Routine Outcome Monitoring in the Forensic Psychiatric Centre "The Kijvelanden"

Frida van der Veeken

University of Tilburg

1e supervisor: Prof. dr. S. Bogaerts

2e supervisor: Drs. B. Koudstaal.

Thanks to: T. Kanters M.Sc. & Drs A. Zwets M.Sc.

Abstract

Routine outcome monitoring is a methodology which is increasingly used in mental health care institutions in order to study treatment progress. Several instruments have been developed for the benefit of ROM, however the instruments to study a decrease in risk factors in the forensic mental health care are limited. Therefore the instrument for forensic policlinic treatment evaluation (IFpBE) is designed. In this study 34 participants rated 2 self reports together with 4 different therapists who rated a practitioners report for every participant with a period of three months between the two measurements. The intra correlation coefficients, the deviance from the therapist for psychotic and non psychotic patients and the changeability have been studied. Results show reasonable through good intra correlation coefficients, a limited difference in deviance and a limited changeability. The results show an implication for the use of ROM, whit the multidisciplinary rating and consideration of the most important risk factors as a large advantage. However future studying is important. The period of three months is rather short to show some changeability.

Routine Outcome Monitoring in the Forensic Psychiatric Centre "The Kijvelanden"

Introduction

The government and health insurance companies want to evaluate the costs and benefits to gain insights whether and to what extent financial input contributes positively to treatment outcomes (Mulder, Staring Loos, Buwalda, Kuipers, Sytema & Wierdsma, 2004). Mental health institutions pursue a different objective and want to evaluate treatment progress for clinical decision. In the last decade, evaluating treatment programs and periodically measuring treatment progress is increasingly common in mental health organizations, even in the Dutch context (Mulder, Graag, Bruggeman, Delespaul, Dries, Faber, de Haan, van de Heijden, Kempen, Morgendorff, Slooff, Sytema, Wiersma, Wunderink & van Os, 2004). Normally, randomized controlled trials comprising an experimental and control group (TAU), are used to examine the effectiveness of treatment programs. However, (quasi-)experimental studies are usually cross-sectional which means that the entire period of treatment is not included. This limited design only gives professionals insight into effects at one given time (snapshot). Also it is difficult to generalize a reported effect to the entire clinical period. In order to measure the clinical effect for the entire treatment, it is necessary to evaluate everyday clinical practice. Routine Outcome Monitoring (ROM) is a methodology that is increasingly used for measuring patients' treatment progress, risks and needs periodically so that adjustments can be made if necessary (Beurs & Zitman, 2007). ROM is more and more common in mental health institutions and several criteria are established in order to apply ROM in a more standardized way (Beurs & Zitman, 2007). For example a ROM test battery has to be validated and reliable and its application should not take too long. ROM has to measure the patient's perspective and it has to be sensitive for progression or decrease (Mulder et al., 2004). Three performance outcomes (PO) are important: PO1: change in seriousness mental health PO2: changes in daily functioning, PO3: Change in quality of life (QOL).

Several questionnaires (self-report and/or practitioners report) have been developed for the benefit of ROM. For example the Camberwell Assessment of Needs Short Appraisal Schedule (CANSAS-P) has been developed in order to measure the care needs of the patient through an interview and a rating list (Busschbach, 2002). The Manchester Short Assessment of quality of life (MANSA) measures the QOL through an interview or a questionnaire (Van Nieuwenhuizen, 2000) and the Health of the Nations Outcome Scale (HoNOS) measures patient outcome on areas like behavior, impairment, symptoms and social problems (Preston, 2000) with a

practitioners report (Mulder, Staring, Loos, Buwalda, Kuijpers, Sytema & Wierdsma, 2004). The HoNOS is developed in the 1990's, commissioned by the UK government in order to apply ROM and is now widely used in several countries (Mulder et al., 2004). Salvi, Leese and Slade (2005) studied several ROM instruments, they concluded the HoNOS and CANSAS provided the most information about the individual patient and the patient's met and unmet needs. The GAF score is considered adequate to measure QOL. However Priebe, Huxley, Knight and Evans (1999) showed the MANSA also appears to be valid and reliable to measure QOL. With the use of the MANSA, the patient is allowed to determine his own experienced QOL. Several studies have shown that a self report for the benefit of the QOL is far more effective than when the therapist or practitioner determines the QOL for a patient (Douma, Kersten, Koopman, Schuurman & Hoekman, 2001). The HoNOS appears to be a rather common instrument which can be used in several mental health institutions (Salvi et al., 2005).

The majority of the ROM instruments, used in regular mental health care institutions are not suitable enough to be solely applied in forensic psychiatry. One of the primary goals of forensic psychiatry is to reduce the risk of recidivism. Obtaining insights into risk and protective factors is important to reduce the likelihood of recidivism. Before the eighties, practitioners relied mainly on unstructured clinical assessment for important clinical decisions such as probation or unconditional dismissal. The unstructured clinical judgment relies on human judgment and on clinical expertise of a mental health professional, this is a rather subjective, non-consistent method and there is no empirical support to validate this method (De Vogel, 2005). Afterwards, actuarial risk assessment instruments were developed based on empirical driven algorithmic methods. However these instruments had shortcomings because they only predict violence, but not the risk of violence and gave no insight into the dynamic mechanisms that give rise to more violence, such as a lack of empathy, hostility and lack of coping skills. Moreover, these first-generation risk assessment instruments, such as the STATIC-99, consisted only of static and irreversible factors. Dynamic factors were not included in these instruments. Therefore, the structured professional judgment has been designed. The SVR-20 and the HCR-20 are the most widely used instruments to predict future deviant (respectively sexual and aggressive) behavior. In addition to the HCR-20, the thirteen Dutch forensic psychiatric centers developed the HKT-30 risk assessment instrument (DJI, 2003). This instrument was developed in the Netherlands by and for psychiatrists and psychologists and includes historic (H), clinical (K), and future (T) risk factors. The HKT-30 has shown to be an instrument with more dynamic items than the HCR-20 (De

Beurs et al., 2010) and several studies have shown that the predictive value is rather high and even better than the HCR-20 (Spreen, Ter Horst, Bogaerts, Lammers, Hochstenbach & Soe-Agnie, 2009; Lammers, 2007). However because the HKT-30 is only scored once a year in order to estimate the likelihood of recidivism and shows a rather small change of functioning in a year (Cohen's d = .33) and a larger effect in two years (Cohen's d = .66) (Drieschner, 2010; De Jonge, Nijmans & Lammers, 2009), this instrument is in its original form not suitable for specific ROM purposes.

