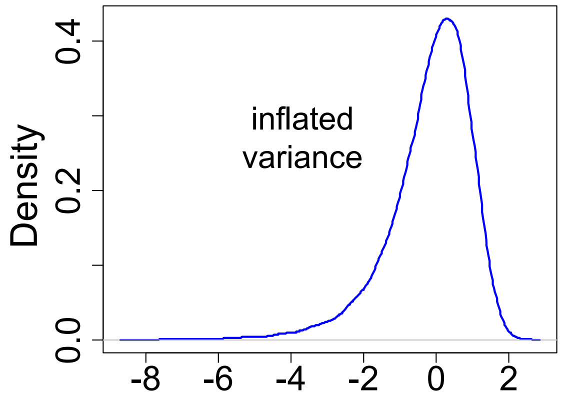
λ = 1	normal
λ = 0	log
λ > 1	*left* skewed
λ < 1	*right* skewed

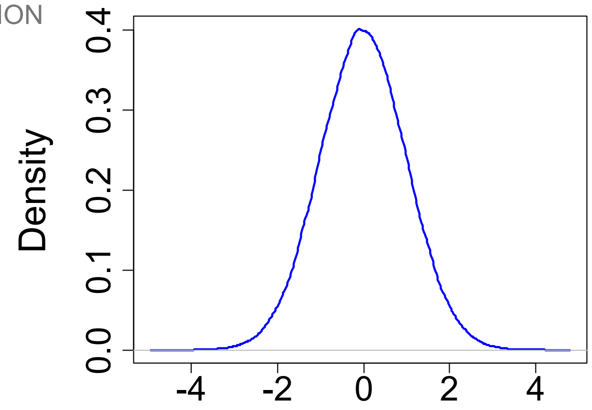
In NLMEM $\quad h(DV) = h(IPRED) + \varepsilon$

Untransformed scale:
skewed residuals

TRANSFORMATION

Transformed scale:
normal residuals





UPPSALA
UNIVERSITET

Dynamic Transform Both Sides
Likelihood calculation given the transformation

- Likelihood of *untransformed* data– **allows OFV comparison**[1,2]

$$p(y) = \int p(y|b)p(b)db$$

$$p(y|b) = \prod \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} \times e^{-\frac{(\boldsymbol{h}(\boldsymbol{y})-\boldsymbol{h}(\hat{y}))^2}{2\sigma^2}}}_{\text{Likelihood of transformed data}} \times \underbrace{\frac{\boldsymbol{d\ h}(\boldsymbol{y})}{\boldsymbol{dy}}}_{\text{First derivative of Box Cox: } y^{\lambda-1}}$$

[1] Carroll & Ruppert. *Transformation and Weighting in Regression*. Chapman and Hall. 1998
[2] Oberg & Davidian. *Estimating Data Transformations in Nonlinear Mixed Effects Models*. Biometrics. 2000

Dynamic Transform Both Sides

Implementation in NONMEM

- **Dynamic** estimation of λ[1]

- Additional files[1,2]
  – data transformed during estimation
  – redefinition of likelihood

$$\underbrace{-2LL'}_{\text{Likelihood of untransformed data}} = \underbrace{-2LL}_{\text{Likelihood of transformed data}} - \underbrace{2(\lambda - 1)\log(y)}_{\text{Contribution of transformation}}$$

[1] Frame B. *Within Subject Random Effect Transformations with NONMEM VI*. Wolverine Pharmacometrics Corporation. 2009.
[2] Robert Bauer, adaptation courtesy for NM7

