Psychological Medicine, 2006, **36**, 1183–1191. © 2006 Cambridge University Press
doi:10.1017/S0033291706007811 First published online 17 May 2006 Printed in the United Kingdom

Pre-morbid IQ and response to routine outcome assessment

MIKE SLADE[1]*, MORVEN LEESE[1], MATT GILLARD[1], ELIZABETH KUIPERS[2] AND GRAHAM THORNICROFT[1]

[1] *Health Services Research Department, Institute of Psychiatry, King's College London, London, UK*;
[2] *Department of Psychology, Institute of Psychiatry, King's College London, London, UK*

ABSTRACT

Background. Routine use of standardized outcome measures within adult mental health services is compulsory in many countries, but not common in clinical practice. As well as political and professional factors, one reason may be that there is little empirical evidence of benefit for patients. It is therefore important to identify predictors of response. Pre-morbid intellectual functioning has predicted response to other interventions. This paper tests the hypothesis that pre-morbid IQ impacts on the effectiveness of routine assessment and feedback of outcome measures.

Method. The Feedback of Outcomes to Users and Staff (FOCUS) single-blind randomized controlled trial (RCT) investigated routine completion and feedback of outcome measures for 160 (101 intervention, 59 control) adult mental health service patients in Croydon, South London (International Standard RCT Number 16971059).

Results. Patients in the top quarter (pre-morbid IQ > 110) differentially improved in patient-rated unmet need (adjusted difference 3·4, 95% CI 0·8 to 5·9, $p = 0{\cdot}012$) and in the top half (pre-morbid IQ > 99) in quality of life (adjusted difference −0·6, 95% CI −1·1 to −0·1, $p = 0{\cdot}02$). The top quarter result remained when controlling for the influence of baseline ($p = 0{\cdot}004$) and baseline plus follow-up variables ($p = 0{\cdot}047$).

Conclusions. Feedback of routine outcome measures may improve outcome for patients with higher pre-morbid IQ. There is a need to understand more how routine outcome information is used by staff and patients to inform care. If the goal of routine use of outcome measures is to benefit patients directly, then targeted policies are required.

INTRODUCTION

The routine use of outcome measures is widely recommended (Department of Health, 2001; Health Research Council of New Zealand, 2003) but most psychiatrists do not use standardized assessments (Gilbody *et al.* 2002). Clinicians prefer to be evaluated on access, utilization, process or satisfaction measures than on outcome indicators (Valenstein *et al.* 2004).

* Address for correspondence: Dr Mike Slade, Health Services Research Department (Box P029), Institute of Psychiatry, King's College London, Denmark Hill, London SE5 8AF, UK.
(Email: m.slade@iop.kcl.ac.uk)

Why is this? A range of political and professional factors have been proposed (e.g. NIMHE, 2005), but one reason may be the evidence base. Evaluations of routine outcome measurement in adult mental health services both in the UK (Ashaye *et al.* 2003; Marshall *et al.* 2004) and in a cross-Europe study (Van Os *et al.* 2004) have involved single patient-rated needs assessments being fed back to staff, and the impact of routine use of outcome measures in adult mental health services on subjective outcomes has been limited and mainly negative. Consistent with wider reviews of the use of mental health outcomes (Gilbody *et al.* 2001),

https://doi.org/10.1017/S0033291706007811 Published online by Cambridge University Press

the evidence base for the beneficial impact of routine use of measures in an untargeted manner is weak.

If routine use of outcome measures with all patients is ineffective, this raises the question of whether there are characteristics of services, staff or patients that predict response to routine outcome assessment. There has been no specific research in this area. In the UK700 Study it was found that case management models had differential effects for patients on the basis of their pre-morbid IQ (Hassiotis *et al.* 2001). Therefore, we investigated whether there was an interaction between pre-morbid IQ and response to routine outcome measurement.

Following a systematic review of routine outcome measurement (Slade, 2002*b*), a model was developed of the intended effects of completion of outcome measures by staff and patients accompanied by identical feedback to both (Slade, 2002*a*). In summary, the model postulates that both outcome measure completion and receipt of feedback may have beneficial impacts on process and content of care, leading to improved outcomes. The model was then tested in an exploratory randomized controlled trial (RCT) (International Standard RCT Number 16971059). The results of the main study have been published elsewhere (Slade *et al.* in press). We report here a hypothesis-driven pre-planned subgroup analysis.

METHOD

The method is summarized here (for more details, see Slade *et al.* in press). The sample size was based on the comparison between the trial arms.

Hypothesis

The hypothesis tested was that pre-morbid IQ impacts on the response to the intervention of routine outcome assessment.

Participants

Patients had been on the caseload of any of the eight Community Mental Health Teams (CMHTs) in Croydon, South London on 1 May 2001 for at least 3 months, and were aged between 18 and 64 inclusive. Stratified random sampling was used for sample selection.