Because several studies have shown that the HKT-30 is dynamic (De Jonge et al., 2009), Schuringa, Spreen and Bogaerts (2010) developed a new Dutch Instrument derived from the HKT-30 for Forensic Treatment Evaluation (Instrument for Forensic-policlinic treatment evaluation: [Instrument voor Forensich-poliklinische Behandel Evaluatie, IFpBE]). This instrument could be useful for the benefit of ROM in forensic institutes with consideration for the dynamic risk factors.

Another instrument suitable for ROM is the HoNOS. The HoNOS has been extended with a forensic subscale namely the HoNOS secure. This subscale adds seven domains to the original HoNOS. These domains include, for example, the need for security regarding a relapse or the risk of harming other people (Sugarman & Walker, 2004). These domains are useful for the ROM in forensic institutes. However they do not include some very important risk factors that are supposed to be reduced in forensic psychiatry. One of these risk factors is problem understanding. This item is considered important for the psychosocial functioning of the patient, medication compliance and the understanding of the criminal pathway (DJI, 2003). Two other relevant items are impulsivity and coping which are related to violent behavior and recidivism (DJI, 2003). These two factors are effective in reducing the risk of recidivism and therefore important for treatment, it seems desirable to include these items in ROM measurements in the forensic psychiatry.

The goal of this study is to investigate if the IFpBE applies to a few of the important requirements of ROM. One important requirement is sensibility to change (Mulder et al., 2004). We want to consider if the IFpBE is sensitive to change. The IFpBE includes a practitioners report (pr) and a self report (sr) for the patient. A self report list is important to gain insight in the patient's perceptions and experiences of problem understanding and behavior and can be used during treatment (Mulder et al., 2004; Wing, Beevor, Curtis, Park, Hadden & Burns, 1998). Several studies have showed that psychotic patients show lower problem understanding than non-psychotic patients

(Matton, Wampers, De Hert & Peuskens, 2004). Raffard, Bayard, Capdevielle, Garcia, Boulenger and Gelly-Nargeot (2008) showed that 50 to 80% of the patients with schizophrenia have a limited sense of insight into disease what can negatively affect treatment compliance. Otherwise, a lack of insight into disease can also cause more psychotic problems (Bell, Fiszdon, Richardson, Lysaker & Bryson, 2007). Therefore this study also considers if psychotic patients show a higher discordance with their caregivers than non-psychotic patients. Because the IFpBE is scored by several practitioners, the inter rater reliability will also be studied. Several studies have shown that the interrater reliability (.77) and the predictive value (AUC = .72) of the original HKT-30 are good (Hildebrand, Hesper, Spreen & Nijman, 2005).

Method

Participants

Eighty-seven male patients from the Forensic Psychiatric Centre de Kijvelanden have been asked to participate in our study. Thirty eight patients agreed and eventually 34 patients filled in the questionnaire at T0, 5 patients refused to rate the questionnaire at T1. The participants are spread over seven wards within the forensic institute. Three wards include patients with a mental illness who require more structure. Three wards include patients with a personality disorder who are required to take more responsibilities over their daily functioning. The last ward concerns the rehabilitation department where patients have more freedom and responsibilities in order to rehabilitate. On three wards, patients have been personally approached; on the other wards patients have been approached during group-meetings. The population in this study is very heterogeneous, such as diversity in personality disorders, psychiatric disorders, years of followed therapy etc. The mean age of the studied population is 38.74 (SD=9.3). In this study a distinction will be made between patients who suffer a psychotic disorder and patients who do not suffer a psychotic disorder. For each participating patient, four disciplines filled in the questionnaire, the head of treatment and risk management (HrB), the psychologist connected to the ward, the personal coach of the patient and the art- or psychomotor therapist (PMT). In one ward the HrB was not available in the period of this study, therefore the ward psychologist scored in replacement of the HrB and the coordinator of the ward's social therapists scored in replacement of the psychologist. For one ward the

psychiatrist in training filled in the questionnaire, no art- or psychomotor therapist was connected to this ward. For T0, 4 practitioner reports have not been returned, for T1 3 practitioner reports have not been returned.

Questionnaire

The Instrument for Forensic Policlinic Treatment Evaluation (IFpBE) is developed by Schuringa, Spreen and Bogaerts (2010: practitioners report) and Schuringa, Bogaerts and Spreen (2011: self report). Both instruments have several advantages compared to some other ROM instruments mentioned above. De IFpBE is based on the HKT-30 and contains most observed risk factors in forensic psychiatry related to recidivism such as hostility, impulsivity and coping skills. After a short introduction, the IFpBE can be filled in by several disciplines and is not restricted to psychologists and psychiatrists. This is important for clinical perspectives because a (forensic) psychiatric patient is more than the sum of psychological and psychiatric characteristics. For example, the social context and social skills are important features and add to the overall picture of the patient (Preston, 2000). The use of a practitioners (pr) and self report (sr) has another advantage, both patient and caregiver are involved in the treatment program and evaluation. Both perspectives can improve the caregiver-patient communication which can lead to a higher consensus considering the treatment plan (Van Os & Trifaux, 2008). The IFpBE is supposed to be rated periodically in order to evaluate treatment. Both self report and practitioner report consist of 20 general items focused on general dynamic risk and protective factors, such as working skills, degree of manipulative behavior and the extent to which the patient takes responsibility for his offence. In addition to the general items, three subscales are developed. The first subscale considers substance abuse, the second physical aggressive behavior and the third includes sexually transgressive behavior. All subscales include the question whether the patient acknowledges the problem, whether the patient is motivated to handle the problem and to which extend the patient shows skills to handle the problem. The subscale substance abuse includes an extra item which considers possible recent use of substances; the subscale sexual transgressive behavior includes the question if the patient recognizes implicit thought and cognition in replacement of the item motivation. The factors in these subscales are important risk factors however they are not applicable for every patient. In total the subscales include ten items so the entire questionnaire consists of 30 items. All items are scaled on a five point rating scale. Between every two possible scores, there are three sub choices, which give the items a 17 point rating scale. All items have the extra answering possibility; not enough information (N.E.I.).

