CHAPTER 2

The use of the prospective NIMH Life Chart Method as bipolar mood assessment method in research: a systematic review of different methods, outcome measures and interpretations

Koenders MA, Nolen WA, Giltay EJ, Hoencamp E, Spijker AT (2015). The use of the prospective NIMH Life Chart Method as bipolar mood assessment method in research: a systematic review of different methods, outcome measures and interpretations. Journal of Affective Disorders; 136 (1): 260-268

Abstract

The severity of bipolar disorder can be assessed using the daily prospective National Institute of Mental Health's Life Chart Method (LCM-p). Also for scientific research the LCM-p, has been used frequently. However, processing and analysing the LCM-p for research purposes, is challenging because of the multitude of complex measures that can be derived from the data. In the current paper we review the different LCM-p course variables (mood episodes, average severity, proportion of time ill and mood switches) and their definitions. Strengths and limitations and the impact of the use of different LCM-p course measures and definitions on the research results are described.

A systematic review of original papers on the LCM was conducted using 9 electronic databases for literature between January 1996 and January 2014. Papers using other prospective charting procedures were not evaluated in the current study. The initial literature search led to 1319 papers of which 21 were eventually selected. A relatively wide variety of definitions of LCM-p course variables was used across the studies. Especially for the calculation of number of episodes and mood switch no univocal definition seems to exist. Across studies several different duration and severity criteria are applied to calculate these variables. We describe which variables and definition are most suitable for detecting specific bipolar disease course characteristics and patterns. In the absence of a golden standard for the calculation of LCM-p course variables, researchers should report the exact method they applied to their LCM-p data, and clearly motivate why this is their method of first choice considering their research aim.

2.1 Introduction

Bipolar disorder is a common mood disorder, with a life time prevalence of 2.4% (58). The natural course of BD is characterized by an always present risk of recurrences even when patients receive treatment according to contemporary practice guidelines (12). Over 90% of patients with bipolar disorder experience recurrences during their lifetime (14). Furthermore, while that symptomatic recovery is attained by 90% of the patients within 2 years after a severe episode, only 30% attain functional recovery (16), implicating the major impact the disorder has on functioning and daily life of bipolar patients.

For the monitoring of symptom severity there are several measures such as the Hamilton Depression Rating Scale (59) for depressive symptoms and the Young Mania Rating Scale (60) for manic symptoms. However these instruments only allow for a cross-sectional assessment of symptom severity and do not evaluate the longitudinal course.

The National Institute of Mental Health's Life Chart Method (NIMH LCM) (20) was developed as a tool for longitudinal monitoring of chronic cyclic affective disorders, such as bipolar disorder (BD) and enables patient and clinician to visualize and obtain insight in the mood course. Several variants of this Life Chart Method (LCM) are used in clinical practice and research, including clinician and self-rated versions taken retro- and prospectively at daily to monthly intervals. The prospective LCM (LCM-p) consists of a clinician version that is rated by the treating physician, and a self-rated patient version. In both versions scoring of severity of mania or depression is based on the level of mood (i.e. mania or depression) associated functional impairment, which is considered to simplify the process of rating. The LCM-p uses four levels of severity: mild, moderate low, moderate high, and severe. The LCM also gives the opportunity to score mixed mood states, when patients experience impairment caused by manic and depressed symptoms at the same time. The LCM proved its major clinical value by giving both clinician and patient an immediate graphical insight in the patient?s individual bipolar mood course over longer periods of time. Both the clinician and patient version of the LCM-p have now been validated, showing that the daily assess-

ment of mood and episode severity based on the degree of mood associated functional impairment correlated highly with cross-sectional measurements of mood symptoms and global functioning (61-63). The retrospective LCM (R-LCM) has not yet been validated.

