Predicting power of future studies in the presence of uncertainty about the variance

John Sherington
Pfizer, Sandwich, UK.

NCSC 2010 – Lyon – France
27-29 September 2010

Outline

- Introduction & motivating example
- Analysis of historical data
  - Estimate important parameters
- Simulations and comparison with classical power/sample size results

Overview

- Need to design safety studies with adequate power/precision.
- Classical methods assume a known variance.
  - Allow only for sampling variation
  - In reality, the variance may vary between studies.
- Historical data can be used to explore the distribution of the study variances
- This uncertainty can be incorporated into power/sample size calculations

Motivating example

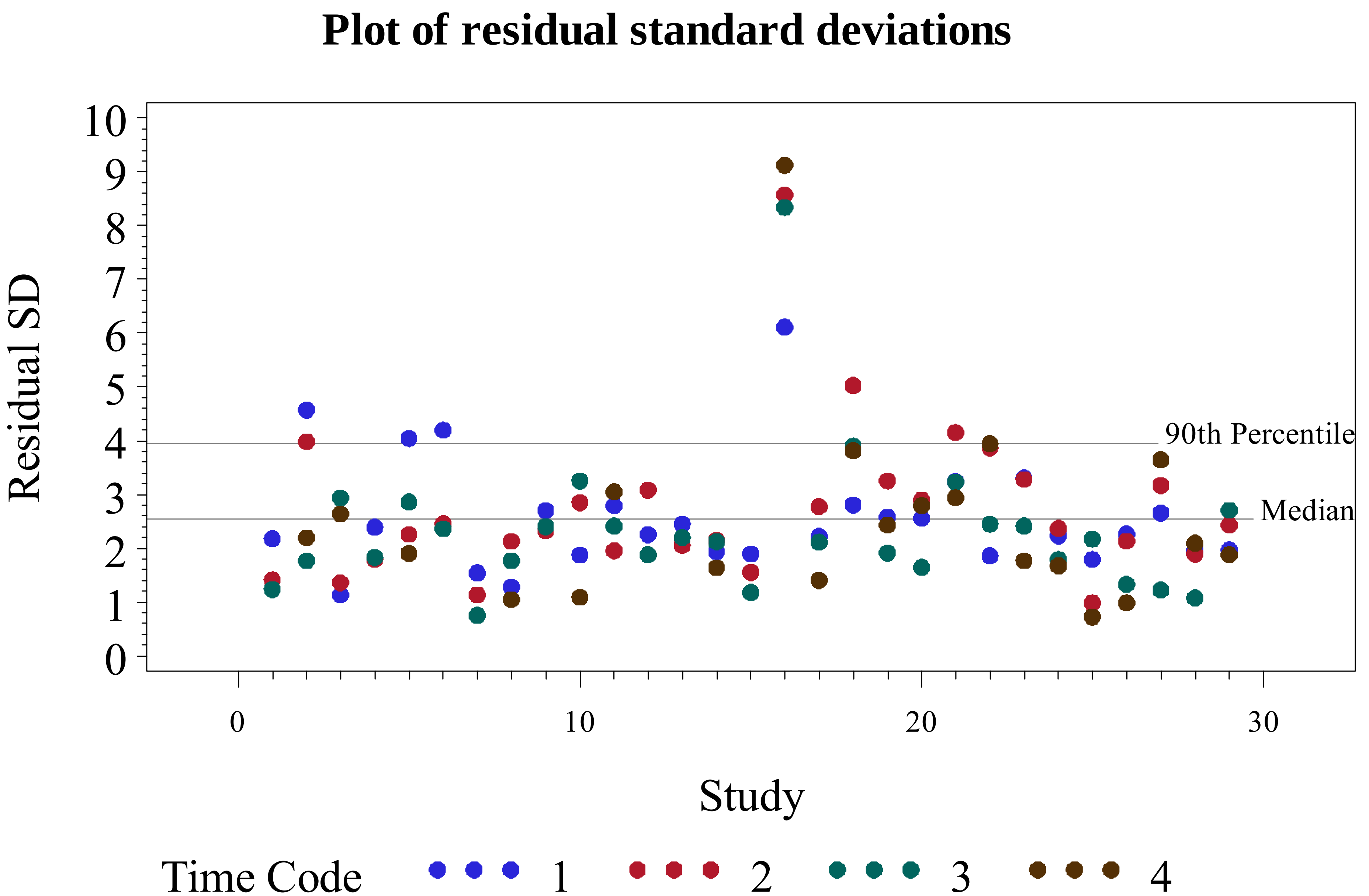
- Telemetered Dog CV Study
- Typically 4×4 crossover
  - Vehicle + 3 doses of compound
- 24-hour telemetry of cardiovascular parameters
  - QT_c, heart rate, blood pressure....
  - 3 or 4 summary time bins
    - E.g. 1-3hr; 4-6hr; 7-14hr; 14-24hr
- Analysed using linear model (ANOVA) for each time bin
  - (fit Animal, Period, Treatment)
  - Precision determined by the residual variance, s^2.

