Corresponding author: Henry Brodaty

Email: h.brodaty@unsw.edu.au

Title: The longitudinal course of post-stroke apathy over five years

Authors: Henry Brodaty, M.D., D.Sc., Adrienne Withall, Ph.D., Perminder S. Sachdev, M.D., Ph.D.

Supplementary Material

Supplemental Methods

Supplementary Tables: 3

Supplementary References

Supplemental Methods

Statistical Analyses

Linear mixed models (PROC MIXED) were used to describe the change of apathy across four assessments and to test which baseline variables were predictors of apathy. In order to overcome the inherent problem in apathy research of maintaining subjects in longitudinal follow-up, missing data were imputed using multiple imputation (PROC MI).

The imputation was based on demographic and clinical variables (at index: age, gender, education, vascular risk factors, stroke severity, interval cerebrovascular events between baseline and index), and variables which were associated with the AES and also predicted missingness (time-varying: dementia, ADL/IADL, GDS, AES).

Linear mixed models (PROC MIXED) as well as multiple imputation (PROC MI) assume that data are missing at random (MAR; i.e. missingness can depend on observed outcomes, but not on missing outcomes). For this study, missingness was tested as a binary outcome variable using logistic regression. Overall the analyses suggested that more disabled patients were more likely to have missing AES data at 5 years (see 'sample attrition and missing data'). Variables, such as ADL/IADL, dementia and GDS which predicted missing AES data (and were also associated with AES levels in the not imputed data) were used in the multiple imputation rendering the data – at least theoretically – MAR.

Multiple imputation (PROC MI) assumes multivariate normality between variables when using the Markov Chain Monte Carlo (MCMC) method, but not when using propensity scores for data imputation. In particular the ADL/IADL scores were left skewed and the GDS scores were right skewed. Transformations of the Box Cox family including log and square root transformations did not normalise the distributions. The data were left untransformed and an imputation method was used based on a mixed approach using the MCMC method to impute enough values to make the missing data pattern monotone with a nonparametric method based on propensity scores (1). Imputed values for categorical variables (i.e. dementia, interval cerebrovascular event, gender) were rounded.

As multiple imputations for missing data can be unreliable under certain conditions (2-3) in particular when only few (e.g. 5-10) imputations are used (2), we tested univariate and multivariate results with up to 1000 imputations. Most univariate results were stable with 20 imputations, multivariate results stabilized only with 1000 imputations. The large number of imputations in the multivariate model was due to the variable GDS, which had

the tendency to fluctuate around the significance level (p=0.05). We imposed an ad hoc criterion that a result was stable when the variability due to the imputation process in 50 replications did not affect the conclusion. Results in table 3 and 4 are based on 1000 imputations.

While the main results are given for imputed data, key results such as AES slopes are also provided for actual (not imputed) data for comparison. As additional information about the quality of the data imputation, mean AES scores for the group with complete AES data and the group, which had at least one missing AES score are provided in Table 1. In the first group AES data were not imputed, but in the second group missing AES data were imputed. The higher AES means for the group with imputed data can be considered as bias correction in the expected direction, because the missing data analysis showed that more disabled patients were more likely to have missing AES data.

Supplemental References

1. SAS online manual: support.sas.com/rnd/app/papers/miv802.pdf: The MI Procedure – last accessed 16th June, 2010.
2. Spratt M, Carpenter J, Sterne JAC, Carlin JB, Heron J, Henderson J, et al. Strategies for Multiple Imputation in Longitudinal Studies. Am J Epidemiol. 2010;172:478–87.
3. Allison PD. Multiple imputation for missing data—a cautionary tale. Sociol Methods Res. 2000;28:301–9.

Table S1. Baseline Demographic Characteristics of Controls and Stroke Patients.

	Control (n =101)	Patients (n = 152)	Statistic	p
Age, years (Mean, SD)	71.05 (6.09)	72.13 (8.88)	*t*(250) = 1.15	0.25
Education, years (Mean, SD)	11.78 (3.27)	10.15 (2.75)	*t*(189.54) = 4.12	<.001
Male (N, %)	49 (48.5)	88 (57.9)	OR 1.46 (95% CI 0.88-2.42)	

Table S2. Baseline Demographic Characteristics of Included and Excluded Patients.

	Included (n = 152)	Excluded (n = 50)	Statistic	p
Age, years (Mean, SD)	72.1 (8.88)	71.9 (9.19)	*t*(81.23) = 0.16	0.88
Education, years (Mean, SD)	10.15 (2.75)	9.23 (2.97)	*t*(195) = 1.99	0.048
Male (N, %)	88 (57.9)	28 (56.0)	OR 0.93 (95% CI 0.49-1.76)	

Table S3. Demographic and Clinical Characteristics of Apathetic versus not Apathetic Patients at Index Assessment and at 5 Years (Not Imputed Data)

	Apathy at Index (N=36)	No Apathy at Index (N=99)

	Mean (s.d.)	Mean (s.d.)	t	d.f.	*P*
Age (years)	75.17 (7.01)	71.13 (9.21)	-2.71	81.25	0.01
Education (years)	9.80 (2.00)	10.37 (2.99)	1.27	89.65	0.21
ADL+ IADL*	9.72 (3.98)	12.39 (2.92)	3.69	49.40	0.001
Vasc. Risk Factors	2.74 (1.58)	2.61 (1.11)	-0.39	34.99	0.70
	N (%)	N (%)	OR	95%CI	
Gender (male)	22/36 (61.1)	60/99 (60.6)	0.98	0.45-2.14	
Dementia	16/35 (45.7)	8/99 (8.1)	9.58	3.59-25.58	
	Estimate (s.e.)	Estimate (s.e.)	χ^2†	d.f.	*P*
NH Placement‡	354.65 (30.87)	440.00 (17.76)	3.36	1	0.07
Mortality§	321.27 (24.94)	395.28 (16.99)	5.22	1	0.02
	Apathy at 5 Years (N=27)	No Apathy at 5 Years (N=43)			
	Mean (s.d.)	Mean (s.d.)	t	d.f.	*P*
Age (years)	80.93 (8.74)	79.59 (8.89)	-0.62	68.0	0.54
Education	9.65 (2.13)	10.86 (2.87)	1.86	67.0	0.07
ADL/IADL at 5yrs*¶	8.67 (5.78)	12.95 (2.15)	3.69	30.69	0.001
	N (%)	N (%)	OR	95%CI	
Gender (male)	14/27 (51.9)	24/43 (55.8)	1.17	0.45-3.08	
Dementia	11/25 (44.0)	2/42 (4.8)	15.71	3.10-79.80	

* Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL)

† Log Rank (Mantel-Cox) Chi-Square

‡ Mean weeks since stroke. In the group with apathy 11/30 patients were placed in a nursing home during time of observation. This compares to 19/96 in the group without apathy. When the analysis was corrected for dementia at index the comparison was no longer significant (OR=1.34, 95%CI=0.56-3.19).

§ Mean weeks since stroke. In the group with apathy 22/35 patients died during time of observation. This compares to 37/99 in the group without apathy. When the analysis was corrected for dementia at index the comparison was no longer significant (OR=1.14, 95%CI=0.61-2.15)

¶ Sample size N=27/27 in the group with apathy and N=42/43 in the group without apathy.