This is applicable when a caregiver has no information on the factor. For example not every caregiver has knowledge of the patients' finances. Some items also contain the extra answering possibility 'not applicable' (N.A.). Some questions are not applicable for all patients, such as psychotic symptoms. A high score on an item means an adequate way of functioning on the item. On the items 10, 11, 12, 13, 14, 16, 18 and 24 a low score meant an adequate way of functioning, the scores on these items have been recoded into opposite scores. The high rating scale makes the instrument more sensitive so that treatment progress can be easily observed. In this study the period between the two observed measurements is three months. T0 took place in January and February, T1 took place in May and June.

Statistical analyses

The scores N.E.I. and N.A. were recoded into missing values. Firstly, intraclass correlation coefficients were calculated six times, because the questionnaire was answered by multiple raters. A correlation of .75 and higher is judged very well, values between 0.60-0.75 are considered good, values between 0.40-0.60 are considered reasonable and values of 0.40 and lower are considered moderate or insufficient. Secondly, to consider whether psychotic patients scored in higher discordance with their caregivers than non-psychotic patients, an independent sample T-test was used. Because most psychotic patients use medication, their sickness and problem understanding should be increased. Therefore, an ANOVA has been calculated in order to compare patients who showed no psychotic symptoms with patients who did show psychotic symptoms and patients who suffered a psychotic episode in the last year. A paired sample T-test has been computed in order to determine whether the IFpBE shows a change of functioning on the items between T0 and T1 (a period of three months). Results have been entered in SPSS 18.0 for analysis.

Results

Characteristics of the rated patients

Table 1: Psychopathology, offenses and treatment period of the observed patients

Characteristics	*n*	Percentage
Psychopathic patients		
Psychopathy	3	9 %
Signs of psychopathy	10	29%
No psychopathy	21	62%
Psychotic patients		
Psychotic	17	50%
Non psychotic	17	50%
Psychotic symptoms in the past year		
No symptoms	19	56%
Symptoms	11	32%
Psychotic episode	4	12%
Offence		
Violent property offence	5	15%
Assault	5	15%
Homicide	12	35%
Sexual crime	8	23%
Arsoning	3	9%
Restraint	1	3%
Personality disorder		
No signs	8	24%
Pathological features	3	9%
1< PSD no cluster B*	13	38%
1< PSD incl. Cluster B*	10	29%
Years of institutional treatment		
1	2	6%
2	10	29%
3	9	27%
4	5	15%
5	4	12%
6	2	6%
9	1	3%
13	1	3%

**Criteria from historic risk factors of the HKT-30*

As shown in table 1, the patient group is heterogeneous. Fifty percent of all patients have a psychotic disorder and 24 percent shows no sign of a personality disorder, 38 percent has one or more personality disorder but no cluster B personality disorder and 29 percent of the patients have one or more personality disorder with a cluster B personality disorder. These numbers show diversity in combinations of personality and psychiatric disorders. Patients also differ in types of crimes and years of hospitalization. The largest group of crime is homicide with 35%, followed by sexual crime (23%). Arsoning and restraint contain the smallest groups of patients (9% and

3%). Thirty-eight percent of the patients remain longer than three years in FPC de Kijvelanden, most patients remain two or three years within the forensic institute (29% and 27%). However not shown in table 1, it is important to consider that most patients have also stayed in prison before entering the current institution and some patients have stayed in more than one forensic psychiatry centre.

Intraclass correlation coefficients

Table two shows the intraclass correlation coefficients (ICC) for all separate items. The table also shows the rating frequencies per item. For each patient, not all items were scored by each discipline because of the answering possibilities N.A. and N.E.I. The largest differences have been found for the items sexually transgressive behavior, manipulative behavior, antisocial or criminal subcultures and psychotic symptoms, apart from the subscales which have not been answered consistently. The ICC of most scored items can be valued as reasonable to very good. The ICC of the items daily schedule, working skills, impulsivity, rule compliance and the patient's recognition of his own implicit thought and cognitions are very good. The items, treatment cooperation, independent living skills, financial/administrative skills, anti social behavior, manipulative behavior adequate social network and motivation for a problem with physical aggression show a good ICC. The items sexually transgressive behavior, acknowledge a problem with substances, motivated for a problem with substance abuse, recent use of substances and skills to prevent physical aggression, show an insufficient ICC. All other items show a reasonable ICC. For some items, that do show a good ICC, the range is very large. For example item 5 shows an ICC of .814, but the lowest ICC for this item is .589 which is considered reasonable. The item acceptable partner relationship has only been answered once; therefore it was not possible to compute the ICC for this item.

