

Statsconsultancy Ltd

Freelance Statistical Consultancy

Address: 40 Longwood Lane, Amersham, Bucks. HP7 9EN

Email: paul@statsconsultancy.co.uk

Phone: 07905 530446

Web: www.statsconsultancy.co.uk

Statistical Report

Author: Paul Bassett

Client: Chloë Hood

Date: 5th April 2024

Report Number: 1

Description: Analysis examining the repeatability of a tool for auditing memory services

Introduction

As part of the National Audit of Dementia, an audit of memory services was performed. As part of the audit process, each patient was assessed twice using the same tool. The aim of the analysis was to examine the repeatability of the questions contained within the data collection tool.

Statistical Methods

The data collection tool consisted of a series of questions. The majority of questions were categorical in nature, with a finite number of different responses. However, some response relating to dates were considered continuous in nature.

The repeatability of the measurements for the categorical variables were assessed using kappa method. These measures the agreement between repeat measurements over and above that which would be expected due to chance. It is measured on a scale ranging up to a maximum agreement of 1. Most categorical variables were either binary in nature, or unordered (nominal) variables. These were analysed using the standard (unweighted) kappa method. A small number of variables were ordinal in nature (e.g. age category). To allow for the extra information and the ordering of categories, the weighted kappa statistics was used for these outcomes.

Agreement between repeat measurements for the continuous variables were assessed using the intra-class correlation (ICC) method. This divides the total variation in all measurements combined into variation between patients, and variation within patients (due to the repeat measurements). The ICC the proportion of the total variation between patients. If there is good agreement between repeat measurements, the within-patient variation will be small, and thus the ICC value will be close to 1.

Although the kappa and ICC methods both score agreement on a scale up to 1, the interpretation of the values produced is different. A suggested interpretation of the strength of agreement based on the kappa and ICC values is suggested in Table 1. The kappa interpretation is fairly well documented, whilst there is less agreement on how to interpret the ICC values.

Table 1: Interpretation of Kappa and ICC values

Strength of agreement	Kappa	ICC
Poor	< 0.20	<0.60
Fair	0.21 – 0.40	0.61 – 0.70
Moderate	0.41 – 0.60	0.71 – 0.80
Good	0.61 – 0.80	0.81 – 0.90
Very Good	0.81 – 1.00	0.91 – 1.00

For all analyses, the kappa/ICC values from the patient sample were calculated, along with corresponding confidence intervals, indicating the level of uncertainty in the calculated values.

Results

Analyses were performed to examine the repeatability of the repeat measurements on the same patients.

The kappa method was used to examine the repeatability of the categorical variables. A summary of the analysis results for the demographic factors are given in Table 2. The figures are the number of patients on which the analyses were based, along with summaries of the number and percentage of responses in each category, both on the original and repeat measurements. Also reported are the calculated kappa values, presented with corresponding confidence intervals.

Table 2: Agreement for categorical variables - demographics

Quest.	Question Details	N	Category	Summary – n (%)		Kappa (95% CI)
				Original	Repeat	
Q1	Age (*)	565	≤ 75	154 (27%)	152 (27%)	0.97 (0.92, 1.00)
			76 – 80	119 (21%)	25 (22%)	
			81 – 85	137 (24%)	131 (23%)	
			86 – 90	104 (18%)	105 (19%)	
			91+	51 (9%)	52 (9%)	
Q2	Gender	565	Female	239 (42%)	240 (42%)	0.97 (0.89, 1.00)
			Male	326 (58%)	325 (58%)	
Q4	Sexual orientation	285	Heterosexual	284 (99%)	284 (99%)	1.00 (0.88, 1.00)
			Gay/lesion	1 (<1%)	1 (<1%)	
Q5	Ethnicity	487	White	432 (89%)	437 (90%)	0.89 (0.83, 0.95)
			Asian	23 (5%)	23 (5%)	
			Black	16 (3%)	15 (3%)	
			Mixed	7 (1%)	7 (1%)	
			Other	9 (2%)	15 (1%)	
Q6	1 st Language	509	English	471 (93%)	469 (93%)	0.88 (0.79, 0.97)
			Not English	35 (7%)	37 (7%)	
Q7	Interpreter required	547	No	521 (95%)	522 (95%)	0.94 (0.85, 1.00)
			Yes	26 (5%)	25 (5%)	
Q8	Lives alone	4	No	349 (64%)	356 (66%)	0.91 (0.82, 0.99)
			Yes	193 (36%)	186 (34%)	