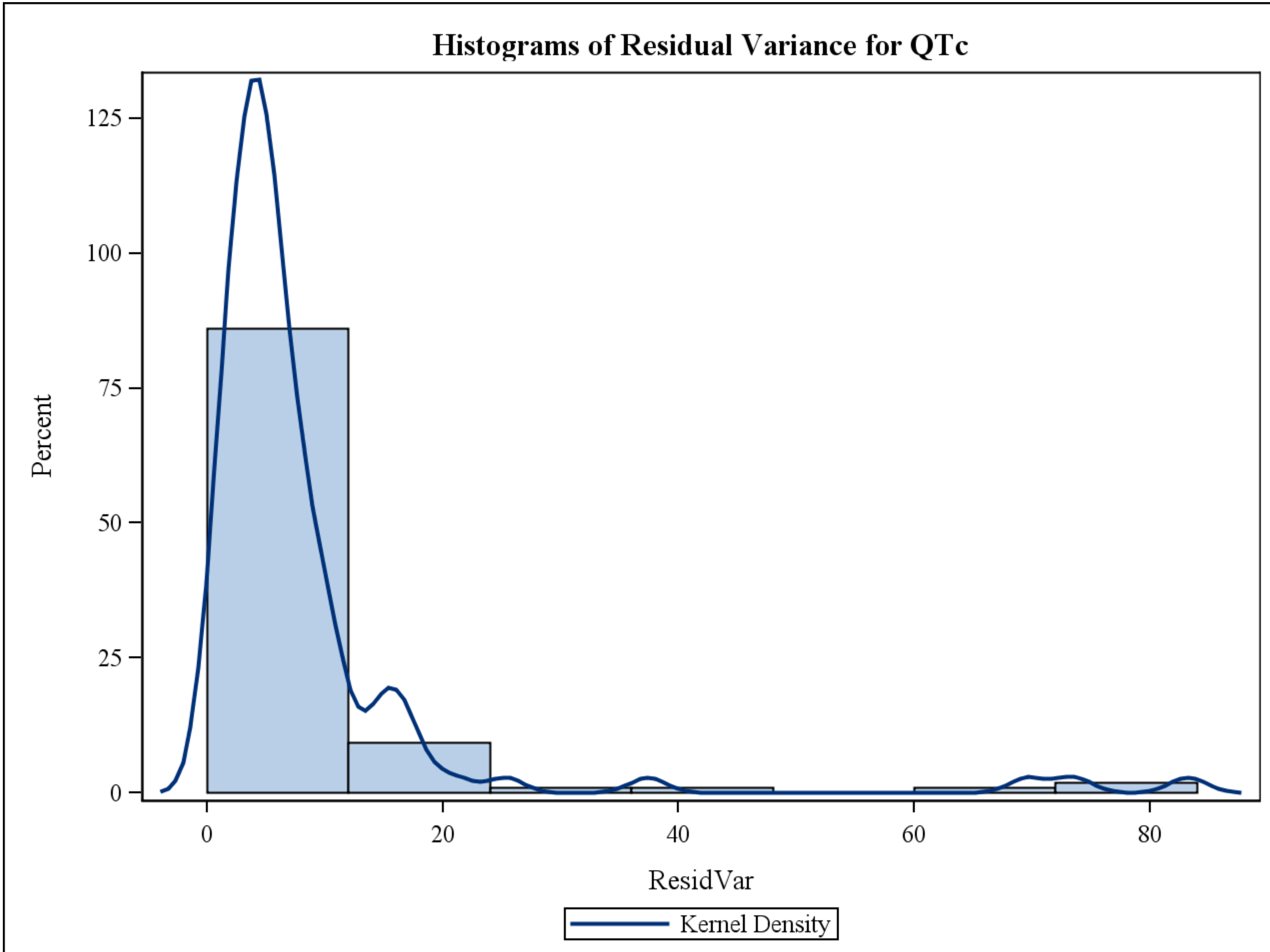
Historical Data

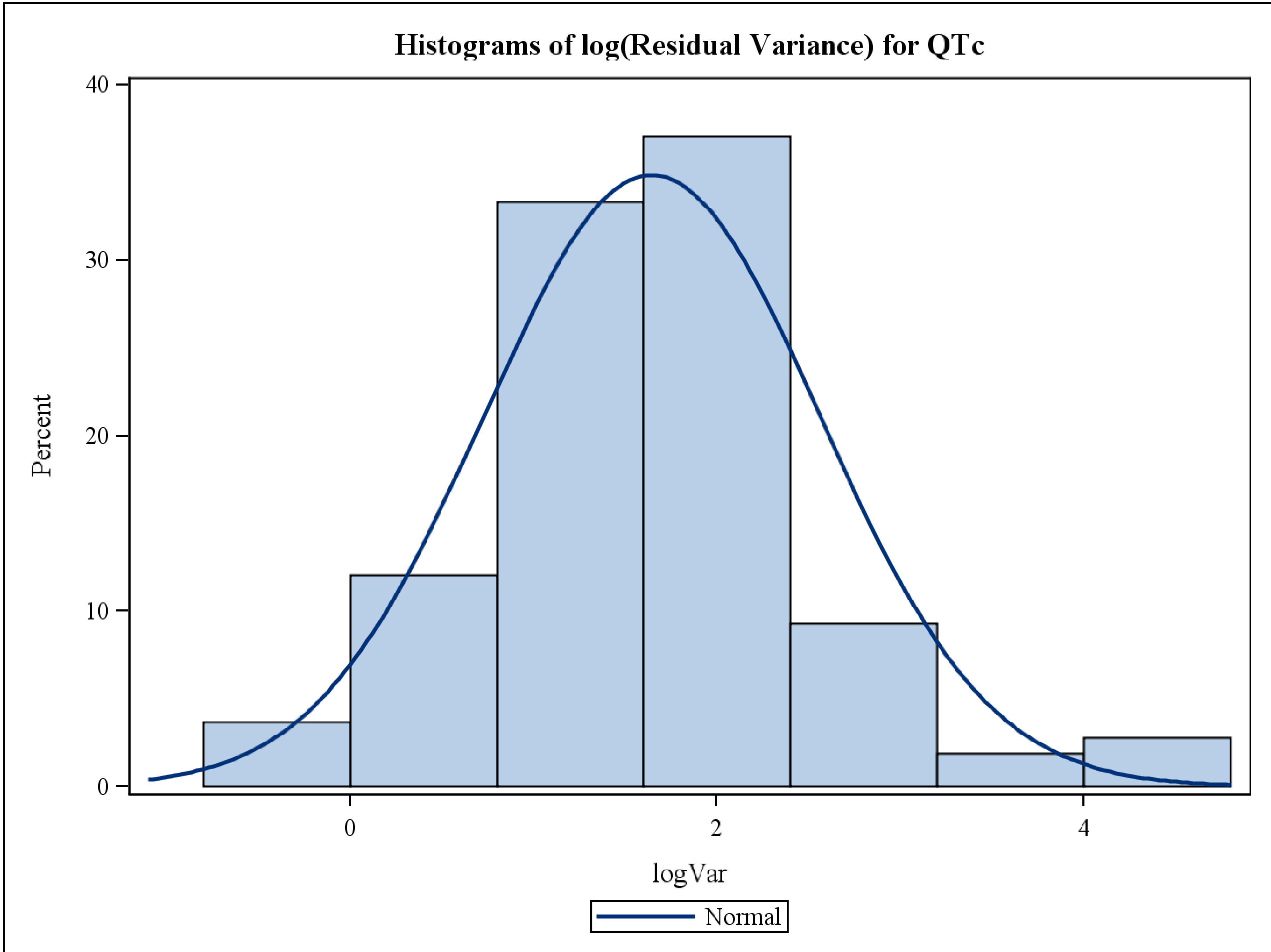
- 29 studies
- Each study has 3 or 4 time bins
- This presentation just considers QTc

