Guideline concordance and outcome in long-term naturalistic treatment of bipolar disorder-a one-year longitudinal study using latent change models

Renes, Joannes W.; Maciejewski, Dominique F.; Regeer, Eline J.; Hoogendoorn, Adriaan W.; Nolen, Willem A.; Kupka, Ralph W.

Published in:
Journal of Affective Disorders

DOI:
10.1016/j.jad.2020.12.106

Publication date:
2021

Journal of Affective Disorders 283 (2021) 395–401

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Research paper

Guideline concordance and outcome in long-term naturalistic treatment of bipolar disorder - a one-year longitudinal study using latent change models

Joannes W. Renes [a,*], Dominique F. Maciejewski [b], Eline J. Regeer [a], Adriaan W. Hoogendoorn [c], Willem A. Nolen [d], Ralph W. Kupka [a,c]

[a] Altrecht Institute for Mental Health Care, Lange Nieuwstraat 119, 3512 PG, Utrecht, the Netherlands
[b] Radboud University, Department of Development Psychopathology, Behavioural Science Institute, Montessorilaan 3, 6525 HR Nijmegen, the Netherlands
[c] Amsterdam University Medical Center / Vrije Universiteit, Department of Psychiatry & Amsterdam Public Health research institute, Oldenaller 1, 1081 HJ, Amsterdam, the Netherlands
[d] University Medical Center Groningen, University of Groningen, Department of Psychiatry, Hanzeplein 1, 9713 GZ Groningen, the Netherlands

ARTICLE INFO

Keywords:
Bipolar
Guideline
Concordance
Outcome

ABSTRACT

Background: Only few studies investigated the relation between concordance with treatment guidelines and treatment outcome in everyday treatment of bipolar disorder (BD). Prospective studies are scarce.

Methods: A nationwide, naturalistic, prospective study on the relation between guideline concordance and treatment outcome in the long-term outpatient treatment of patients with BD. Participants completed a survey on treatments received and various outcome measures at baseline and after one year.

Results: Of 839 patients who completed the baseline survey, 615 (73.3%) also completed the follow-up survey. Consistent with our a priori hypothesis, cross-sectional analyses at baseline showed correlations between guideline concordance with quality of life ($r = .17$, $p < .001$), treatment satisfaction ($r = .17$, $p < .001$), and impaired functioning ($r = -.10$, $p = .04$). At follow-up, guideline concordance was correlated with severity of illness ($r = -.10$, $p = .05$), quality of life ($r = .18$, $p < .001$), and treatment satisfaction ($r = .15$, $p < .001$). Concerning three additional hypotheses on longitudinal relations between concordance and outcome measures, only a positive relation was found between change in guideline concordance and change in quality of life.

Limitations: Selection bias may have occurred by inclusion of patients with neither a very severe nor a very mild course of illness.

Conclusions: Although guideline concordance was high throughout the study, change in guideline concordance was positively associated with change in quality of life, suggesting that especially in long-term treatment, continuous efforts to optimize ongoing treatment is essential.

1. Introduction

Over the last decennia many clinical practice guidelines for the treatment of psychiatric disorders have been developed, including practice guidelines for the treatment of bipolar disorder (BD) (Fountoulakis et al., 2005; Parker et al., 2017). However, studies have shown that concordance with guidelines in everyday clinical practice varies considerably, and various factors have been found to be associated with better or worse concordance, as reviewed elsewhere (Renes et al., 2018). Moreover, it has been relatively understudied whether implementing treatment guidelines leads to better patient outcomes. A review of the literature published between January 1996 and March 2006 concluded that high-quality evidence on the effects of implementation of specific psychiatric guidelines on health care providers performance and patient outcomes was largely lacking (Weinmann et al., 2007). However, in a secondary analysis of data from three studies that evaluated a program designed to encourage collaborative care for depression-based quality improvement in primary care practices, Hepner et al. (2007) showed that greater adherence to practice guidelines predicted more favorable outcomes on measures for depression. These findings have been replicated in other studies of depression (Köhler et al., 2012) and eating disorders (Alañón Pardo et al., 2017), although others failed to find such

* Correspondening author.
E-mail addresses: j.renes@altrecht.nl (J.W. Renes), d.maciejewski@pwo.ru.nl (D.F. Maciejewski), e.regeer@altrecht.nl (E.J. Regeer), a.hoogendoorn@amsterdamumc.nl (A.W. Hoogendoorn), w.a.nolen@umcg.nl (W.A. Nolen), r.kupka@amsterdamumc.nl (R.W. Kupka).

https://doi.org/10.1016/j.jad.2020.12.106
Received 23 September 2020; Received in revised form 25 November 2020; Accepted 25 December 2020
Available online 29 December 2020

associations (Prins et al., 2011).

To the best of our knowledge, so far only a few studies specifically addressed the relation between concordance with guidelines and outcome in BD. In Australia, Fang et al. (2019) performed a file-audit study of 67 participants with BD referred to a tertiary youth mental health service. A proportional concordance score was made based on adult guidelines. Results showed that participants with higher guideline-concordant care had less favorable symptomatic and functional outcomes at 18 months prospective follow-up. Similarly, Altinbas et al. (2011) found no positive effect of adherence to the Turkish treatment guideline for bipolar disorders on time to remission in patients with bipolar depression being treated in a specialized out-patient department of a psychiatric training hospital. However, Kessing et al. (2013) found a decrease in readmission to a psychiatric hospital, and increase in satisfaction with care, when patients with bipolar disorder were treated more in concordance with treatment guidelines in specialized outpatient mood disorder clinics, compared with patient treated with standard out-patient treatment, after hospitalization for a mood episode early in the course of illness.