Dynamic Transform Both Sides

Power term for dynamic heteroscedasticity

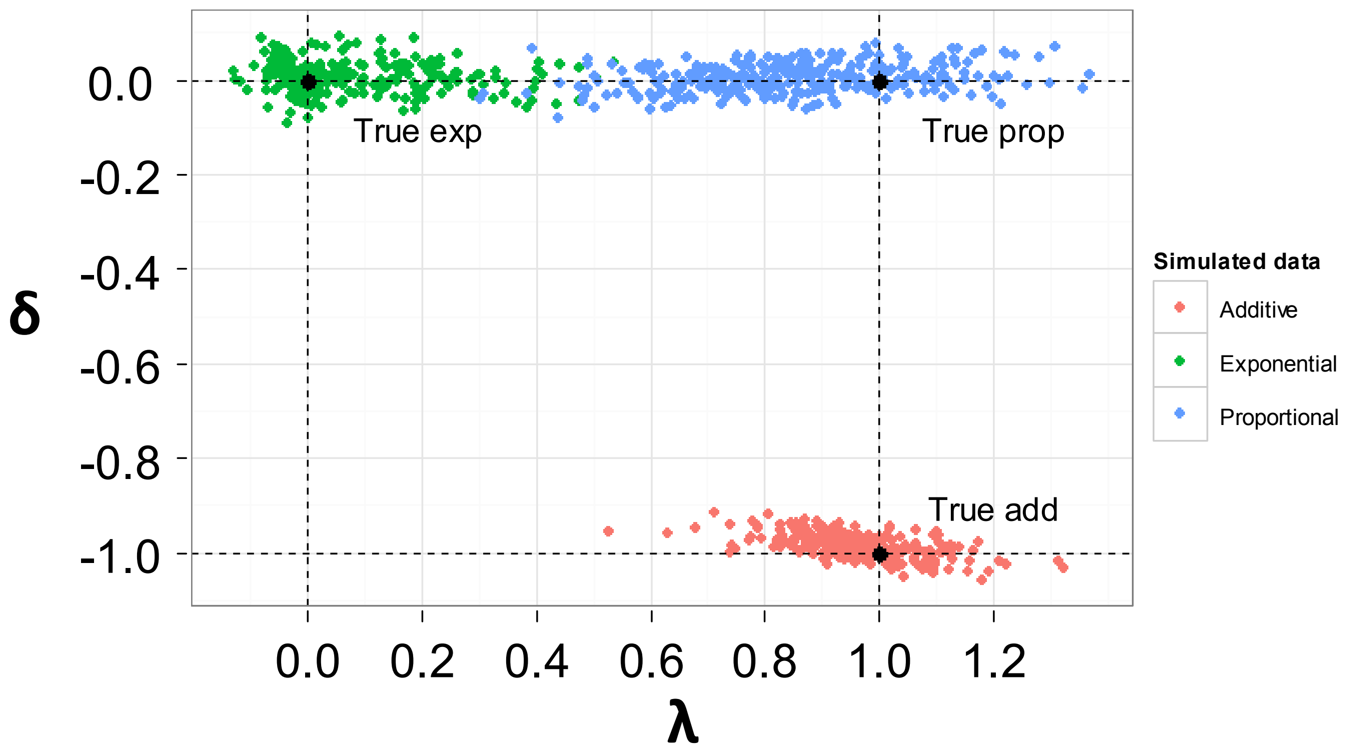
UPPSALA UNIVERSITET

- **Power dTBS**: $\frac{\mathrm{DV}^{\lambda}-1}{\lambda} = \frac{\mathrm{IPRED}^{\lambda}-1}{\lambda} + \mathrm{IPRED}^{\zeta} * \varepsilon$
  $\varepsilon \sim N(0, \sigma^2)$
- Dynamic heteroscedasticity
- Parameterization: $\zeta = \lambda + \delta$

Model	**Equation**	$\widetilde{sd}$ **(ε) untransformed scale**	**sd (ε) transformed scale**
Power	$DV = IPRED + IPRED^{\zeta} * \varepsilon$	$\sigma \times IPRED^{\zeta}$	NA
Box-Cox	$h(DV) = h(IPRED) + \varepsilon$	$\sigma \times IPRED^{1-\lambda}$	σ
dTBS	$h(DV) = h(IPRED) + \mathrm{IPRED}^{\zeta} * \varepsilon$	$\sigma \times IPRED^{1-\lambda+\zeta}$	$\sigma \times IPRED^{\zeta}$

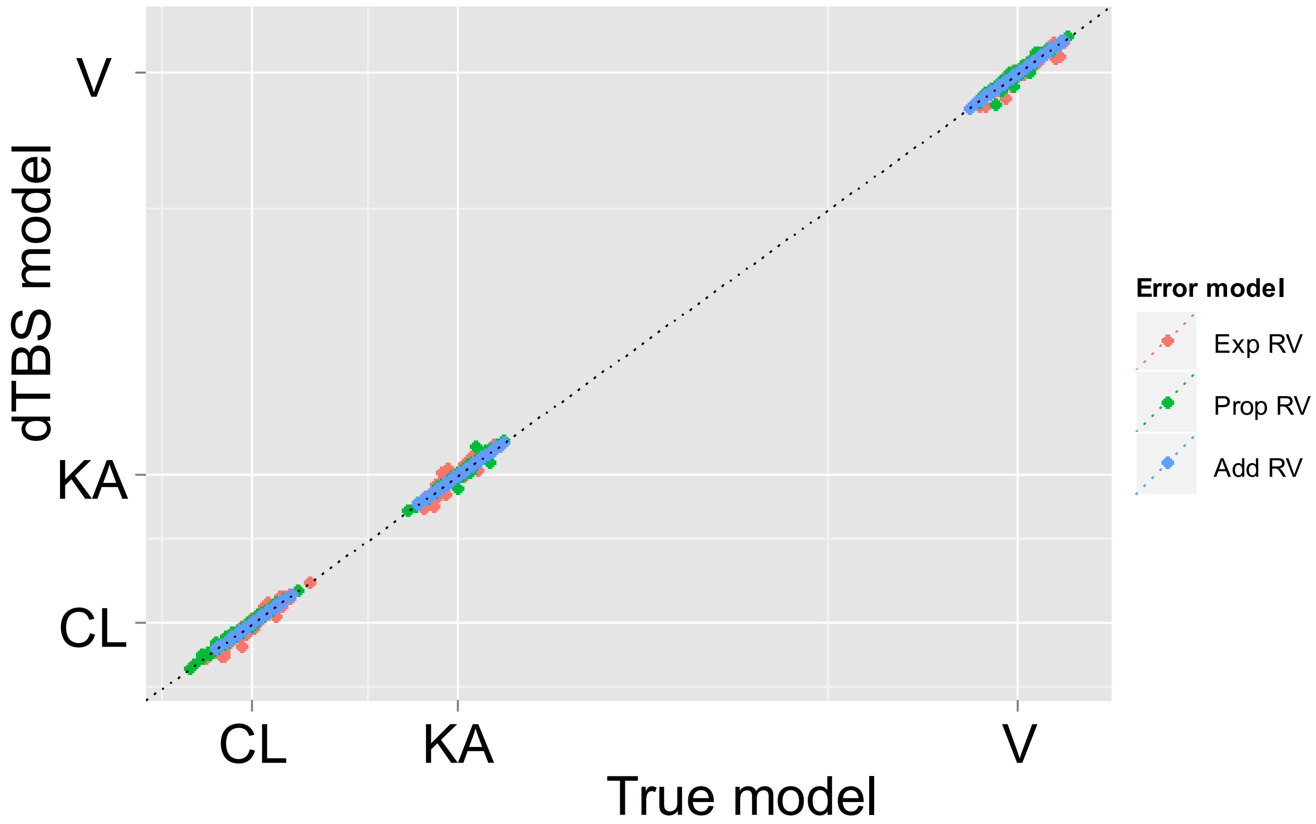
Dynamic Transform Both Sides
Simulation : estimation of λ and δ

- Simulation **additive, proportional and exponential** error models
- Estimation dTBS + true model