QTc: Residual Variance (s^2 - 6 d.f.)						
n	**Mean**	**Median**	**Min**	**Max**	**90%**	**IQR**
108	8.40	5.04	0.53	83.2	15.6	3.23 – 8.31

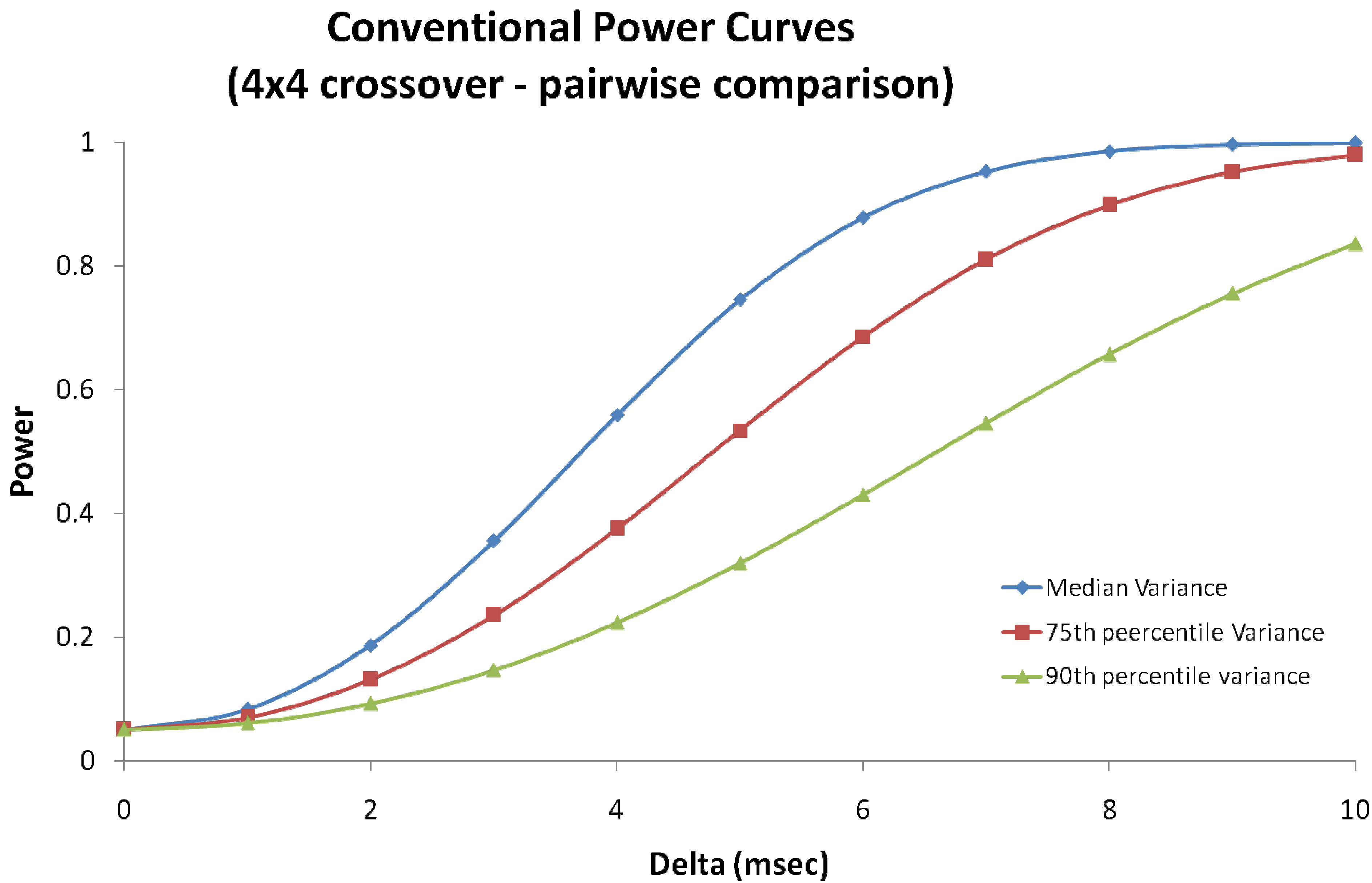
- The assumption of constant σ^2 may not be reasonable