Overall, studies on concordance with guidelines and patient outcomes show at best mixed results. Obviously, one would expect that for patients with an unfavorable course of illness health care providers optimize treatment as much as possible according to the current guidelines, which in turn would improve patient outcomes. For patients with an already favorable outcome (i.e., low severity, high quality of life, good overall functioning and life satisfaction), a common treatment goal is to preserve stability, and therefore concordance with those parts of the guideline addressing maintenance treatment should not decline.

Since little is known about the relation between concordance with guidelines and outcomes in the treatment of BD, more studies are needed to elaborate this relation for all recommended treatment modalities, thereby reflecting treatment in everyday practice. Such studies will improve understanding for which patients guideline-concordant treatment is effective, for whom guideline concordance can be improved, or if a guideline itself could be improved. For instance, Fang et al. (2019) concluded that 'youth with early stage BD need specific evidence-based guidelines that are in-turn predicated on high-quality intervention studies in that population'.

We performed a prospective study to assess the relation between outcome of long-term treatment in outpatients with BD and concordance of their treatment with the Dutch guideline for BD (Nolen et al., 2008). Treatment results were measured at baseline and after one-year follow-up, focusing on four outcome measures: severity of illness, quality of life, psychosocial functioning, and satisfaction with care. We tested the original a priori hypothesis of the study that higher concordance scores would be associated with a better outcome, both at baseline and at follow-up. Results at baseline have been published elsewhere (Renes et al., 2018). However, since the relation between concordance of treatments with the guideline and treatment outcome was expected to be complex as both may influence each other, we also tested the following three additional longitudinal hypotheses that were developed at a later stage, prior to the analysis of the current study; (1) that there would be a positive relation between change in concordance and change in outcome from baseline to follow-up; (2) that poorer treatment outcome at baseline would be associated with an improvement in concordance from baseline to follow-up; and (3) that higher concordance at baseline would be associated with greater improvements in outcomes from baseline to follow-up.

2. Methods

2.1. Study design and participants

The Treatment of Bipolar Disorder in the Netherlands study (TBDN) is a nationwide, naturalistic cohort study of guideline concordance in the long-term outpatient treatment of patients with BD or schizoaffective disorder, bipolar type (SZA), in mental health settings. Psychiatrists were asked to invite all patients in treatment for BD or SZA, aged eighteen or older, to participate in the study by a letter provided by the researchers. Only patients unable to fill out the questionnaires were excluded from the study. Subsequently, patients willing to participate sent their informed consent form to the researchers. At baseline and after one-year, patients completed a questionnaire. Questionnaires were returned directly to the researchers, blind to the treating psychiatrist. Psychiatrists provided a DSM-IV-TR diagnosis for the patients (American Psychiatric Association, 2000). The study took place between December 2009 and June 2014. The selection of psychiatrists and patients has been described in detail elsewhere (Renes et al., 2014). The study was approved by the Medical Ethical Committee of the University Medical Center Utrecht, The Netherlands.

2.2. Measurement of concordance with the Dutch guideline

The patients' questionnaire addressed previous and current illness characteristics, treatments received, and various other outcome measures. The treatment questions concerned maintenance pharmacotherapy (a list of medication was provided with the possibility to add medications), whether or not the patient had participated in a group psychoeducation program, whether psychotherapy or supportive treatment was part of the treatment, whether there was at least once yearly an appointment with a psychiatrist, whether patients had an emergency plan on how to deal with emerging mood symptoms, and whether they regularly monitored their mood, e.g. by completing a prospective Life-Chart using the NIMH Life-Chart Method that is widely available in the Netherlands (Denicoff et al., 2000). The method we used to measure concordance of these treatments with the Dutch guideline is fully described elsewhere (Renes et al., 2018). Since recommendations in the guideline vary for patients with different clinical profiles, we distinguished four clinical profiles. For each profile it was described whether or not a particular treatment modality was recommended to be part of the overall treatment. An overview of the guideline recommendations concerning these treatment modalities is provided in the supplementary material. To be able to measure concordance with the guideline as a composite score including the concordance of all treatment modalities together, points were assigned to each treatment modality taking into account an assumed impact factor on treatment outcome as determined by consensus among four of the authors (JR, ER, WN, RK). These points were assigned as follows: pharmacotherapy 40 points; group psychoeducation 20; psychotherapy 20; and 5 points each for participation of a psychiatrist, having an emergency plan on how to deal with early symptoms, mood monitoring, and supportive treatment. For every treatment modality that was in accordance with the guideline points were added to the total score for that patient. This resulted in a composite score from 0 to 100. If a treatment modality was not recommended in the guideline, and indeed not part of the treatment, this was considered to be guideline concordant. If for instance maintenance pharmacotherapy would have been absent in the treatment of a patient who is asymptomatic, and with no recommendation for maintenance pharmacotherapy according to the guideline, this would have been in concordance with the guideline. This is based on the assumption that guideline concordance not only means that recommended treatments should be part of the treatment, but also that treatments not recommended in the guideline, should not be part of the treatment, since this may unnecessarily burden the patient. If treatment modalities were part of the treatment or missing in the treatment, and this was *not* in accordance with the guideline, no points were added, i.e., no points were subtracted from the total score. Concordance with the guideline was assessed at baseline and after one-year follow-up.