(*) Analysis using weighted kappa

The results suggested very good agreement for all demographic factors, with high kappa values for each factor.

Similar analyses were performed for categorical variables relating to referral and patient evaluation. The results are summarised in Table 3.

Table 3: Agreement for categorical variables – referral and evaluation factors

Quest.	Question Details	N	Category	Summary – n (%)		Kappa (95% CI)
				Original	Repeat	
Q10	Referrer	565	GP	529 (94%)	531 (94%)	0.57 (0.51, 0.63)
			Acute hospital	11 (2%)	9 (2%)	
			CMHT	8 (1%)	5 (1%)	
			Day hospital	1 (<1%)	1 (<1%)	
			Other	16 (3%)	19 (3%)	
Q13	Assessed face to face	565	No	10 (2%)	10 (2%)	0.90 (0.82, 0.98)
			Yes	555 (98%)	555 (98%)	
Q13	Assessed virtually	565	No	548 (97%)	548 (97%)	0.82 (0.74, 0.90)
			Yes	17 (3%)	17 (3%)	
Q15	Alcohol units per week ^(*)	434	0 units	273 (63%)	267 (62%)	0.86 (0.79, 0.92)
			1 – 4 units	91 (21%)	95 (22%)	
			5 – 9 units	32 (7%)	26 (6%)	
			10 – 14 units	14 (3%)	20 (5%)	
			> 14 units	24 (6%)	26 (6%)	
Q16a	Eyesight recorded	565	No	378 (67%)	367 (65%)	0.78 (0.70, 0.87)
			Yes	187 (33%)	198 (35%)	
Q16b	Hearing recorded	565	No	357 (63%)	358 (63%)	0.76 (0.68, 0.4)
			Yes	208 (37%)	207 (37%)	
Q17	OT assessment	80	No	52 (65%)	52 (65%)	1.00 (0.78, 1.00)
			Yes	28 (35%)	28 (35%)	
Q18	Falls history	565	No	466 (82%)	457 (81%)	0.64 (0.56, 0.72)
			Yes	99 (18%)	108 (19%)	

(*) Analysis using weighted kappa

The agreement between repeat measurements for the referral factors ranged from moderate (Q10 referrer) up to very good (Q13 assessed face to face, Q17 OT assessment).

The results for measurements relating to scanning and investigations are summarised in Table 4.

Table 4: Agreement for categorical variables – Scanning and investigations

Quest.	Question Details	N	Category	Summary – n (%)		Kappa (95% CI)
				Original	Repeat	
Q19	Neuropsych. referral	565	No Yes	52 (9%) 513 (91%)	51 (9%) 514 (91%)	0.84 (0.76, 0.92)
Q20	Brain scan requested	565	No Yes	300 (53%) 265 (47%)	293 (52%) 272 (48%)	0.91 (0.83, 0.99)
Q22	Requester of brain scan	256	Memory service GP Hospital	246 (96%) 6 (2%) 4 (2%)	248 (97%) 4 (2%) 4 (2%)	0.66 (0.56, 0.75)
Q23	Brain scan performed	256	No Yes	13 (5%) 243 (95%)	12 (5%) 244 (95%)	0.96 (0.84, 1.00)
Q25	Type scan	243	MRI CT	72 (30%) 171 (70%)	72 (30%) 171 (70%)	0.96 (0.83, 1.00)
Q27	Special investigations	565	Not performed Performed	555 (98%) 10 (2%)	553 (98%) 12 (2%)	0.72 (0.64, 0.80)

There was very good agreement for some of the measures assessed, such as whether a brain scan was requested (Q20), whether the scan was performed (Q23) and the type of scan performed (Q25). Agreement was good for other measures, including who requested the scan (Q22) and special investigations (Q27).