Plot of residual standard deviations
Residual SD
0
1
2
3
4
5
6
7
8
9
10
90th Percentile
Median
0
10
20
30
Study
Time Code
1
2
3
4


Histograms of Residual Variance for QTc
Percent
125
100
75
50
25
0
0
20
40
60
80
ResidVar
Kernel Density



Histograms of log(Residual Variance) for QTc
40
30
20
10
0
Percent
0
2
4
logVar
Normal



Conventional Power Curves
(4x4 crossover - pairwise comparison)
Power
1
0.8
0.6
0.4
0.2
0
0
2
4
6
8
10
Delta (msec)
Median Variance
75th peercentile Variance
90th percentile variance

Study to Study variation

- If underlying variance is constant ($= \sigma^2$) then the observed variance, s^2, has a scaled chi-squared distribution

$$\nu . s^2/\sigma^2 \sim \chi^2_{(\nu)}$$

- If some studies have higher variability than others (σ^2 is not constant), then conventional power/sample size calculations may be misleading.

Fitting distributions

Sample variance has scaled chi-squared distibution, based on true variance for that study

$$s_i^2 \sim \frac{\sigma_i^2}{\nu}\chi_\nu^2 \equiv \mathrm{Gamma}\left(\frac{\nu}{2}, \frac{\nu}{\sigma_i^2}\right)$$