2.3. Outcome measures

Severity of illness course over past 12 months (further: Severity): to

assess the severity of the course of BD we designed a 3-point scale based on the occurrence (yes or no) of mood episodes and hospital admissions in the previous twelve months. Severity was scored as: 1 (mild) if no mood episodes and no hospital admissions occurred in the previous twelve months, 2 (moderate) when there was least one mood episode, but no hospital admissions, or 3 (severe) when there was at least one hospital admission for BD.

Functioning: was measured with a modified self-rated version of the Functioning Assessment Short Test (FAST) to assess problems with functioning in six specific areas. The clinician-rated FAST is described and validated in bipolar patients by Rosa et al. (Rosa et al., 2007). Higher scores on the FAST are associated with greater impairments in functioning.

Quality of life: the WHOQOL-BREF, an average score of four domains (physical health, physiological, social relationships, and environment) of the short version of the original 100 item self-rating scale was used for the measurement of health related quality of life (Trompenaars et al., 2005). Higher scores corresponded to higher quality of life.

Satisfaction with care: patients were asked to rate their satisfaction with the treatment they currently received between 0 (very unsatisfied, worst possible treatment) to 10 (very satisfied, best possible treatment).

2.4. Statistical analysis

We used latent change score models to analyze the longitudinal association between concordance and severity, quality of life, functioning, and treatment satisfaction. Analyses were conducted in R version 3.6.2 (R Core Team, 2019) using the lavaan package (version 0.6-5). All code and output is available in the supplementary material (a link to the online material is provided in the section on availability of data and materials at the end of this article).

Latent change score models combine aspects of two structural equation models: (1) autoregressive cross-lagged panel models, in the sense that they capture moment-to-moment associations between variables and allow to examine the direction of effects, as well as (2) growth curve models in the sense that they capture both intraindividual (i.e., within-person) change and interindividual (i.e., between-person) differences in change (Grimm et al., 2016). The combination of cross-lagged paths to test for direction of effects, and interindividual differences in intraindividual change, which is often disregarded in conventional longitudinal models (e.g. cross-lagged models alone), was important for our hypotheses, because we were interested in whether concordance with the guideline predicted interindividual differences in intraindividual change in outcome measures and/or vice versa.

To better understand the latent change score models we will explain the model specifications and the estimates derived from the model for a bivariate latent change score model (see Fig. 1). For more information see Grimm et al. (2016); Kievit at al. (2018).

In this model, two variables (Y and X, e.g. Illness severity and the degree of concordance) are measured at two time-points (T0 and T1). In the latent change score model, the score of the construct at T1 is a linear combination of the score of the construct at T0 and the latent change score of the construct from T0 to T1. In the model, the autoregressive effect from T0 to T1 and the factor loading of the construct at T1 on the change score is set to 1. This creates a latent factor that captures the true amount of change (d) between two occasions for an individual (also referred to as within-person score). One can also examine how the score of a construct at T0 predicts change in the same (β) or another construct (ϒ). Additionally, one can examine the covariance between initial scores (ϕ) and change scores (ρ).

We extended this model to incorporate concordance (X) and four outcome measures (Y1, Y2, Y3 and Y4). Using the extended model, we tested the associations between concordance and four outcome measures in one model. Specifically, we tested for three possible relations: (1) Whether the degree of concordance at T0 was related to changes in severity, quality of life, functioning, and treatment satisfaction (denoted as four γ-parameters, one for each outcome variable); (2) whether severity, quality of life, functioning, and treatment satisfaction at T0 were related to changes in the degree of the concordance (denoted as four γ-parameters, one for each outcome variable); and (3) whether changes in the degree of the concordance were related to changes in severity, quality of life, functioning, and treatment satisfaction (denoted as four ρ-parameters, one for each outcome variable). We used Maximum Likelihood Estimation with robust standard errors (MLR) to correct for missing data and to account for non-normal distributions. The model was fully saturated, thus no model fit indices could be calculated.

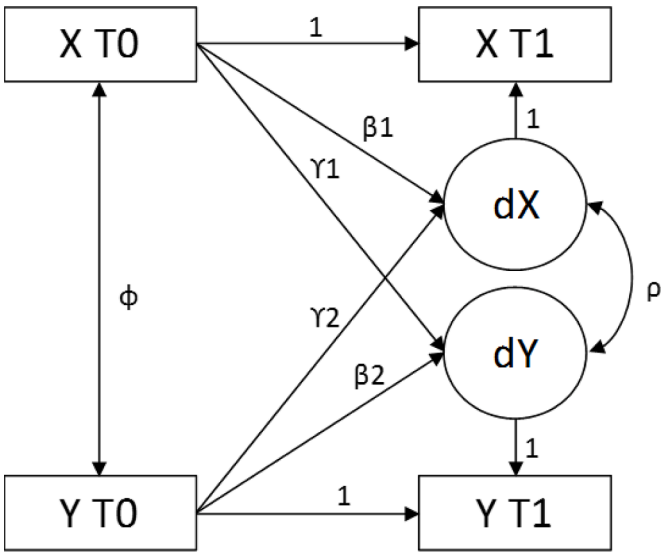

Fig. 1. Bivariate latent change model.

