A Non-Randomized Controlled Clinical Study on Lung-Volume-Reduction Surgery (LVRS) in Patients with Severe Emphysema

On the sensitivity of results and conclusions to the type of missing value treatment

Jochem König,
Institute of Medical Biometry, Epidemiology and Medical Informatics, University of Saarland, Homburg

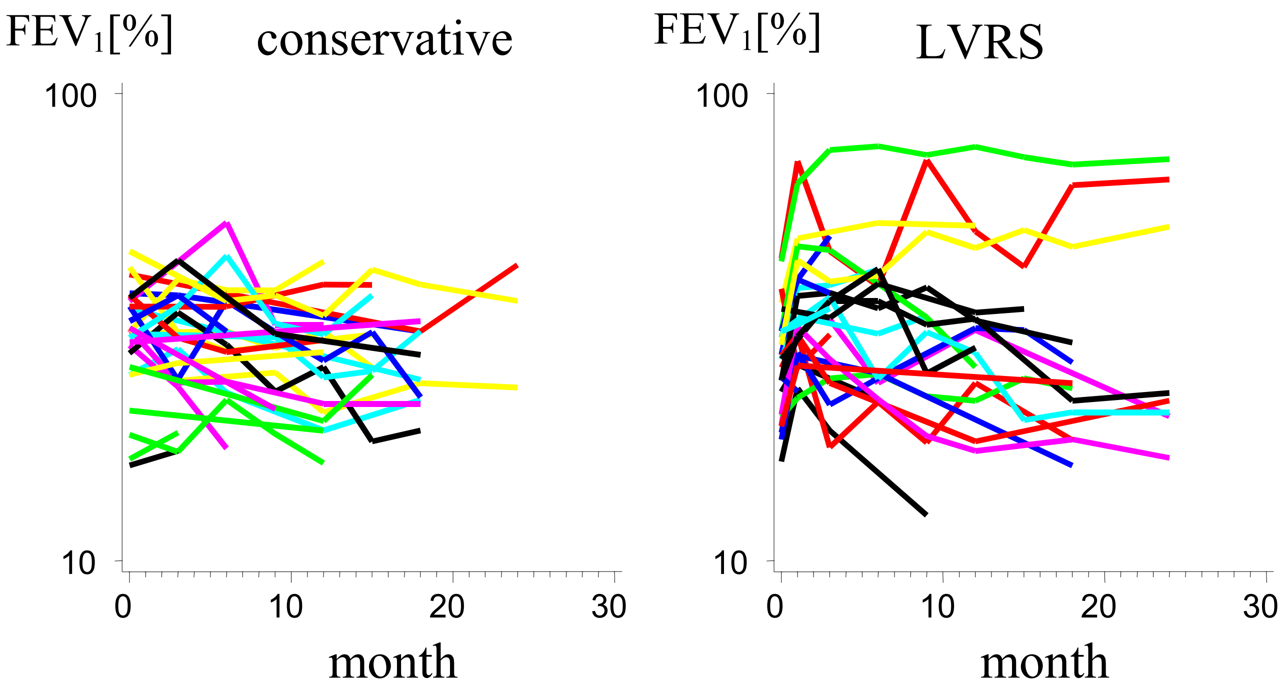
Contents

- LVRS-study, REML-Analysis
- Patterns of missingness
- REML vs. Bayes
- Modelling the dropout pattern
- Lessons learned

Design

- In order to assess the benefit of **lung volumen reduction surgery (LVRS)** in patients with severy **lung emphysema** a non - randomized comparative study was conducted 57 patients.
- Patients eligible for operation, and satisfying the inclusion criteria, were asked, whether they would do a conservative rehabilitation or to postpone surgery.
- 29 patients were operated, 28 aggreed to conservative treatment.
- **Lung function tests as well as the** modified **MRC dyspnea score and the 6 min walking distance** were observed at **entry and in 3 months‘ intervals over 18 months**. In the LVRS-group a control visit 4-6 weeks post surgery was added.
- **Primary endpoint was one second forced expiratory volume** (**FEV1**) measured in percent of reference.

individual curves by treatment



Homogeneity of Treatment Groups

		LVRS	Control	p-value
Age	yrs	58.8±1.7	58.5±1.8	
		(40-72)	(33-77)	
Females/males	n	8/21	5/23	
α1-AT deficiency	n	4	3	
Oxygen supplementation*	n	16	15	
MMRC dyspnoea score		3.5±0.1	3.1±0.15	<0.04
6-min walking distance	m	236±34	326±36	0.06

Data are presented as absolute values or mean±SEM with or without range in parentheses. *: continuous or intermittent.

α1 :α1 -antitrypsin; MMRC: modified Medical Research Council.

Homogeneity
Preoperative lung function and gas exchange

	LVRS	Control	p-value
FEV_1 l	0.80±0.04	0.895±0.04	NS
% pred	27.6±1.3	30.8±1.4	0.085
TLC	8.52±0.26	8.33±0.28	NS
% pred	137±2.5	133±2.1	NS
RV	6.2±0.25	5.8±0.26	NS
% pred	286±10.5	263±10	NS
FVC	2.29±0.12	2.7±0.2	NS
% pred	60±3.1	67±3.9	NS
MIP kPa	4.86±0.44	5.5±0.42	NS
P_{a,O_2} kPa	8.7±0.3	8.6±0.3	NS
P_{a,CO_2} kPa	5.4±0.2	5.41±0.144	NS
$D_{L,CO}$%pred	42±3.2	43±4.6	NS

TLC: total lung capacity; RV: residual volume; FVC: forced vital capacity; MIP: maximal inspiratory mouth pressure; Pa,O 2 : arterial oxygen tension; Pa,CO 2 : arterial carbon dioxide tension; DL,CO: diffusing capacity of the lung for carbon monoxide;