Measures

Staff completed three measures in their postal questionnaire. The Threshold Assessment Grid (TAG) is a seven-item assessment (range 0–24, low score better) of the severity of a person's mental health problems (Slade *et al.* 2000). The Camberwell Assessment of Need Short Appraisal Schedule Staff version (CANSAS-S) assesses the patient's needs from the staff perspective (Phelan *et al.* 1995). CANSAS assesses the presence of an unmet need (current serious problem, regardless of any help received) or a met need (no/moderate problem due to help given) in 22 domains of health and social need. The summary measures (range for both 0–22, low score better) are total number of met needs and total number of unmet needs. The Helping Alliance Scale – Staff (HAS-S) assesses therapeutic alliance from the staff perspective (range 0–10, high score better) (McCabe *et al.* 1999).

Patients completed three measures in their postal questionnaire. The CANSAS-P assesses the patient's met and unmet needs from the patient's perspective (summary scores as for CANSAS-S) (Phelan *et al.* 1995). The Manchester Short Assessment (MANSA) assesses quality of life (Priebe *et al.* 1999). Twelve subjective domains of life satisfaction are assessed on a seven-point satisfaction scale. The summary score (range 1–7, high score better) is the mean of the 12 items. The HAS-P assesses therapeutic alliance from the patient perspective (summary score as for HAS-S) (McCabe *et al.* 1999). Patient-rated unmet need (assessed using CANSAS-P) and quality of life (assessed using MANSA) were the primary outcomes.

Intellectual functioning was measured at baseline only using the second edition of the interviewer-rated National Adult Reading Test (NART; Nelson, 1982). The NART error score can range from 0 to 50, and converts into an estimate of pre-morbid full-scale IQ. A NART error score of 25 indicates a pre-morbid full-scale IQ of 100.

Two measures were assessed at baseline and follow-up only. The Brief Psychiatric Rating Scale (BPRS) is an 18-item interviewer-rated assessment of symptomatology (summary score range 0–126, low score better) (Overall & Gorham, 1988). The Health of the Nation Outcome Scale (HoNOS) is a 12-item staff-rated

https://doi.org/10.1017/S0033291706007811 Published online by Cambridge University Press

assessment of social disability (summary score range 0–48, low score better) (Wing *et al*. 1998). At follow-up, self-report data were collected from staff and patients on the cognitive and behavioural sequelae of the intervention for the process and content of care.

Procedures

Following baseline assessments, patients were randomly allocated by a purpose-written STATA program (StataCorp., College Station, TX, USA) using minimization to ensure balance on prognostic factors.

The control group received treatment as usual. This involved mental health care from the multi-disciplinary CMHT (comprising psychiatrists, psychiatric nurses, occupational therapists, clinical psychologists and social workers) focused on informally assessed mental health and social care needs, along with care from the general practitioner for physical health problems.

The intervention group received the same care as the control group, but in addition patients and staff (i) were asked to complete a postal questionnaire on a monthly basis; and (ii) were provided with identical feedback by post at 3-monthly intervals. The postal questionnaire element of the baseline data was used for the round 1 information. One month after allocation, the first postal questionnaire was sent (round 2). Postal questionnaires were then sent monthly. Feedback was sent 2 weeks after the round 3 and the round 6 postal questionnaires to both staff and patients. Feedback comprised colour-coded graphic and text: charts compared summary ratings over the preceding 3 months, and text described item-level responses for the most recent rating, highlighting areas of disagreement. No instruction was given about how the information was to be used. The intervention thus comprised two cycles of three rounds of monthly postal questionnaires followed by feedback. Patients were paid £5 for each round of assessments. Follow-up assessments were made 6 weeks after the second feedback was sent (i.e. at 7 months).

Analysis

Interaction between pre-morbid IQ and response to the intervention was investigated using regression analysis. The dependent variables were patient-rated unmet need (CANSAS-P) and quality of life (MANSA). Categorical factors were allocation status (Control *versus* Intervention), gender, CMHT (1–8), ethnicity (Non-White *versus* White), diagnosis (non-psychosis *versus* psychosis) and pre-morbid NART IQ subgroup indicator dichotomized (to maximize clinical utility) into (i) the top half *versus* the rest and (ii) the top quarter *versus* the rest for two alternative analyses. Continuous explanatory variables were age, baseline TAG, CANSAS-S met needs, CANSAS-S unmet needs, HAS-S, HAS-P, BPRS, HoNOS, CANSAS-P met needs and whichever of CANSAS-P unmet or MANSA was not the dependent variable. All variables were entered as main effects, and the two-way interaction was tested between the NART subgroup indicator and allocation status.

Where there was evidence for interaction, separate models were fitted for the IQ subgroups. For each outcome measure (unmet need and quality of life), two models were estimated using a hierarchical strategy for entering variables in blocks. Preliminary analysis indicated that age, gender and CMHT did not contribute to the model, so these were omitted. Model 1 entered variables in five blocks: ethnicity, pre-morbid IQ (actual score), diagnosis, all baseline clinical measures, and allocation status. Model 2 had six blocks, blocks 1–4 the same as Model 1, block 5 comprised follow-up variables and block 6 comprised the allocation status. Data analysis was undertaken using SPSS for Windows version 10 (SPSS Inc., Chicago, IL, USA).