Due to the high number of tests, we applied a False Discovery Rate (FDR) correction for the p-values of the hypothesized relations. FDR-corrected p-values $< .05$ were considered statistically significant. In total, 20 relations that were tested, namely (1) cross-sectional relations between concordance and four outcomes at baseline (i.e., four tests) and follow-up (i.e., four tests); (2) relations between concordance and change in four outcomes (i.e., four tests); (3) relations between four outcomes and change in concordance (i.e., four tests); and (4) relations between change in concordance and change in four outcomes (i.e., four tests).

3. Results

At baseline, 839 patients completed the first questionnaire, of which 639 were in treatment at centers specialized in the treatment of mood disorders, and 200 in general psychiatric outpatient facilities. At one-year follow-up, 643 completed the second questionnaire; of these, three were excluded from the study because diagnosis had been changed to another mood disorder, one was excluded due to logistical problems, eight patients were excluded because they were no longer in psychiatric treatment (only treated by a general practitioner), and sixteen because the date of completion was too deviant from one year follow-up (less than 40 weeks or more than 78 weeks), resulting in $n = 615$. The mean time to follow-up was 56 weeks ($SD = 4.7$). Demographics and clinical characteristics are shown in Table 1.

Of the treatments received by the participants, the use of maintenance medication was fairly constant during study. At follow-up 441 participants (71.8%) used lithium, 209 (34.1%) used anticonvulsants, and 265 (43.1%) used antipsychotics as maintenance medication. These were used as monotherapy by 326 (53%) participants and in any combination by 264 (42.9%). One participant did not provide information on maintenance medication at T1. The frequency in which psychosocial

Table 1
Demographics and clinical characteristics at baseline and follow-up.

	T0	T1
Female, N (%), *(T0, N = 839; T1, N = 615)*[1]	555 (66.2)	417 (68)
Age[2], mean years (s.d.), *(T0, N = 839; T1, N = 615)*[1]	49.5 (11.2)	49.6 (11.1)
Diagnosis[3]: N (%), *(T0, N = 839; T1, N = 615)*[1]		
BD I	551 (65.7)	402 (65.4)
BD II	211 (25.1)	153 (24.9)
BD NOS	32 (3.8)	26 (4.2)
SZA	45 (5.4)	34 (5.5)
Years since first medication for depressive symptoms[2], mean (s.d.), *(T0, N = 702; T1, N = 508)*[1]	17.4 (10.9)	17.3 (10.7)
Years since first medication for manic symptoms[2], mean (s.d.), *(T0, N = 750; T1, N = 554)*[1]	13.5 (10.5)	13.8 (10.4)
Number of life time episodes[2], N (%), *(T0, N = 755; T1, N = 561)*[1]		
1	9 (1.2)	6 (1.1)
2	28 (3.7)	21 (3.7)
3	34 (4.5)	24 (4.3)
4	68 (9.0)	51 (9.1)
5 or more	616 (81.6)	459 (81.8)

[1] Sample size differs among variables, depending on missing data points. Total number of participants included at T0 (baseline) was 839, total number of participants included at T1 (follow-up) was 615.
[2] Until T0 (baseline).
[3] BD I: bipolar I disorder; BP II: bipolar II disorder; BP NOS: bipolar disorder NOS; SZA: schizoaffective disorder, bipolar type.

treatments were part of the treatment at follow-up ranged from 52 in 614 participants (8.5%) for psychotherapy, to 577 in 611 participants (94.9%) for the participation of a psychiatrist in the treatment. Detailed overviews of the medication used as maintenance pharmacotherapy, psychosocial treatments, and the concordance of each treatment modality, are provided in the supplementary material.

Means and standard deviations of the study variables (i.e., concordance, severity, quality of life, functioning, treatment satisfaction) are shown in Table 2. Paired t-tests indicated that severity significantly improved from T0 to T1, while the other three outcomes and concordance were overall relatively stable over time. However, when plotting the pattern of change of the study variables, a highly heterogeneous picture emerged. Fig. 2 depicts the pattern of change of all study variables, indicating that there is considerable heterogeneity in both the intercept and the slopes. Specifically, individuals differed considerably in their overall levels as well as in their pattern of change (with some increasing, some decreasing, and some staying stable). This is also supported by the significant variances of the intercept and change scores (for all study variables *ps* $< .001$; see supplementary material).

Table 3 depicts the correlations between the study variables. Results indicated that concordance was significantly associated with all clinical variables cross-sectionally (hypothesis 1). After FDR correction, concordance T0 correlated positively with quality of life at T0 ($r = .17$, $p < .001$) and treatment satisfaction at T0 ($r = .17$, $p < .001$) and negatively with impaired functioning ($r = -.10$, $p = .04$). Moreover, concordance at T1 correlated negatively with severity ($r = -.10$, $p = .05$) and positively with quality of life ($r = .18$, $p < .001$) and treatment satisfaction at T1 ($r = .15$, $p < .001$).

