UNIVERSITEIT HASSELT

CENTER FOR STATISTICS

Nocturnal wandering among cognitively impaired elderly inpatients in France: cross-sectional and longitudinal analyses of predictors

By Soutrik BANERJEE

Internal supervisor: Pr. dr. Roel BRAEKERS
External supervisor: Pr. dr. Pascal COUTURIER

Thesis submitted in partial fulfilments of the requirements for the degree of Master of Science in Biostatistics

October 2008

Universiteit Hasselt

Censtat

Nocturnal wandering among cognitively impaired elderly inpatients in France: cross-sectional and longitudinal analyses of predictors

By: _____; date: _____
Soutrik BANERJEE

Internal supervisor: _____; date: _____
Pr. dr. Roel BRAEKERS

External supervisor: _____; date: _____
Pr. dr. Pascal COUTURIER

Reader: _____; date: _____

ABSTRACT

INTRODUCTION: Abnormal nocturnal wandering is well known among the elderly having varying degrees of cognitive function, although its objective assessment using modern technology is seldom carried out. This study was carried out on geriatric inpatients in a French University Hospital situated at Grenoble, France. METHODS: Patient data were collected from March 2003 until November 2004. Hospitalised elderly patients were monitored using eight passive infra-red sensors placed at strategic locations at a height of 2-3 metres above the floor. Nocturnal patient activity (00:00-06:00) was captured by the sensors and stored as a data file. A validated algorithm analysed the stored data and estimated the duration of activity. Nocturnal activity was the outcome variable (N = 21) measured in patients after ≥ eight nights of hospital stay. The principal exposure variable was cognitive function (Mini-Mental State Examination (MMSE), cut-off = 24/25). Median nocturnal activity (of seven consecutive nights) was modelled cross-sectionally to determine if cognitive function and other covariates were significant predictors of high nocturnal activity. Nocturnal activity was also modelled longitudinally to determine, in addition to the variables used for cross-sectional analysis, if nocturnal activity changed with time (over eight consecutive nights). RESULTS: Significant heteroscedasticity was observed with nocturnal activity. Cross-sectional analysis revealed 'cognitively impaired' group (MMSE < 25) had 6.1 (2.3-16.4, 95% CI) times the median nocturnal activity compared to the 'normal' cognitive function group. In addition, patients having old age depression (Geriatric depression Scale (GDS), 0-30 score, continuous) and less fragile patients (Waterlow score < 11) were identified as other significant factors of greater nocturnal activity in a multivariate model. Longitudinal analysis revealed that time (nights) was not significantly (P = 0.76) associated with nocturnal activity in a model that included subject-specific random-effects; no interaction was observed between time and cognitive function. 'Cognitively impaired' group had significantly 3.4 (1.3-8.6, 95% CI) times nocturnal activity compared to the 'normal' cognitive function group. No other variable was significant in the random-effects model. Significant inter-patient heterogeneity was noted (P = 0.05). CONCLUSION: Poor cognitive function among the elderly is significantly associated with abnormal nocturnal wandering in hospitalised patients, although short-term (eight days) treatment (after at least eight days of hospitalisation) does not result in improvement of nocturnal wandering among such patients.

ACKNOWLEDGEMENTS

I would personally like to thank all the patients and their family, whose participation in this study helped to create this dataset. I thank Pr. Pascal Couturier for being my guide during the entire period for this study. In addition, I would like to thank Pr. Alain Franco, Mr. François Steenkeste, Mr. Philippe Moulin, Ms. Josette DiPasquale, Ms. Christiane Rteil, Late Mr. René Culet, Ms. Gina Corte-Franco, Ms. Vanina Bongard, Ms. Johanne Garçon, Ms. Véronique Moulin, Ms. Sylvie Charbonnier, Pr. Jean-François Le Bas, Pr. Jean-Luc Bosson, Late Pr. Michel Frossard, and Pr. Françoise Forette for their cooperation and help in my research in France. I am also very much grateful to the nurses / night staffs of the Élisée Chatin at the CHU, Grenoble, who recorded patients' behaviour every night. At Censtat, Diepenbeek, I would like to thank Pr. Roel Braekers for his valuable guidance and advice in finishing this project. In addition, I wish to thank Pr. Geert Molenberghs, Pr. Noël Veraverbeke, Pr. Paul Janssen, Pr. Marc Aerts, Pr. Geert Verbeke, Pr. Marc Buyse, Pr. Tomasz Burzykowski, Pr. Ziv Shkedy, Pr. Clarice Démetrio, Pr. Christophe Lange, Dr. Thierry van Effelterre, Dr. Herbert Thijs, Dr. Niel Hens, Ms. Martine Machiels, and Ms. Viviane Mebis for their wonderful teaching and support during my Master's programme Diepenbeek. Besides, I earnestly thank my classmates, buddies, and colleagues from Belgium: Wannes Rosius, Steven Moors, Kris Lauwers, Steven Nijs, Nyankomo Marwa, Pryseley Assam Nkouibert, Anchang Benidict Nchang, Kasim Adetayo, Serena Arima, Nuria Porta, Annemie Hoogstoel, Teele Tamme, François Haguinet, Florent Richy, Erik Spaepen, Liesbeth Boeckx, Patricia Grossman, Celestina Arrigo, Dorothy Keininger, Maeva Germé, Oana Purcaru, Simon Borghs, Christophe Gervasoni, Laure Delbeque, Károly Kulich, Enkeleida Nikai, Martin Brown, Lance Richard, Seng Tan, Lawrence Anochie, Angela Blackwell, Sabine Bouchard, Manuela Iseni, Deidre Mitchell, Jerry Welkenhysen-Gybels, Jean-Yves Pirçon, Stefanie Larmuseau, Vincent Calomme, Nelly Sekirin, Ada Sekirin, Peter van Reusel, Laurie Frenkel, Raphaël Martinelli, Patrick Dorland, and many others whose names are omitted by mistake. Finally, I thank my parents, my wife Claire, and my in-laws family for all their support for me. The fundings received from the *France Alzheimer association et maladies apparentées*, the *bourse Agrica* helped us to carry out this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADL	Activities of Daily Living
CI	Confidence Interval
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, 4th edition
FMM	Finite Mixture Model
G.A.R.D.I.E.N.	*Gérontologie Assistée par la Recherche et le Diagnostic des Incidents et des Errances Nocturnes*
GDS	Geriatric Depression Scale
GEE	Generalised Estimating Equation
GLMM	Generalised Linear Mixed Model
GLM	Generalised Linear Model
IADL	Instrumental Activities of Daily Living
MMSE	Mini-Mental State Examination
MQL	Marginal Quasi-Likelihood
PQL	Penalised Quasi-Likelihood
SD	Standard Deviation
SE	Standard Error
UTI	Urinary Tract Infection

1 INTRODUCTION

Abnormal nocturnal wandering phenomenon is well known among elderly subjects with varying degrees of cognitive function, although its objective assessment using modern technology to characterise the normative patterns is seldom carried out in the domiciled, hospitalised, or institutionalised elderly subjects.

With the advancement of new patient-monitoring technologies, measuring patient activity by different methods have gained popularity [*Brown et al. 1990; Hilten et al. 1993; Miskelly et al. 2001; Sadeh et al. 1995; Someren et al. 1997; Yesavage et al. 1998*]. Most commonly employed methods consist of different types of wrist 'actimeters' worn on the non-dominant wrist of the subjects [*Acebo et al. 1999; Aharon-Peretz et al. 1991; Lavie et al. 1992; Lemke et al. 1999*]. In recent times, new methods of distant patient-monitoring are also beginning to make an impression in the community and institutionalised patients by use of environmental (distant) sensors, *e.g.*, infrared movement detectors, installed in the patient's habitat [*Ohta et al. 2002*] or hospitals [*Banerjee et al. 2003; Couturier et al. 2002*]. There is a distinct advantage of patient-monitoring by environmental sensors, where measurement of activity can continue for long durations either at home or in the institutions, since the patient will not have to wear any device. Moreover, in demented patients, the compliance of wrist 'actimeters' is questionable [*Banerjee et al. 2003; Ohta et al. 2002*], where passive environmental monitoring may offer a possible remedy. Although wrist actimetry has been growing in use since the late seventies [*Sadeh et al. 1995*], there is a lack of reports describing the pattern of nocturnal activity in patients suffering from dementia. Monitoring by environmental sensors is still at an early stage as far as actimetric data obtained by observing patients either at home or in the institution is concerned.

The objective of this study is twofold. First, to observe in a cross-sectional manner, if median nocturnal activity measured in time units (minutes or seconds) could be explained by clinically measured patient's characteristics using standard tools for assessment. Second, to determine in a longitudinal manner using repeated measurements, if there was a time-trend in nocturnal activity as well as to determine which patient characteristics could significantly predict nocturnal activity. Of particular interest in this study was to examine the effect of the cognitive function-based two exposure groups categorised as: 'cognitively impaired' and 'normal' cognitive function.

2 METHODS

2.1 Nocturnal activity measurement in patients with G.A.R.D.I.E.N.© system

This study was carried out in a French Geriatric University Hospital situated in the Grenoble agglomeration region in France. Patient data were collected from March 2003 until November 2004 for this study. The system and methods are described in detail elsewhere [*Banerjee et al. 2003*]. In brief, hospitalised elderly patients were monitored using eight passive infrared sensors placed at strategic locations at a height of 2-3 metres above the floor. Patient movements were captured by the sensors and stored as a data file in a computer (by cable network) placed in a separate room. A validated algorithm analysed the stored data at the end of the monitoring period everyday and produced an output of the duration of activity of the monitored patient. The patient-monitoring system, called G.A.R.D.I.E.N.©, used two thresholds: one, a minimum number of sensor activation (≥ five times), and two, a time duration limit (≥ 30 seconds) to define a 'valid' movement in order to eliminate activities, which were of small duration, like occasional leg movements, fine tremors, or jerks, etc. It only translated longer complex movements, which indicated a real displacement or activity for some duration on the part of the patient. The monitoring system in addition, could eliminate with high accuracy, the external visits by personnel or visitors by an algorithm in order that only activities or displacements by the patient alone were taken into account. Since patient activity was more likely to be confounded by the presence of visitors during the daytime or grossly underestimated due to the frequent leaving from the patient's room, we were primarily interested in analysing the stored data during the period when most patients were 'normally' asleep and minimally disturbed by visitors or night personnel within a hospital set-up for elderly patients. This observation period ranged from 00:00 to 06:00 every night (six-hour span) [*Virone et al. 2002; Paavilainen et al. 2005*]. On summing the distinct episodes of patient activities during this observation period, it was possible to extract the total or cumulative nocturnal activity for a patient measured in time units. It was also possible to classify the measured patient activity into: activity in the toilet, activity in bed, and activity in the room, although some overlap was naturally present among these classified patient activities.