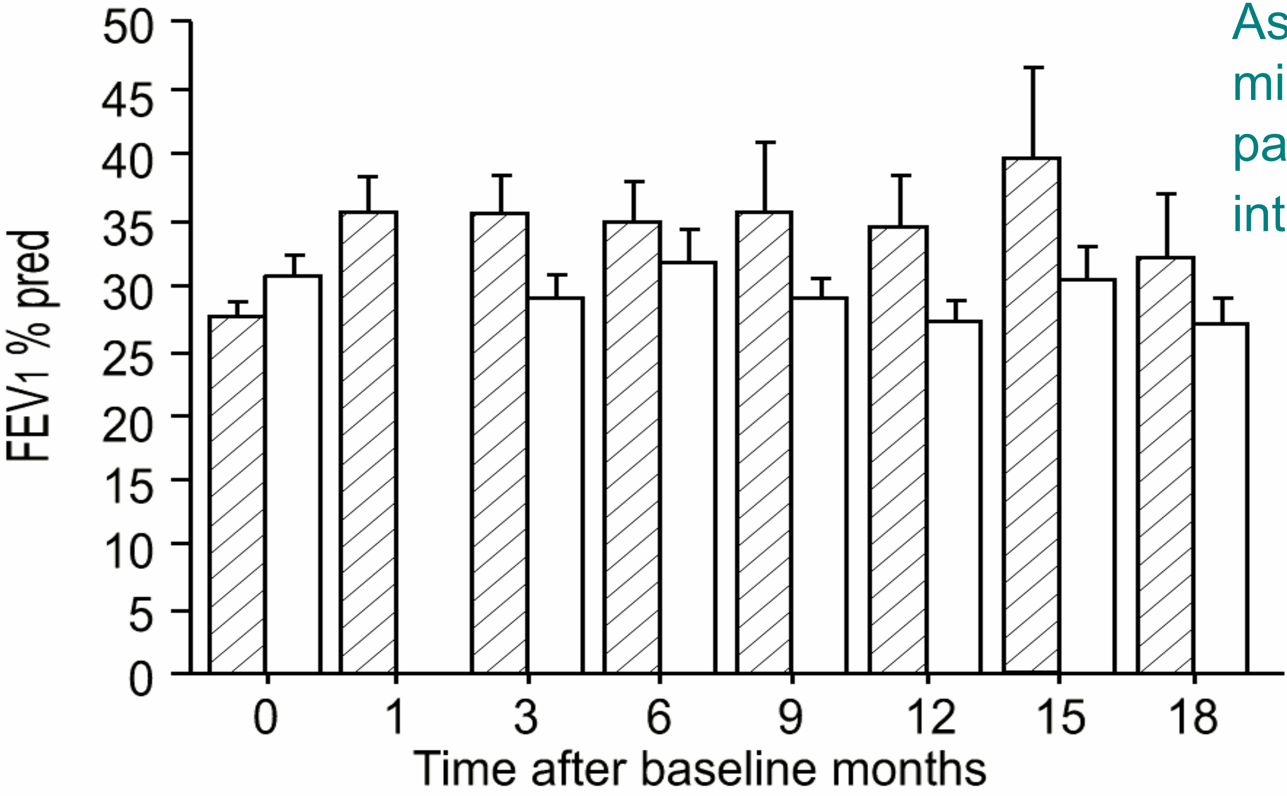
Mortality

An adverse effect of surgery was neither suggested nor could it be excluded. Bias on primary analysis probably small.

Table 3. – Mortality in the lung volume reduction surgery and control groups

Patient No.	Age yrs	Survival months	Cause of death
Lung volume reduction surgery			
12	43	21	Colon perforation, respiratory failure
19	69	14	Respiratory failure
22	55	8	Pneumonia
23	57	12	Sepsis after lung transplantation
Control			
6	61	29	Oesophageal carcinoma
11	51	24	Respiratory failure
14	55	6	Pneumonia
20	62	3	Cardiac arrest

Raw Analysis



Asynchroneous missing value pattern inhibits interpretation

Fig. 1. – Unadjusted time course of forced expiratory volume in one second (FEV_1) in the lung volume reduction surgery (LVRS, ▨) and control group (☐). Data are presented as mean±SEM. Number of patients (LVRS/control): 0 months: 29/28; 1 month: 25/0; 3 months: 21/19; 6 months: 18/17; 9 months: 13/15; 12 months: 16/18; 15 months: 8/10; and 18 months: 13/13. % pred: percentage of the predicted value.

Mixed Linear Models

1. time varying treatment effect $$Y_{hij} = \mu + \alpha_{hj} + \vec{\beta}\vec{x}_{hi} + \gamma_j + e_{hij}$$
2. treatment effect as linear trend $$Y_{hij} = \mu + \alpha_h + \delta_h t_j + \vec{\beta}\vec{x}_{hi} + \gamma_j + e_{hij}$$
3. time constant treatment effect $$Y_{hij} = \mu + \alpha_h + \vec{\beta}\vec{x}_{hi} + \gamma_j + e_{hij}$$

$Y = \log_{10} FEV_1[\%]$ e_{ijk} *random error*

α_h bzw α_{hj} treatment effect (fixed) $\alpha_0 = \alpha_{0j} = 0$

δ_h trend of treatment effect $\delta_0 = 0$

β... effects of baseline variables (fixed): 6-min walking distance, lung function tests, MMRC