Table 2
Descriptive statistics of concordance and outcome measures at T0 and T1.

	T0		T1		Paired sample t-test	
	Mean	SD	Mean	SD	t	p
Concordance (range 0-100)	72.69	17.32	73.24	16.95	0.87	.38
Severity (range 1-3)	1.68	0.66	1.58	0.62	3.46	<.001
Quality of Life (range 0-100)	64.36	14.29	64.57	14.09	-0.62	.54
Functioning (FAST) (range 0-72)	20.71	15.91	20.39	15.59	-0.95	.34
Treatment satisfaction (1-10)	7.90	1.56	7.97	1.39	-0.11	.91

We subsequently investigated the longitudinal relations between concordance and the four outcome variables using the latent change score models.

Table 4 presents the results from the latent change score models with regard to the association between concordance and the other outcome variables (severity, quality of life, functioning, treatment satisfaction). All other results (association between clinical variables, intercepts, variances etc.) can be found in the online supplement.

We found no evidence that concordance at T0 was longitudinally associated with subsequent changes in severity, quality of life, functioning, and treatment satisfaction. Also, we found no evidence that severity, quality of life, functioning, and treatment satisfaction at T0 were longitudinally associated with changes in concordance. However, increase in concordance from T0 to T1 was associated with significant (after FDR-correction) increase in quality of life from T0 to T1 ($B = 17.59$, $SE = 5.90$, $p = .012$, $\beta = .15$). Associations between changes in concordance and changes in the other clinical variables were not significant.

4. Discussion

The aim of the current study was to investigate the relation between concordance with the Dutch treatment guideline for BD and four outcome measures (i.e., severity of BD, quality of life, functioning, and treatment satisfaction) in the long-term outpatient mental health care. In line with our a priori hypothesis, in the cross-sectional analysis, both at baseline and at one-year-follow-up, guideline concordance was positively associated with three of the four outcome measures. In the longitudinal analyses, we found a positive association between an increase in guideline concordance from baseline to follow-up and an increase in quality of life, partially supporting the first additional hypothesis. However, we did not find such a relation for the other three outcome measures, nor did we find evidence for the second and third additional hypothesis that outcome at baseline would predict change in guideline concordance, or vice versa.

These mixed results are in line with findings in other studies. Some studies found a positive association between concordance and clinical outcomes (Alañón Pardo et al., 2017; Hepner et al., 2007; Kessing et al., 2013; Köhler et al., 2012), whereas others did not (Prins et al., 2011). For instance, and in contrast with our hypothesis, Fang et al. (2019) even found that adolescent patients with higher guideline concordance scores had greater odds of remaining unwell at follow-up. In contrast, and in line with our results, Prins et al. (2011) did not report a relation between patients with or without guideline concordant care, and improvement in symptoms of depression or anxiety over one-year follow-up in general practices.

There are several possible explanations for these divergent findings and the results of our study. First and foremost, it may reflect the highly heterogeneous longitudinal course of BD, and hence the complexity of individual treatment decisions as a result. Second, differences in methodology, study population, and treatment settings make comparison of various studies difficult. Third, in our study we found that overall, guideline concordance was high both at baseline and at one-year-follow-up, and on average participants preformed fairly well on the four outcome measures, resulting in less variance, and possible ceiling and floor effects. The clinical importance of the significant difference we found between the mean severity at baseline and at one-year-follow-up, seems to be limited. The difference was small, and the fact that the measure only had three categories means that the variability was limited, which may have impacted the finding of significant change. Moreover, as shown in Fig. 2, 67% did not change in severity, and of the 33% that did change (14% decreased, and 19% increased), only 5% changed more than one category. Despite the overall stability and high