RESULTS

One hundred and sixty patients (101 intervention, 59 control) were recruited. Overall, there was no evidence for differences between groups in mean follow-up patient-rated unmet need (mean difference 0·15, 95% CI −1·20 to 1·49, $p=0·83$) or in mean follow-up quality of life (mean difference −0·07, 95% CI −0·44 to 0·31, $p=0·72$) (Slade *et al*. in press).

One hundred and fifty-three patients had baseline NART data. Mean NART error score was 25·8 (S.D. = 12·2) in the intervention group and 24·9 (S.D. = 11·5) in the control group, and 25·4 (S.D. = 11·9) overall. This converts into

https://doi.org/10.1017/S0033291706007811 Published online by Cambridge University Press

Table 1. *Baseline clinical and social characteristics of intervention and control group patients, by pre-morbid IQ groups*

	Intervention group			Control group		
	All ($n=101$)	Top half ($n=50$)	Top quarter ($n=23$)	All ($n=59$)	Top half ($n=27$)	Top quarter ($n=15$)
Age, mean (S.D.)	41·8 (11·4)	44·6 (10·9)	46·4 (8·9)	40·2 (10·8)	40·5 (10·5)	43·4 (10·1)
Male, *n* (%)	48 (48)	26 (52)	12 (52)	30 (51)	14 (52)	8 (53)
White, *n* (%)	77 (76)	39 (78)	20 (87)	45 (76)	22 (82)	13 (87)
Education						
No formal qualification, *n* (%)	38 (38)	12 (24)	6 (26)	23 (39)	6 (22)	2 (13)
GCSE/GCE, *n* (%)	28 (28)	15 (30)	7 (30)	19 (32)	11 (41)	8 (53)
A levels or higher, *n* (%)	34 (34)	19 (38)	10 (43)	17 (29)	6 (22)	3 (20)
Diagnosis						
Schizophrenia, *n* (%)	40 (40)	20 (40)	10 (44)	20 (34)	9 (33)	6 (40)
Bipolar affective disorder, *n* (%)	8 (8)	6 (12)	2 (9)	9 (15)	5 (19)	3 (20)
Other psychoses, *n* (%)	12 (12)	6 (12)	3 (13)	7 (12)	2 (7)	0 (0)
Affective disorder, *n* (%)	27 (27)	14 (28)	5 (22)	16 (27)	9 (33)	5 (33)
Years since first contact, mean (S.D.)	14·2 (12·6)	16·5 (13·8)	17·8 (14·3)	11·1 (9·8)	11·1 (10·1)	13·5 (10·5)
Years in this episode of care, mean (S.D.)	4·3 (4·7)	3·8 (3·1)	3·8 (3·0)	3·9 (3·3)	3·4 (3·1)	3·4 (3·2)
Staff-rated						
TAG	5·4 (3·6)	4·6 (3·3)	3·7 (2·7)	4·8 (3·7)	5·1 (4·3)	4·1 (2·8)
CANSAS-S unmet	3·2 (3·3)	3·0 (3·2)	2·6 (2·5)	2·5 (2·9)	3·0 (3·0)	2·3 (2·4)
CANSAS-S met	5·1 (3·3)	4·7 (3·1)	4·3 (3·4)	5·0 (3·7)	4·4 (3·4)	3·3 (2·4)
HAS-S	7·5 (1·6)	7·6 (1·4)	7·7 (1·4)	7·1 (1·6)	6·7 (1·9)	6·6 (2·1)
HoNOS	9·2 (6·6)	8·2 (6·6)	7·0 (4·8)	8·4 (6·1)	8·9 (5·1)	7·5 (4·5)
Patient-rated						
MANSA	4·3 (1·0)	4·1 (1·0)	4·2 (0·9)	4·3 (1·1)	4·0 (0·9)	4·0 (1·0)
CANSAS-P unmet	4·4 (3·4)	4·2 (3·5)	3·7 (3·1)	5·0 (4·1)	5·3 (4·4)	5·1 (4·7)
CANSAS-P met	4·2 (2·8)	4·4 (2·8)	4·7 (3·0)	4·2 (3·0)	4·1 (3·4)	5·1 (3·5)
HAS-P	8·2 (1·8)	8·0 (1·8)	8·3 (1·2)	7·5 (2·1)	7·6 (2·2)	7·9 (2·4)
Interviewer-rated						
BPRS	33·4 (9·0)	33·4 (9·5)	32·7 (7·0)	33·8 (9. 8)	33·5 (8·7)	32·3 (9·9)

TAG, Threshold Assessment Grid; CANSAS, Camberwell Assessment of Need Short Appraisal Schedule (P, Patient version; S, Staff version); HAS, Helping Alliance Scale (P, Patient version; S, Staff version); HoNOS, Health of the Nation Outcome Scale; MANSA, Manchester Short Assessment; BPRS, Brief Psychiatric Rating Scale.

a mean estimated full-scale IQ of 99 in the intervention group, 100 in the control group, and 99 (range 69–128) overall.