Table 2. Rating frequencies of the items and Intraclass correlation coefficients

Item	HrB	N	Psy	N	Coach	N	Art/pmt	N	Range ICC	ICC
1. Problem awareness		33		32		32		29	.415 - .982	.553
2. Treatment cooperation		33		32		32		31	.400 - .918	.683
3. responsibility offence		33		32		31		27	.662 - .919	.517
4. Coping skills		32		32		31		31	-.192 - .995	.478
5. Daily schedule		32		31		32		26	.589 - .993	.814
6. working skills		21		25		27		15	.540 - .975	.764
7. Social skills		33		32		32		30	.368 - .589	.471
8. Independent living skills		31		32		32		28	.160 - .869	.602
9. Financial/administrative skills		26		17		32		9	.133 - .988	.632
10. Impulsivity		33		32		32		31	.658 - .992	.860
11. Anti social behavior		33		32		32		31	.490 - .917	.646
12. Hostility		33		30		32		31	.107 - .722	.453
13. Sexually transgressive behavior		32		31		31		27	-.200 - .774	.144
14. Manipulative behavior		31		32		32		26	.221 - .941	.688
15. Rule compliance		33		32		32		30	.269 - .971	.751
16. Antisocial or criminal subcultures		32		25		29		25	.228 - .914	.474
17. Medication compliance		29		23		28		26	.131 - .645	.478
18. Psychotic symptoms		28		31		26		19	.160 - .806	.582
19. Acceptable partner relationship		1		0		2		1	/	/
20. Adequate social network		24		25		27		12	.357 - .835	.672
21. Acknowledges SUD		24		24		22		16	.217 - .507	.263
22. motivated		23		25		21		16	.437 - .837	.399
23. Skills to prevent substance use		23		25		17		14	.287 - .851	.464
24. Recent use.		27		31		32		25	.123 - .969	.355
25. acknowledge problem physical aggression		16		21		17		14	.288 - .950	.498
26. Motivated		17		22		15		16	.476 - .907	.735
27. Skills to prevent physical aggression		17		23		15		17	-.541 - .321	-.077
28. Acknowledges problem sexual behavior		8		8		9		6	.159 - .818	.444
29. recognizes implicit theories		6		6		9		5	.779 - .986	.770
30. Skills to prevent behavior		6		7		10		3	/	/

Differences between psychotic and non-psychotic patients

Table 3 shows the mean differences between psychotic and non psychotic patients. A significant more negative deviance of psychotic patients stands for a more positive rating of the patient himself compared to non psychotic patients. A significant higher positive score stands for a more negative rating of the patient himself. Psychotic patients differ on a small sample of items, as can be seen (Table 3). Remarkably, most psychotic patients rate more psychotic symptoms (item 18) than their caregivers. For other items, psychotic patients differ mostly from their coach, with whom they spend most hours, on the items independent living skills, antisocial or criminal subcultures, psychotic symptoms, acknowledgment of a problem with substance abuse, motivation for the problem with substance abuse and skills to prevent substance abuse. Psychotic patients differ least from the art or psychomotor therapist on two items, psychotic symptoms and skills to prevent physical aggression. Psychotic patients differ on the items working skills, acknowledges problem with physical aggression and skills to prevent

physical aggression from the head of treatment and on the items social skills, psychotic symptoms and acknowledges a problem with substance abuse, psychotic patients differ more from their psychologist, compared to non psychotic patients. Also remarkably is that several deviations from the caregivers are the result of a more negative rating of the patient compared to the caregiver. This accounts as mentioned before for the psychotic symptoms, were psychotic patients report more psychotic symptoms than their caregivers do, this also accounts for the item antisocial or criminal subcultures. For some significant items, there is a difference in deviance, however for these items, the non-psychotic patients show a more positive rating. This considers the items independent living skills and social skills.