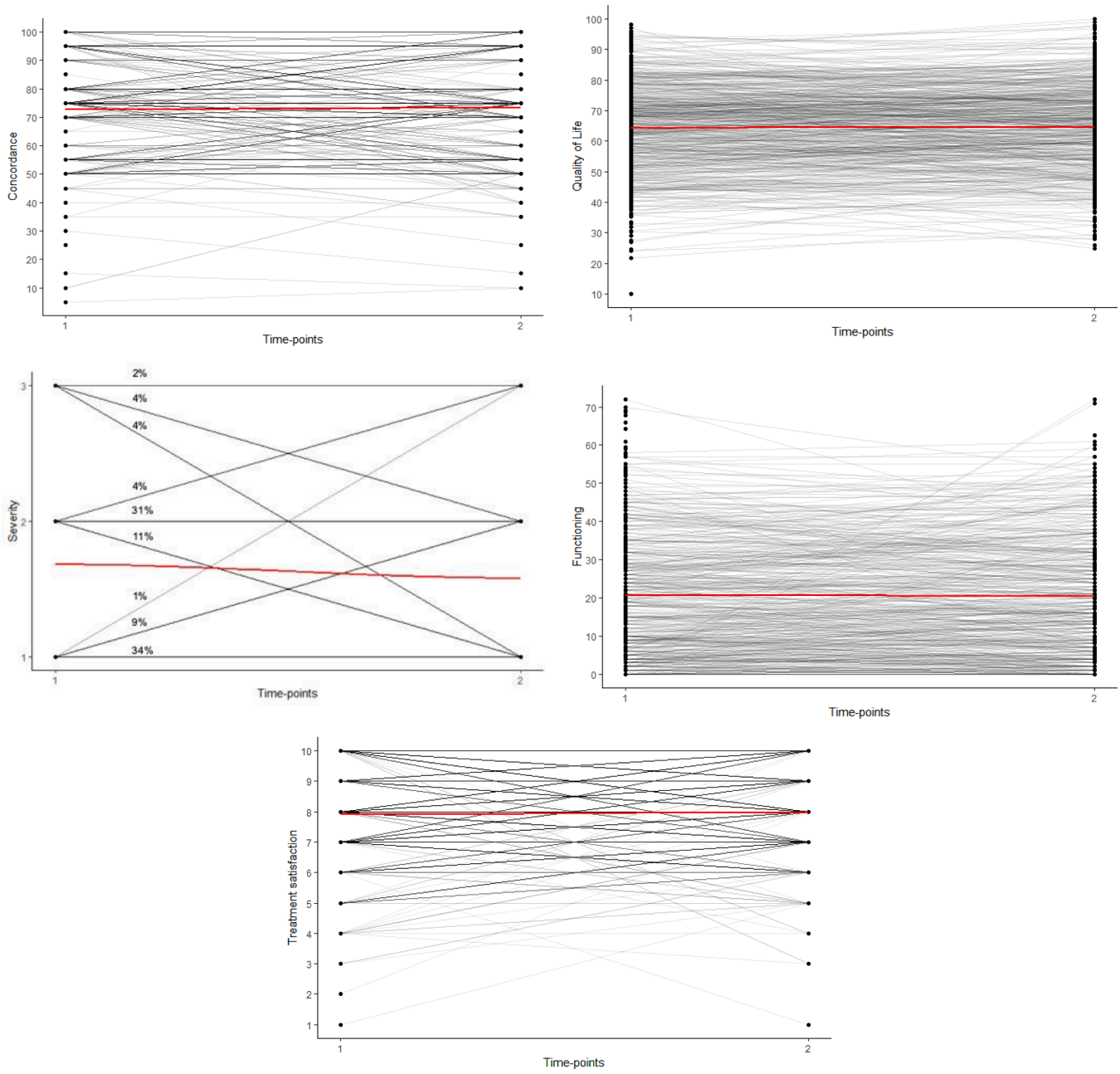

Fig. 2. Change from baseline to follow-up of guideline concordance, severity of BD, quality of life, functioning, and treatment satisfaction for individual participants.

levels of concordance and outcome measures, in the latent change score models, significant interindividual differences were found between overall levels, and intraindividual changes of study variables over time. However, with these models we were unable to find covariance relations between change in concordance from baseline to follow-up, and changes in severity, functioning, and satisfactions with care (additional hypothesis 1), or to predict change in concordance from baseline to follow-up by the outcome measures at baseline (additional hypothesis 2), or to predict change in outcome measures from baseline to follow-up by concordance at baseline (additional hypothesis 3).

Our study has several limitations. First, most participants that were included in the study were being treated in specialized centers for mood disorder. This may have resulted in the inclusion of patients receiving an already optimized treatment. In addition, it is likely that better functioning patients will especially participate in a study like this. Adjustments of medication for the treatment of acute mood episodes were not taken into account in the guideline concordance score, therefore we were unable to detect possible influences of these adjustments on outcomes. Secondly, the sensitivity of the composite score for the measurement of the severity of illness was limited, since the number and duration of mood episodes and admissions were not included. And finally, as in many studies in BD, we included mainly participants with a later stage BD, often with many years since their first treatment with medication, and with multiple lifetime mood episodes, for whom improvement may be more difficult to achieve. Studies in populations with an earlier stage of BD may be more likely to show the effects of guideline concordance on clinical outcome, as for example was shown in the study by Kessing et al. (2013).

Our study has also several strengths. Attrition was relatively low, and there were hardly any differences between participants at baseline and follow-up. We used a comprehensive composite score for the measurement of guideline concordance, that included most aspects of the long-term treatment of BD. And to the best of our knowledge, this is the first study on this subject in BD, that used latent change models for

Table 3
Correlations of concordance and outcome measurements at T0 and T1.

	1	2	3	4	5	6	7	8	9	10
1 Concordance T0	1									
2 Severity T0	-.07*	1								
3 Quality of Life T0	.17***	-.30***	1							
4 Functioning (FAST) T0	-.10*	.34***	-.71***	1						
5 Treatment satisfaction T0	.17***	-.12***	.31***	-.25***	1					
6 Concordance T1	.64***	-.09*	.18***	-.12*	.19***	1				
7 Severity T1	-.05	.42***	-.32***	.38***	-.12**	-.10*	1			
8 Quality of Life T1	.09*	-.24***	.75***	-.62***	.28***	.18***	-.37***	1		
9 Functioning (FAST) T1	-.10*	.33***	-.57***	.73***	-.15***	-.11*	.41***	-.69***	1	
10 Treatment satisfaction T1	.13**	-.11**	.31***	-.19***	.50***	.15***	-.17***	.41***	-.18***	1

Note. ***$p < .001$, **$p < .01$, *$p < .05$. Note that p-values in this table are not FDR-corrected.

longitudinal analyses. This made it possible to test multiple possible relations in one model, including both intraindividual change, and interindividual differences in change.