Dynamic Transform Both Sides

Simulation : satisfactory estimation of model parameters



Dynamic Transform Both Sides

Real data examples

Compound	Data	Model	RV model	Fixed transfo.	Obs.
ACTH & Cortisol	PD	turnover	combined	-	364
Cladribine	PK	IV 3 CMT	combined	-	488
Cyclophosphamide & metabolite	PK	4 CMT, CL induction	combined	-	383
Ethambutol	PK	2 CMT, transit	combined	-	1869
Moxonidine	PK	Oral 1 CMT	additive	Log	1021
Moxonidine	PD	E_{max}	additive	Log	1364
Paclitaxel	PD	Neutrophil	additive	Box-Cox	530
Pefloxacin	PK	IV 1 CMT	proportional	-	337
Phenobarbital	PK	IV 1 CMT	proportional	-	155
Prazosin	PK	Oral 1 CMT	proportional	-	887

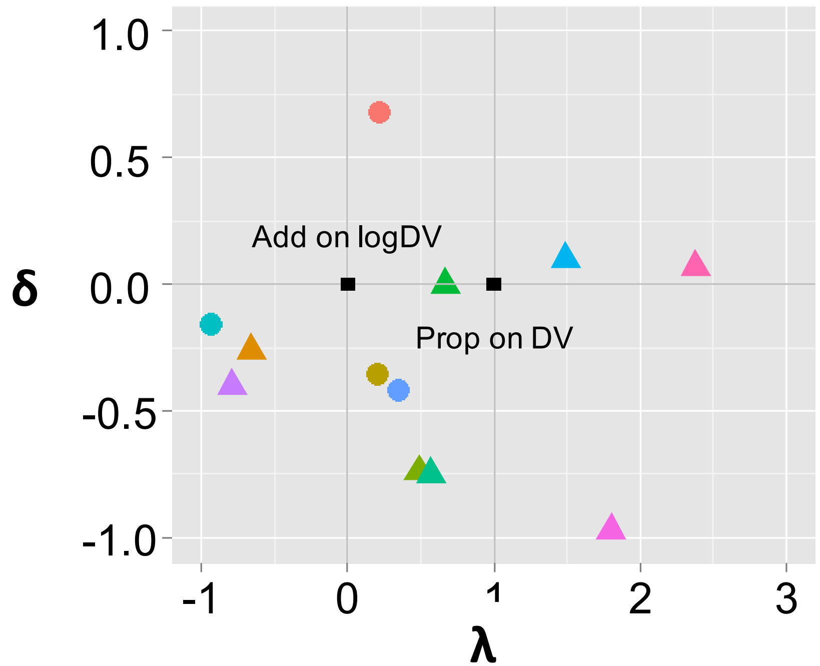
Dynamic Transform Both Sides

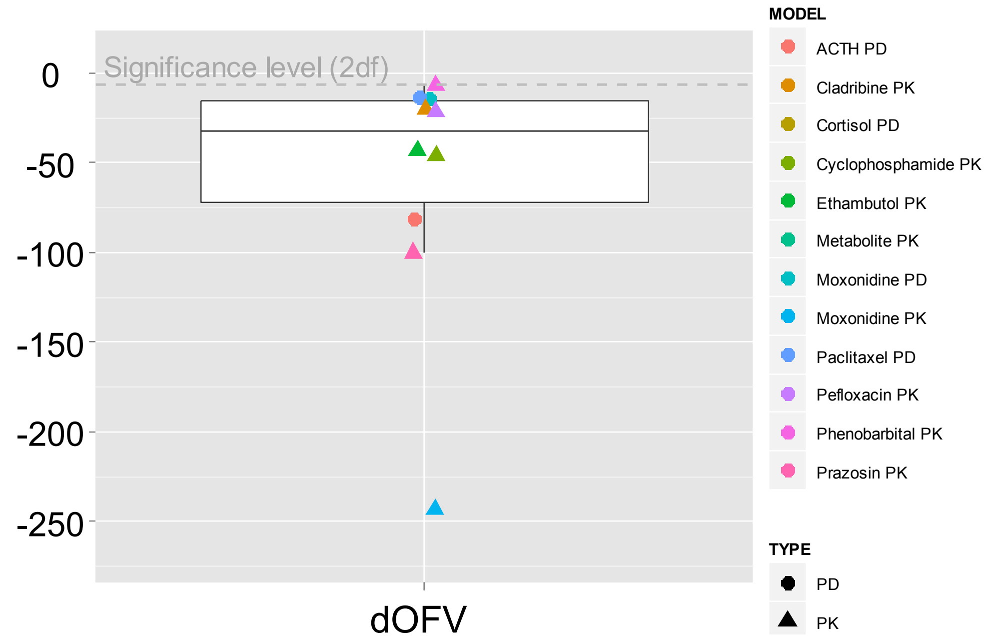
Real data examples: dTBS estimates & OFV drop

UPPSALA UNIVERSITET

λ	δ
[-0.94; 2.38]	[-0.97; 0.68]