Table 3: *Differences in deviance between psychotic and non psychotic patients*

Item	HrB			Psychologist			Coach			Art- or psychomotor-therapist		
	Non vulnerable - vulnerable			Non vulnerable - vulnerable			Non vulnerable - vulnerable			Non vulnerable - vulnerable		
	Mean deviation	Mean diff.	t-value	Mean deviation	Mean diff.	t-value	Mean deviation	Mean diff.	t-value	Mean deviation	Mean diff.	t-value
1. Problem awareness	-4.13 - -3.00	-1.13	-.70	-4.33 - -3.44	-0.90	-.58	-3.73 - -2.81	-0.92	-.70	-3.54 - -3.75	0.21	.14
2. Treatment coöperation	-5.56 - -1.06	4.50	-2.65	-3.87 - -3.94	0.07	.05	-4.00 - -2.31	-1.69	-1.20	-2.14 - -3.38	1.23	.15
3. responsibility offence	-4.38 - -1.00	3.38	-2.24	-3.93 - -4.31	0.38	.26	-4.81 - -4.73	-0.08	-.05	-2.21 - -3.15	0.94	.83
4. Coping skills	-5.56 - -3.69	-1.88	-1.36	-3.38 - -3.31	-0.06	-.05	-4.81 - 3.80	-1.01	-.66	-2.33 - -4.25	1.92	.59
5. Daily schedule	-3.56 - -2.00	-1.56	-1.18	-2.06 - -3.40	1.34	1.05	-3.44 - -2.44	-1.00	-.70	-4.00 - -3.00	-1.00	1.07
6. working skills	-.50 - -3.50	3.00 **	1.16	.77 - -.73	1.50	.69	-.81 - .063	-0.88	-.42	1.13 - -7.27	8.40	-.67
7. Social skills	-3.77 - .94	-4.70	-3.17	-3.31 - -1.75	-1.56 **	1.35	-3.63 - .94	-2.69	-2.09	-2.71 - -2.63	-0.09	2.54
8. Independent living skills	-1.73 - ,13	-1.86	-1.63	-2.38 - -1.56	-0.81	-.510	-2.38 - -.50	-1.88 *	-1.87	-1.77 - -2.60	0.54	-.06
9. Financial/ administrative skills	-1.27 - .18	-1.45	-.91	-.38 - -2.11	1.74	.601	-.81 - -3.13	2.31	1.69	-1.50 - -1.20	-0.30	.64
10. Impulsivity	-3.50 - 1.88	-5.38	-2.44	-.80 - .06	-0.86	-.40	-2.50 - -.50	-2.00	-1.03	-2.29 - 1.25	-3.54	-.18
11. Anti social behavior	-1.69 - .94	-2.63	-1.50	-.06 - -.75	0.69	.41	-1.27 - -1.25	-0.02	-.01	-.21 - -.19	-0.03	-.02
12. Hostillity	.35 - 1.75	-1.40	-.91	-.27 - -.87	0.60	.31	1.13 - -.38	1.50	.88	1.13 - .13	1.01	.51
13. Sexually transgressive behavior	-.56 - 1.20	-1.76	-2.80	-.19 - 1.07	-1.26	-2.30	-.63 - -.64	-1.27	.65	-1.00 - .67	-1.67	-2.18
14. Manipulative behavior	-3.13 - -.73	-2.40	-1.62	-2.19 - -2.06	-0.13	-.10	-3.60 - -1.94	-1.66	-1.12	-2.85 - -1.50	-1.35	-.91
15. Rule compliance	.24 - .38	-0.14	-.12	.56 - .00	0.56	.39	-.56 - -.06	-0.50	-.33	.00 - -.07	0.07	.05
16. Antisocial or criminal subcultures	-2.63 - 2.27	-4.89	-2.90	-2.00 - -1.00	-1.00	-.43	-1.27 - 2.36	-3.63 **	-2.30	-5.42 - 1.25	-6.67	-2.93
17. Medication compliance	-2.06 - .19	-2.25	-.92	-1.00 - -1.23	0.23	.17	-.69 - -.81	0.13	.06	-10.33 - .43	-10.76	-1.73
18. Psychotic symptoms	-.33 - 1.50	-1.83	-1.16	.79 - 2.56	-1.78 **	-1.10	.90 - 3.06	-2.16 *	-1.42	-.33 - 3.83	-4.17 **	-2.66
19. Acceptable partner relationship	1.76 - .56	1.20	.67	/	-1.00	/	1.27 - .64	0.62	.34	2.63 - 1.00	1.63	1.03
20. Adequat social network	-3.64 - -.54	-3.10	-1.17	-2.92 - -1.92	-1.00	-.59	-1.67 - -3.21	1.55	.93	-1.83 - -4.17	2.33	.22
21. Acknowledge a problem with substances	-1.94 - -2.50	0.56	.28	-2.40 - -4.08	1.68 *	.77	-.87 - -4.06	3.20 *	1.79	-2.08 - -1.33	-0.74	-.34
22. motivated	-3.63 - -2.06	-1.56	-.86	-5.64 - -1.83	-3.80	-1.81	-2.33 - -2.67	0.33 *	.16	-2.39 - -2.17	-0.22	-.10
23. Skills to prevent substance use	-3.75 - -2.33	-1.42	-.77	-5.10 - -.75	-4.35	-2.41	-3.13 - -6.90	3.77 **	1.76	-4.14 - -3.67	-0.48	-.23
24. Recent use.	.31 - .62	-0.31	-.44	-.57 - .69	-1.26	-1.58	-.50 - .56	-1.06	-1.60	.308 - -.18	0.49	1.42
25. acknowledge problem physical agression	.47 - 4.14	-3.67 *	-1.38	-.63 - .91	-1.53	-.49	.27 - -2.50	2.78	1.57	.08 - 1.73	-1.65	-.63
26. Motivated	-2.77 - -1.20	-1.57	-.60	-5.29 - -1.00	-4.29	-1.79	-2.13 - -2.92	0.78	.27	-3.07 - -2.58	-0.49	-.15
27. Skills to prevent physical agression	-1.44 - -2.23	0.79 *	.27	-2.71 - -.50	-2.21	-1.15	-1.73 - -1.50	-0.23	-.08	-1.08 - -2.50	1.42 **	.428
28. Acknowledges problem sexual behavior	.41 - .930	-0.52	-.32	-2.50 - -6.00	3.50	.63	-.44 - -1.50	1.06	.53	5.78 - 5.65	-1.38	-.59
29. recognizes implicit theories	-.47 - .77	-1.24	-.67	-2.0 - -11.0	9.00	1.97	-.73 - -.62	-0.12	-.05	8.06 - 5.41	-4.13	-1.43
30. Skills to prevent behavior	1.50	2.88	-1.31	-3.75 - -1.00	-2.75	-.99	-1.69 - .62	-2.30	-1.01	7.60 - 5.34	-2.90	-.98

*sig<,05. **sig<.01.

Differences between patients with or without a psychotic episode or symptoms in the past year.

Because not all psychotic patients have shown psychotic symptoms in the past year, it is possible that these patients report no symptoms and their problem and sickness insight could be improved because of their medication. Therefore three groups were compared; patients with psychotic symptoms in the last year, patients with a psychotic episode in the past year and patients who did not suffer from either of these in the past year. An Anova has been conducted in order to compare these three groups. No more significant differences came forward compared to the independent sample T-test. However, it is rather notable that the patients who suffered a psychotic episode in the past year rate significant (sig <.05, F_{Hrb} (28)= 7.32, M = 3.00, $F_{Psych(27)}$ = 9.60, M =3.33, $F_{Coach}(27)$= 3.76, M = 2.00, $F_{art\ or\ PMT(23)}$ = 8.90, M = 3.33) less sexual transgressive behavior than their caregivers do. This difference is large for this patient group compared to the other patients. It is important to take into account that there are only three patients who suffered a psychotic episode in the last year.

Changes in risk and need factors over time: repeated measures

Differences were examined between the two measurement points (see Table 4). The head of treatment reports significant progress on one item between the first and second assessment, namely acknowledgement of a problem with substance abuse. The psychologist reports significant progress on treatment cooperation and acknowledgement of a problem with substance abuse. For the coach, four items show significant progress: responsibility for the offence, independent living skills, financial and administrative skills and anti social behavior towards the caregiver. The items rated by the creative therapist show progress on problem awareness, impulsivity and motivation for the problem with substance abuse. The item medication compliance shows a decrease. The patient scores show progress on financial and administrative skills and hostility. We find little agreement between the practitioner and patient scores which is certainly no problem because the self-perception of the patient differs from the professional clinical judgment and any form of reporting is associated with response bias.

Table 4. Differences between T0 and T1.