Also for scientific research on the course of illness, the LCM-p, especially the daily version, has been used frequently. However, processing and analysing the LCM-p for research purposes, is challenging because of the multitude of complex measures that can be derived from the data. The largest benefit of the LCM-p is that it summarizes the course of mania or depression associated functional impairment over time without diagnostic bias as to what constitutes for example a manic or depressive episode. However, to interpret these complex data, researchers need to transform the raw LCM-p data into course variables. The main course variables that are derived from the LCM are: number of episodes (e.g. 64), number of mood switches (e.g. 65), average mood severity over time (e.g. 30) or proportion of time impaired (66). Although clinical definitions of several bipolar course variables (e.g. recurrence, relapse, remission, response) are previously described by a taskforce of the International Society of Bipolar Disorders (ISBD), criteria to derive these course variables from the LCM were not always clear and differed across studies. For example, for the definition of an episode some researchers have used the Diagnostic and Statistical Manual IV (DSM-IV) (4) duration criteria (e.g. 67) while others used more strict criteria (e.g. 62). One reason for these differences is the fact that different aspects of the mood course are relevant for different research questions. Another reason might be that there is a lack of concise and unequivocal criteria for the calculation of p-LCM course variables.

The aim of the current paper is to review the four main course variables utilized across studies that use the daily LCM-p. We will first describe definitions of these LCM-p course variables and the impact of definition differences on the research results. Subsequently we will amplify on the strengths and limitations of the different LCM-p course measures with respect to specific research aims. We will further suggest recommendations that may help with the interpretation and analysing of life charts.

2.2 Method

To create an overview of the different interpretation methods of the LCM-p across the literature, we conducted an electronic literature search in several databases that all let to unique hits (unique hits are presented between brackets): ScienceDirect (396), LWW (333), Academic Search Premier (294), PubMed (169), PsycINFO (68), Wiley (61), Embase (22), Cochrane (5), Web of Science (4). We selected documents containing the following descriptors in title, abstract or free text: ("Bipolar Disorder" OR "bipolar disorder" OR "bipolar disorders? OR "Mania" OR "Manias" OR "Manic State" OR "Manic States" OR "hypomania OR "Bipolar Depression" OR "Manic Disorder" OR "Manic Disorders" OR "bipolar" OR bipolar* OR "manic depressive" OR "manic depression" OR "manic" OR "Manic-Depressive Psychosis" OR "Manic Depressive Psychosis" OR "Bipolar Affective Psychosis" OR "Manic-Depressive Psychoses" OR "Cyclothymic Disorders" OR "Cyclothymic Disorder" OR "Cyclothymic Personality" OR "Cyclothymic Personalities" OR Cyclothym* " OR "hypomanic" OR "Affective Disorders, Psychotic" AND ("life chart" OR "life charts" OR "life charting" OR "lifechart" OR lifechart* OR "mood chart" OR "mood charts" OR "mood charting" OR "moodchart" OR moodchart* OR "NIMH-LCM-p" OR "LCM" OR "mood rating" OR "mood ratings").

Abstract and titles were used to determine the relevance of the reference. When potentially relevant, full texts were retrieved to assess the article for inclusion. The search was limited to studies published in peer reviewed journals in English available between January 1996 and December 2014. The articles were selected if they met the following criteria: original paper, use of the original daily prospective NIMH-LCM, within adult bipolar I and/or II sample, $N \geq 30$, follow-up time ≥ 3 months. Articles that missed a description of the use and interpretation of the NIMH-LCM were excluded, as well as studies that modified the LCM data by anchoring the life-chart data to other measures such as cross-sectional mood questionnaires (e.g. 68, 69-71) or to other functioning scales (e.g. 72). The initial literature search led to a total of 1352 hits. Figure 2.1 shows the flow chart of the selection process, which led to 21 original articles (presented in Table 2.1) that met our inclusion criteria.