2.2 Patient inclusion / exclusion, consent & ethics

Patients were required to spend at least three nights prior to data capturing for acclimatisation with the surroundings. Following this, patients were observed for at least eight nights, except in four patients in the cognitively impaired group, who did not stay this minimum required period for the study; two stayed for seven nights, one stayed for six nights and the other stayed for five nights. For these missing observations of the outcome variable (see later), comprising of 3% of full data, the mean nocturnal activity (calculated over their period of observation) was imputed in each of the four patients (*i.e.*, seven observations out of 216 observations).

The patients, who were not mobile, needed support for their displacements, having acute illness of any kind (*i.e.*, cardio-respiratory illness, infection with multi-resistant organisms, stroke, etc.), having a fear of displacement due to falls, and bed-ridden patients were excluded from this study.

The methodology was approved by the ethics committee of the University Hospital of Grenoble, France. Informed consents were obtained from all the patients or from their legal guardian / next-of-kin, where the patient was unable to sign the consent form due to advanced cognitive impairment.

2.3 Variables

The exposure variable of interest was cognitive function, measured by Mini-Mental State Examination (MMSE) [*Folstein et al. 1975*]. It was dichotomised into two categories, namely, 'cognitively impaired' group were patients with MMSE scores < 25, and 'normal' cognitive function group were patients with MMSE score ≥ 25. In the latter group, no patient was clinically diagnosed as dementia. Trained medical doctors obtained the MMSE scores from all patients.

The covariates measured were age; gender; Geriatric Depression Scale (GDS) ranging from 0-30 with higher scores indicating more depressed state [*Yesavage et al. 1983*]; Activities of Daily Living (ADL) ranging from 0-6 with scores greater than 3 indicating better autonomy [*Katz et al. 1970*]; Instrumental Activities of Daily Living (IADL) ranging from 0-14 with higher scores indicating better preserved instrumental activities of daily living [*Lawton & Brody 1969*]; Waterlow score with scores greater than 10 indicating risk of pressure sore (as a proxy for 'immobility') [*Waterlow 1985*]; psychoactive medications use (classified into: anti-dementia drugs, hypnotic / sedative drugs, antidepressants, and

neuroleptics); neoplasia; hip / knee prostheses; diarrhoea / urinary tract infection (UTI); extra-pyramidal signs; history of falls in the last six months; and history of inability to get up from the ground after fall in the last six months. In this analysis, data obtained while the patient was being monitored by an ambulatory holter device, were excluded. In some patients in the 'cognitively impaired' group, a door barrier was put in order that the patient did not go out of the room. These patients had a tendency to go out of the room very often. In such instances, the measured patient activity might have been seriously underestimated by the system compared to that without a barrier.

In the cross-sectional analysis, the outcome variable of interest was the duration of median nocturnal activity measured over seven (*i.e.,* odd no. of) consecutive nights; in case of missing data, the duration of median nocturnal activity was calculated over the available no. of observations. Median was chosen over the mean, because the former was more stable in a sense that an occasional peak in nocturnal activity would have had a less influential impact on its value, although in this data, both median and mean showed a very high degree of correlation (> 95%) [*Banerjee 2005*]. For the longitudinal analysis, the duration of nocturnal activity measured every night (for eight consecutive nights) was the outcome variable of interest, including imputed observations.

2.4 Statistical analysis

First, univariate associations between the outcome variable (median nocturnal activity) and the exposure variable (cognitive function) and covariates were tested. In addition to cognitive function, among other covariates the GDS and Waterlow score showed significant associations ($P \leq 0.10$) with the logarithm of median nocturnal activity (see below). Therefore, it was decided to keep these variables for fitting the models for cross-sectional analyses (and later also for longitudinal analyses), although there were six missing observations out of 27 for the GDS variable (five from the 'cognitively impaired' group and one from the 'normal' cognitive function group). Second, the distribution of the outcome variable (median nocturnal activity) showed a significant heteroscedasticity (*i.e.,* a very right-skewed distribution with the mean approximating the standard deviation). In such cases, a logarithmic transformation is often considered a solution to stabilise the variance and make the distribution of the outcome variable fit a normal distribution [*Bland & Altman 1996*]. Therefore, a log-transformation was performed and only four variables were selected for model fitting, given the small sample size. These variables were: cognitive function

(exposure), GDS, and Waterlow score; in addition, ADL was chosen as it was considered to be an important predictor of (nocturnal) patient activity. A stepwise backward elimination procedure was employed and only significant variables were retained in the final fitted model. Model fit was checked by plotting the residuals against the predicted and covariates (to check for non-constancy of variance or trends), normal probability plot of the residuals (for normality assumption), index plot of Cook's distance (for influential observations), and studentised residuals (for outliers). On account of the apparent mean-variance association, a generalised linear model (GLM) approach was applied (stepwise backward elimination), first, using a Poisson distribution; second, using a quasi-Poisson distribution with scale parameter (= deviance, Pearson); and third, using a negative binomial distribution. The last two options were employed, because the model fitted with a Poisson distribution showed significant overdispersion. In all these three cases, Pearson's residuals were checked for normality assumption, and studentised residuals (for outliers).

Longitudinal analysis started first by exploring the empirical distribution of nocturnal activity for the two exposure groups over time (nights) in all patients. The (Spearman's) correlations of the nocturnal activity between night 1 through night 8 were analysed. In addition to cognitive function as the exposure variable, the same three covariates (GDS, ADL, and Waterlow score) used in cross-sectional analyses were selected for model fitting. Statistical analyses were performed on the log-transformed nocturnal activity (continuous) as outcome, but the results were not presented in this report due to its departure from normality on most time-points. Instead a GLM approach was employed, where one could model the mean of the nocturnal activity to explanatory variables using a known link function. Using generalised estimating equations (GEE) [*Liang & Zeger 1986*], a marginal model assuming Poisson, quasi-Poisson and negative binomial distributions of the outcome variable, respectively, were fitted (stepwise backward elimination) to obtain a 'rough' estimate of the mean function. The last two options were employed, because the model fitted with a Poisson distribution showed significant overdispersion. The unstructured covariance matrix did not converge (most likely due to the high number of parameter estimates, relative to the small sample size). Independent, autoregressive, and exchangeable (compound symmetry) working correlation structures were fitted; model-based and empirical SEs were compared for each of the working correlation structures. Next, a generalised linear mixed model (GLMM) including subject-specific random-effects was fitted using penalised quasi-likelihood (PQL) procedure. Although the random slope model assuming Poisson distribution converged (*i.e.*, compound symmetry covariance), the negative binomial model failed to converge in this

procedure. In addition, other covariance structures did not converge assuming Poisson distribution. Finally, an adaptive Gaussian quadrature approach was used (size = 20) to fit a GLMM (stepwise backward elimination) including subject-specific random-effects. As previously, first, a Poisson model was fitted with a random slope model. This model revealed overdispersion. Second, a negative binomial model (after specifying the log-likelihood equation for the negative binomial distribution [*Gelman*]) was fitted. Outlying individuals were checked by quantile-quantile plot and histogram of the empirical Bayes estimates.

For cross-sectional (linear and GLM fits) and longitudinal (GEE and PQL fits) analyses, bootstrapping was performed to check for bias and distribution of parameter estimates. The cross-sectional models were further analysed by Jackknife-after-bootstrap to detect significant influence of any observation.

A finite mixture model (FMM) fitting of the outcome variable using the R package 'FlexMix' [*Leisch 2004*] was additionally performed. A mixture of normal distributions was fitted to cross-sectional and longitudinal data, respectively, to identify latent clusters. FlexMix provides the facility to take into account the repeated / clustered nature of the data. Principally, two clusters were identified to be of clinical interest in both longitudinal and cross-sectional models. Furthermore, an attempt was made to identify the differentiating characteristics of these two clusters (N = 17 and N = 4) by fitting cross-sectional (GLM) and longitudinal (GEE) models, with binary outcome variable (*i.e.*, those with 'high nocturnal activity' versus those with 'low nocturnal activity').

Missing data analysis was performed comparing six patients without the GDS covariate and 21 patients with the GDS covariate, who were included in all analyses, with respect to population characteristics.

2.5 Softwares

All statistical analyses were performed with SAS 9.1.3 [*SAS Institute Inc.*] and R 2.7.1 [*CRAN*].

3 RESULTS

Table 3-1 Demographic and clinical characteristics of *all* study patients; N = 27:

Characteristics	Exposure groups		*P*-values
	'Cognitively impaired' (MMSE ≥ 25) (N = 21)	**'Normal' cognitive function (MMSE < 25) (N = 6)**	
Age (years), mean (SD)	82.4 (6.0)	81.7 (7.2)	0.91
Men, N (%)	9 (43%)	2 (33%)	1.00
MMSE, mean (SD)	13.0 (5.2)	27.0 (1.3)	**< 0.01**
ADL, mean (SD)	4.0 (1.2)	4.8 (1.7)	0.16
IADL, mean (SD)	4.1 (4.2)	7.2 (4.6)	0.21
GDS, mean (SD)	14.2 (5.7)[1]	12.8 (7.1)[2]	0.77
Waterlow, mean (SD)	11.8 (3.7)	15.8 (4.5)	**0.05**
Hip / knee prosthesis, N (%)	2 (10%)	2 (33%)	0.20
Psychoactive medications, N (%)	20 (95%)	4 (67%)	0.11
Anti-dementia drugs, N (%)	13 (62%)	0 (0%)	**0.02**
Hypnotic / sedative drugs, N (%)	17 (81%)	4 (67%)	0.59
Antidepressants, N (%)	8 (38%)	3 (50%)	0.63
Neuroleptics, N (%)	13 (62%)	1 (17%)	0.08
Neoplasia, N (%)	2 (10%)	1 (17%)	0.54
Continent, N (%)	8 (38%)	5 (83%)	0.13
Presence of UTI / Diarrhoea, N (%)	2 (10%)	1 (17%)	0.54
Holter monitoring, N (%)	2 (10%)	0 (0%)	1.00
Extra-pyramidal signs, N (%)	7 (33%)	2 (33%)	1.00
Door barrier use at night, N (%)	6 (29%)	0 (0%)	0.28
Falls in the last six months, N (%)	14 (67%)	4 (67%)	1.00
Unable to get up after fall in the last six months, N (%)	12 (57%)	3 (50%)	1.00
Nocturnal activity (minutes), mean (SD)[3]	22.0 (21.3)	4.9 (3.3)	**0.01**

MMSE = Mini-Mental State Examination; ADL = Activities of Daily Living; IADL = Instrumental Activities of Daily Living; GDS = Geriatric Depression Scale; UTI = Urinary Tract Infection;
[1] N = 16; [2] N = 5; [3] Nocturnal activity averaged over 8 successive nights;
Continuous outcomes between the two groups were compared by Mann-Whitney-Wilcoxon test and binary outcomes between the two groups were compared by Fisher test.