- 100% dOFV > 5.99
- **Mean drop -60**



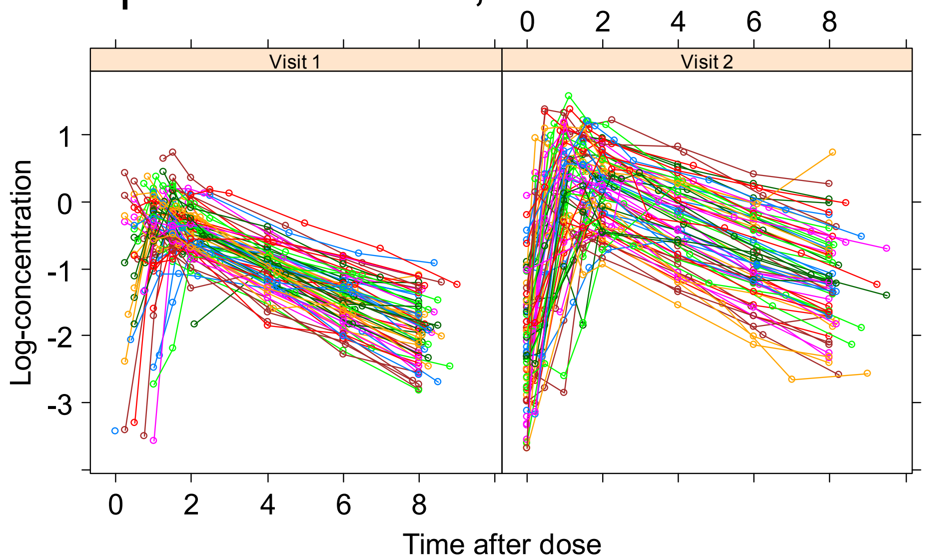


UPPSALA UNIVERSITET

Dynamic Transform Both Sides

Moxonidine PK example: data and model

- **Data:** rich
  - 74 patients, 1021 observations
  - 3 possible doses, 2 occasions



- **Model**:
  - Additive error on log-transformed data
  - 1 CMT, first-order absorption, lag time

UPPSALA UNIVERSITET

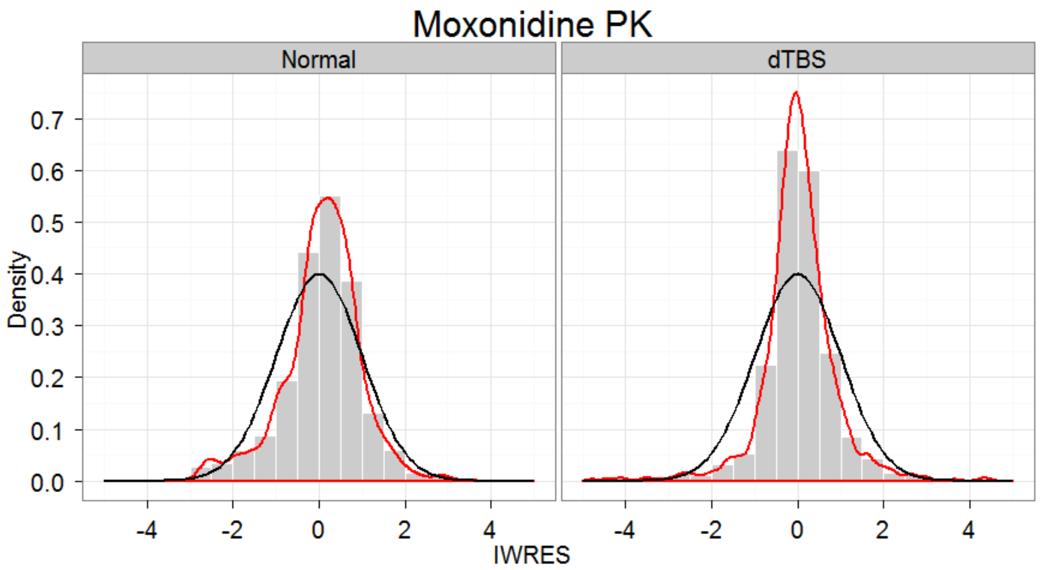
Dynamic Transform Both Sides

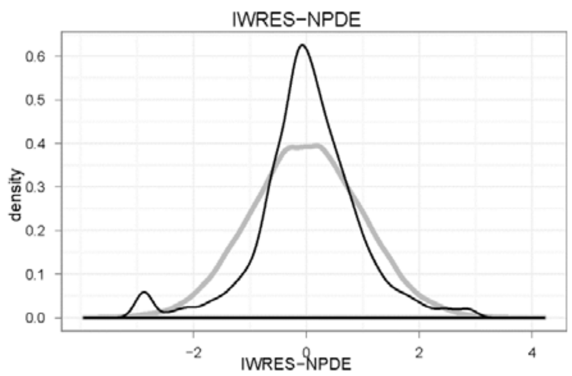
Moxonidine example: OFV and parameter estimates

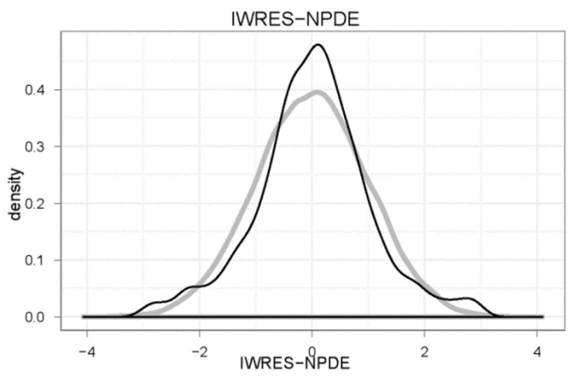
Parameter	Fixed log Value	Fixed log RSE(%)	dTBS Value	dTBS RSE(%)
OFV	-2173	-	-2416	-
dOFV	0	-	**- 243**	-
λ (-)	0	-	**1.5**	*6.3*
ζ (-)	0	-	**1.6**	-
Δ (ζ-λ)	0	-	0.1	*39*
CL ($L.h^{-1}$)	27	*1.2*	26	*3.4*
V (L)	110	*2.6*	108	*3.3*
KA (h^{-1})	4.5	*9*	4.9	*18*
LAG (h)	0.24	*1.3*	0.23	*2.3*
RV (na)	0.33	*1.3*	0.25	*4.1*
IIV CL (%)	27	*10*	25	*11*
Cor IIV CL-V (-)	0.74	-	0.70	-
IIV V (%)	24	*13*	24	*10*
IIV KA (%)	168	*12*	136	*14*
IOV CL (%)	12	*18*	17	*11*
IOV KA (%)	71	*18*	88	*16*

- Important OFV drop
- λ ≈ ζ ≈ 1.5: left skewness
- Parameter estimates changed (RV on different scales)