Item	HrB	t-value	Psychologist	t-value	Coach	t-value	Art or PMT	t-value	Self report	t-value
1. Problem awareness	10.24 - 10.36	-.34	9.87 - 9.13	1.09	10.48 - 11.11	-.86	10.44 - 9.22 *	2.14	13.89 - 14.33	-.66
2. Treatment cooperation	10.48 - 10.58	-.22	9.53 - 10.73 *	-2.58	10.39 - 10.43	-.05	10.87 - 10.33	.89	13.57 - 12.89	1.58
3. responsibility offence	11.28 - 11.44	-.33	9.93 - 11.00	-1.96	9.46 - 11.08 *	-2.45	12.10 - 11.70	.73	13.36 - 13.57	-.44
4. Coping skills	8.50 - 8.81	-.94	9.13 - 9.03	.28	8.93 - 9.70	-1.39	9.59 - 9.34	.49	12.79 - 12.90	-.12
5. Daily schedule	10.22 - 10.97	-1.65	10.24 - 10.24	.00	9.93 - 10.79	-1.12	10.22 - 10.81	-.59	12.57 - 12.54	.07
6. working skills	11.65 - 11.00	1.21	11.32 - 11.64	-.52	10.32 - 10.84	-.42	10.17 - 10.17	.00	11.61 - 11.52	.08
7. Social skills	10.76 - 10.61	.61	9.83 - 9.57	.50	9.96 - 10.46	-.73	9.10 - 9.66	-1.04	12.14 - 12.90	-1.47
8. Independent living skills	12.65 - 12.65	.00	11.83 - 11.40	.91	12.07 - 13.29 *	-2.08	11.27 - 11.58	-.64	12.97 - 14.24	-2.64
9. Financial/administrative skills	12.80 - 12.80	.00	10.17 - 10.50	-.44	10.46 - 12.11 *	-2.95	12.00 - 11.13	.96	12.41 - 13.66 **	-2.11
10. Impulsivity	10.27 - 10.49	-.415	10.97 - 10.87	.21	8.68 - 9.89	-1.50	10.08 - 11.96 **	-2.84	11.86 - 10.86	1.21
11. Anti social behavior	11.39 - 11.88	-.919	11.53 - 11.27	.51	10.43 - 12.29 *	-3.15	11.93 - 12.03	-.25	11.82 - 12.46	-.84
12. Hostility	12.42 - 11.73	1.52	11.24 - 11.03	.29	11.36 - 12.29	-1.60	12.67 - 12.96	-.63	10.81 - 13.48 **	-3.15
13. Sexually transgressive behavior	15.56 - 15.66	-.55	15.69 - 15.59	1.00	15.22 - 14.89	.89	15.00 - 14.70	.40	15.21 - 15.43	-.65
14. Manipulative behavior	11.68 - 11.90	-.66	11.37 - 11.23	.260	10.32 - 11.11	-1.50	11.72 - 11.48	.39	13.93 - 13.56	1.13
15. Rule compliance	12.79 - 13.18	-1.43	12.60 - 12.10	1.03	11.96 - 11.82	.34	12.75 - 12.54	.53	12.00 - 11.86	.170
16. Antisocial or criminal subcultures	13.81 - 14.44	-1.61	13.48 - 13.96	-.59	14.24 - 14.4	.31	11.17 - 11.94	-1.25	14.14 - 14.00	.18
17. Medication compliance	11.71 - 12.08	-.47	13.11 - 12.33	.79	11.80 - 11.80	.00	14.00 - 12.86 *	2.49	13.25 - 12.75	.41
18. Psychotic symptoms	12.17 - 12.56	-.40	15.08 - 15.00	.19	14.16 - 13.47	1.01	13.42 - 12.92	.59	12.76 - 12.48	.48
19. Acceptable partner relationship		/	/		7.50 - 10.00	-5.00	/		7.25 - 11.00	-1.54
20. Adequate social network	11.21 - 11.79	-.94	10.52 - 9.87	.93	10.59 - 10.36	.19	11.64 - 11.73	-.13	13.70 - 14.81	-1.28
21. Acknowledge a problem with substances	9.29 - 10.67 *	-2.18	8.71 - 10.38 *	-2.20	9.71 - 8.41	1.45	11.00 - 12.30	-1.30	12.24 - 11.35	1.12
22. motivated	9.64 - 11.14	-1.73	8.50 - 8.86	-.56	9.94 - 10.19	-.21	10.45 - 11.64 **	-3.14	12.22 - 12.28	-.07
23. Skills to prevent substance use	11.22 - 11.48	-.53	10.85 - 10.45	.57	9.93 - 10.27	-.34	10.13 - 11.13	-1.53	13.47 - 13.71	-.24
24. Recent use.	15.81 - 15.44	.82	15.38 - 15.08	.44	15.30 - 15.26	.09	15.81 - 15.94	-1.00	15.10 - 15.35	-.44
25. acknowledge problem physical aggression	7.33 - 7.47	-.17	8.17 - 9.22	-1.01	5.45 - 7.27	-1.37	8.11 - 9.44	-1.54	5.80 - 7.07	-.88
26. Motivated	8.50 - 7.69	1.03	10.79 - 10.84	-.08	8.55 - 9.18	-1.41	10.00 - 9.82	.30	12.19 - 11.94	.18
27. Skills to prevent physical aggression	10.50 - 10.75	-.40	12.21 - 12.11	.24	12.10 - 10.10	1.79	9.62 - 10.00	-.39	13.43 - 13.79	-.47
28. Acknowledges problem sexual behavior	11.25 - 10.63	.76	6.25 - 7.50	-.63	6.33 - 7.67	-.61	9.00 - 7.00	1.00	9.43 - 8.71	.22
29. recognizes implicit theories	10.00 - 8.33	1.27	2.50 - 2.00	.33	8.67 - 8.33	.38	/	/	8.63 - 12.00	-1.04
30. Skills to prevent behavior	9.80 - 10.80	-.79	7.00 - 5.75	1.32	8.00 - 8.67	-1.00	15.00 - 27.00	/	8.50 - 10.89	1.00

