

# Statsconsultancy Ltd

Freelance Statistical Consultancy

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

**Email:** [paul@statsconsultancy.co.uk](mailto:paul@statsconsultancy.co.uk)

**Phone:** 07905 530446

**Web:** [www.statsconsultancy.co.uk](http://www.statsconsultancy.co.uk)

## **Statistical Report**

**Author:** Paul Bassett

**Client:** Chloë Hood

**Date:** 26<sup>th</sup> April 2021

**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 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 (*)	505	≤ 75	139 (28%)	141 (28%)	0.99 (0.93, 1.00)
			76 – 80	97 (19%)	95 (19%)	
			81 – 85	123 (24%)	122 (24%)	
			86 – 90	100 (20%)	99 (20%)	
			91+	46 (9%)	48 (10%)	
Q2	Gender	505	Female	297 (59%)	295 (58%)	0.98 (0.90, 1.00)
			Male	208 (41%)	210 (42%)	
Q4	Sexual orientation	271	Heterosexual	269 (99%)	268 (99%)	0.80 (0.71, 0.89)
			Gay/lesion	1 (<1%)	1 (<1%)	
			Other	1 (<1%)	2 (<1%)	
Q5	Ethnicity	428	White	394 (92%)	397 (92%)	0.89 (0.83, 0.95)
			Asian	10 (2%)	9 (2%)	
			Black	9 (2%)	10 (2%)	
			Mixed	7 (2%)	6 (1%)	
			Other	8 (2%)	6 (1%)	
Q6	1 <sup>st</sup> Language	456	English	441 (97%)	440 (96%)	0.90 (0.81, 0.99)
			Not English	15 (3%)	16 (4%)	
Q7	Interpreter required	486	No	479 (99%)	479 (99%)	0.86 (0.77, 0.94)
			Yes	7 (1%)	7 (1%)	
Q8	Lives alone	472	No	322 (68%)	319 (68%)	0.93 (0.84, 1.00)
			Yes	150 (32%)	153 (32%)	

(\*) 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	504	GP	460 (91%)	464 (92%)	0.82 (0.77, 0.88)
			Acute hospital	15 (3%)	14 (3%)	
			CMHT	9 (2%)	11 (2%)	
			Day hospital	12 (2%)	8 (2%)	
			Other	8 (2%)	7 (1%)	
Q13	Assessed face to face	505	No	149 (29%)	152 (30%)	0.91 (0.81, 0.99)
			Yes	356 (71%)	353 (70%)	
Q13	Assessed virtually	505	No	338 (67%)	333 (66%)	0.92 (0.84, 1.00)
			Yes	167 (33%)	172 (34%)	
Q14	Video call facilities	54	No	10 (19%)	6 (11%)	0.71 (0.45, 0.96)
			Yes	44 (81%)	48 (89%)	
Q15	Alcohol units per week <sup>(*)</sup>	365	0 units	240 (66%)	234 (64%)	0.91 (0.83, 0.99)
			1 – 4 units	58 (16%)	61 (17%)	
			5 – 9 units	19 (5%)	15 (4%)	
			10 – 14 units	13 (4%)	23 (6%)	
			> 14 units	35 (10%)	32 (9%)	
Q16a	Eyesight recorded	505	No	218 (43%)	211 (41%)	0.75 (0.67, 0.84)
			Yes	287 (57%)	294 (58%)	
Q16b	Hearing recorded	505	No	239 (47%)	235 (47%)	0.82 (0.73, 0.90)
			Yes	266 (53%)	270 (53%)	
Q16c	Health discussion <sup>(+)</sup>	8	No	0 (0%)	0 (0%)	1.00 (#)
			Yes	8 (100%)	8 (100%)	
Q17	OT assessment	505	No	437 (87%)	432 (86%)	0.86 (0.77, 0.95)
			Yes	68 (13%)	73 (14%)	
Q18	Falls history	505	No	120 (24%)	120 (24%)	0.74 (0.65, 0.82)
			Yes	385 (76%)	385 (76%)	

(\*) Analysis using weighted kappa

(+) Data collected from patients in Wales only

(#) Unable to calculate confidence interval as all responses on both occasions in same category

The agreement between repeat measurements for the referral factors were at least good, and very good in some instances.