Dynamic Transform Both Sides
Moxonidine example: IWRES distribution

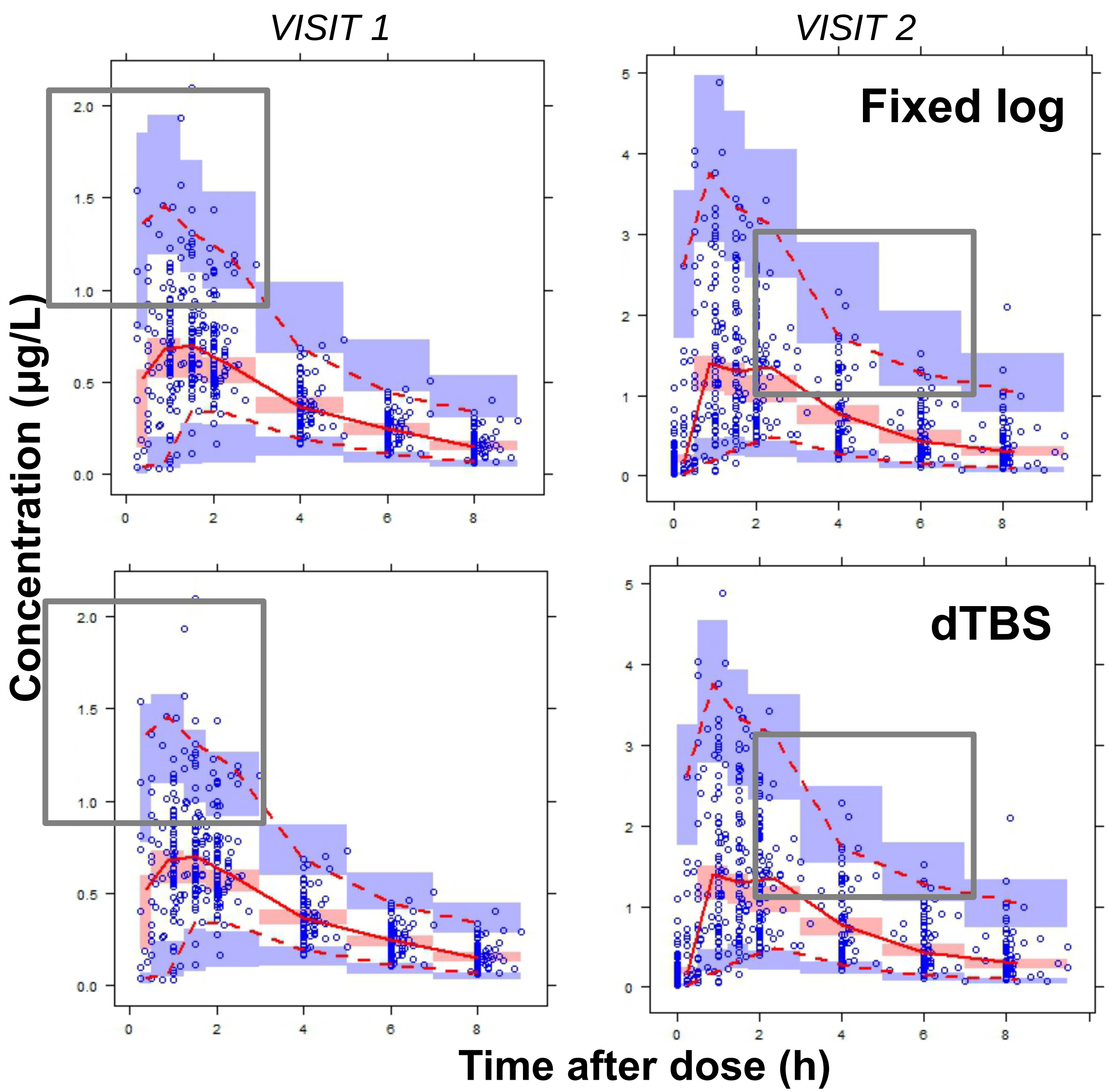






Dynamic Transform Both Sides

Moxonidine example: prediction properties



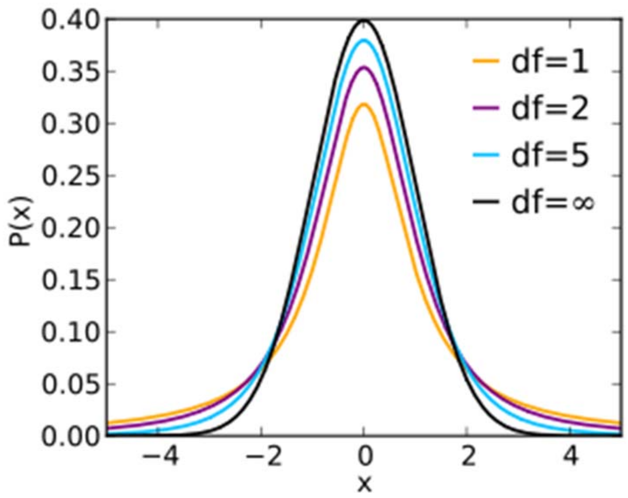
- **Narrower CI** with dTBS
- E.g. highest upper bound of CI_{95} 95^{th} perc. = **2 vs 1.6** (log vs dTBS)
- ***more precise predictions***