Assume true study variance has either logNormal or Gamma distribution

$$\log\left(\sigma_i^2\right) \sim \mathrm{Normal}\left(\mu, \theta^2\right)$$

$$\sigma_i^2 \sim \mathrm{Gamma}(r, \mu)$$

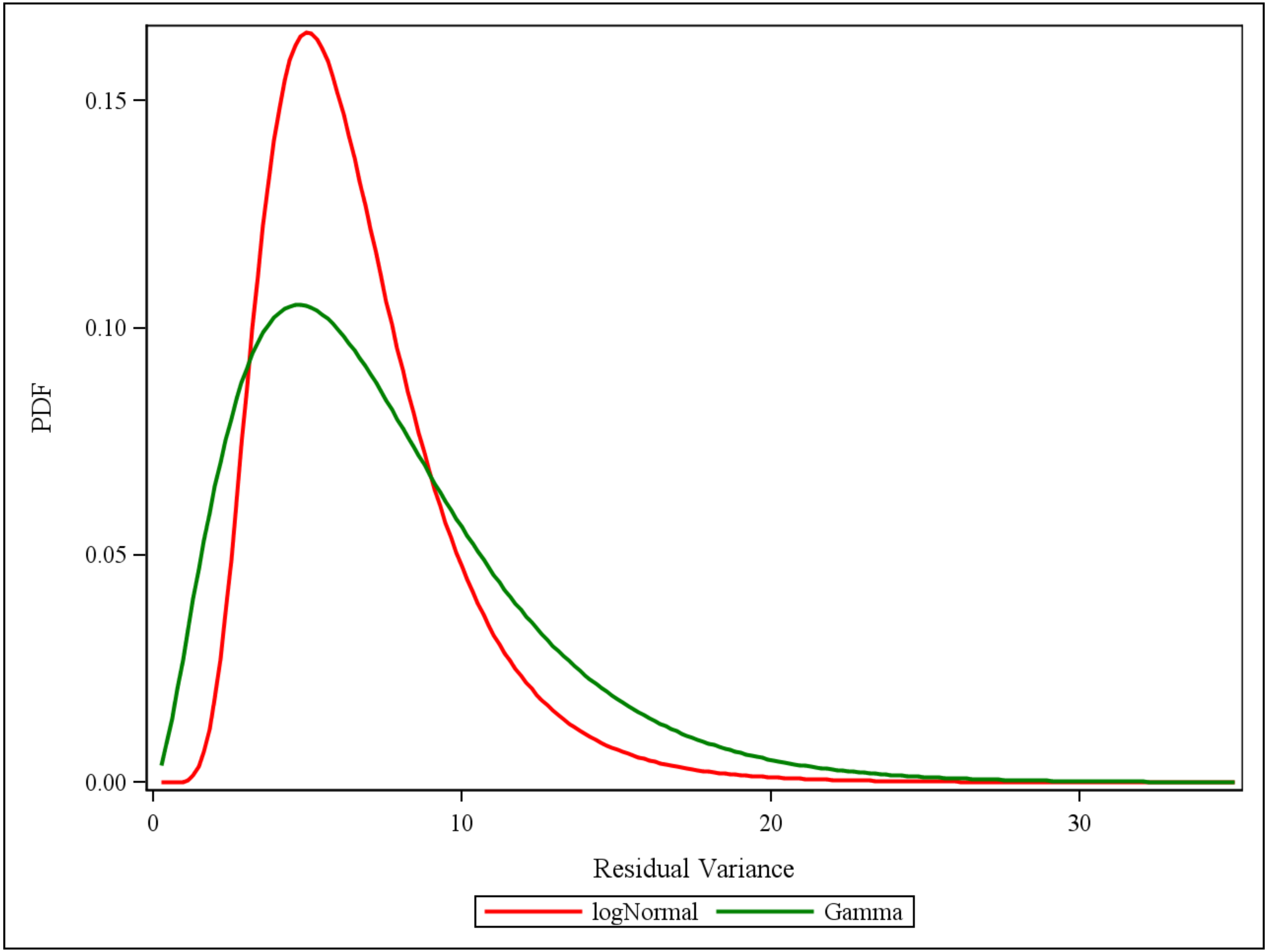
Estimate parameters (using WinBUGS) with vague priors

Model comparisons

- LogNormal and Gamma models fitted for σ_i^2
  1. Variance constant within a study (i=1…29)
  2. Variance unconstrained (i=1…108)
- Deviance information criterion (DIC) used to compare models

				σ^2	DIC
Constant				8.40	757
		μ	**θ**		
LogNormal	**Study-based**	1.80	0.44	6.05	576
	Unconstrained	1.81	0.50		596
		r	**μ**		
Gamma	**Study-based**	2.69	0.36	7.44	580
	Unconstrained	2.51	0.33		599