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	505	No	474 (94%)	452 (89%)	0.10 (0.01, 0.18)
			Yes	31 (6%)	53 (11%)	
Q20	Brain scan requested	505	No	271 (54%)	274 (54%)	0.92 (0.83, 1.00)
			Yes	234 (46%)	231 (46%)	
Q22	Requester of brain scan	214	Memory service	200 (93%)	196 (92%)	0.80 (0.69, 0.91)
			GP	11 (5%)	11 (5%)	
			Hospital	3 (1%)	7 (3%)	
Q23	Brain scan performed	214	No	24 (11%)	24 (11%)	0.91 (0.77, 1.00)
			Yes	190 (89%)	190 (89%)	
Q24	Type scan	188	MRI	56 (30%)	56 (30%)	0.94 (0.81, 1.00)
			CT	132 (70%)	132 (70%)	
Q25	Reason no scan done	25	Not required	10 (40%)	12 (48%)	0.75 (0.47, 1.00)
			Patient declined	10 (40%)	8 (32%)	
			Previous scan	5 (20%)	5 (20%)	
Q27	Special investigations	505	Not performed	497 (98%)	497 (98%)	1.00 (0.91, 1.00)
			Performed	8 (2%)	8 (2%)	

There was good agreement for all scanning variables, with very good agreement for whether a brain scan was requested, whether the scan was performed and the type of scan performed.

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	
Q29	Diagnosis	505	Alzheimer's	152 (30%)	148 (29%)	0.94 (0.90, 0.99)
			Vascular dem.	63 (12%)	65 (13%)	
			Front lob dem.	8 (2%)	8 (2%)	
			Mix/other dem.	117 (23%)	120 (23%)	
			MCI	84 (17%)	87 (17%)	
			Non-dementia	81 (16%)	77 (15%)	
Q30	Diagnosis confirmed	432	Confirmed	393 (91%)	393 (91%)	0.73 (0.64, 0.82)
			Working	39 (9%)	39 (9%)	
Q32	Medication prescribed	434	No	245 (56%)	245 (56%)	0.97 (0.88, 1.00)
			Yes	189 (44%)	189 (44%)	
Q33	Type of medication	186	Donepezil	98 (53%)	100 (54%)	0.96 (0.86, 1.00)
			Galantamine	3 (2%)	2 (1%)	
			Memantine	66 (35%)	64 (34%)	
			Rivast. oral	9 (5%)	10 (5%)	
			Rivast. patch	10 (5%)	10 (5%)	
Q34	CST offered - Face to face	499	No	456 (91%)	447 (90%)	0.24 (0.16, 0.33)
			Yes	43 (9%)	52 (10%)	
Q34	CST offered - Virtual	498	No	449 (90%)	449 (90%)	0.23 (0.14, 0.32)
			Yes	49 (10%)	49 (10%)	
Q34	CST offered - Any	498	No	434 (87%)	430 (86%)	0.25 (0.16, 0.34)
			Yes	64 (13%)	68 (14%)	
Q35	Dementia advice service	452	No	192 (42%)	221 (49%)	0.13 (0.03, 0.22)
			Yes	260 (58%)	231 (51%)	
Q35a	Dementia advice <sup>(+)</sup>	8	No	1 (12%)	1 (12%)	1.00 (0.31, 1.00)
			Yes	7 (88%)	7 (88%)	

(+) Data collected from patients in Wales only

The results suggested very good agreement between the two sets of measurements for diagnosis, the prescription of medication and type of medication.

However, there was only fair or poor agreement for variables relating to CST being offered (either face to face, virtually or of any type), and also poor agreement for the use of the dementia advice service. For all these variables, the kappa values were 0.25 or lower.

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	505	0.98 (0.97, 0.98)
Q12	Date of initial assessment	505	1.00 (0.99, 1.00)
Q21	Date brain scan requested	215	0.99 (0.98, 0.99)
Q26	Date of brain scan	188	0.97 (0.96, 0.98)
Q31	Date of diagnosis	434	0.93 (0.91, 0.94)

The results suggested very high levels of agreement between repeat measurements for all date variables.