5. Implications and conclusions

Overall, in this naturalistic cohort of patients in long term treatment for BD, concordance of treatments with the guideline was high, and on average clinical outcome was fairly stable over time. Although we were not able to show all the expected relations between guideline concordance and outcome measures, we did find that an increase in guideline concordance during the study was significantly associated with an increase in quality of life. This is a hopeful finding since it points in the direction that although patients maybe in psychiatric treatment for many years, improving treatments, especially psychosocial treatments, may still have added value even if symptomatic improvement is limited. To gain further insight in the relations between treatments, outcomes, and guidelines, more studies are needed. For naturalistic cohort studies it would be very helpful when computerized data on treatments received by patients become better accessible for research. These data can then be combined with routine outcome measurements, patients' characteristics and illness characteristics. Moreover, given the high heterogeneity in the pattern of change of the outcome measures, it could be useful to use growth mixture models that tease apart different clusters of change to see how these differentially relate to concordance.

Table 4
Regression and covariance coefficients concerning the association between concordance with severity, quality of life, functioning and treatment satisfaction.

	Estimate	SE	p	Standardized estimate
Effects from concordance on latent change scores of outcomes				
Concordance T0 → change in severity	0.001	0.002	0.738	0.016
Concordance T0 → change in Quality of Life	-0.023	0.024	0.517	-0.040
Concordance T0 → change in functioning (FAST)	-0.001	0.031	0.987	-0.001
Concordance T0 → change in Treatment satisfaction	0.002	0.003	0.641	0.025
Effects from outcomes on latent change scores of concordance				
Severity T0 → change in concordance	0.291	0.901	0.830	0.013
Quality of Life T0 → change in concordance	-0.001	0.068	0.987	-0.001
Functioning (FAST) T0 → change in concordance	-0.043	0.062	0.640	-0.047
Treatment satisfaction T0 → change in concordance	0.747	0.420	0.150	0.081
Associations between latent change scores				
change in concordance <-> change in severity	-0.479	0.321	0.247	-0.068
change in **concordance** <-> change in **Quality of Life**	**17.591**	**5.904**	**0.012**	**0.148**
change in concordance <-> change in functioning	-5.752	6.581	0.546	-0.042
change in concordance <-> change in Treatment satisfaction	0.676	0.700	0.517	0.044

Note. p-values are FDR-corrected.

Authors' contributions

J.W. Renes drafted this paper, D.F. Maciejewski, and A.W. Hoogendoorn performed the analysis. The paper was modified by all authors. W.A. Nolen, and R.W. Kupka conceived the study, J.W. Renes, and E.J. Regeer contributed to the design and study protocol. All authors approved the final manuscript.

Funding

R.W. Kupka received an unrestricted research grant for this study by AstraZeneca.

Declaration of Competing Interests

J.W. Renes and E.J. Regeer received speaker's fees from AstraZeneca, Bristol-Myers Squibb, and Ely Lilly. A. W. Hoogendoorn, and D.F. Maciejewski report no potential conflicts of interest. W.A. Nolen has received grants from the Netherlands Organisation for Health Research and Development, the European Union; has received honoraria/speaker's fees from Lundbeck and Aristo Pharma and has served as consultant for Daleco Pharma. R.W. Kupka received speaker's fees for lectures on symposia sponsored by AstraZeneca, Bristol-Myers Squibb, Lundbeck, Sanofi, and Janssen. W.A. Nolen and R.W. Kupka were members and chairs of Dutch bipolar disorder guideline committees in 2008 and 2015.

Acknowledgments

The authors thank the participating psychiatrists and patients for their contribution to the study.

Availability of data and materials

All code and output of the analysis conducted in R version 3.6.2 using the lavaan package is available at https://osf.io/35zrg/?view_only=de76a015928c4e799f5f6ea86f8ffc3a. The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jad.2020.12.106.

References

Alañón Pardo, M., del, M., Ferrit Martín, M., Calleja Hernández, M.Á., Morillas Márquez, F., 2017. Adherence of psychopharmacological prescriptions to clinical practice guidelines in patients with eating behavior disorders. Eur. J. Clin. Pharmacol. 73, 1305–1313. https://doi.org/10.1007/s00228-017-2287-2.

Altinbas, K., Smith, D., Oral, E.T., 2011. Adherence to Turkish psychiatric association guideline for bipolar depression treatment in a specialized mood disorders outpatient unit. Psychiatry Danub. 23, 189–193.

American Psychiatric Association, 2000. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890423349.

Denicoff, K.D., Leverich, G.S., Nolen, W.A., Rush, A.J., McElroy, S.L., Keck, P.E., Suppes, T., Altshuler, L.L., Kupka, R., Frye, M.A., Hatef, J., Brotman, M.A., Post, R. M., 2000. Validation of the prospective NIMH-life-chart method (NIMH-LCM(TM)-p) for longitudinal assessment of bipolar illness. Psychol. Med. 30, 1391–1397. https://doi.org/10.1017/S0033291799002810.

Fang, M., Filia, K., Davey, C., McGorry, P., Berk, M., Cotton, S., 2019. Poster. Are current guidelines appropiate for youth with early stage bipolar disorder? A prospective observational study of the impact of guideline concordant care on 18-month symptomatic and functional outcomes. Bipolar Disord. 21, 66–147. https://doi.org/10.1111/bdi.12746.