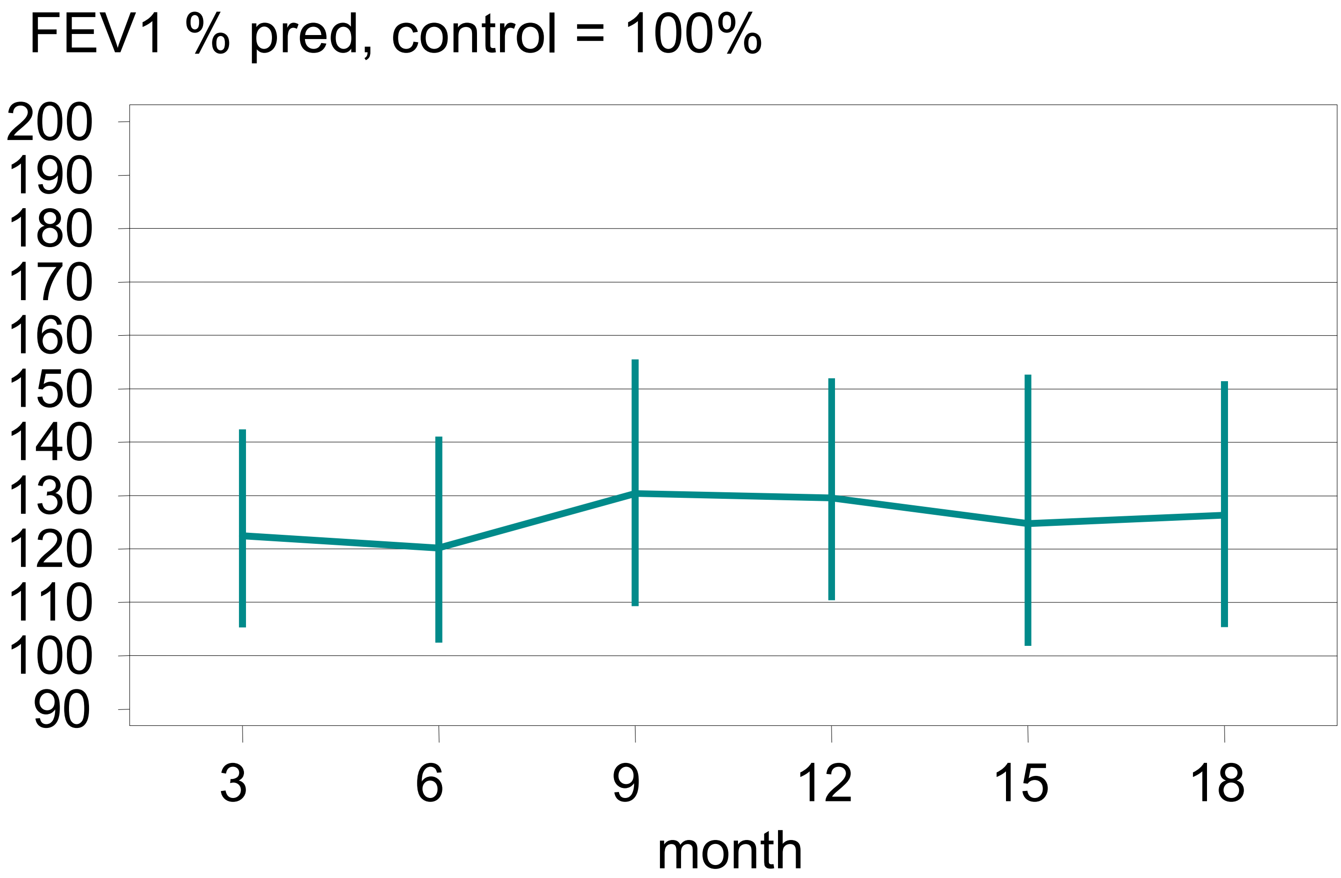
γ_j time effect

covariance structure: compound symmetry.

treatment $h = 0, 1$ (conservativ/ surgery), patient $i = 1, .., n_h$; $n_0 = 28$, $n_1 = 29$

times $j=1,...,6$, $t_j ∫ \{3, 6, 9, 12, 15, 18\}$

Model1: Time varying treatment effect and 95%-confidence interval



baseline variable: logFEV1 [% predicted]

Models 2 & 3

Model 2: Test for time trend of treatment effect not significant.

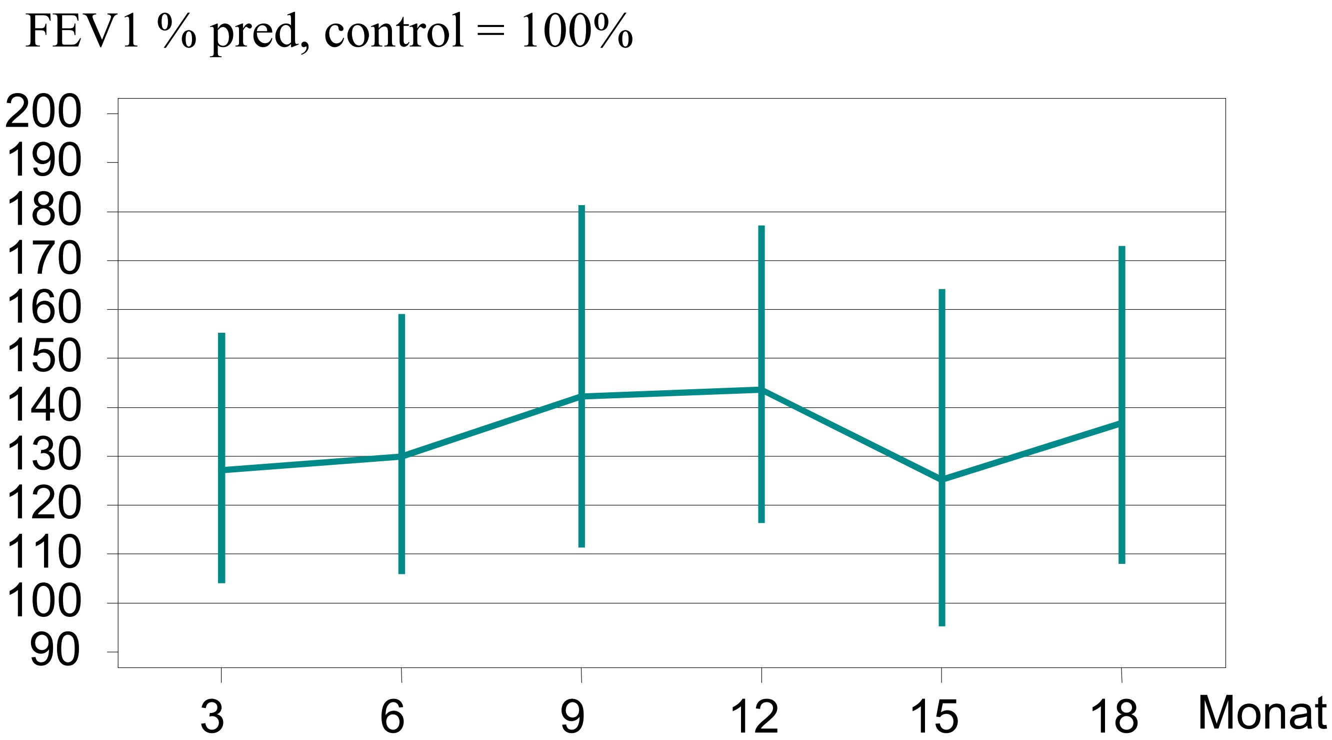
Modell3: Estimation of time constanten treatment effect and 95%-confidence interval

FEV1% post LVR-surgery on average

130% (116% - 145%) of the control group mean.

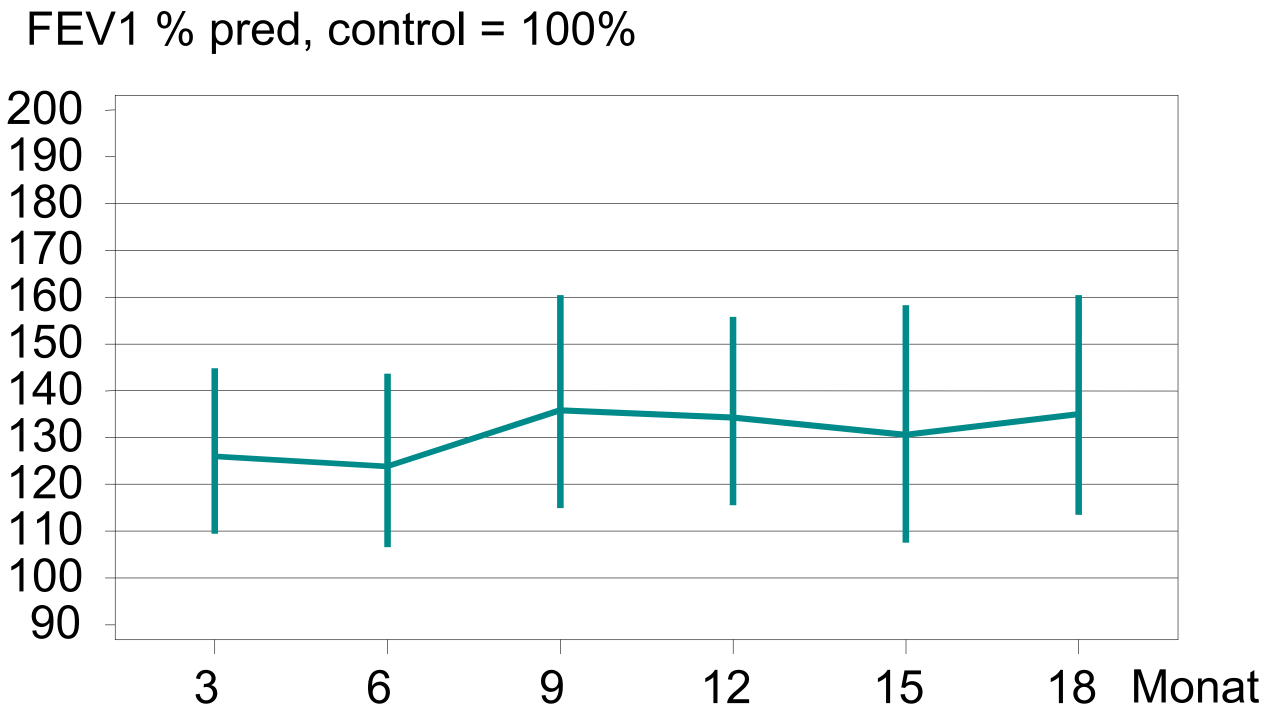
(exponentiation of effect estimates of model 3)