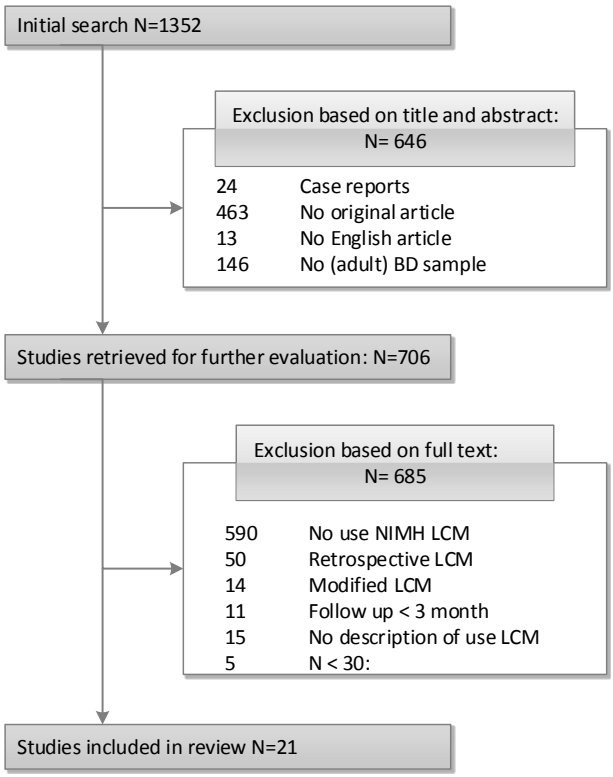

Figure 2.1 | Flow chart of systematic article selection.

Study	N	Follow-up period	LCM	Study aim	Outcome measure	Compliance
Denicoff et al., 1997a	52	1 year	Clinician and patient version	Efficacy psychotropic drugs	- episodes (leapfrog) - proportion of time impaired - average severity	83% compliance to 1 year treatment intervention (no specific LCM compliance information)
Denicoff et al. 1997b	30	2 year	Clinician and patient version	Validation of the prospective clinician NIMH-LCM	- episodes (leapfrog) - average severity	unknown
Denicoff et al. 2000	270	6 months	Clinician and patient version	Validation of the prospective clinician NIMH-LCM	- average severity	27% incomplete data
Post et al., 2001	64	1 year	Clinician and patient version	Effect of antidepressant on induction of mania	- episodes (DSM-IV and shorter 'episodes') - switch rate	unknown
Denicoff et al., 2002	52	1 year	Clinician and patient version	Utility of LCM in clinical trials	- episodes (no definition) - proportion of time impaired - average severity - switch rate	unknown
Joffe et al., 2002	69	1 year	Unknown	Effect of antidepressant on induction of mania	- episodes (DSM-IV) - switch rate	unkown
Post et al., 2002	258	1 year	Clinician and patient version	Ratio depression and mania	- episodes (leapfrog) - proportion of time impaired	unkown
Post et al., 2003	258	1 year	Clinician and patient version	Clinical predictors of mood course	- episodes (DSM-IV + leapfrog) - proportion of time impaired - average severity	unkown
Nolen et al., 2004	258	1 year	Clinician and patient version	Clinical predictors of mood course	- episodes (leapfrog) - average severity	unknown

Table 2.1 | Selected studies using the prospective NIMH LCM.