A total of 216 nights (161 nights for the 'cognitively impaired' group and 48 nights for the 'normal' cognitive function group) was monitored in the full dataset (N = 27). In the 'cognitively impaired' group, the different clinical diagnoses were: 10 patients diagnosed as Alzheimer's disease (DSM-IV of Mental Disorders: probable type), five patients with dementia of Lewy body type, four patients with vascular dementia, one patient with fronto-temporal dementia, and one patient having mild cognitive impairment of uncertain origin.

The demographic and clinical characteristics were similar across the two groups in Table 3-1. Waterlow score was significantly different, which was not unexpected given the

number of univariate comparisons. The difference noted in the use of anti-dementia drugs was expected, because the 'cognitively impaired' group mainly consisted of dementia patients due to various aetiologies. Furthermore, the MMSE scores were of principal interest in this study; that is, to classify the two exposure groups on the basis of MMSE scores and subsequently study the effect of each exposure group on nocturnal activity. Finally, we also observed that the outcome variable: mean nocturnal activity, was different among the two groups. It has to be borne in mind that this finding was based on a non-parametric comparison of the distributions of nocturnal activity between the two exposure groups without adjusting for potential confounders that might arise in an observational study. In the next sections, we modelled the outcome variable by different methods, adjusting for covariates, both in cross-sectional and longitudinal set-ups, to determine the significant predictors for increased nocturnal activity in a heterogeneous group of geriatric inpatients having different levels of cognitive function.

All the above tests were non-parametric given the small sample size in this study and non-normal distribution of the outcome and several other parameters. This meant that the inferences derived from these comparisons were more conservative compared to standard parametric approaches.

Table 3-2 Cross-sectional analysis between the outcome (= median nocturnal activity, from night 1 to night 7, from 00:00 to 06:00) and significant predictor variables; N = 21 (6 patients excluded due to missing data for the GDS variable):

Parameters	Log-transformed outcome variable		Poisson model		Negative binomial model	
	Estimates (SE)	***P*-values**	**Exponentiated estimates (95% CI)**	***P*-values**	**Exponentiated estimates (95% CI)**	***P*-values**
Intercept	4.58 (0.63)	< **0.01**				
Cognitive function						
MMSE < 25	2.18 (0.54)	< **0.01**	5.41 (2.91-10.04)	< **0.01**	6.12 (2.28-16.42)	< **0.01**
MMSE ≥ 25	0.00 (referent)		1.0 (referent)		1.0 (referent)	
GDS (0-30)	-0.16 (0.04)	< **0.01**	0.88 (0.86-0.90)	< **0.01**	0.88 (0.83-0.94)	< **0.01**
ADL						
≤ 3			1.74 (1.38-2.19)	< **0.01**		
> 3			1.0 (referent)			
Waterlow						
> 10			0.30 (0.24-0.38)	< **0.01**	0.38 (0.19-0.77)	< **0.01**
≤ 10			1.0 (referent)		1.0 (referent)	
Dispersion parameter[1]					0.39 (0.08-0.71)	< **0.01**

[1] Dispersion parameter (> 0) (95% CI) was not exponentiated;
Only significant variables are retained in the models.

The main outcome variable of interest was median nocturnal activity (7 nights), from 00:00 to 06:00.

The main predictor variable of interest was binary: those with MMSE < 25 ('cognitively impaired') and the others with MMSE ≥ 25 ('normal' cognitive function). The former consisted of geriatric patients having varying degrees of cognitive function loss, mostly due to dementia of various aetiologies and various levels of severity, whereas the latter consisted of geriatric patients not suffering from any form of clinically detectable dementia.

The covariates cognitive function (exposure group), GDS, Waterlow score, and ADL were entered to fit the general linear model (as previously described). Other covariates, which did not show a significant association ($P < 0.10$) with the logarithm of median nocturnal activity, *i.e.*, age, gender, antidepressants, neuroleptics, hypnotics / sedatives use, IADL, continence (urinary or faecal or both), extra-pyramidal signs, hip / knee prosthesis, diarrhoea / UTI, neoplasia, history of falls in the last six months, and history of inability to get up from the ground after fall in the last six months, were not included to fit the models. This conservative strategy was employed because in our study, the final sample size to fit the model was small (N = 21). Therefore, we had to be parsimonious in choosing the number of covariates to fit the model.

In Table 3-2, only the variables significant ($P \leq 0.5$) in the models are retained in the final model. First, a log-transformed outcome variable was fitted. Only the exposure group and GDS were the significant variables (ADL and Waterlow score were not significant). Further exploration of the fit revealed a good model fit (normality assumption of the residuals and linear normal probability plot, $R^2_{adjusted} = 0.57$), no influential observation (Cook's distance < 0.4), no non-constancy of variance or systematic pattern of the residuals plotted against the covariates or predicted, and no outliers.

In order to provide more flexibility in our modelling approach, a GLM framework was employed. In this instance, the outcome variable was not transformed, but a link function was employed to model the mean of the outcome variable of interest with the covariates. The most natural choice was to use a Poisson approach, as activities can be considered to be counts and always greater than or equal to zero [*Sadeh et al. 1995; Sadeh & Acebo 2002*]. This revealed overdispersion (deviance / degrees of freedom = 6.15). As a consequence, a negative binomial approach was used to correct for this overdispersion (for this data, this approach was found to be better than a quasi-Poisson approach). The dispersion factor = 0.39 (0.08-0.71, 95% CI; $P < 0.01$) and deviance / degree of freedom = 1.30. It can be observed in this case that the CIs of all the (exponentiated) mean parameter estimates became wider as compared to the Poisson model. In addition, the ADL was not anymore significant in the negative binomial model. In the fitted negative binomial model, the Pearson's residuals were normally distributed ($P = 0.13$; Kolmogorov-Smirnov's test), and there were no outliers.

Table 3-3 Longitudinal analysis between the outcome (= nocturnal activity, from night 1 to night 8, from 00:00 to 06:00) and significant predictor variables; N = 21 (6 patients excluded due to missing data for the GDS variable):

Marginal GEE models

Parameters	Poisson model (exchangeable working correlation)		Negative binomial model (exchangeable working correlation)	
	Exponentiated estimates (95% CI; empirical)	***P*-values**	**Exponentiated estimates (95% CI; empirical)**	***P*-values**
Time (nights)	0.99 (0.95-1.02)	0.39	0.99 (0.94-1.03)	0.54
Cognitive function				
MMSE < 25	2.89 (1.39-6.00)	**< 0.01**	3.22 (2.11-4.90)	**< 0.01**
MMSE ≥ 25	1.0 (referent)		1.0 (referent)	
GDS (0-30)	0.93 (0.88-0.98)	**0.01**		
ADL				
≤ 3	1.83 (1.18-2.85)	**< 0.01**	2.05 (1.13-3.72)	**0.02**
> 3	1.0 (referent)		1.0 (referent)	
Waterlow				
> 10	0.43 (0.26-0.70)	**< 0.01**	0.49 (0.23-1.01)	**0.05**
≤ 10	1.0 (referent)		1.0 (referent)	
Dispersion parameter[1]			2.42 (2.00-2.84)	**< 0.01**

[1] Dispersion parameter (> 0) (95% CI) was not exponentiated.

Exploratory longitudinal analyses (Figure 6-1, Figure 6-2, Figure 6-3, & Figure 6-4 in the APPENDIX) indicated towards a possibility of a mean-variance relationship for the distribution of the outcome variable. In addition, from the results obtained from the previous cross-sectional analyses (Table 3-2), it was decided to retain the same four covariates, namely, cognitive function (exposure group), GDS, ADL, and Waterlow score for longitudinal data analysis.

Marginal GEE models were first fitted to explore the mean function only. First on fitting a Poisson GEE model (Table 3-3), all the four covariates were significant. However, overdispersion was observed (deviance/degrees of freedom = 1075.75). Therefore, a negative binomial GEE model (Figure 6-5) was fitted, where deviance / degrees of freedom = 1.27, correcting for the overdispersion to a large extent; the dispersion factor = 2.42 (2.00-2.84, 95% CI; $P < 0.01$). The CIs (empirical) of the (exponentiated) mean parameter estimates became wider than those obtained using the Poisson GEE model, except for cognitive function (exposure group). GDS was no longer significant in this model. On trying different correlation structures, namely, independent, exchangeable (= compound symmetry), and auto-regressive, the model-based SEs were found to be somewhat different from the empirical SEs for all the three correlation structures. The unstructured correlation structure did not converge, given the small sample size (N = 21) relative to the large number of parameters that

needed to be estimated. The empirical SEs obtained by fitting the three correlation structures were not too different from one another, although in the given situation, an exchangeable structure was considered the most 'appropriate' (because the time gap separating the measurements were not distant in an individual), which assumes equal correlation between any two outcome measures.

In Table 3-3, only significant covariates were retained in the respective models, with the exception of the time (nights) variable (see also Figure 6-5 in the APPENDIX). It can be noted that time (nights) had no impact on the outcome. In addition, the interaction between the time (nights) and exposure group was not significant (result not shown).

Table 3-4 Longitudinal analysis between the outcome (= nocturnal activity, from night 1 to night 8, from 00:00 to 06:00) and significant predictor variables; N = 21 (6 patients excluded due to missing data for the GDS variable):

Mixed-effects PQL model (GLMM)

Parameters	**Poisson model with random slope & intercept (compound symmetry covariance)**	
	Exponentiated (REML) estimates (95% CI)	***P*-values**
Time (nights)	0.99 (0.95-1.03)	0.62
Cognitive function		
MMSE < 25	3.77 (1.61-8.84)	**< 0.01**
MMSE ≥ 25	1.0 (referent)	
GDS (0-30)	0.96 (0.90-1.02)	0.23
ADL		
≤ 3	2.97 (1.16-7.61)	**0.02**
> 3	1.0 (referent)	

Usually PQL estimates are higher than their GEE counterparts, which is the case here. In addition, more heterogeneity (variability) in the PQL Poisson model compared to the GEE Poisson model can be noted in the wider CIs of (exponentiated) parameter estimates.

The unstructured option for the variance structure did not converge, given the small sample size (N = 21) and the large number of parameters that needed to be estimated. The other options like autoregressive, Toeplitz, and simple did not converge either. None of the negative binomial models converged, for which reason, the parameter estimates were not provided in Table 3-4. As with previous analyses from marginal GEE models, a compound symmetry covariance option appeared 'reasonable' for model fitting in this case. In addition, as random-effects are not null in this model, the Poisson marginal quasi-likelihood (MQL) model was not tried given its poorer performance in the presence of random-effects [*Molenberghs & Verbeke 2005*].