Some Sensitivity Analysis



7 baseline variables: logFEV1 [% predicted], age, gender, GEH6m,RVP,PMAX, MRC,
average treatment effekt 133 (113-158)

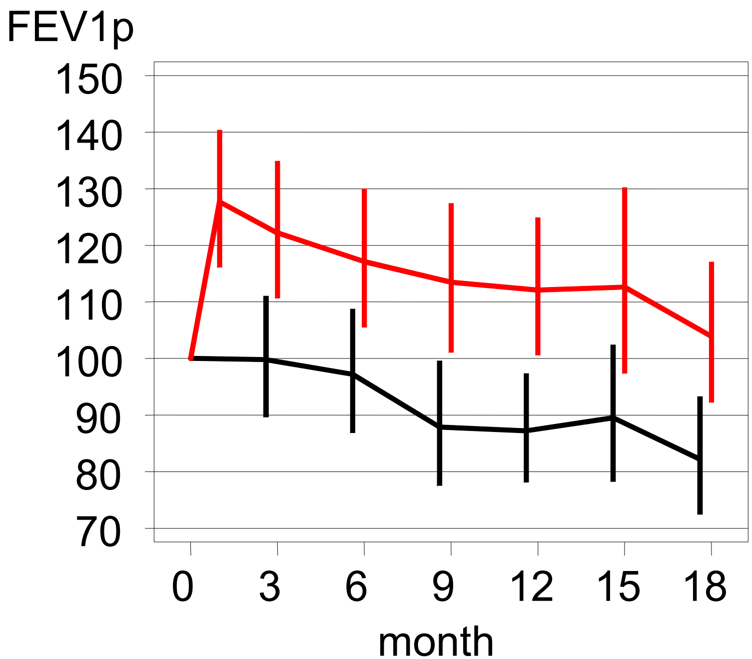
Some Sensitivity Analysis



2 baseline variables: logFEV1 [% predicted], RVP

Durchschnittseffekt 125 (111-141)

Change of Lung Function since Entry



FEV1 in percent of baseline, estimates of populaton means and 95%-confidence intervals.

mixed model (compound symmetry) for log(FEV1p(t)/FEV1p(0)).

syntax

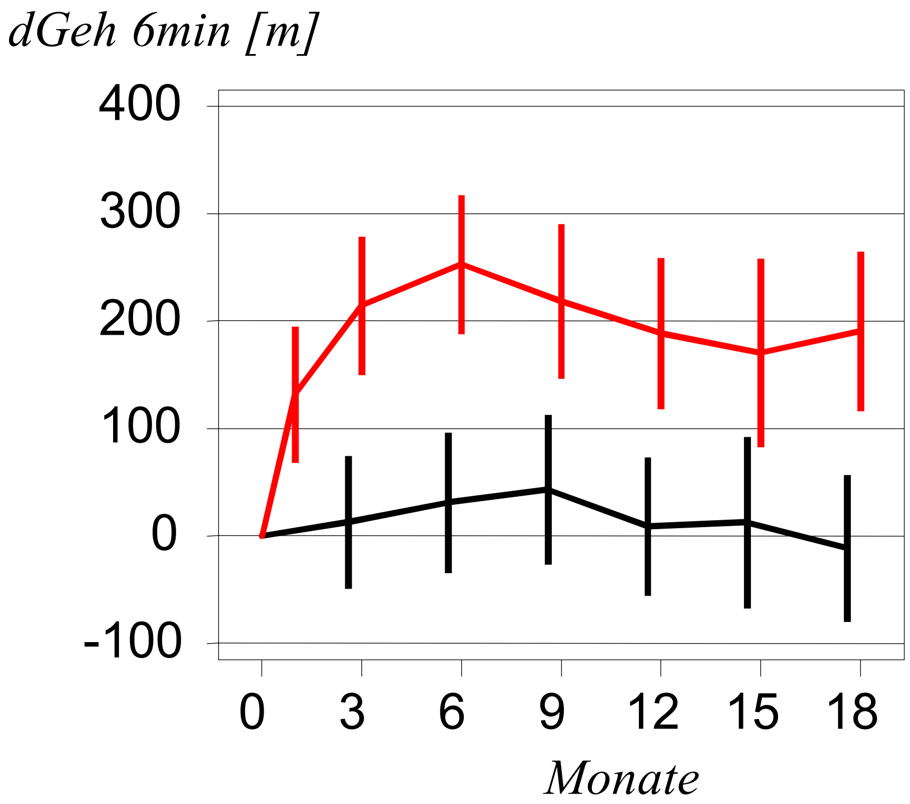
... class time op;

model dlFEV1p = TIME*OP/NOINT;

repeated /type=VC;