Kupka et al., 2005	539	1 year	Clinician and patient version	Risk factors for rapid cycling	- episodes (DSM-IV + leapfrog) - proportion of time impaired - average severity	unknown
Leverich et al., 2006	159	1 year	Clinician and patient version	Effect of antidepressant on induction of mania	- episodes (DSM-IV and shorter 'episodes') - switch rate	unknown
Kupka et al., 2007	507	1 year	Clinician and patient version	Ratio depression and mania	- episodes (DSM-IV + leapfrog) - proportion of time impaired - average severity	unknown
Leverich et al., 2007	480	1 year	Clinician and patient version	Age of onset and course severity	- episodes (DSM-IV) - proportion of time impaired - average severity	unknown
Goldberg et al., 2008	182	6 months	Patient version	Effect of lamotrigine on mood stability	- proportion of time impaired	>90% drop out
Langosch et al., 2008	38	1 year	unknown	Effect of quetiapine and sodium valproate on rapid cycling	- proportion of time impaired - switch rate	unknown
Shivakumar et al., 2008	41	3 months	Clinician and patient version	Effect of menstrual cycle on bipolar mood course	- average severity	66% drop out
Born et al., 2009	49	1 year	Clinician and patient version	Comparison of the BD I and II mood course	- proportion of time impaired	unknown
Post et al., 2010	529	1-4 year	Clinician and patient version	Age of onset BD and longitudinal course	- episodes (DSM-IV) - proportion of time impaired - average severity	91% finished year 1, 55% year 2, 34% year 3, 26% year 4.
Zaane et al., 2010	137	1 year	Patient version	Effect of Alcohol use on course of bipolar disorder	- episodes (DSM-IV + leapfrog) - proportion of time impaired - average severity	34% drop out

Table 2.1 | Selected studies using the prospective NIMH LCM (cont.)

Zaane et al., 2014	137	1 year	Patient version	Effect of Alcohol use on course of bipolar disorder	- average severity - switch rate	34% drop out
Born et al. 2014	108	1-4 year	Patient version	Validation of the prospective patient NIMH-LCM	- average severity	unknown

Table 2.1 | Selected studies using the prospective NIMH LCM (cont.)

2.3 Results

Of the 21 selected articles 12 studies used the LCM-p to investigate the bipolar mood course and the effect of clinical predictors (e.g. age of onset, alcohol use, bipolar subtype, menstrual cycle) on the course, 6 studies investigated medication effects on the mood course, and 3 studies investigated the utility/validity of the LCM-p in clinical trials. These 21 studies are based on (subsets of) data collected in 7 different study populations.

A combination of the clinician and patient version of the LCM-p was most often used in the selected studies (76%). In studies using both versions the clinician version of the LCM-p was based on the patient self-rated LCM-p and only the data from the clinician version were used in the main analyses. Further in the majority of the studies the compliance was not reported, however based on the ones that did report this it became clear that compliance rates over the observed period varied strongly between 10% and 91% (Table 2.1). Goldberg et al. (73) mentioned worsening of affective symptoms as a reason for drop-outs. In the study by van Zaane et al. (74) reasons for drop-outs are well documented and reported, with the most important reason being aversion against daily registration of the LCM-p. Further, it is notable that studies that reported higher drop-out rates used the patient version (self-report) rather than the clinician version of the LCM-p (73-75).

In Table 2.1 the different LCM-p course variables used in the different studies are also depicted. Table 2.2 describes the 4 most frequently used course variables of the LCM-p data across the included 21 studies, and the most common definitions of these variables. Most studies used several different course variables within one study. In the subsequent sections, we will describe the use of these variables in detail, as well as their advantages and disadvantages. Figure 2.2 (a fictitious life chart mood course) aims to illustrate the effects of different course variable definitions.

	Used in studies N (%)	Most commonly used definition	Advantages	Disadvantages
Number of manic and depressed episodes (DSM criteria)	9 (43%)	Depressed episode requires a duration of 2 weeks, a hypomanic episode requires at least 4 days without functional impairment, and a manic episode requires 1 week with functional impairment.	Represents global cycling pattern Can be used to distinguish different course types	May underestimate the number of episodes Short episodes with clinical relevance are missed No specific criteria for euthymia duration between two mood episodes of same polarity
Number of manic and depressed episodes (leapfrog rule)	8 (38%)	A manic episode requires 1 day of moderate or severe mania, a hypomanic episode requires 2 days of mild mania, and a depressive episode requires 2 days of moderate depression or 1 day of severe depression. Episode ends at switch or 2 week euthymia.	Prevents counting false-positive episodes after ultra-brief periods of euthymia Represents detailed cycling pattern, with shorter episodes Can be used to distinguish different course types	May overestimate the number of episodes due to short duration criteria for episodes.
Proportion of time impaired	12 (57%)	Number of days or percentage of time with mania/depression related impaired during the observed period.	Little data manipulation Represents disorder impact over specific period Consistent definition across literature	No consensus in handling of mixed states Does not provide information on cycling pattern
Average severity	14 (66%)	Mean level of severity of depression and mania during the observed time period.	Little data manipulation Represents impairment severity over specific time period Consistent definition across literature	No consensus in handling of mixed states Does not provide information on cycling pattern
Mood switch	6 (29%)	Either any switch into a sustained mood period, or any crossing of the midline on the LCM.	Represents mood instability Identification of rapid cycling patterns	No consensus about definition of switch Switches into very brief periods of mood episodes might be clinically less relevant