To allow investigation of the impact of pre-morbid IQ, the 153 patients were partitioned into quartiles, comprising those with pre-morbid IQ scores ranging from 69 to 89 ($n=38$), from 90 to 99 ($n=38$), from 100 to 110 ($n=39$), and from 111 to 128 ($n=38$). They were also partitioned into halves, comprising those ranging from 69 to 99 ($n=76$) and from 100 to 128 ($n=77$). The baseline characteristics of the patients are shown in Table 1.

The intervention was offered to all 101 patients in the intervention group. Return rates for patients from rounds 2 to 6 were 85, 84, 76, 76 and 76%. Return rates for staff were 78, 71, 67, 59 and 58%. Three-monthly summary feedback was sent after round 3 to 96 staff–patient pairs (95% of intervention group), and after round 6 to 93 staff–patient pairs (92% of intervention group). During the study, two intervention group patients died from natural causes and one intervention group patient withdrew consent. Follow-up data were collected for 93 (92%) patients in the intervention group and 49 (83%) patients in the control group, a total of 142 (89%) out of 160 patients in the study. Follow-up staff data were collected for 56 (95%) control and 95 (94%) intervention group patients, a total of 151 (94%) out of 160 patients in the study. Follow-up case-note data were collected for 58 (98%) control and 100 (99%) intervention group patients, a total of 158 (99%) out of 160 patients in the study. All follow-up data were collected by researchers who were blind to allocation status.

To test for an interaction between allocation status and pre-morbid IQ scores, models including interaction terms were fitted. Of the 153

https://doi.org/10.1017/S0033291706007811 Published online by Cambridge University Press

Table 2. *Outcome results by pre-morbid IQ status*

	Allocation status						
	Intervention		Control		Difference[a]		
	Mean (S.D.)	*n*	Mean (S.D.)	*n*	Estimated	95% CI	*p*
Follow-up patient-rated unmet need ($n=136$)							
Top half	3·3 (3·3)	47	4·7 (4·1)	23	1·4	−0·4 to 3·2	0·13
Lower half	4·6 (3·8)	43	4·0 (4·7)	23	−0·7	−2·8 to 1·5	0·55
Top quarter	2·5 (2·6)	21	5·9 (4·7)	13	**3·4**	**0·8 to 5·9**	**0·012**
Lower three-quarters	4·4 (3·7)	69	3·7 (4·1)	33	−0·7	−2·3 to 1·0	0·43
Follow-up quality of life ($n=135$)							
Top half	4·4 (1·0)	46	3·8 (1·0)	23	**−0·6**	**−1·1 to −0·1**	**0·02**
Lower half	4·1 (1·1)	43	4·4 (1·1)	23	0·3	−0·3 to 0·8	0·37
Top quarter	4·5 (0·7)	21	4·0 (1·1)	13	−0·5	−1·2 to 0·1	0·09
Lower three-quarters	4·2 (1·1)	68	4·2 (1·1)	33	0·0	−0·5 to 0·4	0·83

CI, confidence interval.
$p<0·05$ shown in bold.
[a] Difference = control score – intervention score.

patients with pre-morbid IQ scores, 135 (88%) had baseline and follow-up quality of life data, and 136 (89%) had baseline and follow-up patient-rated unmet need data. The interaction terms were between allocation status and being either in the top half or in the top quartile patient subgroups. The dependent variable was either follow-up patient-rated unmet need or quality of life. Four models were therefore estimated, each including all sociodemographic and baseline clinical variables. An interaction was found for follow-up patient-rated unmet need between allocation status and the top quarter of patients ($p=0·002$). An interaction was found for quality of life between allocation status and the top half of patients ($p=0·015$). As this analysis provided some *post hoc* evidence that the intervention may be differentially effective on the basis of pre-morbid IQ, further analysis was undertaken.

To investigate differences in response to the intervention between patients with high pre-morbid IQ compared to other patients, the top half and the top quarter of patients were compared with the remaining patients. For patients in the top half, the mean patient-rated unmet need rating at follow-up ($n=70$) was 3·7 (S.D. = 3·6) and the mean quality of life rating ($n=69$) was 4·2 (S.D. = 1·0). For the 34 patients in the top quarter, the mean patient-rated unmet need rating was 3·8 (S.D. = 3·9) and the mean quality of life rating was 4·3 (S.D. = 0·9). The results by intervention group are shown in Table 2.