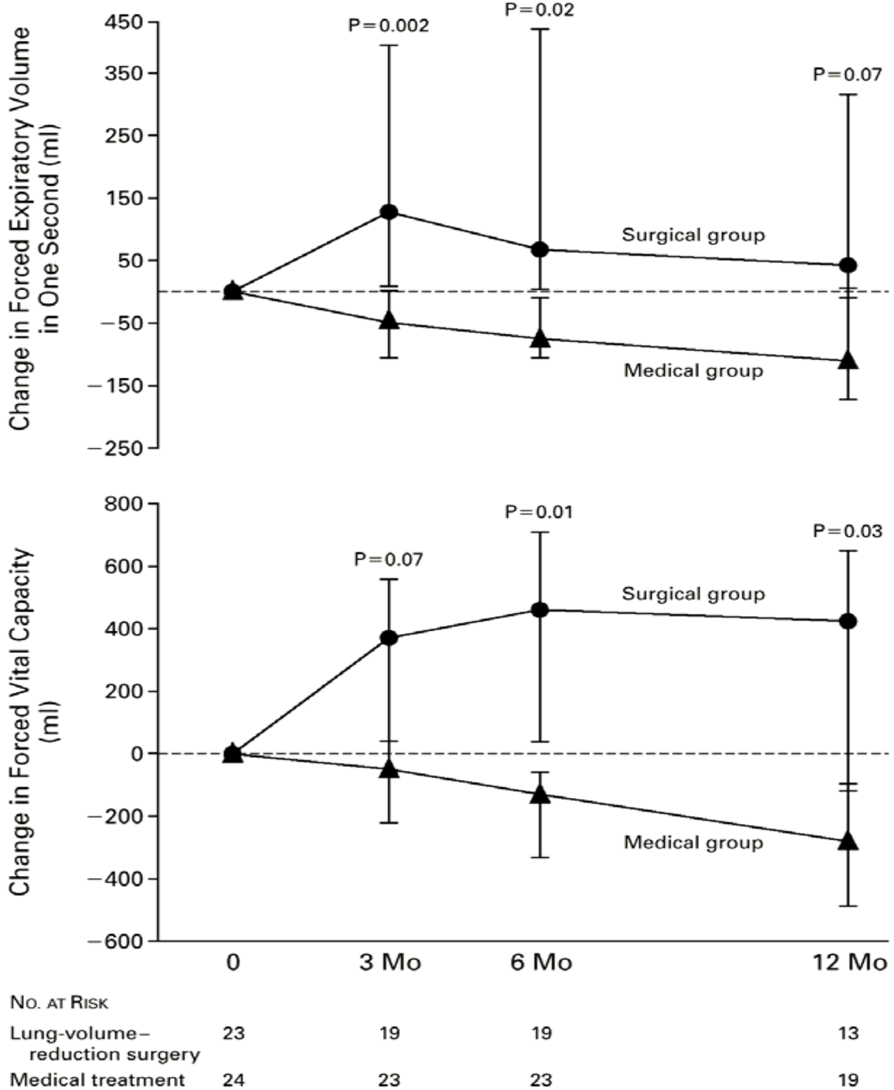
The **rise at surgery of 28 %,** is followed by a **decrease of** 15%/year resulting in a **time gain of 22 months**.

Change in 6 min Walking Distance



Differences from baseline, estimates of population means, 95% confidence intervals.

Geddes‘ Randomized Study NEJM 343:239-245



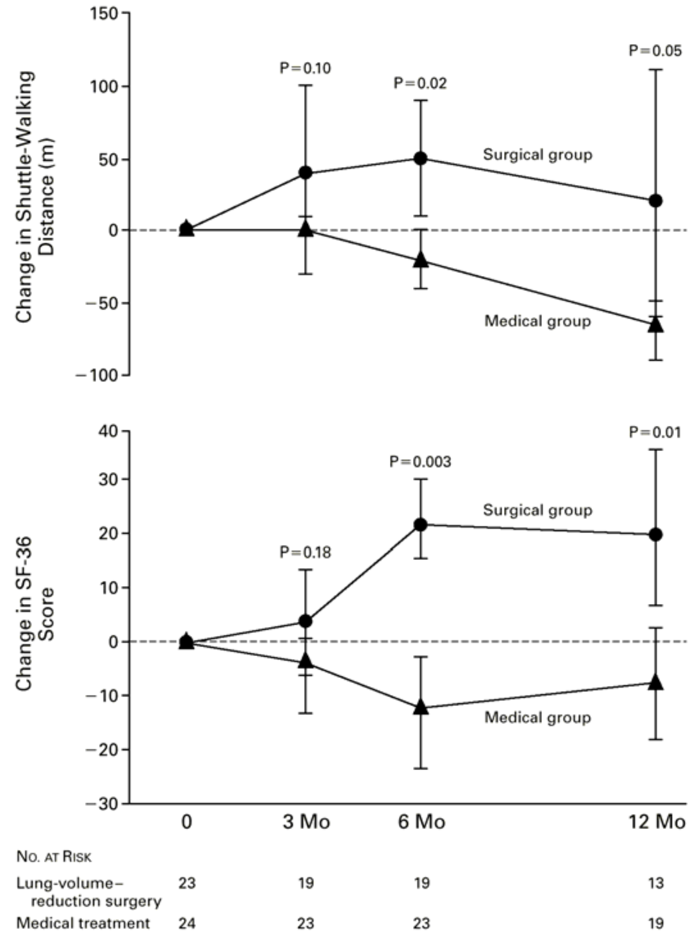
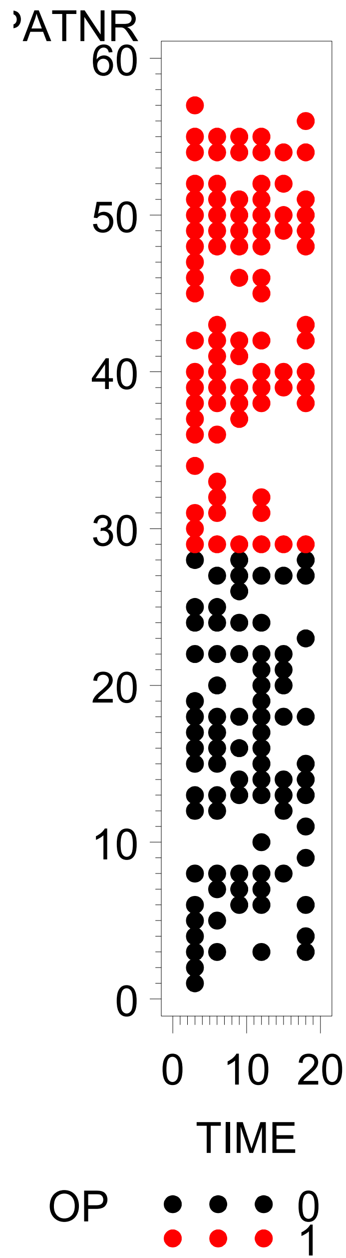


Figure 2. Median Changes in Shuttle-Walking Distance and Score on the 36-Item Short-Form Questionnaire (SF-36) Measuring Quality of Life in the Groups Receiving Surgical and Medical Treatment.
The median changes were obtained by comparing the responses of each subject with his or her baseline values, and they therefore differ from the values shown in Table 1. P values are for the comparison between the two groups at each time point. The SF-36 score ranges from 0 (indicating a low quality of life) to 100 (indicating a high quality of life). I bars show 95 percent confidence intervals.

Missing Value Pattern

- 29+28 treated medically or surgically
- 28+26 have FEV1% observed at least once on months 3,6, ...,18.
- 53/54 observed on baseline FEV
- 40/54 observed on all of 8 baseline vars.
- 170 of 324 observations available.
- months 9 and 15 frequently missing.
- 12+12 obs on month 18.
- some 2 or 3 who are missing on month 18 are observed on months 24.