Table 2.2 | NIMH Life Chart outcome variables.

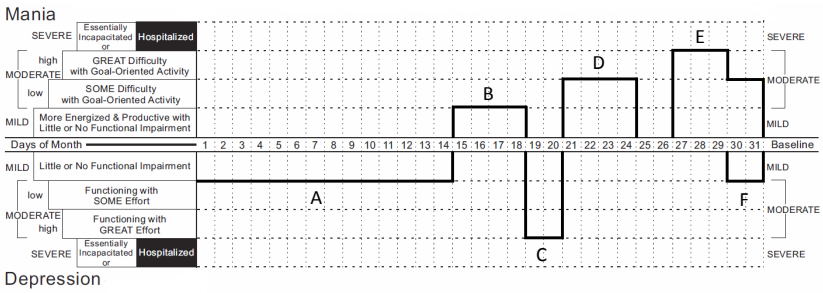

Figure 2.2 | Fictitious mood course on the prospective LCM.

2.3.1 Course variable: number of mood episodes

The number of episodes over the observed period is the most often used LCM-p course variable. However, the criteria for a LCM-p depressive and manic mood episode are not consistent across the literature. Basically there seem to be two different methods to define an episode, the first follows the DSM (DSM-III-R of DSM-IV) duration criteria and the second is the so called 'leapfrog rule' (62). In total 13 studies used number of episodes as an course variable, with 5 studies applying the DSM-IV criteria (76-80), 4 studies the leapfrog rule (30, 62, 81, 82), and 4 studies both the DSM-IV and the leapfrog criteria (64, 67, 74, 83).

Mood episodes based on DSM criteria

The DSM based method (DSM-III-R or DSM-IV) is the most frequently used method. However the scoring of the LCM-p is not based on DSM bipolar episode criteria, but on the severity of mood associated functional impairment. This means it is not possible to define a truly DSM defined bipolar mood episode from the LCM-p. Therefore researchers combined the DSM criteria for the duration of mania, hypomania or depression with severity criteria derived from the LCM-p ratings. However, the actual criteria that have been used across the studies are somewhat different as is depicted in Table 2.3. Further, not all authors specified how they exactly applied DSM criteria to the LCM data (77, 79).

According to the DSM-III-R or DSM-IV duration criteria, a depressive epi-

Study	Depression	Hypomania	Mania	
Kupka et al., 2005; Kupka et al 2007; Post et al. 2010; Zaane et al. 2010	≥ 2 weeks ≥ low moderate	≥ 4 days mild	≥ 1 week ≥ low moderate	
Post et al. 2003	≥ 2 weeks ≥ low moderate	≥ 4 days mild	unclear	
Post et al. 2001	-	≥ 1 week mild	'Sustained period' ≥ low moderate	Additionally distinguished shorter duration (hypo-) manias
Leverich et al., 2006	-	≥ 1 week mild	≥ 2 days ≥ low moderate	Additionally distinguished shorter duration (hypo-) manias

Table 2.3 | Applied duration and severity criteria for DSM definitions of mood episodes.

sode has a duration of at least 2 weeks, a hypomanic episode at least 4 days and a manic episode of at least 1 week. Further, the DSM-IV requires that a hypomanic episode is not accompanied by functional impairment (comparable with severity level 1 (mild) on the LCM-p), while manic and depressed episodes do affect functioning (comparable with severity level ≥2 (low moderate) on the LCM-p). So applying the criteria for a depressed episodes (≥ 2 weeks with ≥ mild moderate symptoms) period A is not considered a true DSM defined depressive episode, as the severity criterion is not fulfilled.