There were significant differences between the subgroups in terms of their response to the intervention – quality of life was improved by the intervention for patients with higher IQ (statistically significant for the top half of patients) and patient-rated unmet needs was reduced for those patients by the intervention (statistically significant for the top quarter). The difference expressed as an effect size was moderate to large (0·8), and the significance remained after Bonferroni correction to take account of the two categorizations of the data.

The predefined criteria for improvement were a reduction of 1·0 in CANSAS-P or an increase of 0·25 in MANSA from baseline to follow-up. For patients in the top quarter of pre-morbid IQ, 13 (62%) of the 21 intervention group reduced their patient-rated unmet need by at least 1·0 need, as did four (31%) of the 13 control group. This gives a risk reduction of 32%, equivalent to a number need to treat (NNT) of 3, indicating that three patients with a pre-morbid IQ of at least 111 would need to receive the intervention for one to experience a reduction of at least 1·0 patient-rated unmet need. An NNT of 2 or 3 indicates that a treatment is quite effective (McQuay & Moore, 1997). For patients in the top half of pre-morbid IQ (at least 100), 25 (54%) of the 46 intervention group improved in MANSA rating of quality of life by at least 0·25, as did seven (30%) of the 23 control group. This gives a risk reduction of 24% and an NNT of 4.

https://doi.org/10.1017/S0033291706007811 Published online by Cambridge University Press

Table 3. *Regression on follow-up patient-rated unmet need for pre-morbid IQ top quarter of patients (n = 34), with (Model 2) and without (Model 1) follow-up variables*

	Model 1 (baseline only) $r^2 = 0{\cdot}79$, adjusted $r^2 = 0{\cdot}63$			Model 2 (baseline and follow-up) $r^2 = 0{\cdot}97$, adjusted $r^2 = 0{\cdot}89$		
	B	*β*	*p*	*B*	*β*	*p*
Intervention group	**−3·33**	**−0·41**	**0·004**	**−3·88**	**−0·48**	**0·047**
White ethnicity	**3·53**	**0·33**	**0·020**	2·01	0·19	0·16
Pre-morbid IQ	0·00	0·00	0·97	−0·18	−0·22	0·13
Psychosis	1·96	0·24	0·10	**3·20**	**0·39**	**0·026**
Baseline						
CANSAS-P unmet	**0·72**	**0·70**	**0·002**	**0·62**	**0·59**	**0·008**
CANSAS-P met	0·09	0·08	0·58	−0·13	−0·11	0·37
MANSA	1·40	0·34	0·12	**3·73**	**0·88**	**0·005**
HAS-P	−0·47	−0·22	0·14	−0·72	−0·34	0·09
CANSAS-S unmet	−0·29	−0·18	0·41	−0·50	−0·31	0·32
CANSAS-S met	−0·31	−0·25	0·09	−0·28	−0·22	0·18
HAS-S	**0·93**	**0·43**	**0·026**	−0·01	−0·01	0·97
TAG	0·51	0·35	0·18	−0·56	−0·39	0·24
BPRS	0·12	0·26	0·22	0·00	0·01	0·97
HoNOS	0·06	0·07	0·83	**0·78**	**0·97**	**0·046**
Follow-up						
MANSA				**−3·32**	**−0·76**	**0·011**
CANSAS-P met				**0·70**	**0·57**	**0·018**
HAS-P				−0·52	−0·28	0·09
CANSAS-S unmet				−0·84	−0·41	0·06
CANSAS-S met				0·01	0·01	0·94
HAS-S				**1·38**	**0·59**	**0·049**
TAG				−0·06	−0·04	0·81
BPRS				−0·13	−0·35	0·08
HoNOS				**0·41**	**0·55**	**0·021**

CANSAS, Camberwell Assessment of Need Short Appraisal Schedule (P, Patient version; S, Staff version); MANSA, Manchester Short Assessment; HAS, Helping Alliance Scale (P, Patient version; S, Staff version); TAG, Threshold Assessment Grid; BPRS, Brief Psychiatric Rating Scale; HoNOS, Health of the Nation Outcome Scale.
$p < 0{\cdot}05$ shown in bold.

The impact of including baseline and follow-up variables was investigated. Two linear regression models were estimated. Model 1 included allocation status, sociodemographic and baseline clinical measures. Model 2 included all independent variables from Model 1 plus follow-up variables. One positive result from Table 2 was not supported. Allocation status did not predict follow-up quality of life for the top half (Model 1: adjusted difference 0·3, $p = 0{\cdot}10$; Model 2: adjusted difference 0·2, $p = 0{\cdot}21$), indicating that the relationship found in Table 2 is not maintained when controlling for the influence of baseline or follow-up variables. The other positive result – allocation status predicts follow-up patient-rated unmet need for the top quarter – was confirmed, as shown in Table 3.

Receiving the intervention was associated with reduced patient-rated unmet need for the top quarter IQ patients ($n = 38$), when controlling for baseline and follow-up variables. To investigate this positive result for top-quarter IQ patients, the intervention group ratings of the impact of the intervention were compared for the top quarter and the lower three-quarters pre-morbid IQ groups. This is shown in Table 4.