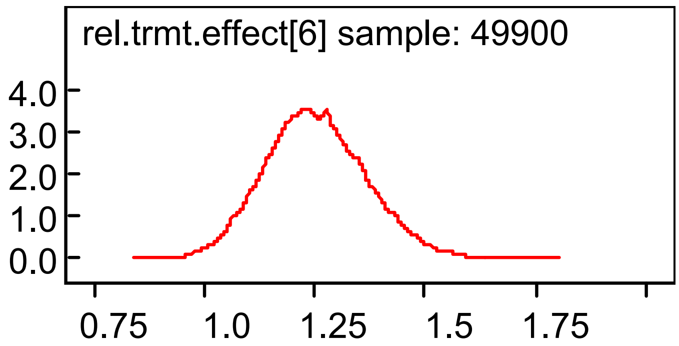
A WinBugs Model

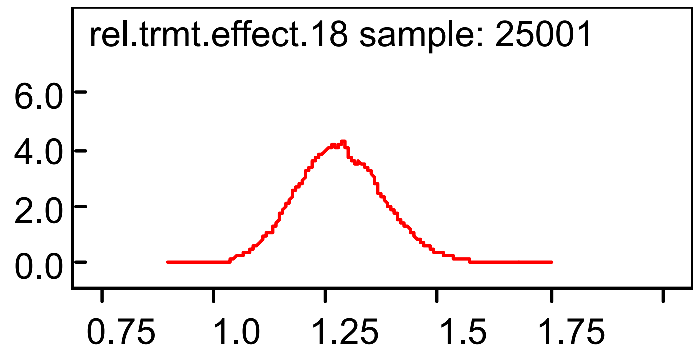
```
model{   for( i in 1 : NPAT ) {
              for( j in 1 : T ) {
                     Y[i , j] ~ dnorm(mu[i , j],tau.c) #Y=log fev1p
                     mu[i , j] <- alpha[i] + beta.op*op[i] +beta.time[j]+ beta.x0 * lfev1p0[i]
              }
              alpha[i] ~ dnorm(alpha.c,tau.alpha)  # random intercept
              lfev1p0[i]~dnorm(alpha.x0,tau.x0)
         }
         beta.op ~ dnorm(0.0,1.0E-6)              # fixed effects priors
         for( j in 1 : T ) {
              beta.time[j]~dnorm(0.0,1.0E-6)
         }
         beta.x0 ~ dnorm(0.0,1.0E-6)
         alpha.x0 ~ dnorm(0.0,1.0E-6)             # prior of covariate distrib.parameters
         tau.x0 ~ dgamma(0.001,0.001)             # for missing covariates
         sigma.x0<-1/sqrt(tau.x0)

         tau.c ~ dgamma(0.001,0.001)              # residual prec
         sigma <- 1 / sqrt(tau.c)                 #residual SD
         sigma.alpha~ dunif(0,100)                # prior of random effects variances
         tau.alpha<-1/(sigma.alpha*sigma.alpha)
         rel.trmt.effect<-exp(2.302585*beta.op)    # relative treatment effect
    }
```

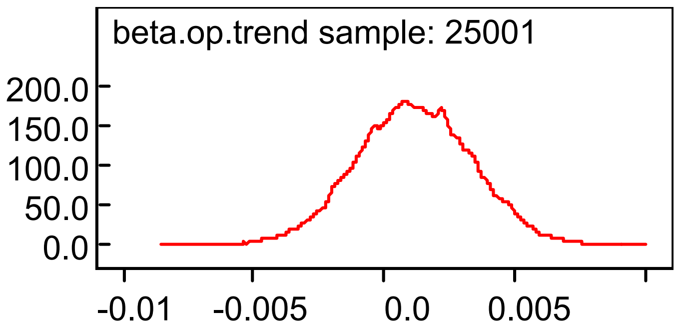
Bayes vs REML

	Treatm. Param.			covariable	N, ddf	Beta(op)
Model 2	level t=18	Bayes		LFEV1P0	54	0.1072±0.0334
	level t=18	REML	Satterth Sandwich	LFEV1P0	53, 104	0.1098±0.0332 ±0.0409
	trend	REML	Satterth Sandwich	LFEV1P0	53, 126	0.0169±0.0271/year ±0.0329
	trend	Bayes		LFEV1P0		0.0132±0.0271/year entspr. 3.1%±6.4%
Model 3	level	REML	satterth	8 baseline vars	40, 35.2	0.1086±0.0305
	level	REML	Satterth **KenwardRo** Sandwich	LFEV1P0	53, 51.5	0.0969±0.0262 ±0.0262 ±0.0266
	level	Bayes		LFEV1P0	54	0.0972±0.0266
	level	Bayes	covars complete	LFEV1P0	53	0.0968±0.0270

Some posterior densities

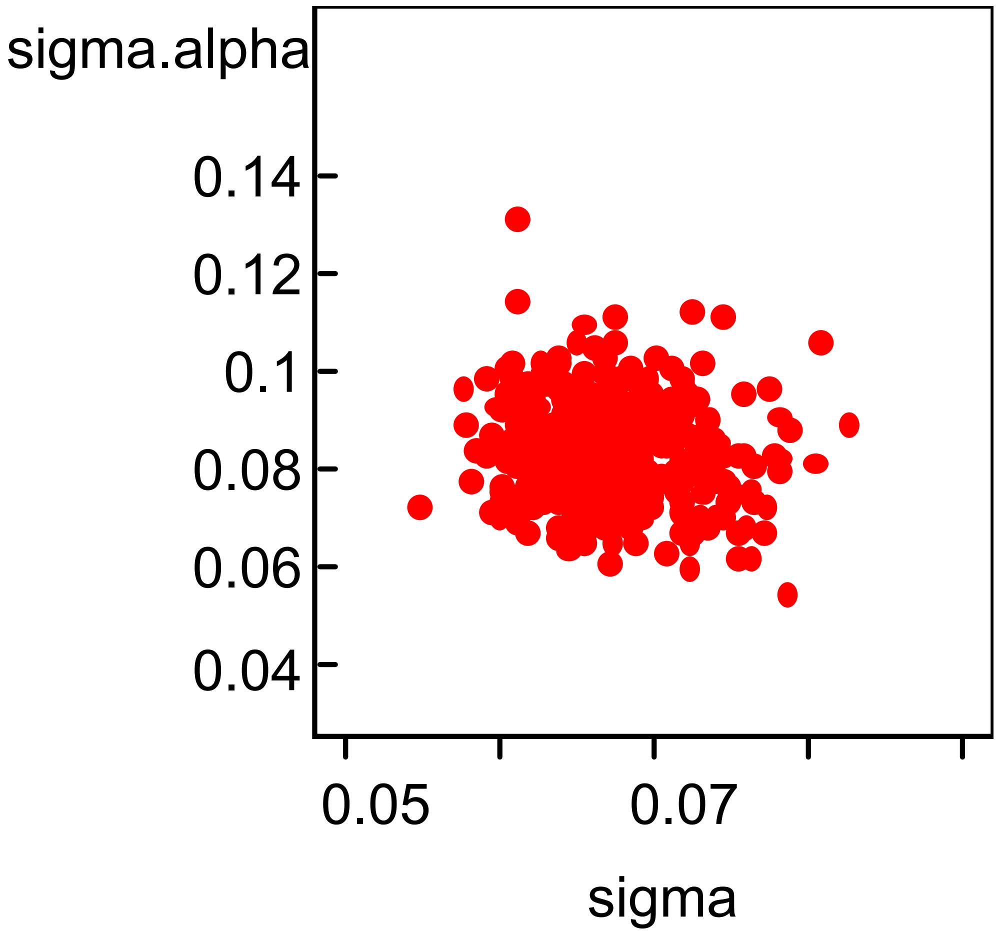






Rel.effect(t=18) in model 1 (95% interval, median 1.044, 1.249,1.492) and model 2: (95% interval, median: 1.099, 1.28,1.489),
trend parameter in model 2,