Most authors (64, 74, 80, 83) applied the above mentioned DSM criteria to calculate mood episodes. However, in the studies by Post et al. (76) and Leverich at al. (78) different duration criteria are applied to hypomanic and manic episodes. Both studies investigated the effects of antidepressants on mood changes from depression into (hypo-)mania. For this purpose also more subtle and brief periods of hypomanic symptoms of at least 1 day were relevant to evaluate based on a method previously proposed by Angst et al. (84). To be able to clearly distinguish the brief hypomanic episodes from full duration episodes, the duration criteria for a true hypomanic episode were set at 7 days instead of 4 days. Further, in these two studies manic episodes are rather based on severity criteria (≥ low moderate) than on strict duration

criteria.

So when these different criteria are applied to the mood course in Figure 2.2, the period B would be defined as a true DSM hypomanic episode by the studies that applied the 4 days criteria but would not be counted as such in the studies by Post et al. and Leverich at al. in which this would be counted as a brief hypomania. Additionally, because of the shorter duration criteria of Post and Leverich period E in Figure 2.2 would be counted as a manic episode, while according to the DSM duration criteria this short period with substantial manic symptoms and clinical relevance would not be counted as a true DSM manic episode.

Mood episodes: the leap-frog rule

Besides using the DSM-IV duration criteria as a method for defining episodes, another method has also been widely used. This method is called the 'leapfrog rule', which is suitable only for the daily LCM-p (62). This method does not only set a duration criterion, but also a severity criterion. According to this rule, mood fluctuations are not counted as an episode unless one of the following criteria are met:

1. a *manic* episode requires at least 1 day of moderate or severe mania;
2. a *hypomanic* episode requires at least 2 days of mild mania;
3. a *depressive* episode requires at least 2 days of moderate depression or at least 1 day of severe depression.

   This means that LCM ratings of mild depressed mood in the absence of ratings of at least moderate depressed mood within the same episode as depicted in period A in Figure 2.2 are not counted as an episode, while two days of moderate depressed would be considered an episode as depicted in period C.
4. In addition to the above, an episode is considered ended when the patient switches polarity or when there are 2 weeks of euthymic mood. In case of fewer than 2 euthymic weeks but longer than the longest contiguous duration of the adjacent mania or depression, the episode is

> also considered ended. Consequently, if the euthymic period is less or equal to the longest duration of adjacent depression or mania ratings, then the episode length is extended from the previous episode into the subsequent one, but only when the set of mood ratings does not change in polarity. This prevents false positive episodes when there are only (ultra-)brief periods of euthymia.

So when applying this rule to Figure 2.2, period D and period E should be considered as one single manic episode. By applying this last euthymia duration criterion, the leapfrog rule claims to be more conservative and prevent overestimation of the number of episodes in comparison to applying only the DSM-IV duration criteria.