Fountoulakis, K.N., Vieta, E., Sanchez-Moreno, J., Kaprinis, S.G., Goikolea, J.M., Kaprinis, G.S., 2005. Treatment guidelines for bipolar disorder: a critical review. J Affect Disord. 86, 1–10. https://doi.org/10.1016/j.jad.2005.01.004.

Grimm, K.J., Ram, N., Estabrook, R., 2016. Growth modeling: structural equation and multilevel modeling approaches. Methodol. Soc. Sci. Guilford Publications, New York, NY.

Hepner, K.A., Rowe, M., Rost, K., Hickey, S.C., Sherbourne, C.D., Ford, D.E., Meredith, L. S., Rubenstein, L.V., 2007. The effect of adherence to practice guidelines on depression outcomes. Ann. Intern. Med. 147, 320–329. https://doi.org/10.7326/0003-4819-147-5-200709040-00007.

Kessing, L.V, Hansen, H.V, Hvenegaard, A., Christensen, E.M., Dam, H., Gluud, C., Wetterslev, J., 2013. Treatment in a specialised out-patient mood disorder clinic v. standard out-patient treatment in the early course of bipolar disorder: randomised clinical trial. Br. J. Psychiatry 202, 212–219. https://doi.org/10.1192/bjp.bp.112.113548.

Kievit, R.A., Brandmaier, A.M., Ziegler, G., van Harmelen, A.L., de Mooij, S.M.M., Moutoussis, M., Goodyer, I.M., Bullmore, E., Jones, P.B., Fonagy, P., Lindenberger, U., Dolan, R.J., 2018. Developmental cognitive neuroscience using latent change score models: a tutorial and applications. Dev. Cogn. Neurosci. 33, 99–117. https://doi.org/10.1016/j.dcn.2017.11.007.

Köhler, S., Hoffmann, S., Unger, T., Steinacher, B., Fydrich, T., 2012. Adherence to guidelines and effectiveness of inpatient treatment for unipolar depression. Int. J. Psychiatry Clin. Pract. 16, 103–112. https://doi.org/10.3109/13651501.2011.638142.

Nolen, W.A., Kupka, R.W., P.F.J., S., Knoppert-van der Klein, E.A.M., Honig, A., Reichart, C.G., Goossens, P.J.J., Daemen, P., Ravelli, D.P., 2008. Richtlijn Bipolaire stoornissen. Nederlandse Vereniging voor Psychiatrie. De Tijdstroom, Utrecht.

Parker, G.B., Graham, R.K., Tavella, G., 2017. Is there consensus across international evidence-based guidelines for the management of bipolar disorder? Acta Psychiatry Scand. 135, 515–526. https://doi.org/10.1111/acps.12717.

Prins, M.A., Verhaak, P.F.M., Hilbink-Smolders, M., Spreeuwenberg, P., Laurant, M.G.H., van der Meer, K., van Marwijk, H.W.J., Penninx, B.W.J.H., Bensing, J.M., 2011. Outcomes for depression and anxiety in primary care and details of treatment: a naturalistic longitudinal study. BMC Psychiatry 11, 180. https://doi.org/10.1186/1471-244X-11-180.

Core Team, R, 2019. R: a language and environment for statistical computing. r found. Stat. Comput. Vienna, Austria.

Renes, J.W., Regeer, E.J., Hoogendoorn, A.W., Nolen, W.A., Kupka, R.W., 2018. A nationwide study on concordance with multimodal treatment guidelines in bipolar disorder. Int. J. Bipolar Disord. 6, 22. https://doi.org/10.1186/s40345-018-0130-z.

Renes, J.W., Regeer, E.J., van der Voort, T.Y., Nolen, W.A., Kupka, R.W., 2014. Treatment of bipolar disorder in the Netherlands and concordance with treatment guidelines: study protocol of an observational, longitudinal study on naturalistic treatment of bipolar disorder in everyday clinical practice. BMC Psychiatry 14, 58. https://doi.org/10.1186/1471-244X-14-58.

Rosa, A.R., Sánchez-Moreno, J., Martínez-Aran, A., Salamero, M., Torrent, C., Reinares, M., Comes, M., Colom, F., Van Riel, W., Ayuso-Mateos, J., Kapczinski, F., Vieta, E., 2007. Validity and reliability of the functioning assessment short test (FAST) in bipolar disorder. Clin. Pract. Epidemiol. Ment. Heal. 3, 5. https://doi.org/10.1186/1745-0179-3-5.

Trompenaars, F.J., Masthoff, E.D., Van Heck, G.L., Hodiamont, P.P., De Vries, J., 2005. Content validity, construct validity, and reliability of the WHOQOL-Bref in a population of Dutch adult psychiatric outpatients. Qual. Life Res. 14, 151–160. https://doi.org/10.1007/s11136-004-0787-x.

Weinmann, S., Koesters, M., Becker, T., 2007. Effects of implementation of psychiatric guidelines on provider performance and patient outcome: systematic review. Acta Psychiatry Scand. 115, 420–433. https://doi.org/10.1111/j.1600-0447.2007.01016.x.