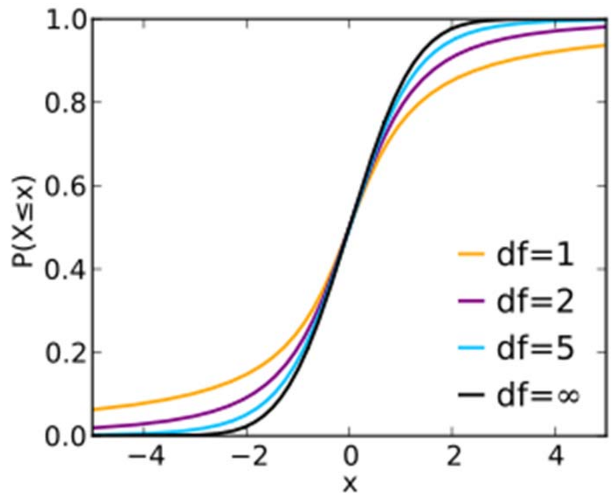
UPPSALA UNIVERSITET

Residual error modeling with Student's t-distribution

Student's t-distribution
Function characteristics

- Parameter: **degrees of freedom** ν (df)
- Symmetric around 0, heavy-tailed
- Variance = $\frac{\nu}{\nu-2} > 1$ for $\nu > 2$
- Approaches normal distribution when $\nu \to \infty$





UPPSALA UNIVERSITET

Student's t-distribution

Implementation: redefinition of likelihood

- **Estimation**

- Likelihood based on pdf of t-distribution[1]

$$\boldsymbol{p}(\boldsymbol{y}|\nu,\sigma,\mu) = \frac{\Gamma(\frac{\nu+\mathbf{1}}{\mathbf{2}})}{\Gamma(\frac{\nu}{\mathbf{2}})\sqrt{\boldsymbol{\pi}\nu\boldsymbol{\sigma}^{2}}}\left(\mathbf{1}+\frac{\mathbf{1}}{\nu}\frac{(\boldsymbol{x}-\boldsymbol{\mu})^{2}}{\boldsymbol{\sigma}^{2}}\right)^{-\frac{\nu+\mathbf{1}}{\mathbf{2}}}$$

- Function Γ: Nemes approximation[2]
- ν ≥ 3: stability and full distribution definition

[1] Jackman, Simon. *Bayesian Analysis for the Social Sciences*. Wiley. 2009
[2] Nemes, Gergő. *New asymptotic expansion for the Gamma function*. Archiv der Mathematik . 2010

UPPSALA UNIVERSITET

Student's t-distribution
Real data examples

Compound	Data	Model	RV model	Fixed transfo.	Obs.
Moxonidine	PK	Oral 1 CMT	additive	Log	1021
Pefloxacin	PK	IV 1 CMT	proportional	-	337
Phenobarbital	PK	IV 1 CMT	proportional	-	155
Prazosin	PK	Oral 1 CMT	proportional	-	887

Compound	dOFV	ν
Moxonidine	- 400	3
Pefloxacin	- 16	4.7
Phenobarbital	0	∞
Prazosin	- 169	3

UPPSALA UNIVERSITET

Student's t-distribution

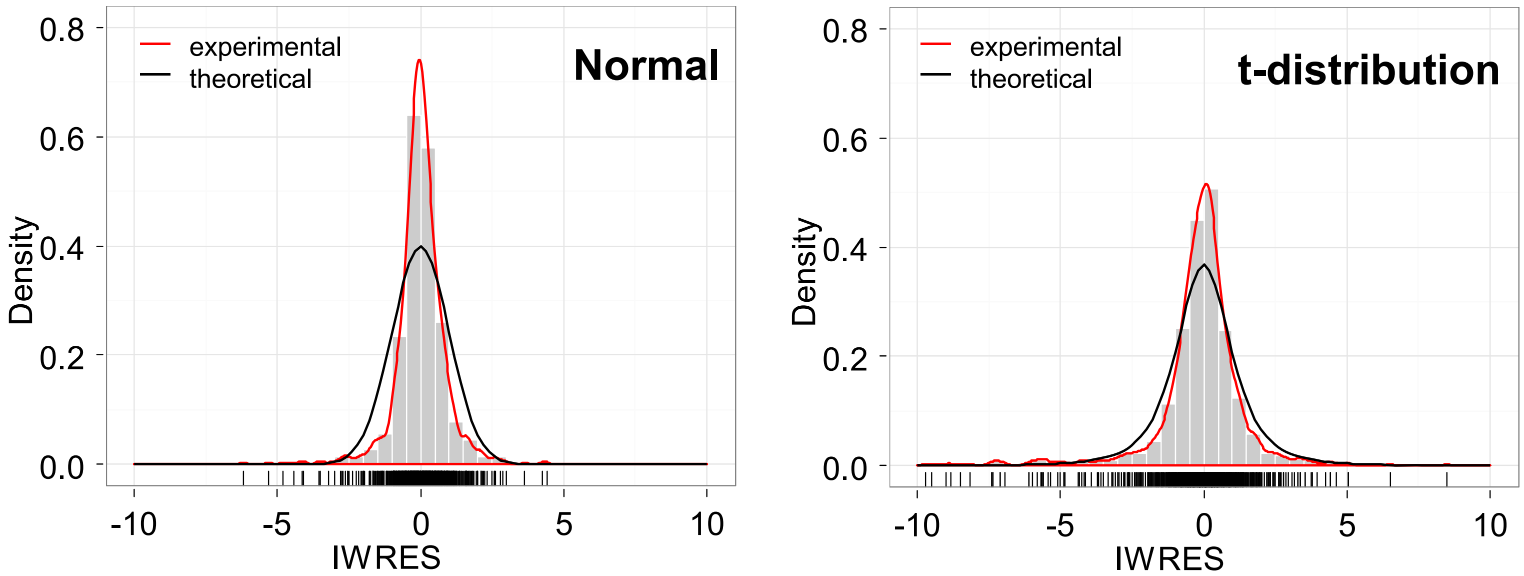
Moxonidine example: OFV and parameter estimates

Parameter	Normal	t-distribution
OFV	1221	820
dOFV	0	**- 400**
DF	∞	**3**
CL ($L.h^{-1}$)	26	26
V (L)	107	102
KA (h^{-1})	5.4	4.7
LAG (h)	0.24	0.24
RV (na)	0.33	0.17
IIV CL (%)	27%	28%
Cor IIV CL-V (-)	0.74	0.77
IIV V (%)	24%	23%
IIV KA (%)	162%	150%
IOV CL (%)	12%	12%
IOV KA (%)	63%	20%

