MRC | Biostatistics Unit

How large does my study have to be?

Fiona Matthews and Ardo van den Hout
MRC Biostatistics Unit
Cambridge, UK

Background

- Estimating transitions and life expectancies using longitudinal data is being encouraged
- Many longitudinal studies exist, but how large do they have to be to estimate parameters well

Study size questions

- Irreversible diseases
- Reversible diseases
- Interval between interviews
- Total study length
- “Doctors Care”
- Variable study lengths

Overview

- Simulation models
- Fixed interval models
- Variable study lengths
- Discussion

Models

- Simulation models of reversible disease
  - Transition between health and ill-health and back allowed
  - Transitions between health, ill-health and death
- All individuals disease free at start of study
- Total study length (10 years and 5 years)
- Interval between waves (1 year and 2 years)

Parameters of the model

- Transitions (for example)

- Values for simulation

$$\alpha_{12} = 0.002 \quad \beta_{12} = 0.05$$

$$\alpha_{13} = 0.007 \quad \beta_{13} = 0.05$$

$$\alpha_{21} = 0.01 \quad \beta_{21} = 0.05$$

$$\alpha_{23} = 0.02 \quad \beta_{23} = 0.05$$

Methods

- Continuous time Markov Chain models with piecewise constant intensities
  - (similar to IMaCh which uses discrete time and piecewise constant probabilities)
- Simulations, transitions and life expectancies undertaken using R and C
- Life expectancy measured using the integrals of the probabilities

Results for fixed interval

Study length	Ten years		Five years
Interview interval	One year	Two years	One year
Study size n=250			
Actual values α_{21}**=0.01**	0.0098 (0.01)	0.0113 (0.02)	0.0122 (0.03)
α_{23}**=0.02**	0.0207 (0.01)	0.0219 (0.02)	0.0203 (0.02)
β_{21}**=0.05**	0.0599 (0.09)	0.0524 (0.18)	0.0897 (0.30)
β_{23}**=0.05**	0.0542 (0.04)	0.0500 (0.06)	0.0757 (0.21)
Study size n=500			
Actual values α_{21}**=0.01**	0.0106 (0.01)	0.0107 (0.01)	0.0102 (0.01)
α_{23}**=0.02**	0.0205 (0.01)	0.0206 (0.01)	0.0206 (0.01)
β_{21}**=0.05**	0.0486 (0.04)	0.0521 (0.06)	0.0611 (0.15)
β_{23}**=0.05**	0.0524 (0.03)	0.0528 (0.03)	0.0547 (0.05)

Detail on study size

Study length ten years interview every 24 months

	n=250		n=500		n=1000	
α_{12}=0.002	0.0021	(0.01)	0.0020	(0.01)	0.0020	(0.01)
β_{12}=0.05	0.0506	(0.05)	0.0514	(0.02)	0.0511	(0.02)
α_{13}=0.007	0.0069	(0.01)	0.0070	(0.01)	0.0070	(0.01)
β_{13}=0.05	0.0511	(0.01)	0.0507	(0.01)	0.0502	(0.01)
α_{21}=0.01	0.0113	(0.02)	0.0107	(0.01)	0.0103	(0.01)
β_{21}=0.05	0.0524	(0.18)	0.0521	(0.06)	0.0511	(0.04)
α_{23}=0.02	0.0219	(0.02)	0.0206	(0.01)	0.0202	(0.01)
β_{23}=0.05	0.0500	(0.06)	0.0528	(0.03)	0.0516	(0.02)