It can be noted in the Poisson PQL model that time (nights) had no impact on the outcome. In addition, the interaction between the time (nights) and exposure group was not significant (result not shown). A significant random slope in this model implies that variance changes with time (by -2loglikelihood difference compared to the random intercept model).

The fitted model in Table 3-4 is only a preliminary trial to fit a hierarchical model including random-effects. In the next subsection, an approach with adaptive Gaussian quadrature was used to fit hierarchical models.

Table 3-5 Longitudinal analysis between the outcome (= nocturnal activity, from night 1 to night 8, from 00:00 to 06:00) and significant predictor variables; N = 21 (6 patients excluded due to missing data for the GDS variable):

Mixed-effects model with adaptive Gaussian quadrature (size = 20) (GLMM)

Parameters	Poisson model		Negative binomial model	
	Exponentiated estimates (95% CI)	***P*-values**	**Exponentiated estimates (95% CI)**	***P*-values**
Fixed-effects estimates				
• Time (nights)	0.99 (0.95-1.03)	0.63	0.98 (0.88-1.10)	0.76
• Cognitive function				
MMSE < 25	3.25 (1.29-8.21)	**0.02**	3.38 (1.33-8.58)	**0.01**
MMSE ≥ 25	1.0 (referent)		1.0 (referent)	
Random-effects estimates[1]				
• d11	1.03 (0.36-1.70)	**< 0.01**	0.48 (0.00-0.96)	**0.05**
• d12	-0.05 (-0.10-0.00)	**0.05**		
• d22	0.01 (0.00-0.02)	**< 0.01**		
Dispersion parameter[1]			2.02 (1.57-2.47)	**< 0.01**

[1] Random-effects estimates (95% CI) and dispersion parameter (95% CI) were not exponentiated.

The term d11 is the variance of the random intercept, and d22 is the variance of the random slope, and finally d12 is the covariance between d11 and d12.

Finally, an adaptive Gaussian quadrature was employed for model fitting. A mixture model is a flexible way to account for overdispersion. The mean of the distribution is Poisson, but the mean itself varies according to some, most commonly gamma, distribution [*Agresti 2002*]. First, using a Poisson approach (Table 3-5, see also Figure 6-6 in the APPENDIX), one can observe that the -2loglikelihood = 102,943 was high (a lower value implies better model fit). Only the exposure group remained significant, but Waterlow score, ADL, GDS, 'exposure group x time' interaction (not shown) variables as well as the time variable were no more significant. The random slope was significant, implying changing variance with time. Quadrature-based methods use numerical method (grid) compared to the PQL method (Table 3-4), which uses linearisation of the mean to estimate the parameters. As observed previously with the PQL method, the Poisson GLMM (quadrature-based) model also has more heterogeneity than the Poisson GEE model. Marginally, the gamma mixture of the Poisson distributions yields the negative binomial distribution [*Agresti 2002*]. However, one cannot be sure if including subject-specific random-effects to account for the correlation between the repeated measurements 'fully' corrected for the overdispersion [*Agresti 2002; Der & Everitt 2006*]. Therefore, an adaptive Gaussian quadrature was again employed to fit a negative binomial mixed-effects model using the same four covariates and 'exposure group x

time' interaction (not shown) (Table 3-5, see also Figure 6-7 in the APPENDIX). Interestingly, one can observe that essentially this model had the same fixed-effects variable (*i.e.*, only the exposure group) as significant, like it was in the above Poisson mixed-effects model; however, the dispersion parameter was found to be significant (and -2loglikelihood = 2,540.0 was lower than that obtained in the Poisson mixed-effects model). In addition, the random slope, d22, and the covariance between random slope and intercept, d12, were no more significant. This signifies that the random slope variance (& covariance) of the above Poisson mixed-effects model has been 'taken-up' by the significant dispersion parameter ($P < 0.01$) in the negative binomial mixed-effects model. The findings also imply that one must not stop after fitting a hierarchical Poisson model including subject-specific random-effects and assume that this would automatically account for 'all' overdispersion arising from correlation between repeated measurements in the data. Notwithstanding, the significant dispersion parameter of the negative binomial mixed-effects model 'did not fully' capture all the subject-specific random-effects of the above Poisson mixed-effects model; d11 was still borderline significant ($P = 0.05$) in the negative binomial mixed-effects model implying some 'residual' between-subject heterogeneity at baseline; and again, on comparing the -2loglikelihood value of the fitted negative binomial mixed-effects model to that of a negative binomial model without random-effects (*i.e.*, without d11), the G^2 statistic was found to be significant. In the final negative binomial model, no outlying subject was identified (empirical Bayes estimates).

Bootstrap and Jackknife

For cross-sectional (linear model and GLM fits) and longitudinal (GEE and PQL fits) analyses, bootstrapping was performed. No large bias of parameter estimates was found (result not shown).

Additionally, Jackknife-after-bootstrap was performed to look for any influential observation in cross-sectional models (linear model and GLM). No significant influence of any observation was found (result not shown).

Finite Mixture Models

Given the highly (right) skewed distribution of the data, an FMM approach was employed to identify and characterise any latent clusters. In the first instance, median nocturnal activity was the outcome, and in the second instance by taking into account the clustered nature of the data, nocturnal activity was the outcome.

Table 3-6 Two FMMs were fitted using the R package 'FlexMix' (N = 21). Prior and posterior probabilities and parameter estimates (= mean (SD)) are shown for each FMM by components:

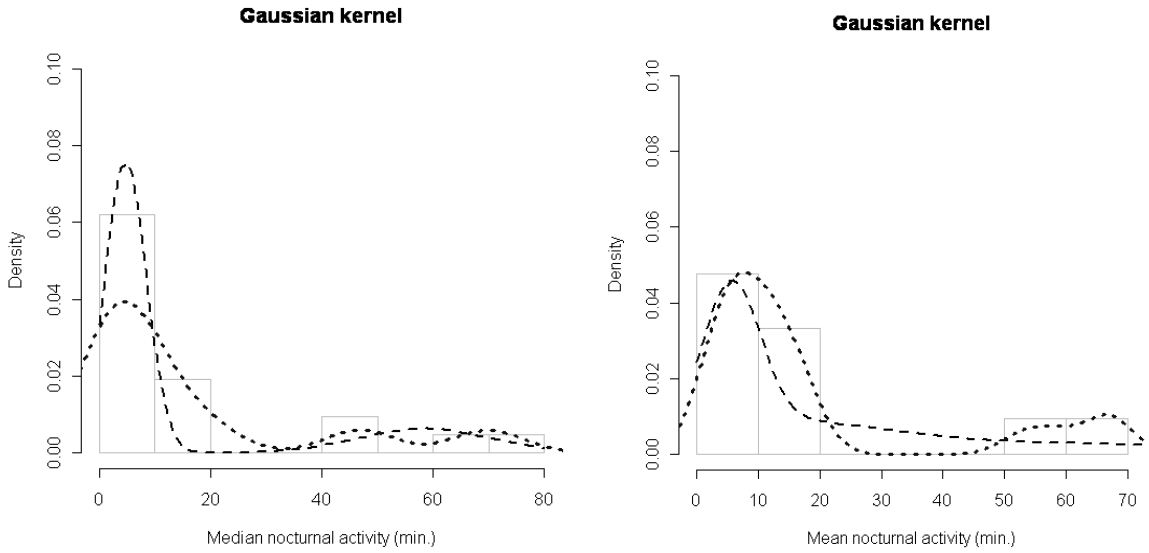
Median nocturnal activity (min.)			**Nocturnal activity (as clusters of repeated measures) (min.)**		
Prior probability	**Posterior probability**	**Mean (SD)**	**Prior probability**	**Posterior probability**	**Mean (SD)**
0.71	0.71	4.77 (3.74)	0.45	0.48	5.73 (4.68)
0.10	0.10	19.66 (0.04)	0.36	0.33	13.86 (18.02)
0.19	0.19	58.41 (11.95)	0.19	0.19	61.47 (27.92)

It can be seen in Table 3-6, for both cases, that the last component has consistent proportions (= 19%). This corresponds to 4 patients having excessive nocturnal activity. In addition, the prior and posterior probabilities were very close for each component in each FMM. In order to simplify analysis and be consistent for both cross-sectional and longitudinal analyses, it was decided to combine patients from the first and second component into one single cluster (N = 17), as 'low nocturnal activity' group, and the remaining (N = 4) into another, as 'high nocturnal activity' group. Next, GLM and GEE models were fitted, for binary outcomes (*i.e.*, high versus low nocturnal activity groups), with same covariates as previously, for cross-sectional and longitudinal analyses, respectively; however, GLMM failed to converge in longitudinal analysis.

In the cross-sectional analysis, only GDS was borderline significant ($P = 0.06$). In the longitudinal analysis, only GDS and Waterlow score were significant predictors to distinguish

the two defined clusters, $P = 0.03$ and $P < 0.01$, respectively. This signified that patients with 'high nocturnal activity' were significantly different from patients with 'low nocturnal activity' in having lower depressive mood and lower immobility. Cognitive function was not a significant predictor in cross-sectional or longitudinal analyses.

Figure 3-1 on the *left* shows median nocturnal activity (histogram) with dotted line (density smoothing of the histogram), and dashed line (FMM). Figure on the *right* shows mean nocturnal activity (histogram) with dotted line (density smoothing of the histogram), and dashed line (FMM taking into account the clustered nature of the data); N = 21.



Missing data analysis for the covariate GDS

Patients having (N = 6) and not having (N = 21) the GDS variable were compared in terms of population characteristics in separate univariate analyses. The variables that were compared: MMSE, ADL, Waterlow score, IADL, age, gender, extra-pyramidal signs, hip / knee prosthesis, antidepressants, neuroleptics, hypnotics / sedatives, continence (urinary or faecal or both), diarrhoea / UTI, neoplasia, history of falls in the last six months, history of inability to get up from the ground after fall in the last six months, and the logarithm of median nocturnal activity. None of these variables were significantly different between the two groups.

In addition, all these variables (MMSE and covariates) were entered in a multiple logistic regression model to estimate the probability of having to not having the missing GDS variable explained by these variables (stepwise backward elimination). Again, none of the variables were retained in the final model.

4 DISCUSSION

The current study, in which 21 elderly hospitalised patients were monitored using passive infrared sensors for 163 nights in a French university hospital, provides evidence that globally impaired cognitive function is a significant predictor of increased nocturnal activity, although short-term (eight days) treatment (after at least eight days of hospitalisation) does not result in improvement of nocturnal wandering among such patients.