- Important OFV drop
- DF at lower boundary
- Variance = 3

- Changes in absorption parameters

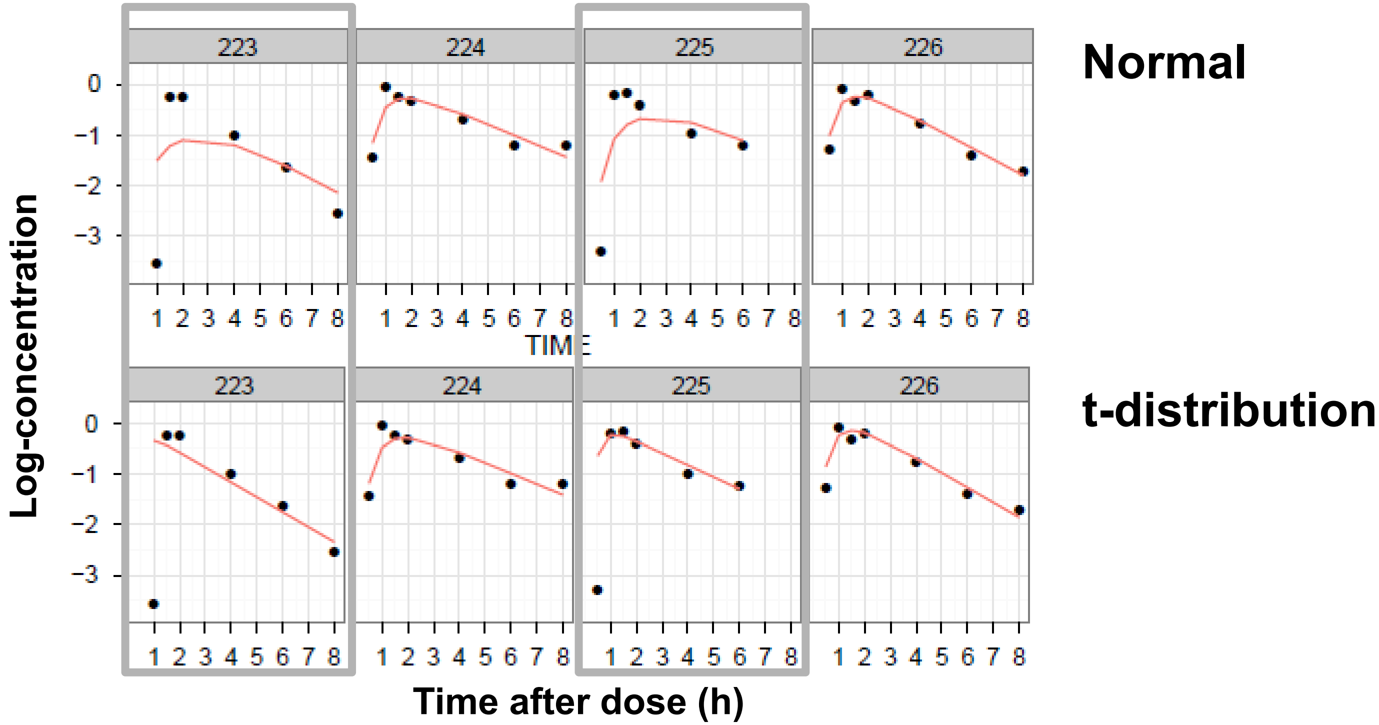
Student's t-distribution
Moxonidine PK: Residual distribution



- Better agreement between experimental and theoretical IWRES distribution for t-distribution

UPPSALA UNIVERSITET

Student's t-distribution
Moxonidine example: Individual fits



- Better global fit by allowing some predictions to be further away from observations

UPPSALA UNIVERSITET

Conclusion

New possibilities for residual error modeling

dTBS

- **Dynamic**
- OFVs comparable
- Handles **skewness** and complex **scedasticity**
- Data: PK, particular assays…
- **Universal** model

t-distribution

- **Dynamic**
- OFVs comparable
- Way to leverage **outlier** information
- Data: multiple baselines, protocol deviations…

- Could be used jointly
- Can improve quality of parameter estimates and predictions

UPPSALA UNIVERSITET

Acknowledgements

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement n° 115156, resources of which are composed of financial contributions from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. The DDMoRe project is also supported by financial contribution from Academic and SME partners. This work does not necessarily represent the view of all DDMoRe partners.

- Uppsala Pharmacometrics Research Group
- Novartis for financial support

Innovative Medicines Initiative

efpia

Drug Disease Model Resources ddmore

UPPSALA
UNIVERSITET

Additional slides

- *dTBS ccontra and contr files*
- *dTBS model file code*
- *dTBS simulation code*
- *t-distribution simulation code*
- *Type I error investigation for dTBS and t-distribution*

UPPSALA UNIVERSITET

Dynamic Transform Both Sides

ccontra and contr files

ccontra

```
 subroutine ccontr (icall,c1,c2,c3,ier1,ier2)
      USE ROCM_REAL,   ONLY: theta=>THETAC,y=>DV_ITM2
      USE NM_INTERFACE,ONLY: CELS
!      parameter (lth=40,lvr=30,no=50)
!      common /rocm0/ theta (lth)
!      common /rocm4/ y
!      double precision c1,c2,c3,theta,y,w,one,two
      double precision c1,c2,c3,w,one,two
      dimension c2(:),c3(:,:)
      data one,two/1.,2./
      if (icall.le.1) return
      w=y(1)

      if(theta(7).eq.0) y(1)=log(y(1))
      if(theta(7).ne.0) y(1)=(y(1)**theta(7)-one)/theta(7)

      call cels (c1,c2,c3,ier1,ier2)
      y(1)=w
      c1=c1-two*(theta(7)-one)*log(y(1))

      return
      end
```