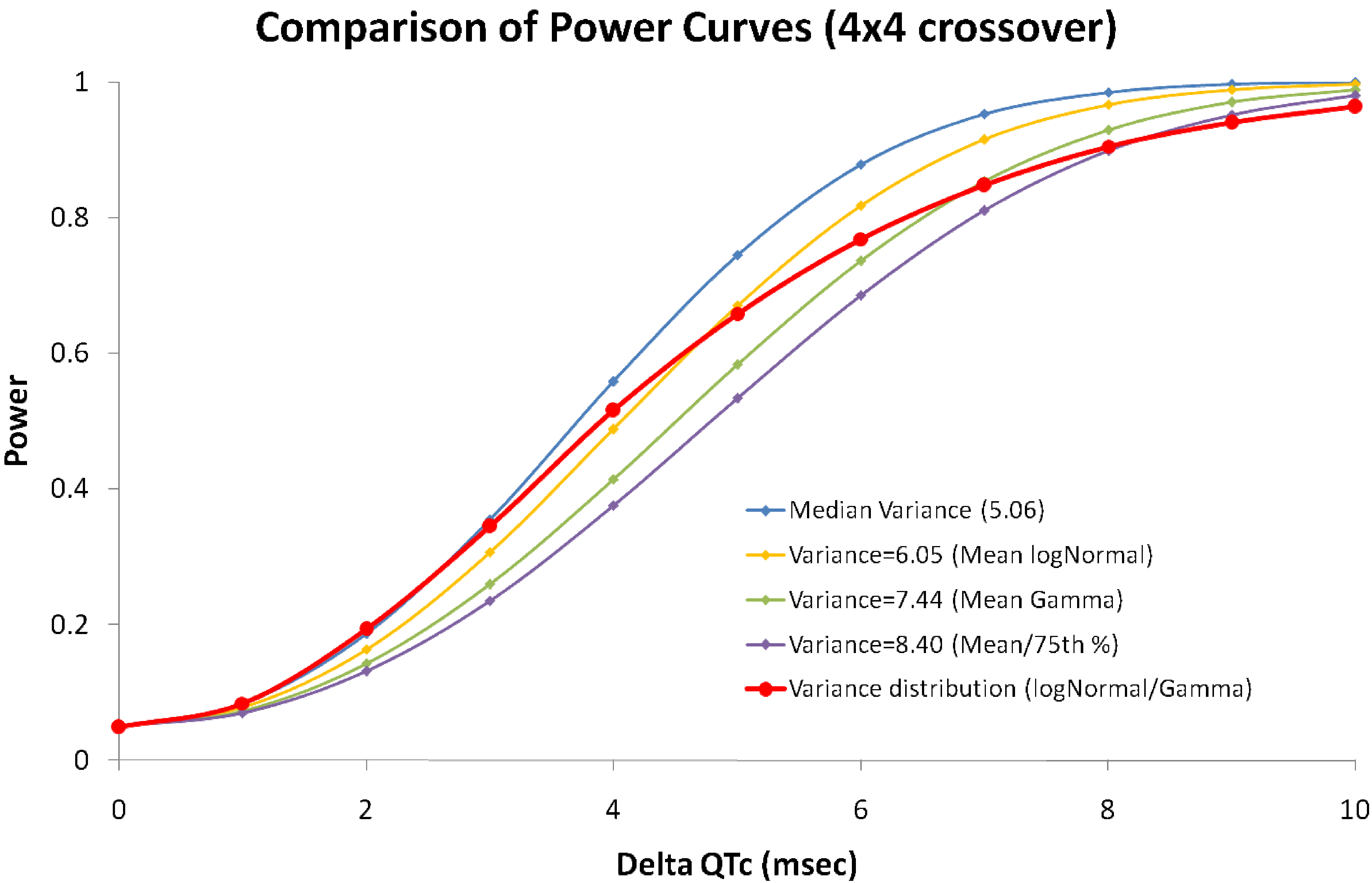
Fitted PDFs



0.15
0.10
0.05
0.00
PDF
0
10
20
30
Residual Variance
logNormal
Gamma

Simulations (pairwise comparison)

- Simulated the true variance, σ^2
  - LogNormal - Mean=1.80; SD=0.44
  - Gamma – r=2.69; μ=0.36
- Simulated sample variance, s^2
  - Chi-squared, ν d.f. (ν=6 for 4×4 crossover)
- Simulated observed difference, d
  - Normal(δ, $2\sigma^2/n$) (n=4 for 4×4 crossover)
  - δ = 0, 1, 2...10
- Calculated SED, t-statistic, P-value, confidence limits
- 10000 simulated studies per combination


Comparison of Power Curves (4x4 crossover)
Power
1
0.8
0.6
0.4
0.2
0
0
2
4
6
8
10
Delta QTc (msec)
Median Variance (5.06)
Variance=6.05 (Mean logNormal)
Variance=7.44 (Mean Gamma)
Variance=8.40 (Mean/75th %)
Variance distribution (logNormal/Gamma)

Results/Conclusions

- Power predictions from logNormal and Gamma distributions are indistinguishable
- Simulated power is less than conventionally calculated power (using median σ^2)
- At power of >75%, simulated power is similar to conventional power using 75th percentile σ^2
- Sample sizes calculated by conventional methods may have less power than anticipated.