*sig<,05 **sig<.01

Discussion

Most items show a reasonable through very good inter rater reliability. However, several items show a large variance in inter rater reliability between caregivers. This variance is largest for the items independent living skills, financial and administrative skills, anti social behavior, manipulative behavior and rule compliance. Several of these variances can be explained by the specification in professional knowledge of the different caregivers. For example, a psychologist or head of treatment has less sight on the financial and administrative skills compared to the personal coach of the patient. Also the coach has more sight on the independent living skills. Schuringa (2010) reports in his study that several behaviors are differently shown on different departments, for example antisocial behavior is presumably more shown in sight of the coaches than when the patient has a meeting with his head of treatment or psychologist. Also sexual transgressive behavior is not always shown by a patient, for

example a creative therapist could observe this behavior less than the coach of the patient, or the psychologist. This variance in inter rater reliability could have a function in a team. When the patient shows different behavior on different departments, this could come forward by the rating on the IFpBE. When the team discusses these different ratings, a consensus could be reached over the patient's behavior and the entire team would have knowledge of these factors (Schuringa, 2010). However, a good inter rater reliability is important, therefore it is important to consider a good introduction or a training to enlarge the understanding of the items, which could enlarge the inter rater reliability. Several items show an insufficient inter rater reliability. Besides the formerly named item sexual transgressive behavior, this insufficient inter rater reliability concerns several items in the subscales. The subscales have not been answered consistently; therefore it is difficult to make conclusions concerning these items. Over all, the IFpBE shows a reasonable inter rater reliability, however in future research, a training or introduction is desirable, to possibly enlarge the found inter rater reliability.

It is difficult to say if the IFpBE displays the patients' perspective. Very few significant effects have been found to underpin this requirement. One significant effect is the difference between reported psychotic symptoms by psychotic patients and their caregivers what means that psychotic patients report more symptoms than their caregivers are aware of. Bank, Delespaul, Hanssen, De Graaf, Vollebergh and Van Os (2004), studied the importance of a self report regarding psychotic symptoms. Eighty patients reported at least one psychotic symptom in this study whereas only 19% of the professionals reported psychotic symptoms which clearly indicate under-reporting and probably a lack of observation by the professional for psychotic symptoms. These unrecognized psychotic symptoms appear to be an important predictor for a future psychotic episode. Several studies have shown that problem understanding decreases when psychotic symptoms such as delusions and hallucinations increase (Matton et al., 2004). Problem and sickness insight can be enlarged by reducing the positive symptoms like delusions and hallucinations (Nijman, van Marle & Kavelaars, 2006). Another issue that can explain the small differences between psychotic and non psychotic patients is the heterogeneity of the studied group. Recently, Bogaerts and Spreen (2011) found three subgroups within an inpatient forensic group of 232 psychotic patients. The first group consists of forensic patients who committed a very serious life offense. However, in this group, hostility, impulsivity, lack of empathy and personality disorders (cluster B) is absent. This group is similar to the group of psychotics we also find in the mainstream health care. The other two subgroups

are very well characterized by impulsivity, hostility, and cluster B personality disorders. The difference between both subgroups relates to the past. The patients in one subgroup suffer from conduct disorders and family abuse and neglect in the past which is not present in the other subgroup. This study also found three subgroups within an inpatient forensic group of 348 patients with a personality disorder. These groups also differ in several risk factors, however they also show a low problem understanding. Therefore, these patients could also differ from their caregivers, which would explain the small differences between psychotic and non psychotic patients.

The results show only small significant changes between T0 and T1 which can be influenced by the small number of participants and the heterogeneity of the studied group. Patients differ in psychopathology, risk factors, type of offense, duration of treatment and treatment program what can influence the outcome of some measured items. Because of the small sample size, it is also difficult to see a global progress on all items. Further, an observation period of three months is a rather short period to show progress in the long term mental health care. The ROM expert group Forensic Psychiatry (2011) claims that a yearly measurement is required. Although the results show limited changes between T0 and T1 it is important to consider that the different disciplines report a change on different factors. For example the coach reports progress for anti social behavior, independent living skills and financial/administrative skills. These are factors which are mostly displayed in the sight of the coach, other disciplines have less sight on these factors, especially independent living skills and financial and administrative skills. These differences in significant progress show the importance of the rating of the IFpBE by a multidisciplinary team. When all disciplines rate the IFpBE, these different (risk)factors could come forward.

This instrument shows possibilities for the ROM of risk factors in forensic psychiatry, for a complete measurement of the ROM as required, the IFpBE could be supplemented with the HoNOS plus the HoNOS secure and the MANSA. With these instruments, the risk factors, symptom severity, daily functioning and quality of life could be measured. The advantage of the IFpBE is the multidisciplinary rating, this means that the risk factors and several daily functioning- and symptom-severity factors are fully exposed. When the differences in rating are discussed in a team, a consensus could be reached so that the entire team has knowledge on all these important risk factors. If necessary these instruments could be extended with the PANNS (positive and negative syndrome scale) for psychotic vulnerable patients. In addition to a yearly measure, the HKT-30 and if indicated the SVR-20 and HCR-

20 could be included in ROM. However it is of importance to study the validity of the IFpBE with acknowledgment of the limits in this study.

Several limits have come forward in this study. The number of participants in this study is very small. This study also includes a very heterogeneous group which makes it more difficult to find corresponding results and to interpret results correctly. The professionals were given two weeks to fill in the questionnaire. However, most professionals needed more time. Most professionals reported some difficulties rating the questionnaire. The IFpBE contains some difficulties and ambiguities that needed to be clarified. Some items were formulated with a double question. For example the highest score on item 20 indicates that the patient's social network is good without any conflict. A low score indicates more conflict. However a conflict-free social network does not automatically mean a good supportive social network for the patient. Therefore this item was regarded as difficult to answer. Another difficulty for the professional is the large rating scale. The questionnaire consists of a seventeen-point rating scale while it seems difficult in practice to distinguish between more than seven categories (Drenth & Sijtsma, 2006). This could explain the rather notable result that the art or psychomotor therapist reports a decrease in medication compliance, while the art of psychomotor therapist has less sight on this factor than other disciplines.