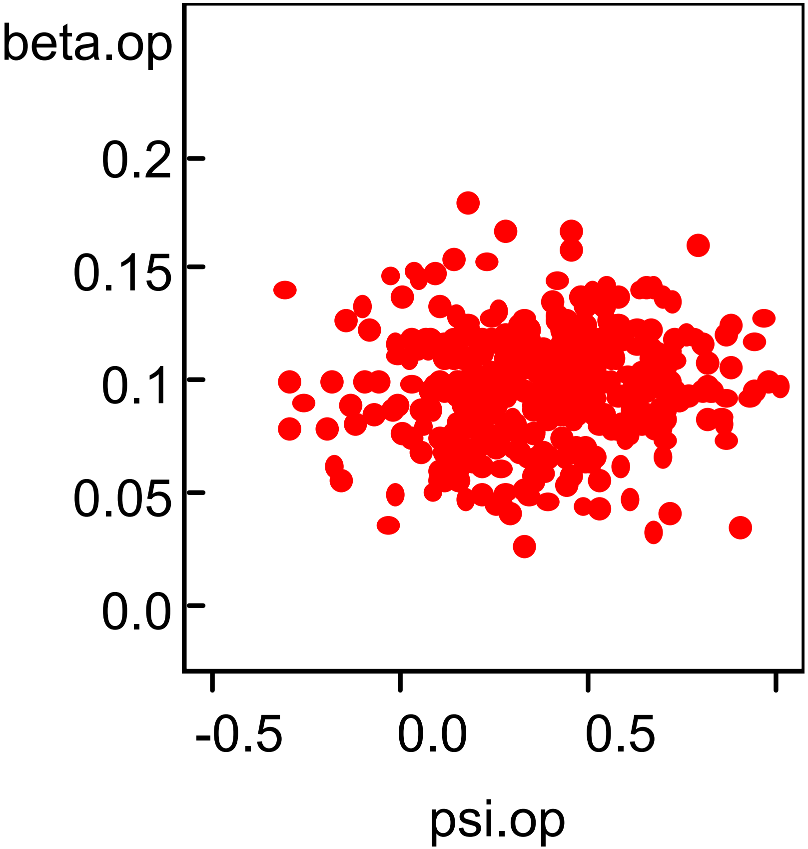
Posterior Density of
Residual SD (sigma) and random intercept SD (sigma.alpha)



A Selection Model

Model 3 and: logit(available)=factor(time)+psi.y0*(y0-mean(y0))+psi.op*op

node	mean	sd
beta.op	0.09819	0.02686
psi.op	0.3855	0.2439
psi.time[1]	0.8067	0.3272
psi.time[2]	0.2797	0.3076
psi.time[3]	-0.3522	0.3045
psi.time[4]	0.2884	0.3045
psi.time[5]	-1.091	0.3333
psi.time[6]	-0.4188	0.3021
psi.x0	2.008	1.046
sigma	0.06718	0.004504
sigma.alpha	0.08331	0.01077



A Selection Model-2

Model 3 and: logit(available)=psi.time[j]+psi.y0*(y0[i]-mean(y0)) + psi.op*op[i]+psi.yl*Y[i,j-1]+RE[i]

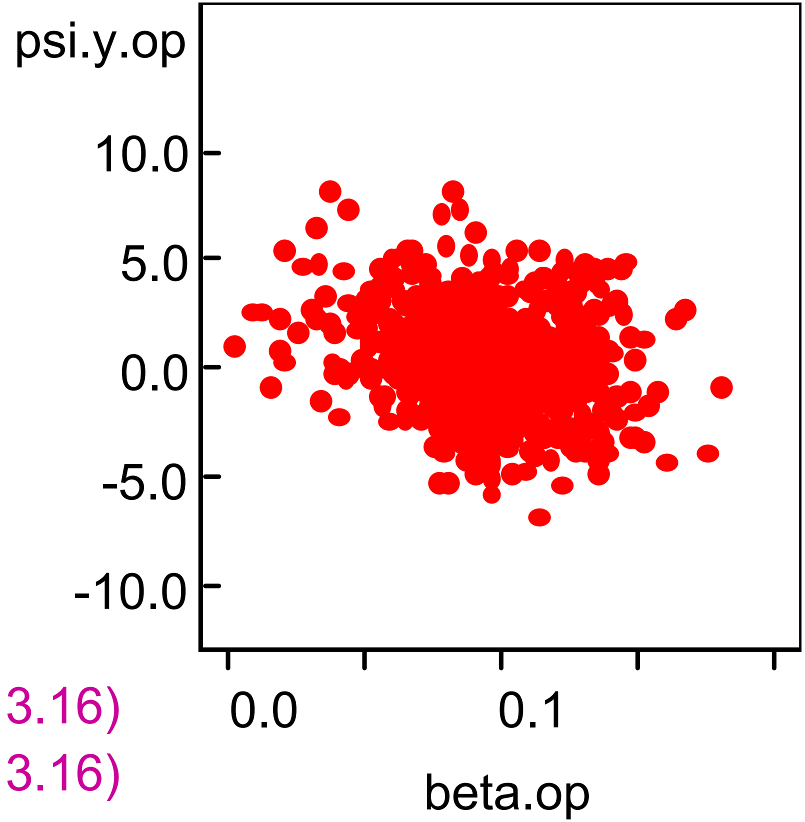
node	mean	sd	MC error
beta.op	0.0958	0.0269	8.73E-4
psi.op	0.1535	0.5442	0.01528
psi.time[1]	1.313	0.5029	0.01165
psi.time[2]	-7.868	3.214	0.2989
psi.time[3]	-8.669	3.202	0.2975
psi.time[4]	-7.579	3.102	0.2882
psi.time[5]	-9.554	3.164	0.2935
psi.time[6]	-8.63	3.157	0.2933
psi.x0	-1.144	2.734	0.1337
psi.yl	5.691	2.187	0.204
sigma	0.0674	0.004559	8.528E-5
sigma.alpha	0.08268	0.0108	1.551E-4
sigma.psi.rand	1.552	0.3002	0.00889