All ratings except that for patients for 'thought about the content of care' were higher for the higher IQ group. There were significant differences between groups for staff assessment of changes in the process of care ($\chi^2_1 = 8{\cdot}90$, $p = 0{\cdot}003$). Staff were more likely for higher pre-morbid IQ patients to assess that the process of care had been impacted on through involvement in the study than for lower pre-morbid IQ patients.

DISCUSSION

The impact of pre-morbid IQ on the ability to benefit from specific feedback of routine

https://doi.org/10.1017/S0033291706007811 Published online by Cambridge University Press

Table 4. *Self-rated impact of the intervention from top-quarter versus lower three-quarters pre-morbid IQ patients and their staff*

	Patients		Staff	
	Top quarter ($n=17$)	Lower three-quarters ($n=65$)	Top quarter ($n=15$)	Lower three-quarters ($n=58$)
Feedback was understandable	15 (88)	53 (84)	14 (93)	44 (85)
Participation promoted thought about the content of care	13 (77)	56 (86)	15 (100)	55 (95)
Participation promoted thought about the process of care	15 (88)	47 (73)	15 (100)	53 (93)
Participation led to discussion about the content of care	8 (47)	28 (29)	9 (64)	26 (48)
Participation led to changes in the process of care	5 (31)	8 (13)	11 (73)	18 (33)*

Data are *n* (% of valid responses).
* $p<0{\cdot}05$.

outcome assessment was investigated. A differential potential to benefit was found for those with a pre-morbid IQ of 111 or more. There were also indications of a weaker effect (possibly not identified as statistically significant due to lack of power) for those with pre-morbid IQ of 100 or more. Pre-morbid IQ is the first potential moderator (Baron & Kenny, 1986) for feedback of routine outcome assessment to be identified.

Research into the association between cognitive abilities and outcome presents mixed findings. For example, lower pre-morbid IQ was associated with better response to case management in terms of care costs and clinical outcomes including satisfaction with care and level of unmet need (Van Minnen *et al.* 1997; Hassiotis *et al.* 2001). By contrast, lower pre-morbid IQ has also been shown to be associated with poorer long-term outcomes (Munro *et al.* 2002), but not to predict response to cognitive behavioural therapy for psychosis (Garety *et al.* 1997). A more differentiating approach was used by Gold *et al.* (2002), who made detailed baseline neuropsychological assessments of 150 patients entering a vocational rehabilitation programme in Baltimore. They found that cognitive performance did not differentiate between the 40 patients who obtained competitive employment at 2-year follow-up and the 110 who did not. However, at follow-up the total number of days worked was associated with 16 of the 26 cognitive indicators, indicating that cognitive ability was not a predictor of the ability to obtain employment, but was predictive of employment tenure. Overall, the extent to which the cognitive abilities of the patient inhibit or promote the effectiveness of health interventions merits future consideration. It may be that future work will indicate the need to provide differentiated interventions for different cognitive ability groups.

Why did higher pre-morbid IQ interact with routine outcome assessment?

At least three possible explanations would account for the findings, each of which has implications for future research. A first potential explanation is that more intellectually able patients may make better use of the intervention, both in reframing their own experience and in being prompted to discuss with staff their care plan or the therapeutic alliance. Furthermore, staff may find it easier to use the feedback to inform future care with more intellectually able patients. This would be consistent with the results, which found more alteration in both staff and patient behaviour in higher than in lower pre-morbid IQ patients subgroups, a difference that reached significance for staff assessments of the impact of the study on the process of care. As this study involved a subgroup analysis (with reduced power), these results are hypothesis generating rather than confirmatory. It is plausible that a larger sample size would have found a significant effect of the intervention on both patient-rated unmet need and quality of life for patients with higher pre-morbid IQs. This could be tested by replicating the Feedback of Outcomes to Users and Staff (FOCUS) RCT with an added inclusion criterion of the patient's pre-morbid IQ being 100 or more. Based on the current results, such a study would require a sample of 74 patients in each group (148 in total) to detect a difference

https://doi.org/10.1017/S0033291706007811 Published online by Cambridge University Press

of 54% of improvers in quality of life in the intervention group compared with 30% in the control group with power of 80% and critical value 0·05.

A second explanation is that lower pre-morbid IQ patients were disadvantaged by the written medium of the feedback, and require other means of information transfer. A pilot study prior to the full RCT investigated the best means of presenting the feedback, but the patient numbers involved were small ($n=5$) and differing preferences were expressed. The combined graphical and text-based feedback that was developed on the basis of this feedback may not have been the most accessible for patients with lower pre-morbid IQ. This hypothesis is not supported by the data presented in Table 4, which found no differences in the self-rated understanding of the feedback by higher and lower pre-morbid IQ patients, although these self-report data may have been inflated by social desirability bias – some patients being too embarrassed to admit that they did not understand the feedback. This explanation could be tested by further pilot study work to compare different modes of information transfer for how comprehensible they are to patients of different intellectual abilities. For example, written feedback could be compared with a face-to-face meeting between the patient, the staff and the feedback provider for high and low pre-morbid IQ patients.