In future research, it is important to consider the studied limitations concerning the research design, questionnaire and observation time. An observation period of three months is too short to expect behavioral changes among forensic psychiatric patients. It is preferably to conduct a design which includes a longer observation period (every half year) to examine the changeability of the IFpBE. To ensure the inter rater reliability of the instrument, a standardized introduction and formulation of the items is required. Also with a large sample, it is possible to define subgroups based on several patient characteristics.

References

Bell, M. Fiszdon, J., Richardson, R., Lysaker, P. & Bryson, G. (2007) Are self-reports valid for schizophrenia patients with poor insight? Relationship of unawareness of illness to psychological self-report instruments. *Psychiatry Research,* 1, 37-46.

De Beurs, E. & Zitman, F.G. (2007). Routine Outcome Monitoring het meten van therapie-effect in de klinische praktijk met webbased software. *MGv,* 62, 13-28

Blok, G.T., De Beurs, E., De Ranitz, A.G.S. & Rinne, T. (2010). Psychometrische stand van zaken van risicotaxatie-instrumenten voor volwassenen in Nederland. *Tijdschrift voor Psychiatrie,* 5, 331-341.

Bogaerts, S. (2011). Dynamische profielen bij psychotische patiënten. Intern werkdocument FPC De Kijvelanden.

Busschbach, J., van. (2002). CANSAS handleiding.

Douma, J.C.H., M.C.O. Kersten, H.M. Koopman, M.I.M. Schuurman en J. Hoekman. (2001) Het 'meten' van kwaliteit van bestaan van mensen met een verstandelijke handicap; een overzicht van de Nederlandse instrumenten, *Nederlands Tijdschrift voor de Zorg aan Verstandelijk Gehandicapten*, 1, 17-36.

Drenth, P.J.D. & Sijtsma, K. (2006). Testtheorie, inleiding in de theorie van de psychologische test en zijn toepassingen. Bohn Stafleu Van Loghum, Houten.

Drieschner, K. & Hesper, B. (2008). DROS dynamic risk outcome scales. Trajectum Hoeve Boschoord, Bosschoord.

Dienst Justitiele Inrichtingen. (2003). HKT-30 Handleiding.

Expergroep Forensische Psychiatrie (2011). ROM forensische psychiatrie adviesdocument. GGZ Nederland.

Expertgroep Forensische Psychiatrie (2011). Bijlage: Checklist ROM-instrementarium voor de doelgroep Volwassenen EPA (ernstig psychiatrische stoornissen). GGZ Nederland.

Hildebrand, M., Hesper, B.L., Spreen, M. & Nijman, H.L.I. (2005). De waarde van gestructureerde risicotaxatie en van de diagnose psychopathie. *Expertise Centrum Forensische Psychiatrie.* Amsterdam: Spinhex en Industrie.

Jonge, E. de, Nijman, H.L.I. & Lammers, S.M.M. (2009). Gedragsverandering tijdens tbs-behandeling: een multicenteronderzoek. *Tijdschrift voor Psychiatrie,* 4, 205-215.

Lammers, S. (2007). ' Blijven stoute jongens stout? Taxatie van het recidiverisico in de tbs', *De Psycholoog,* 4, 194-201.

Matton, C., Wampers, M., de Hert, M. & Peuskens, J. (2004) Ziekte-inzicht bij patiënten met schizofrenie. *Tijdschrift voor de Psychiatrie,* 9, 589 - 599

Mulder, C.L., Staring, A.B.P., Loos, J., Buwalda, V.J.A., Kuijpers, D., Sytetema, S. & Wierdsma, A.I. (2004). De Health of the Nations Outcome Scales (HoNOS) in Nederlandse bewerking. *Onderzoek Centrum GGZ Rijnmond & GGZ Groep Europoort.*

Mulder, C.L., Staring, A.B.P., Loos, J., Buwalda, V.J.A., Kuijpers, D., Sytema, S. & Wierdsma, A.I. (2004). De Health of the Nations Outcome Scales (HoNOS) als instrument voor 'routine outcome assessment'. *Tijdschrift voor de psychiatrie,* 5, 273-284.

Nijman, H., van Mjarle, H. & Kavelaars. (2006). Psychotische patienten in de TBS: achtergronden, delictgedrag en behandeling. *Pedagogiek.* 3. 291-300.

Van Nieuwenhuizen, Ch., Schene, A.H. & Koeter, M.W.J. (2000). Manchester verkorte Kwaliteit van Leven meting. Eindhoven.

Van Os, J. & Triffaux, J-M. (2008). Evidence that the two-way communication checklist identifies patient-doctor needs discordance resulting in better 6-month outcome . *Acta Psychiatrica Scandinavica,* 118, 322-326.

Preston, N.J. (2000). The health of the nations outcome scales: validating factorial structure and invariance across two health services. *Australian and New-Zealand Journal of Psychiatry,* 34,512-519.

Priebe, S., Huxley, P., Knight, S & Evans, S. (1999). Application and results of the Manchester short assessment of quality of life (MANSA). *International Journal of Psychiatry,* 1, 7-12.

Raffard, S., Bayard, S., Capdevielle, P., Garcia, F., Boulengere, J.P. & Gelly-Nargeot, M.C. (2008). Lack of insight in schizophrenia: a review. Part I: theoretical concept, clinical aspects and Amador's model. *L'encephale,* 6, 597-605.

Salvi, G., Leese, M. & Slade, M. (2005) Routine use of mental health outcome assessments: choosing the measure. *The Britisch Journal of Psychiatry,* 186, 146-152.

Schuringa, E. (2010). Routine Outcome Monitoring in het FPC dr. S. van Mesdag. *GGzet Wetenschappelijk,* 14, 27-35.

Spreen, M., Horst, P.R.M. ter, Bogaerts, S., Lammers, S.M.M., Hochstenbach, J & Soe-Agnie, S.E. (2009). Risicotaxatie aan de hand van de HCR-20 en de HKT-30: een vergelijking tussen beide insteumenten. In: Oei, T.I. & Groenhuysen, M.S. (Red.). *Forensische Psychiatrie en haar grensgebieden. Actualiteit, geschiedenis en toekomst.*