Availability part unstable

A Selection Model-3

Model 3 and: logit(available)=psi.time[j]+psi.y0*(y0[i]-mean(y0)) + psi.op*op[i] +psi.y*Y[i,j] +psi.y.op*op[i]*Y[i,j]+RE[i]

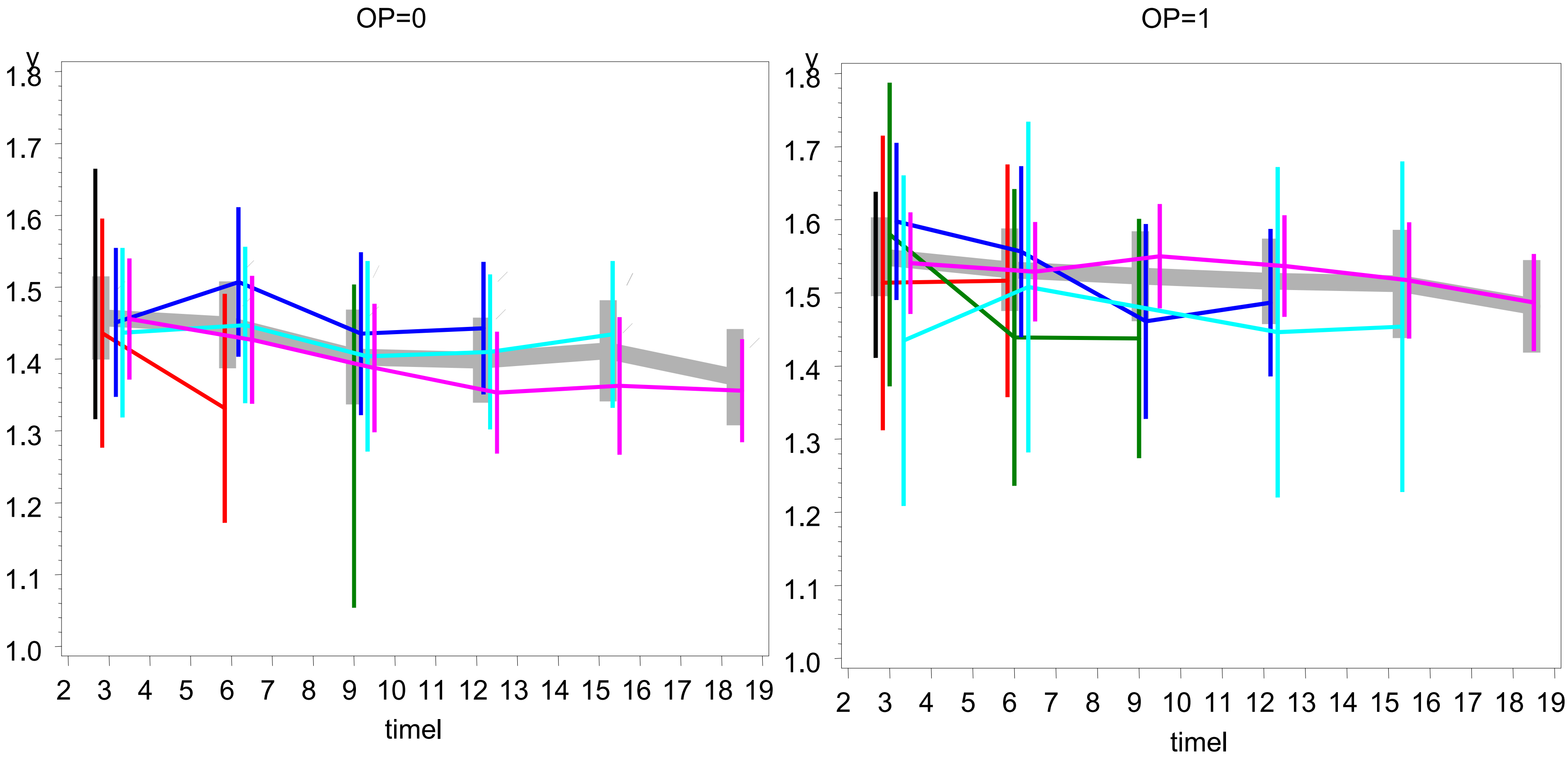
node	mean	sd	
beta.op	0.09555	0.02718	
psi.op	0.4568	0.6007	
psi.time[1]	1.13	0.4918	
psi.time[2]	0.4499	0.4776	
psi.time[3]	-0.3683	0.4924	
psi.time[4]	0.4992	0.5013	
psi.time[5]	-1.416	0.5298	
psi.time[6]	-0.4258	0.5317	
psi.x0	1.56	3.52	
psi.y	1.11	3.102	(prior 0, 3.16)
psi.y.op	0.7409	2.385	(prior 0, 3.16)
sigma	0.06765	0.004634	
sigma.alpha	0.08355	0.0109	
sigma.psi.rand	1.528	0.2928	



Selection Models Critique

- Bayesian selection models for missingness provide a framework for sensitivity analysis of missingness
- Little orientation because of vast multitude of plausible models
- Results were reassuring only to a small degree as compared to sophistication of methods
- How to identify a missingness parameter that is critical with respect to bias and variance of treatment effect(s)?

Stratification by Dropout Pattern



$Y = \log_{10} \mathrm{FEV}_1[\%]$

Stratification by Dropout Pattern

- entered baseline, treatment, time and pattern into a mixed model
- no significant differences in level and trend between patterns but low power
- Treatment effect reproduced
- Interaction between pattern and treatment: No hints, but very low power.

Validity Consideration

Comparability, Homogeneity

- In principle less established than in a randomized study
- Here treatment assignment did not lead to a strong heterogeneity
- Modelling with adjustment for baseline variables necessary, for lack of randomization, which entails ...

Model Uncertainty

- W.resp.to choice of baseline variables,covariance structure and modelling of the treatment effect
- Cannot be removed. It‘s the price to be paid for lack of randomization.

Missing Values

- *LOCF unfairly favours surgery. Available case analysis prone to bias & variance.*
- *Mixed model analysis fairly tolerant*

Conclusions

- The study does help assessing the treatment benefit, as long as no randomized comparisons are available
- Results conform to the randomized study of Geddes
- It is necessary to demonstrate a sustained treatment effect. Therefore a longer and more complete follow up is mandatory. But mixed model analyses allows some extrapolation.
- A lot of information gained from an observational study, but enough question marks remain to ethically justify a subsequent randomized study.

Missing Values

- Longitudinal data appear to tolerate a fairly high amount of missingness
- The endeavour to assess the role of missing values by sensitivity analyses is hampered by model uncertainty. For modeling drop out mechanisms we have more models and less data than for the primary analysis.
- Even if there are indication that MAR is violated, treatment effect is not necessarily biased.
- Is there are no hints against MAR, the treatment effect is not necessarily unbiased.
- Bayesian analysis is a complement that can enhance trustworthiness of REML analysis in the presence of missing values