A third possible explanation is that lower IQ was associated with a more coercive existing style of interaction that was less amenable to shaping through the intervention. Investigating this hypothesis would involve measuring baseline levels of coercion and associated variables (such as empowerment), monitoring the interactions (especially regarding the assessments and feedback) that occur between patient and staff during the study using conversation analysis techniques (Ten Have, 1999), and then analysing whether baseline levels or style of interaction predict response to the intervention, independent of cognitive measures.

Implications for clinical practice

In the main trial paper we showed that the intervention was associated with reduced admission rates and hence costs, but that overall the intervention did not lead to improvements in the primary outcomes (Slade *et al*. in press). Possible reasons were considered, including contamination due to non-cluster design, reduced implementation due to staff turn-over, and lack of prescriptive feedback.

In the current study we report a pre-planned subgroup analysis that raises concerns about over-analysis of results (Freemantle, 2001). Oxman & Guyatt (1992) propose that in subgroup analysis, the strength of inference regarding a proposed difference in treatment effect among subgroups is dependent on: the magnitude of the difference; the statistical significance of the difference; whether the hypothesis preceded or followed the analysis; whether the subgroup analysis was one of a small number of hypotheses tested; whether the difference was suggested by comparisons within or between studies; the consistency of the difference; and the existence of indirect evidence that supports the difference. The current study found a marked and significant difference, was the only subgroup analysis, identified a difference by comparison within a study, showed a consistent difference, and cited indirect evidence to support the difference. Overall, the findings may warrant further consideration, even though they arise from subgroup analysis.

If replicated, then this result indicates the importance of taking account of cognitive functioning when considering the routine use of outcome measures. Such an approach would contrast with the current political climate of targeting specialist mental health services towards *all* people with more severe and complex mental health problems. It also presents logistical challenges for community mental health teams – how in practice should routine outcome assessment be implemented with a heterogeneous case-load, if we know in advance that some patients are unlikely to benefit? Finally, it poses an ethical dilemma – limiting interventions to those who will benefit from them promotes efficiency, but excluding those with lower pre-morbid IQ from access to particular interventions may reinforce social exclusion.

Future research should move beyond pure RCT methodologies, and instead focus on identifying moderators (predictors of the outcome, as in this study) and mediators (variables impacting on the strength or direction of the

https://doi.org/10.1017/S0033291706007811 Published online by Cambridge University Press

outcome) when investigating the effectiveness of interventions (Baron & Kenny, 1986). Identifying moderators will involve theoretically driven, pre-planned and adequately powered subgroup analysis. Investigating mediators will require the 'black box' of interventions to be opened, to give an understanding of *how* interventions are implemented in practice.

ACKNOWLEDGEMENTS

Data were collected by Dr Sharon Cahill, Lisa McNally, Dr Alberto Parabiaghi and Dr Giovanni Salvi. The Trial Steering Committee comprised Prof. Tom Burns, Prof. Stefan Priebe and Dr Rose McCabe. Administrative support was provided by Maura Butcher, Karen Langridge, Sarah Sheppard-Wright, Michelle Wilkins and Roza Wojcik. The generous involvement of Croydon staff and mental health service users is gratefully acknowledged. The study was funded by a Medical Research Council Clinician Scientist Fellowship to M.S. (G108/390).

DECLARATION OF INTEREST

None.

REFERENCES

Ashaye, O. A., Livingston, G. & Orrell, M. W. (2003). Does standardized needs assessment improve the outcome of psychiatric day hospital care for older people? A randomized controlled trial. *Aging and Mental Health* **7**, 195–199.

Baron, R. M. & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology* **51**, 1173–1182.

Department of Health (2001). *Mental Health Information Strategy*. The Stationery Office: London.

Freemantle, N. (2001). Interpreting the results of secondary end points and subgroup analyses in clinical trials: should we lock the crazy aunt in the attic? *British Medical Journal* **322**, 989–991.

Garety, P., Fowler, D., Kuipers, E., Freeman, D., Dunn, G., Bebbington, P., Hadley, C. & Jones, S. (1997). London–East Anglia randomised controlled trial of cognitive-behavioural therapy for psychosis. II: Predictors of outcome. *British Journal of Psychiatry* **171**, 420–426.

Gilbody, S. M., House, A. O. & Sheldon, T. A. (2001). Routinely administered questionnaires for depression and anxiety: systematic review. *British Medical Journal* **322**, 406–409.

Gilbody, S. M., House, A. O. & Sheldon, T. A. (2002). Psychiatrists in the UK do not use outcome measures. *British Journal of Psychiatry* **180**, 101–103.