In the cross-sectional analysis, in which median nocturnal activity measured over seven consecutive nights was the outcome, patients with impaired cognitive function, having less depressive mood, and a lower risk of pressure sore (*i.e.*, less mobile patients due to their frail condition) had an increased propensity for nocturnal wandering. In the longitudinal analysis, in which daily nocturnal activity was the outcome measured over eight consecutive nights, patients with impaired cognitive function, having lower autonomy in basic daily activities, and who were less mobile had an increased propensity for nocturnal wandering. In addition, there was no effect of time, signifying that nocturnal activity did not increase or decrease with time in a systematic way. Furthermore, in another longitudinal analysis, which included subject-specific random effects (*i.e.*, between-subject variability), only impaired cognitive function was a significant predictor of increased day-to-day nocturnal activity. The between-subject variability was found to be borderline significant ($P = 0.05$). There was no difference in the evolution of each exposure group with respect to time.

The nocturnal activity was measured by a previously validated system [*Banerjee et al. 2003*]. The exposure variable was measured by MMSE, which is typically used to measure global cognitive function in elderly subjects. Furthermore, the other covariates used in the analyses, namely, ADL, GDS, and Waterlow score provided proxy for autonomy, depressive mood, and immobility in elderly subjects. The distributions of the outcome variable in both cross-sectional and longitudinal analyses were highly right-skewed. For this reason, GLM approach was used in both instances. However significant overdispersion in the data required fitting negative binomial distribution in place of a Poisson distribution.

On account of the highly skewed data, a finite mixture model approach was also implemented. This revealed, for both cross-sectional and longitudinal analyses, four distinct subjects having an increased nocturnal activity compared to the rest of the group. In one previous study, Satlin et al. (1991) empirically described Alzheimer's disease patients as 'pacers' (those who were restless and moved constantly) and 'non-pacers'. The 'pacers' had increased diurnal activity compared to 'non-pacers', however there was no significant

difference of cognitive function between 'pacers' and 'non-pacers' in that study. They could not determine the characteristics which could distinguish 'pacers' from 'non-pacers'. This motivated our finite mixture model approach to the data in order to explore the latent characteristics of the two groups based on high and low nocturnal activity. The 'high nocturnal activity' group of four subjects differed from the 'low nocturnal activity' group of 17 subjects in having lower depressive mood and a lower immobility, however no significant difference was found between these two groups in terms of cognitive function in our study. The latter finding corroborates with that of Satlin et al. (1991) study. A possible explanation of not finding a difference for cognitive function in our study is that subjects who belonged to the 'high nocturnal activity' group were all demented (with low MMSE), however, subjects from the 'low nocturnal activity' group comprised of both demented (with low MMSE) and non-demented (with high MMSE) subjects.

The strengths of this study were: we measured the outcome variable with a validated, robust, and precise system. The exposure group and covariates were well established instruments in measuring different characteristics of the elderly. To our knowledge, no previous study has been conducted to date to determine the predictors of nocturnal wandering in inpatients. The consistent results in longitudinal and cross-sectional analyses provide corroborating evidence. Finally, we had very few missing data for the outcome variable.

The limits of this study were: low statistical power, lack of generalisability as we focussed only on inpatients and not on subjects who lived at home, and missing covariate information for six subjects for the GDS variable, which was a significant predictor of median nocturnal activity. However, the subjects with missing GDS data did not differ in other measured population characteristics from those with GDS data. In addition, we could not observe patients over a longer period, *e.g.*, over a few weeks, which might have revealed a time trend. Given that we followed patients only after at least eight nights of hospital stay, most of the patients were expected to stabilise to their proper nocturnal rhythm of activity by that time. Thus by excluding the 'acute' initial phase of hospital admission in this study, a possible downward trend in nocturnal activity might have been missed.

From this pilot study, we gained information on the predictors of nocturnal activity of elderly inpatients. In future, it would be interesting to gather more information on more subjects to test the robustness of our results.

5 REFERENCES

Acebo C, Sadeh A, Seifer R et al. Estimating sleep parameters with activity monitoring in children and adolescents: how many nights are necessary for reliable measures? *Sleep* 1999;**22**:95-103.

Agresti A. Categorical Data Analysis, Wiley & sons, 2nd edition, 2002.

Aharon-Peretz J, Masiah A, Pillar T et al. Sleep-wake cycles in multi-infarct dementia and dementia of the Alzheimer type. *Neurology* 1991;**41**:1616-19.

Banerjee S, Steenkeste F, Couturier P et al. Telesurveillance of elderly patients by use of passive infra-red sensors in a 'smart' room. *J Telemed Telecare* 2003;**9**:23-9.

Banerjee S. Téléassistance des malades âgés: le développement de l'actimétrie déambulatoire nocturne, PhD thesis, Université Paul Sabatier, Toulouse, France, 2005.

Bland JM, Altman DG. Measurement error proportional to the mean. *BMJ* 1996;**313**:106.

Brown AC, Smolensky MH, Gilber EDA et al. Actigraphy: a means of assessing circadian patterns in human activity. *Chronobiol Int* 1990;**7**:125-30.

Couturier P, Tyrrell J, Banerjee S et al. Téléassistance en milieu gériatrique: besoins et réponses technologiques. In: Enjalbert M, Rencontres en Rééducation: Aides techniques et matériel d'assistance en gériatrie, Masson edition, 2002, pp63-73.

CRAN – R, version 2.7.1. http://www.r-project.org

Der G, Everitt BS. Statistical Analysis of Medical Data using SAS, Chapman & Hall, 2006.

Folstein MF, Folstein SE, McHugh PR. "Mini-Mental State": a practical method for grading the cognitive state of patients for clinicians. *J Psychiatr Res* 1975;**12**:189-98.

Gelman A. http://www.stat.columbia.edu/~gelman/arm/other.packages/software.pdf

Hilten JJ van, Middelkoop HAM, Braat EAM et al. Nocturnal activity and immobility across aging (50-98 years) in healthy persons. *J Am Geriatr Soc* 1993;**41**:837-41.

Katz S, Downs TD, Cash HR, Grotz RC. Progress in the development of the index of ADL. *Gerontologist* 1970;**10**:20-30.

Lavie P, Aharon-Peretz J, Klein F et al. Sleep quality in geriatric depressed patients: comparison with elderly demented patients and normal controls and effects of Moclobemide. *Dementia* 1992;**3**:360-6.

Lawton MP, Brody EM. Assessment of Older people: self-maintaining and instrumental activities of daily living. *Gerontologist* 1969;**9**:179-86.

Leisch F. FlexMix: a general framework for finite mixture models and latent class regression in R. *J Statistical Software* 2004;**11**.

Lemke MR, Puhl P, Broderick A. Motor activity and perception of sleep in depressed patients. *J Pyschiatr Res* 1999;**33**:215-24.

Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986;**73**:13-22.

Miskelly FG. Assistive technology in elderly health care. *Age Ageing* 2001;**30**:455-8.

Molenberghs G, Verbeke G. Models for discrete longitudinal data, Springer series in statistics, 2005.

Ohta S, Nakamoto H, Shinagawa Y et al. A health monitoring system for elderly people living alone. *J Telemed Telecare* 2002;**8**:151-6.

Paavilainen P, Korhonen I, Lötjönen J et al. Circadian activity rhythm in demented and non-demented nursing-home residents. *J Sleep Res* 2005;**14**:61-8.

Sadeh A, Hauri PJ, Kripke DF et al. The role of actigraphy in the evaluation of sleep disorders. *Sleep* 1995;**18**:288-302.

Sadeh A, Acebo C. The role of actigraphy in sleep medicine. *Sleep Medicine Reviews* 2002;**6**:113-24.

SAS Institute Inc. SAS user's guide, version 9.1.3, Cary, NC, USA, 2002-2005.

Satlin A, Teicher MH, Leiberman HR et al. Circadian locomotor activity rhythms in Alzheimer's disease. *Neuropsychopharmacol* 1991;**5**:115-26.

Someren EJW van. Actigraphic monitoring of movement and rest-activity rhythms in aging, Alzheimer's Disease, and Parkinson's Disease. *IEEE Trans Rehab Eng* 1997;**5**:394-8.

Virone G, Noury N, Demongeot J. A system for automatic measurement of circadian activity deviations in telemedicine. *IEEE Trans Biomed Engg* 2002;**49**:1463-9.

Waterlow J. Pressure sores: a risk assessment card. *Nurs Times* 1985;**81**:49-55.

Yesavage JA, Brink TL, Rose TL et al. Development and validation of a geriatric depression rating scale: a preliminary report. *J Psych Res* 1983;**17**:27.

Yesavage JA, Friedman L, Krämer HC et al. A follow-up study of actimetric measures in home-residing Alzheimer's Disease patients. *J Geriatr Pyschiatry Neurol* 1998;**11**:7-10.

6 APPENDIX

6.1 Parts of the dataset

Cross-sectional analysis

Exposure group & covariates | Outcome

	MMS_bin	GDS	Waterlow_bin	Median_activity_6hours	patid
1	1	13	1	9.33	2
2	1	15	1	7.23	3
3	1	14	1	10.95	4
4	0	4	1	3.23	7
5	1	14	0	45.48	8
6	0	20	1	0.00	9
7	1	15	1	1.08	10
8	0	14	1	1.65	11
9	1	23	1	5.60	12
10	0	19	1	2.65	13
11	1	12	0	19.62	14
12	1	6	0	67.97	15
13	0	7	1	3.95	16
14	1	7	1	9.38	18
15	1	17	1	19.70	19
16	1	14	1	3.18	20
17	1	25	1	0.78	21
18	1	15	0	10.57	23
19	1	11	0	71.92	24
20	1	22	0	2.10	25
21	1	5	1	48.30	27

Longitudinal analysis

Exposure group | Outcome | Covariates

	Patid	Group	N1	N2	N3	N4	N5	N6	N7	N8	ADL_bin	GDS	Waterlow_bin
1	2	1	38	837	3153	981	560	263	136	567	1	13	1
2	3	1	434	0	645	508	526	296	137	411	0	15	1
3	4	1	730	281	1420	598	866	148	657	124	0	14	1
4	7	0	433	194	174	250	519	170	190	308	0	4	1
5	8	1	3820	4821	1475	1488	990	2729	6538	3493	0	14	0
6	9	0	0	1677	0	2354	1157	0	0	0	1	20	1
7	10	1	144	0	1013	65	799	0	0	72	1	15	1
8	11	0	0	51	170	175	99	108	0	197	0	14	1
9	12	1	0	4164	336	77	66	1173	1587	49	0	23	1
10	13	0	159	662	302	105	128	378	131	427	0	19	1
11	14	1	316	1177	1189	1222	383	544	1594	1127	0	12	0
12	15	1	4078	2005	3038	6640	4571	4066	4066	4066	0	6	0
13	16	0	296	252	31	237	562	145	136	953	0	7	1
14	18	1	459	455	867	563	1024	373	573	365	0	7	1
15	19	1	0	904	1640	1184	3575	1825	0	0	0	17	1
16	20	1	193	191	143	270	2891	0	0	905	0	14	1
17	21	1	0	58	28	1296	4934	47	0	640	0	25	1
18	23	1	698	423	823	681	566	634	522	211	0	15	0
19	24	1	3257	6636	6118	4894	1008	420	4315	5195	1	11	0
20	25	1	169	126	114	119	244	0	459	176	0	22	0
21	27	1	4475	2900	2600	5908	4320	2314	2246	3538	1	5	1

6.2 Exploratory longitudinal data analysis (N = 27)

Figure 6-1 on the *left* shows the raw individual patient profiles. The X-axis represents the 8 nights; the Y-axis represents the measured nocturnal activity between 00:00 and 06:00 (seconds). Figure on the *right* shows the individual patient profiles after tracking, i.e., standardisation of each individual profile by centering and scaling.