Differences in course between DSM-IV duration criteria and leapfrog rule

Applying different definition criteria for LCM-p episodes can lead to differences in the number of episodes. Kupka et al. (83) and Van Zaane et al. (74) used in their studies both the leapfrog rule and the DSM duration criteria and reported the differences in course. For the DSM episodes they used the following criteria: a manic episode of at least 7 days and at least (low) moderate severity or hospitalization, a hypomanic episode of at least 4 days and mild severity, and a depressive episode of at least 14 days and at least (low) moderate severity or hospitalization. Their results clearly illustrate the differences between the two methods. In Kupka's study, in which the mood courses of rapid cyclers and non-rapid cyclers were compared, it is demonstrated that applying the leapfrog rule resulted in the detection of on average 9 additional (in addition to DSM defined episodes) mood episodes in the rapid cycling group. In Van Zaane's study applying the leapfrog rule resulted in 2.5 more depressive episodes, 1.5 more hypomanic episodes and 6.4 more manic episodes compared to the number of the DSM defined episodes. This does not support the widely assumed assertion of the leapfrog rule to being more strict and preventing overestimation of the number of episodes (62), but this method does seem to detect more subtle and short mood changes which remain undetected when DSM criteria are applied.

Conclusion number of episodes

In light of the absence of a gold standard methodology, it remains unclear whether the lenient duration criteria of the leapfrog method likely overestimate the number of episodes, or whether the more strict DSM-IV criteria underestimate the number of episodes. What is clear is that the two methods give different perspectives with respect to frequency ratios of the occurrence of different mood states. Based on the research question at hand researchers should decide whether DSM mood episodes are informative or that also brief episodes and subtle mood changes are relevant, for instance when studying (ultra-) rapid cycling patterns (e.g. 82, 83) or when studying the short term (hypo-)mania inducing effects of antidepressants (76, 78).

2.3.2 Course variable: proportion of time ill/impaired

A second indicator of course severity that is used across the literature is the proportion of time ill or impaired. This is defined either as the number of days ill/impaired or the percentage of time ill/impaired during the observed period. Methods of calculating these two course variables are used rather consistently across the literature and simply consist of counting the number of days or the percentage of time (number of days with reported impairment divided by number of days observed) for manic or depressed mood state separately. When calculating the proportion of time impaired for the mood course in Figure 2.2, this results in 13 of the 31 observed days with mania related impairment (proportion 42%) , 18 of the 31 observed days with depression related impairment (proportion 58%), or a total proportion of overall time impaired of 29 days (proportion 93.5%). In some studies the percentage of time impaired was presented for every severity level of the LCM (e.g. 64, 67, 82).

Only when scoring mixed mood states (see Figure 2.2, period F) some inconsistency emerges. The NIMH Life Chart Manual (85) states that in case of mixed mood states the rated mania score is included in the average mania score and the depression score in the depression average. However, Denicoff et al. (86) used the rule that when patients experienced manic and depressive symptoms simultaneously, and met DSM III-R criteria for mania or

hypomania, the day was rated as manic. This means that the rated depressive symptoms of the mixed mood state are not included in the proportion of time impaired calculations.

There are several advantages when using this course variable. First, it provides insight in the long term impact of the disorder in terms of proportion of time that the patient is disabled by the disorder, regardless of whether there occurred strict DSM-IV episodes. Second, methods to calculate these variables are consistent across studies, which improves comparability of these studies. Third, it provides a clear picture of the longitudinal impact of the disorder on daily functioning. There are also some disadvantages however. The cycling pattern is not reflected in this variable, since patients with rapid mood switches cannot be distinguished from patients with prolonged mood episodes when both are impaired for the same proportion of time. However, in some studies the proportion of time ill/impaired is used as a method to distinguish different course types within the patient groups. Post et al. distinguished three different course types (minimally impaired, episodically ill, and chronically ill) based on proportion of time ill/impaired on the LCM-p in combination with a description of their specific cycling pattern (67, 82). Still, this method is most suitable for a more global descriptions of course severity, for instance when long term medication effects are measured (73, 81) or when the effects of specific clinical predictors on overall course severity is measured (e.g. 66, 74, 80, 87).

Conclusions considering proportion of time ill/impaired

Considering the fact that this method requires little manipulation of the LCM-p data and relative consensus exists about the calculation of number of days or percentage of time impaired, this method allows for comparison with previous studies that used this method. Reporting of percentage of time ill/impaired might be preferred over reporting number of days impaired, since it is easier to interpret for the reader. Reporting the proportion of time impaired reflects a more global measure of disease severity, but does not provide specific information about cycling patterns.