Gold, J. M., Goldberg, R. W., McNary, S. W., Dixon, L. B. & Lehman, A. F. (2002). Cognitive correlates of job tenure among patients with severe mental illness. *American Journal of Psychiatry* **159**, 1395–1402.

Hassiotis, A., Ukoumunne, O. C., Byford, S., Tyrer, P., Harvey, K., Piachaud, J., Gilvarry, K. & Fraser, J. (2001). Intellectual functioning and outcome of patients with severe psychotic illness randomised to intensive case management. *British Journal of Psychiatry* **178**, 166–171.

Health Research Council of New Zealand (2003). *Partnerships for Evidence-based Policy and Practice*. HRCNZ: Auckland.

Marshall, M., Lockwood, A., Green, G., Ajac-Roles, G., Roberts, C. & Harrison, G. (2004). Systematic assessments of need and care planning in severe mental illness: cluster randomised controlled trial. *British Journal of Psychiatry* **185**, 163–168.

McCabe, R., Roeder-Wanner, U.-U., Hoffman, K. & Priebe, S. (1999). Therapeutic relationships and quality of life: association of two subjective constructs in schizophrenia patients. *International Journal of Social Psychiatry* **45**, 276–283.

McQuay, H. J. & Moore, R. A. (1997). Using numerical results from systematic reviews in clinical practice. *Annals of Internal Medicine* **126**, 712–720.

Munro, J. C., Russell, A. J., Murray, R. M., Kerwin, R. W. & Jones, P. B. (2002). IQ in childhood psychiatric attendees predicts outcome of later schizophrenia at 21 year follow-up. *Acta Psychiatrica Scandinavica* **106**, 139–142.

NIMHE (2005). *Outcome Measures Implementation Best Practice Guidance*. National Institute for Mental Health in England: Leeds.

Nelson, H. E. (1982). *National Adult Reading Test* (*NART*): *Test Manual*. NFER-Nelson: Windsor.

Overall, J. E. & Gorham, D. R. (1988). The Brief Psychiatric Rating Scale (BPRS): recent developments in ascertainment and scaling. *Psychopharmacology Bulletin* **24**, 97–99.

Oxman, A. D. & Guyatt, G. H. (1992). A consumers guide to subgroup analysis. *Annals of Internal Medicine* **116**, 78–84.

Phelan, M., Slade, M., Thornicroft, G., Dunn, G., Holloway, F., Wykes, T., Strathdee, G., Loftus, L., McCrone, P. & Hayward, P. (1995). The Camberwell Assessment of Need: the validity and reliability of an instrument to assess the needs of people with severe mental illness. *British Journal of Psychiatry* **167**, 589–595.

Priebe, S., Huxley, P., Knight, S. & Evans, S. (1999). Application and results of the Manchester Short Assessment of quality of life. *International Journal of Social Psychiatry* **45**, 7–12.

Slade, M. (2002*a*). Routine outcome assessment in mental health services. *Psychological Medicine* **32**, 1339–1344.

Slade, M. (2002*b*). What outcomes to measure in routine mental health services, and how to assess them – a systematic review. *Australian and New Zealand Journal of Psychiatry* **36**, 743–753.

Slade, M., McCrone, P., Kuipers, E., Leese, M., Cahill, S., Parabiaghi, A., Priebe, S. & Thornicroft, G. (in press). Use of standardised outcome measures in adult mental health services: randomised controlled trial. *British Journal of Psychiatry*.

Slade, M., Powell, R., Rosen, A. & Strathdee, G. (2000). Threshold Assessment Grid (TAG): the development of a valid and brief scale to assess the severity of mental illness. *Social Psychiatry and Psychiatric Epidemiology* **35**, 78–85.

Ten Have, P. (1999). *Doing Conversation Analysis: A Practical Guide*. Sage: London.

Valenstein, M., Mitchinson, A., Ronis, D. L., Alexander, J. A., Duffy, S. A., Craig, T. J. & Barry, K. L. (2004). Quality indicators and monitoring of mental health services: what do frontline providers think? *American Journal of Psychiatry* **161**, 146–153.

Van Minnen, A., Hoodguin, C. A. & Broekman, T. G. (1997). Hospital v. outreach treatment of patients with mental retardation and psychiatric disorders: a controlled study. *Acta Psychiatrica Scandinavica* **95**, 515–552.

Van Os, J., Altamura, A. C., Bobes, J., Owens, D. C., Gerlach, J., Hellewell, J. S. E., Kasper, S., Naber, D. & Robert, P. (2004). Evaluation of the two-way communication checklist as a clinical intervention. *British Journal of Psychiatry* **184**, 79–83.

Wing, J. K., Beevor, A. S., Curtis, R. H., Park, S. B., Hadden, S. & Burns, A. (1998). Health of the Nation Outcome Scales (HoNOS) – research and development. *British Journal of Psychiatry* **172**, 11–18.

https://doi.org/10.1017/S0033291706007811 Published online by Cambridge University Press