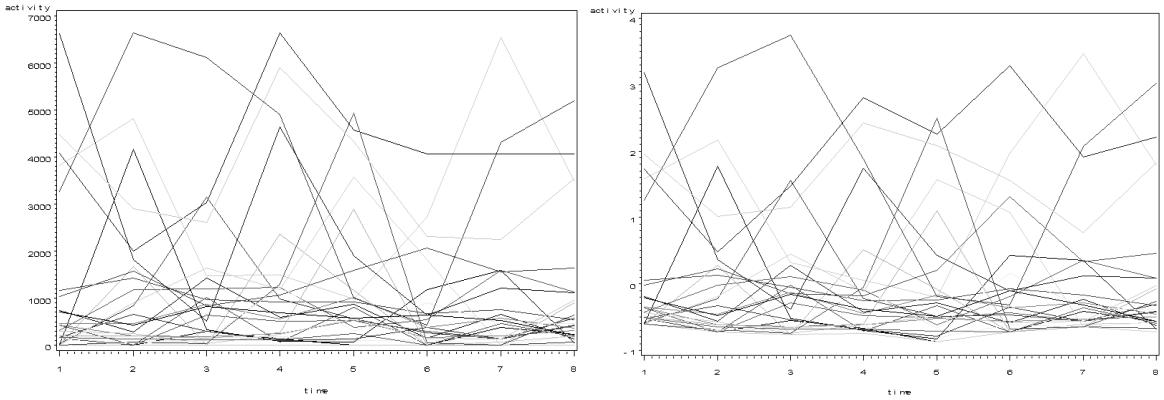


Figure 6-2 shows the observed mean response ± 95% CI, i.e., measured nocturnal activity between 00:00 and 06:00 (seconds); for each group ('cognitively impaired': category '1' & 'normal' cognitive function: category '0') per night.

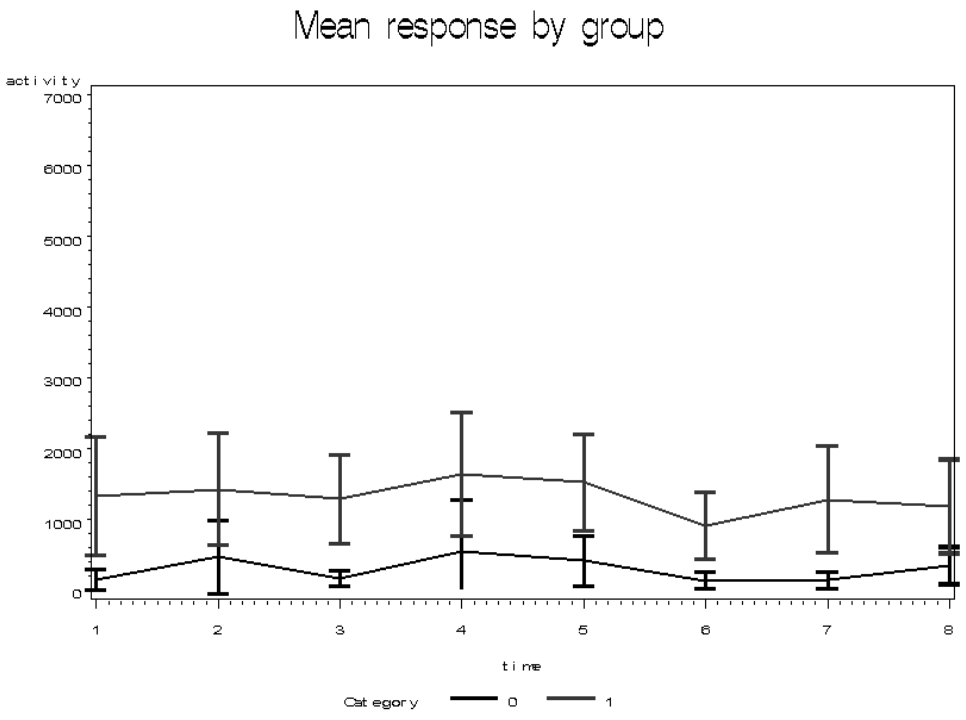


Figure 6-3 shows the boxplot of nocturnal activity (from 00:00 to 06:00) in seconds by time. Each night is again divided into two groups ('cognitively impaired' & 'normal' cognitive function).

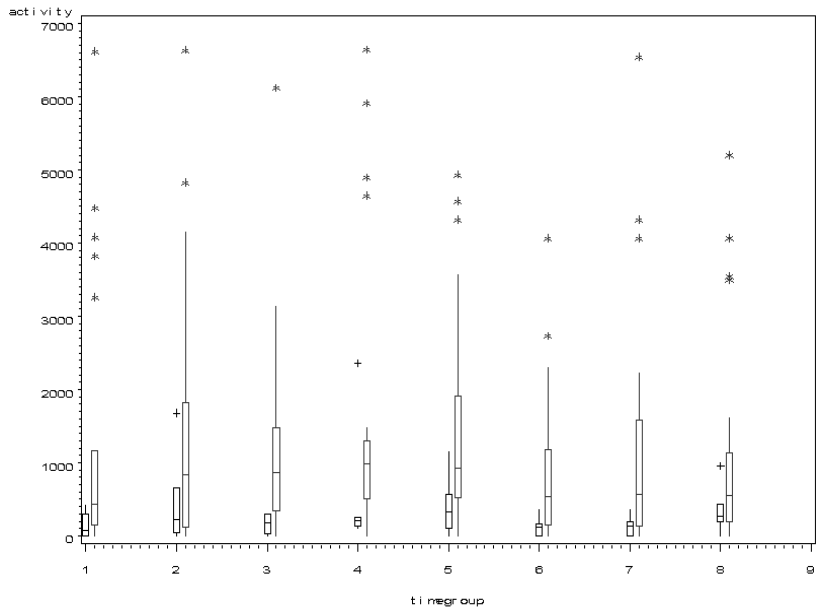
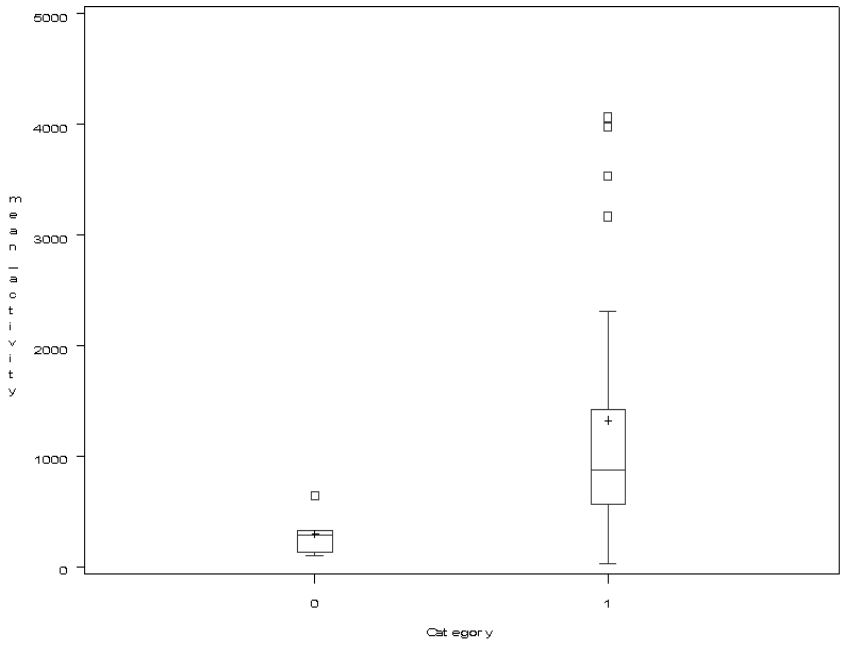


Figure 6-4 shows the boxplot of nocturnal activity (from 00:00 to 06:00) in seconds, by group ('cognitively impaired': category '1' & 'normal' cognitive function: category '0').



6.3 Longitudinal data analysis (N = 21)

Figure 6-5 shows predicted nocturnal activity over 8 nights using marginal (GEE) model with the assumptions of a negative binomial distribution for the outcome variable and exchangeable (compound symmetry) within observations. 'Cognitively impaired' in dotted & 'normal' cognitive function in continuous.

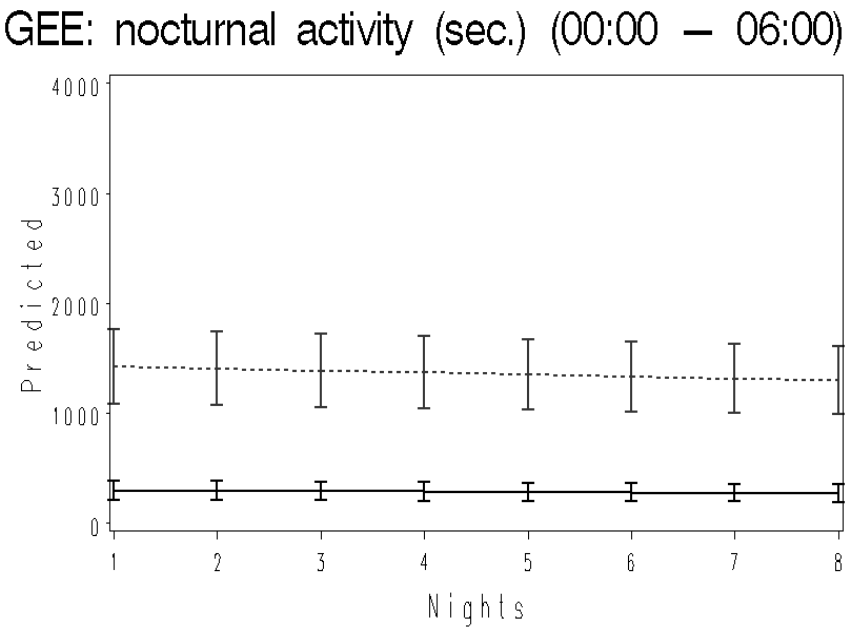


Figure 6-6 shows predicted nocturnal activity over 8 nights using conditional model with adaptive Gaussian quadrature (size = 20) with the assumption of a Poisson distribution for the outcome variable. 'Cognitively impaired' in dotted & 'normal' cognitive function in continuous.