Categorical variables relating to diagnosis and medication were also analysed, with the results summarised in Table 5.

Table 5: Agreement for categorical variables – diagnosis, medication and further treatment

Quest.	Question Details	N	Category	Summary – n (%)		Kappa (95% CI)
				Original	Repeat	
Q30	Diagnosis	565	Alzheimer's	176 (31%)	171 (30%)	0.92 (0.88, 0.97)
			Vascular dem.	58 (10%)	58 (10%)	
			Front lob dem.	5 (1%)	6 (1%)	
			Mix/other dem.	160 (28%)	159 (28%)	
			MCI	87 (15%)	91 (16%)	
			Non-dementia	79 (14%)	80 (14%)	
Q31	Diagnosis confirmed	565	Confirmed	525 (93%)	512 (91%)	0.45 (0.37, 0.53)
			Working	40 (7%)	53 (9%)	
Q33	Medication prescribed	565	Yes	239 (42%)	239 (42%)	0.88 (0.81, 0.95)
			Patient declined	21 (4%)	21 (4%)	
			Not appropriate	288 (51%)	287 (51%)	
			No, contraind.	17 (3%)	18 (3%)	
Q34	Type of medication	233	Donepezil	138 (%)	138 (59%)	0.98 (0.88, 1.00)
			Galantamine	5 (5%)	5 (2%)	
			Memantine	79 (34%)	79 (34%)	
			Rivast. oral	6 (3%)	6 (3%)	
			Rivast. patch	5 (2%)	5 (2%)	
Q35	CST offered	440	Yes	101 (23%)	93 (21%)	0.85 (0.78, 1.00)
			Not appropriate	195 (44%)	195 (44%)	
			Service unavail.	144 (33%)	152 (35%)	
Q35	CST provision	22	Face to face	20 (91%)	20 (91%)	1.00 (0.58, 1.00)
			Virtual	2 (9%)	2 (9%)	
Q36	Dementia advice service	565	No	219 (39%)	232 (41%)	0.75 (0.67, 0.84)
			Yes	346 (61%)	333 (59%)	
Q37	Psychoedu. course	565	Not offered	458 (81%)	462 (82%)	0.77 (0.68, 0.85)
			Offered	107 (19%)	103 (18%)	
Q38	Contacted for Research	321	No	304 (95%)	291 (100%)	0.82 (0.72, 0.91)
			Yes	17 (5%)	0 (0%)	

The results suggested very good agreement between the two sets of measurements for diagnosis (Q30) and type of medication (Q34). The lowest level of agreement was for diagnosis confirmed (Q31)), where the agreement was only moderate.

The ICC method was used to quantify the agreement between repeat measurements for the continuous variables, all of which related to the dates of events. The results are summarised

in Table 6. The figures are the number of patients in the analysis for each variable, and also the calculated ICC value, along with corresponding confidence intervals.

Table 6: Agreement for continuous variables

Question	Question details	N	ICC (95% CI)
Q11	Date of referral	565	0.87 (0.84, 0.89)
Q12	Date of initial assessment	565	0.97 (0.96, 0.97)
Q22	Date brain scan requested	256	0.76 (0.70, 0.80)
Q27	Date of brain scan	243	0.81 (0.76, 0.85)
Q32	Date of diagnosis	558	0.90 (0.88, 0.91)

The results suggested high levels of agreement between repeat measurements for some of the date variables, including date of initial assessment (Q12) and date of diagnosis (Q32). The agreement was only moderate for the date that the brain scan was requested (Q22).