DV transformation

Redefinition of log-likelihood

contr

```
subroutine contr
(icall,cnt,ier1,ier2)
      double precision cnt
      call ncontr
(cnt,ier1,ier2,l2r)
      return
      end
```

Dynamic Transform Both Sides

Model file example: Modification of $ERROR

UPPSALA UNIVERSITET

Power RUV

Transform. of IPRED

```
$ ERROR

IPR1 = A(2)/V
IF(IPR1.LE.0) IPR1 = 0.001

LAMBDA = THETA(7)
ZETA   = LAMBDA + THETA(6)

W    = THETA(5)*IPR1**ZETA

IPRED = IPR1

IPRTR=IPRED
 IF (LAMBDA .NE. 0 .AND. IPRED .NE.0) THEN
    IPRTR=(IPRED**LAMBDA-1)/LAMBDA
 ENDIF
 IF (LAMBDA .EQ. 0 .AND. IPRED .NE.0) THEN
    IPRTR=LOG(IPRED)
 ENDIF
 IF (LAMBDA .NE. 0 .AND. IPRED .EQ.0) THEN
    IPRTR=-1/LAMBDA
 ENDIF
 IF (LAMBDA .EQ. 0 .AND. IPRED .EQ.0) THEN
    IPRTR=-1000000000
 ENDIF
 IPRED=IPRTR
```

Transform. of IWRES

```
 IRES  = IPRED-DV
 IWRES = IRES/W

IWRTR=IWRES

 IF (LAMBDA.NE.0 .AND. DV.NE.0 .AND.
    W.NE.0) THEN
   IWRTR=((DV**LAMBDA-1)/LAMBDA-IPRED)/W
 ENDIF
 IF (LAMBDA.EQ.0 .AND. DV.NE.0 .AND.
    W.NE.0) THEN
   IWRTR=(LOG(DV)-IPRED)/W
 ENDIF
 IF (LAMBDA.NE.0 .AND. DV.EQ.0 .AND.
    W.NE.0) THEN
   IWRTR=(-1/LAMBDA-IPRED)/W
 ENDIF
 IF (LAMBDA.EQ.0 .AND. DV.EQ.0 .AND.
    W.NE.0) THEN
   IWRTR=(-1000000000-IPRED)/W
 ENDIF

IWRES=IWRTR

Y=IPRED+EPS(1)*W
```

UPPSALA
UNIVERSITET

Dynamic Transform Both Sides

Model file : other changes and simulation code

```
$SUBROUTINE ADVAN2 TRANS1  CONTR=contr.txt CCONTR=ccontra_nm7.txt
```

Additional files needed

```
$THETA  0.0001               ;DELTA_ZETA
$THETA  1                    ;LAMBDA

$ESTIMATION METHOD=1 INTER MAXEVALS=9999
```

dTBS SIMULATION ON UNTRANSFORMED SCALE

```
IF (ICALL.EQ.4 .AND. LAMBDA.EQ.0) THEN
  Y=EXP(Y)
ENDIF
IF (ICALL.EQ.4 .AND. LAMBDA.NE.0) THEN
  Y=((Y*LAMBDA)+1)**(1/LAMBDA)
ENDIF
```

UPPSALA UNIVERSITET

Student's t-distribution
Model file example

```
$ERROR
DF    = THETA(5)            ; degrees of freedom of Student distribution
W = THETA(4)*IPRED
SIG1  = W                   ; scaling factor for standard deviation of RUV
IWRES=(DV-IPRED)/SIG1

PHI=(DF+1)/2                ; Approximation of gamma function
INN=PHI+1/(12*PHI-1/(10*PHI))
GAMMA=SQRT(2*3.14159265/PHI)*(INN/EXP(1))**PHI

PHI2=DF/2                   ; Approximation of gamma function
INN2=PHI2+1/(12*PHI2-1/(10*PHI2))
GAMMA2=SQRT(2*3.14159265/PHI2)*(INN2/EXP(1))**PHI2

COEFF=GAMMA/(GAMMA2*SQRT(DF*3.14159265))/SIG1
BASE=1+IWRES*IWRES/DF
IF(BASE.EQ.0) BASE=0.000001
POW=-(DF+1)/2
L=COEFF*BASE**POW
Y=-2*LOG(L)

$EST MAXEVAL=9999 -2LL METH=1 LAPLACE
```

UPPSALA
UNIVERSITET

Student’s t-distribution
Simulation code example

```
$ERROR

IPRED = ((DOSE/V)*(KA/C))*(A-B)

W = THETA(4)*IPRED
IWRES=(DV-IPRED)/W
DF = THETA(5)

IF (ICALL.EQ.4)
Y = IPRED +
W*EPS(1)*(1+((EPS(1)**2+1)/(4*DF))+((5*EPS(1)**4+16*EPS(1)**2+3)/(96*DF**
2))+((3*EPS(1)**6+19*EPS(1)**4+17*EPS(1)**2-15)/(384*DF**3)))

$SIGMA 1FIX
```

UPPSALA UNIVERSITET

Dynamic Transform Both Sides

Type I error rate not increased

Methods	
10,000 simulations 1 CMT PK model, exponential error model (40%)	
Re-estimation true model (λ = δ = 0)	*Re-estimation* dTBS model(λ = δ = est)

↓

Results
dOFV (alternative-true) > 5.99 in **5.04%** of cases

↓

Conclusion
Estimation of dTBS parameters **does not lead to increased type I error rate**

UPPSALA UNIVERSITET

Student's t-distribution

Type I error rate not increased

Methods	
10,000 simulations 1 CMT PK model, normally distributed additive error model	
Re-estimation true normal model	*Re-estimation* t-distributed error model

Results
dOFV (alternative-true) > 3.84 in **4.13%** of cases

Conclusion
Estimation of ν parameter **does not lead to increased type I error rate**