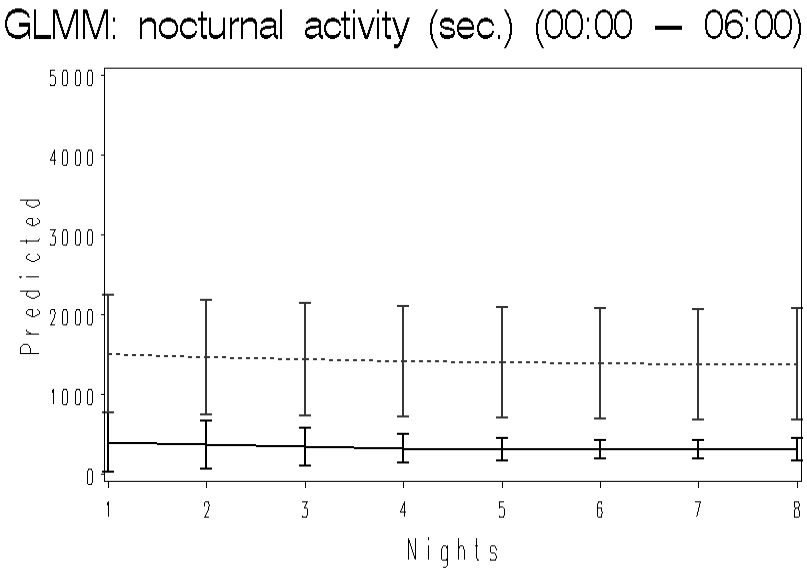
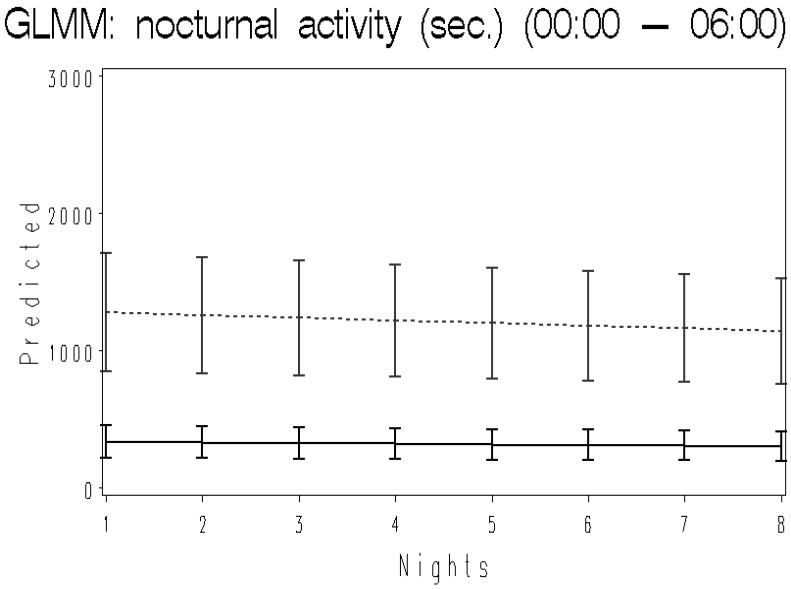


Figure 6-7 shows predicted nocturnal activity over 8 nights using conditional model with adaptive Gaussian quadrature (size = 20) with the assumption of a negative binomial distribution for the outcome variable. 'Cognitively impaired' in dotted & 'normal' cognitive function in continuous.



6.4 SAS CODES

LINEAR MODEL

```
data Actimetry; set Mit.Actimetry_GLM;
 Logactivity=log(median_activity_6hours+0.1);
run;

data Actimetry1; set Actimetry;
 if gds=. then delete;
run;

/*LOG-TRANSFORMED LINEAR MODEL*/
proc glm data=Actimetry1;
 class mms_bin;
 model Logactivity=mms_bin gds;
 output out=out1 p=PredVal r=Residual;
run; quit;
/*'mms_bin x gds' interaction is NOT significant (not shown)*/

/*MODEL DIAGNOSTICS*/
proc reg data=Actimetry1;
 model Logactivity=mms_bin gds;
 plot r.*(mms_bin gds);
 plot r.*p.;
 plot npp.*r.;
 plot cookd.*obs.;
run; quit;
```

GLM

```
/*NEGATIVE BINOMIAL MODEL*/
proc genmod data=Mit.Actimetry_GLM;
 class mms_bin waterlow_bin;
 model median_activity_6hours=mms_bin waterlow_bin gds / dist=nb;
 estimate 'MMSE' mms_bin -1 1 / exp;
 estimate 'Waterlow' waterlow_bin -1 1 / exp;
 output out=nbout1 reschi=rs predicted=pred_val_nb;
run;
/*Deviance: value/DF=1.30~1*/
/*LRT=[-2*(LL(Poisson model)-LL(negative binomial model))]=44.62, P<0.01;
estimated dispersion parameter=0.39, 95% CI=0.08-0.71, which is > 0*/
/*Within covariates interactions are NOT tested due to small sample size*/

/*MODEL DIAGNOSTICS*/
title 'Negative binomial model: normal probability plot';
proc univariate data=nbout1 normaltest; /*P=0.13; Kolmogorov-Smirnov*/
 var rs;
 probplot rs;
run;
title;
```

GEE

```
/*MARGINAL NEGATIVE BINOMIAL MODEL*/
/*EXCHANGEABLE WORKING CORRELATION*/
proc genmod data=Mit.ActimetryL;
 class patid category adl_bin waterlow_bin;
 model activity=time category adl_bin waterlow_bin / d=nb wald type3;
 repeated subject=patid / modelse corrw type=exch;
 output out=nbout_exch1 predicted=pred_val_nb_exch1;
 estimate 'MMSE' category -1 1 / exp;
 estimate 'Waterlow' waterlow_bin -1 1 / exp;
 estimate 'ADL' adl_bin -1 1 / exp;
```

```
run;
/*model-based SEs are somewhat different from empirical SEs*/
/*'time x covariates' interactions are NOT significant (not shown)*/

/*PLOTTING NEGATIVE BINOMIAL MODEL & EXCHANGEABLE WORKING CORRELATION*/
/*BY CATEGORY*/
proc gplot data=nbout_exch1;
 plot pred_val_nb_exch1*time=category / haxis=axis1 vaxis=axis2 nolegend;
 symbol1 c=black v=none i=stdm2jt l=1 w=2 mode=include;
 symbol2 c=red v=none i=stdm2jt l=2 w=2 mode=include;
 axis1 label=(h=2 'Nights') value=(h=1.5) minor=none;
 axis2 label=(h=2 A=90 'Predicted') value=(h=1.5) minor=none;
 title h=3 "GEE: nocturnal activity (sec.) (00:00 - 06:00)";
run; quit;
```

GLMM: PQL (linearisation)

```
/*COMPOUND SYMMETRY COVARIANCE*/
proc glimmix data=Mit.ActimetryL empirical;
 class patid category adl_bin;
 model activity=time category adl_bin gds / s ddfm=satterth d=p;
 random int time / subject=patid type=cs s;
 output out=glimmixout pred(  blup ilink)=PredVal
                       pred(noblup ilink)=PredVal_PA; /*PA: pop.-avg.*/
run;
/*'time x covariates' interactions are NOT significant (not shown)*/
```

GLMM: adaptive Gaussian quadrature

```
/*RANDOM INTERCEPT: NEGATIVE BINOMIAL MODEL*/
proc nlmixed data=ActimetryL qpoints=20 maxiter=50;
 parms int=5.811 T=-0.016 C=1.219 d11=0.481 k=2.017;
 eta=int+b1+T*time+C*category; /*linear predictor*/
 lambda=exp(eta);
 p=lambda/(lambda+1/k);
 ll=lgamma(activity+1/k)-lgamma(activity+1)-
    lgamma(1/k)+activity*log(p)+(1/k)*log(1-p);
 model activity~general(ll);
 random b1~normal(0,d11) subject=patid out=EB; /*Empirical Bayes estimate*/
 predict lambda out=nlmixedout_nb;
run;
/*'time x covariates' interactions are NOT significant (not shown)*/

/*MODEL DIAGNOSTICS*/
/*QQ PLOT: NEGATIVE BINOMIAL (using nqplot macro)*/
%nqplot(data=EB, var=Estimate);
/*HISTOGRAM: NEGATIVE BINOMIAL*/
proc univariate data=EB;
 var Estimate;
 histogram Estimate / normal; /*P=0.15; Kolmogorov-Smirnov*/
run;

/*PLOTTING RANDOM INTERCEPT: NEGATIVE BINOMIAL MODEL*/
/*BY CATEGORY*/
proc gplot data=nlmixedout_nb;
 plot pred*time=category / haxis=axis1 vaxis=axis2 nolegend;
 symbol1 c=black v=none i=stdm2jt l=1 w=2 mode=include;
 symbol2 c=red v=none i=stdm2jt l=2 w=2 mode=include;
 axis1 label=(h=2 'Nights') value=(h=1.5) minor=none;
 axis2 label=(h=2 A=90 'Predicted') value=(h=1.5) minor=none;
 title h=3 "GLMM: nocturnal activity (sec.) (00:00 - 06:00)";
run; quit;
```