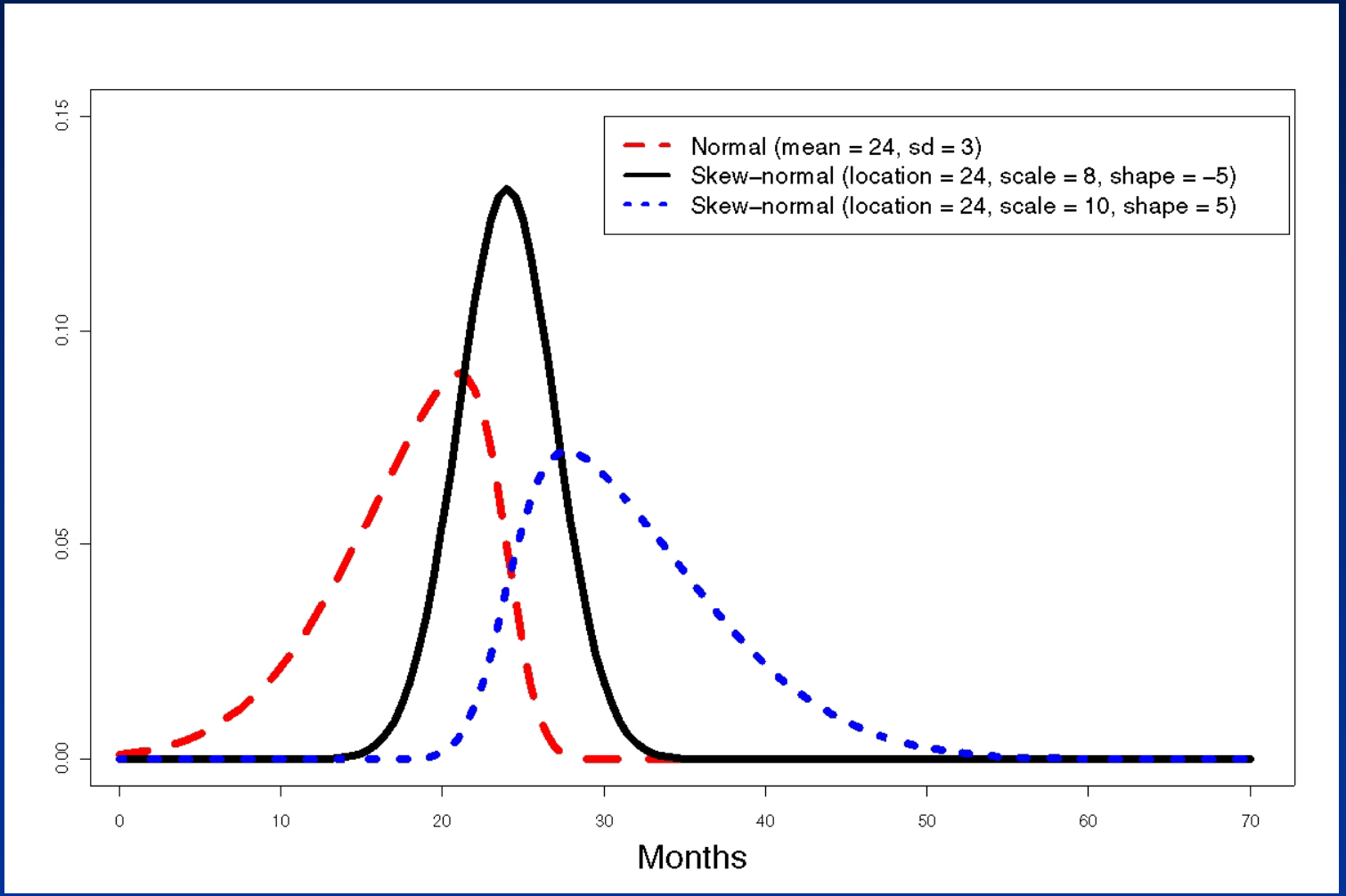
Results (2)

- Coverage of true values is good for all studies, but best for ten years with follow-up every year
- 500 individuals is sufficient for ten years of follow up with interviews every year
- 1000 individuals is sufficient for ten years of follow-up with interviews every two years

Methods (2)

- Fixed interview lengths may be intended but not observed
- Variation based on skewed normal distribution
  - Mean at intended time
  - Length slightly extended

Variable length intervals



Results of variable length

	Fixed interval		Normal	
α_{12}=0.002	0.0020	(0.0101)	0.0020	(0.0101)
β_{12}=0.05	0.0514	(0.0249)	0.0050	(0.0238)
α_{13}=0.007	0.0070	(0.0072)	0.0070	(0.0072)
β_{13}=0.05	0.0507	(0.0086)	0.0504	(0.0086)
α_{21}=0.01	0.0107	(0.0083)	0.1070	(0.0084)
β_{21}=0.05	0.0521	(0.0643)	0.0543	(0.0596)
α_{23}=0.02	0.0206	(0.0138)	0.2090	(0.0143)
β_{23}=0.05	0.0528	(0.0326)	0.0501	(0.0320)
	Negative		**Positive**	
α_{12}=0.002	0.0020	(0.0102)	0.0021	(0.0101)
β_{12}=0.05	0.0507	(0.0211)	0.0517	(0.0285)
α_{13}=0.007	0.0070	(0.0071)	0.0070	(0.0072)
β_{13}=0.05	0.0506	(0.0087)	0.0500	(0.0091)
α_{21}=0.01	0.0105	(0.0079)	0.0106	(0.0090)
β_{21}=0.05	0.0512	(0.0502)	0.0564	(0.0935)
α_{23}=0.02	0.0206	(0.0138)	0.0206	(0.0139)
β_{23}=0.05	0.0505	(0.0285)	0.0534	(0.0358)

Effect of study design on life expectancy (example from in state 1)

	TRUE	Ten years		Five years
		24 mo FI	12 mo FI	12 mo FI
65 years	11.31	11.35 (0.58)	11.36 (0.58)	11.33 (0.78)
75 years	7.88	7.89 (0.30)	7.89 (0.29)	7.88 (0.41)
85 years	5.25	5.23 (0.28)	5.23 (0.29)	5.25 (0.41)

	24 mo Normal	24 mo Negative	24 mo Positive
65 years	11.33 (0.63)	11.33 (0.64)	11.32 (0.59)
75 years	7.90 (0.30)	7.88 (0.32)	7.88 (0.29)
85 years	5.25 (0.27)	5.23 (0.28)	5.21 (0.27)

Conclusions

- Studies do not have to be enormous to measure longitudinal changes
- Irreversible diseases need smaller studies than reversible diseases
- Interval between waves needs to be shorter than latent state
- Variation around a mean for the designed interval does not adversely effect the study size

Other issues

- Different parameters does not change the conclusions
- Biased study designs to subsets of data do not bias the transitions or life expectancy estimates
- Need to investigate study design for mixed disease states at start of study
- Need to investigate different ages at start of study