6.5 R CODES

Bootstrap and Jackknife

LINEAR MODEL

```
# Bootstrap linear: by iteration
linear<-as.matrix(nb_glm)
n<-length(nb_glm$patid) # sample size of bootstrap
N<-1000 # no. of bootstrap
beta2<-rep(0, len=N)
beta3<-rep(0, len=N)
for (j in 1:N){
      mat.b<-linear[sample(n,replace=T),]
      data.b<-as.data.frame(mat.b)
lm.b<-lm(log(Median_activity_6hours+0.1)~MMS_bin+GDS, data=data.b)
beta2[j]<-coef(lm.b)[2]
beta3[j]<-coef(lm.b)[3]
      }
summary(beta2)
summary(beta3)
par(mfcol=c(2,1))
boxplot(beta2)
boxplot(beta3)
par(mfcol=c(2,1))
hist(beta2)
hist(beta3)

# Boostrap linear: using boot function
linear<-as.matrix(nb_glm)
n<-length(nb_glm$patid) # sample size of bootstrap
bsf<-function(linear, n){
      mat.b<-linear[sample(n, replace=T),]
      data.b<-as.data.frame(mat.b)
lm.b<-lm(log(Median_activity_6hours+0.1)~MMS_bin+GDS, data=data.b)
betas<-c(beta2<-coef(lm.b)[2], beta3<-coef(lm.b)[3])
return(betas)
      }
bootpara<-boot(data=nb_glm, bsf, R=50)
plot(bootpara, index=1)
plot(bootpara, index=2)
boot.ci(bootpara, index=1, type="bca", conf=0.95)
boot.ci(bootpara, index=2, type="bca", conf=0.95)

# Jackknife-after-bootstrap: linear
par(mfcol=c(2,1))
jack.after.boot(bootpara, useJ=F, stinf=F, index=1, main="(a) MMSE")
jack.after.boot(bootpara, useJ=F, stinf=F, index=2, main="(b) GDS")
```

GLM

```
# Bootstrap negative binomial: by iteration
nb<-as.matrix(nb_glm)
n<-length(nb_glm$patid) # sample size of bootstrap
N<-1000 # no. of bootstrap
beta2<-rep(0, len=N)
beta3<-rep(0, len=N)
beta4<-rep(0, len=N)
for (j in 1:N){
      mat.b<-nb[sample(n,replace=T),]
      data.b<-as.data.frame(mat.b)
nb.b<-glm(Median_activity_6hours~MMS_bin+GDS+Waterlow_bin,
family=negative.binomial(2), data=data.b)
```

```
beta2[j]<-coef(nb.b)[2]
beta3[j]<-coef(nb.b)[3]
beta4[j]<-coef(nb.b)[4]
      }
summary(beta2)
summary(beta3)
summary(beta4)
par(mfcol=c(3,1))
boxplot(beta2)
boxplot(beta3)
boxplot(beta4)
par(mfcol=c(3,1))
hist(beta2)
hist(beta3)
hist(beta4)

# Boostrap negative binomial: using boot function
nb<-as.matrix(nb_glm)
n<-length(nb_glm$patid) # sample size of bootstrap
bsf<-function(nb, n){
      mat.b<-nb[sample(n, replace=T),]
      data.b<-as.data.frame(mat.b)
nb.b<-glm(Median_activity_6hours~MMS_bin+GDS+Waterlow_bin,
family=negative.binomial(2), data=data.b)
betas<-c(beta2<-coef(nb.b)[2], beta3<-coef(nb.b)[3], beta4<-coef(nb.b)[4])
return(betas)
      }
bootpara<-boot(data=nb_glm, bsf, R=50)
plot(bootpara, index=1)
plot(bootpara, index=2)
plot(bootpara, index=3)
boot.ci(bootpara, index=1, type="bca", conf=0.95)
boot.ci(bootpara, index=2, type="bca", conf=0.95)
boot.ci(bootpara, index=3, type="bca", conf=0.95)

# Jackknife-after-bootstrap: negative binomial
par(mfcol=c(3,1))
jack.after.boot(bootpara, useJ=F, stinf=F, index=1, main="(a) MMSE")
jack.after.boot(bootpara, useJ=F, stinf=F, index=2, main="(b) GDS")
jack.after.boot(bootpara, useJ=F, stinf=F, index=2, main="(c) Waterlow")
```

GEE

```
# Bootstrap Poisson GEE: by iteration; neg. bin. GEE couldn't be done
n<-21 # sample size of bootstrap
N<-100 # small no. of bootstraps
beta2<-rep(0, len=N)
beta3<-rep(0, len=N)
beta4<-rep(0, len=N)
beta5<-rep(0, len=N)
beta6<-rep(0, len=N)
for (j in 1:N){
      data.b<-aw[sample(n,replace=T),]
      data.b$subject.b<-c(1:n)
al.b<-reshape(data.b, idvar="subject.b", v.names="activity",
+ timevar="time", varying=list(c("N1", "N2", "N3", "N4", "N5", "N6", "N7",
+ "N8")), direction="long")
al.gee.b<-gee(activity~time+Group+ADL_bin+Waterlow_bin+GDS, id=subject.b,
+ data=al.b, family=poisson, corstr="exchangeable")
beta2[j]<-coef(al.gee.b)[2]
beta3[j]<-coef(al.gee.b)[3]
beta4[j]<-coef(al.gee.b)[4]
```

```
beta5[j]<-coef(al.gee.b)[5]
beta6[j]<-coef(al.gee.b)[6]
      }
summary(beta2)
summary(beta3)
summary(beta4)
summary(beta5)
summary(beta6)
par(mfcol=c(2,3))
boxplot(beta2)
boxplot(beta3)
boxplot(beta4)
boxplot(beta5)
boxplot(beta6)
par(mfcol=c(2,3))
hist(beta2)
hist(beta3)
hist(beta4)
hist(beta5)
hist(beta6)
```

GLMM: PQL (linearisation)

```
# Bootstrap Poisson GLMM: by iteration; neg. bin. GLMM couldn't be done
# Random slope model
n<-21 # sample size of bootstrap
N<-20 # small no. of bootstraps
beta2<-rep(0, len=N)
beta3<-rep(0, len=N)
beta4<-rep(0, len=N)
for (j in 1:N){
      data.b<-aw[sample(n,replace=T),]
      data.b$ID<-c(1:n)
al.b<-reshape(data.b, idvar="ID", v.names="activity", timevar="time",
+ varying=list(c("N1", "N2", "N3", "N4", "N5", "N6", "N7", "N8")),
+ direction="long")
al.glmmPQL.b<-glmmPQL(activity~time+Group+GDS, random=~time|ID, data=al.b,
+ correlation=NULL, family=poisson, niter=10)
beta2[j]<-mean(coef(al.glmmPQL.b)[2])
beta3[j]<-mean(coef(al.glmmPQL.b)[3])
beta4[j]<-mean(coef(al.glmmPQL.b)[4])
      }
summary(beta2)
summary(beta3)
summary(beta4)
par(mfcol=c(3,1))
boxplot(beta2)
boxplot(beta3)
boxplot(beta4)
par(mfcol=c(3,1))
hist(beta2)
hist(beta3)
hist(beta4)
```

Finite Mixture Models

```
### cross-sectional analysis
fl<-flexmix(Median_activity_6hours~1, data=fmm, k=3)
### poisson model gave a single component model or unstable results
parameters(fl, component=1)
parameters(fl, component=2)
parameters(fl, component=3)
summary(fl) ### AIC=154.00; BIC=162.35; LL=-69.00 (df=8)
x<-seq(from=0, to=80)
d1<-dnorm(x, mean=4.77, sd=3.74) # n=15
d2<-dnorm(x, mean=19.66, sd=0.04) # n=2
d3<-dnorm(x, mean=58.41, sd=11.95) # n=4
f<-(0.71*d1)+(0.10*d2)+(0.19*d3)
hist(fmm$Median_activity_6hours, xlab="Median nocturnal activity (min.)",
+ probability=T, border="gray", main="Gaussian kernel", ylim=c(0,0.1))
lines(x, f, lty=2, lwd=2)
lines(density(fmm$Median_activity_6hours), lty=3, lwd=3, col=4)

### longitudinal analysis
fmm.w$subject<-factor(rownames(fmm.w))
fmm.l<-reshape(fmm.w, idvar="subject", v.names="activity", timevar="time",
+ varying=list(c("N1", "N2", "N3", "N4", "N5", "N6", "N7", "N8")),
+ direction="long")
fl.l<-flexmix(activity~1|subject, data=fmm.l, k=3)
parameters(fl.l, component=1)
parameters(fl.l, component=2)
parameters(fl.l, component=3)
summary(fl.l) ### AIC=2703.6; BIC=2711.0; LL=-1348.8 (df=8)
X<-seq(from=0, to=80)
D1<-dnorm(X, mean=5.73, sd=4.68) # n=10
D2<-dnorm(X, mean=13.86, sd=18.02) # n=7
D3<-dnorm(X, mean=61.47, sd=27.92) # n=4
F<-(0.45*D1)+(0.36*D2)+(0.19*D3)
mean<-c(13.61, 6.16, 10.05, 4.66, 52.82, 10.81, 4.36, 1.67, 15.52, 4.78,
+ 15.73, 67.77, 5.44, 9.75, 19.02, 9.57, 14.59, 9.5, 66.34, 2.93, 58.96)
fmm.w$mean.activity<-mean
hist(fmm.w$mean.activity, xlab="Mean nocturnal activity (min.)",
+ probability=T, border="gray", main="Gaussian kernel", ylim=c(0,0.1))
lines(X, F, lty=2, lwd=2)
lines(density(fmm.w$mean.activity), lty=3, lwd=3, col=4)
```

6.6 APPROVAL

Dear Soutrik,

Thank you for sending me your manuscript. You can plan and defend your thesis.

Best regards.

Pr Pascal Couturier
Clinique Universitaire de Médecine Gériatrique
Pôle pluridiciplinaire de Médecine
Hôpital Albert Michallon
38043 Grenoble Cedex 09
Tel. Secrétariat de la clinique 04 76 76 89 07
Tel. Secrétariat universitaire 04 76 76 56 06

De : Banerjee Soutrik (Business and Decision) [mailto:Soutrik.BANERJEE@UCB-Group.com]
Envoyé : lundi 15 septembre 2008 11:05
À : Couturier, Pascal
Objet : Summer thesis: master biostatistics UHasselt APPROVAL
Importance : Haute

Dear Pascal,

I received the permission from my internal superviser at Univesiteit Hasselt, Pr. Roel Braekers (PS: I forwarded you his email last week). I wish to defend my unfinished summer thesis on 30th October 2008, for which I need to upload it by 17th October 2008 at the censtat server. If I could defend it this year, it would be good for my performance evaluation in 2008 at my company and also my parents are in Belgium, so this is another motivation to speed up the process.

I would like to receive an approval from you by email, which is considered equivalent of a signature approval on the manuscript. This will expedite the process of getting a date to defend the thesis.

On getting the approval from you, I will forward this to our secretary Ms. Martine Machiels.

I have also attached the thesis draft version 1.1. I will add the "discussion" section and also join the abstract.

Thanks and best regards,
Soutrik.

Soutrik BANERJEE (Business & Decision),
Health Outcomes Researcher,
R3 building, office 16,
24, Chemin du Foriest,
UCB Pharma, s.a.,
B-1420 Braine-l'Alleud,
Belgium.
Tel: +32 2 386 3127 (office); +32 473 642 406 (mobile).

---

* Essentially, all models are wrong, but some are useful - George Box. *