2.3.3 Course variable: average severity of illness

The LCM-p assesses the severity of functional impairment due to mood symptoms for a certain day. In research settings, measuring/reporting of illness severity over longer periods of time is required. Calculation of this variable seems rather consistent across the literature. The average of the weighted severity measured of the reported impairment over an certain time period of the LCM-p is used according to the following formula based on Denicoff et al. (62):

[(no. days of severe depression × 10) + (no. days of moderate high × 7.5) + (no. days of moderate low depression × 5) (no. days of mild depression × 2.5)] / # of days observed.

[(no. days of severe mania × 10) + (no. days of moderate high mania × 7.5) + (no. days of moderate low mania × 5) + (no. days of mild (hypo)mania × 2.5)] / # of days observed.

For the mood course in Figure 2.2 this would lead to the following calculation for depression:

(0 × 10) + (2 × 7.5) + (0 × 5) + (16 × 2.5) = 55/31 = 1.8

and for mania:

(0 × 10) + (3 × 7.5) + (6 × 5) + (4 × 2.5) = 62.5/31 = 2.0 Other studies not only evaluated mean LCM-p severity scores for each pole (manic/depressed), but also overall impairment scores by averaging the scores of both poles (30, 74). Furthermore, averages can be calculated over the total observed period, or over specific time periods, such as weekly or monthly averages (62, 63). Regarding mixed episodes, Denicoff et al. (86) used a different method (as described in section 3.2) than was proposed in the NIMH LCM manual. As is described before, when patients experienced manic and depressive symptoms simultaneously, those days should be rated as mania but not depression.

Conclusions considering average severity

Calculating the average severity requires little manipulation of the LCM-p data and calculation methods are consistent among studies. Besides, this

method also allows for a clear representation of the severity of impairment over an observed period and allows for a distinction between impairment severity among patients. This method is comparable to the proportion of time ill/impaired in the sense that it provides a more global severity measures of the disease course. However, again this method does not allow for determining cycling patterns.

2.3.4 Course variable: mood switches

Some researchers were especially interested in the number of mood switches, defined as an immediate change from a depressive episode into a hypomanic or manic episode or vice versa (75, 76, 78). This course variable is mainly used in studies that focussed on the potential risk of antidepressants to induce switches form a depression into hypomania or mania (76, 78). A difficulty in these studies is how to define switch, i.e. how to define an episode from the LCM-p data as discussed above.

This is nicely illustrated by the study of Leverich et al. (78). In addition to switches into full manic episodes (hypomania: ≥7 with mild severity; mania: ≥ 2 days with ≥ moderate severity), they also looked at switches into shorter duration episodes (brief hypomania: ≥ 1 day mild severity). This allows for a distinction in severity of switch (e.g. subthreshold switch into brief episodes versus threshold switches into full duration hypomania or mania) and for a detailed representation of mood instability.

In the study by Post et al. (76) the above mentioned criteria were also applied. They reported that 14% of the patients had switches into full hypomanic or manic episodes, but an additional 18% had ultradian switches or switches into brief hypomanic episodes. This illustrates the additional information about mood instability that can be gained by including subthreshold switches.

Further, besides switch into DSM episodes or shorter episodes a third definition of switch is used, which is any change in mood polarity or any crossing of the midline on the LCM-p (not including ultradian switches) (65, 75, 86).

The different used duration criteria obviously will lead to different numbers of switches within the same course. When the full episode criteria for hypo-

mania and mania of Leverich et al. and Post et al. are applied to Figure 2.2, there is no mood switch in this course. However, when the shorter duration criteria are applied, there are two switches into hypomania: from period A (depression) to period B (hypomania) and from period C to period D. When a switch is defined as any crossing of the midline and also switches into depression are included, three switches in Figure 2.2 can be identified (from period A to period B, from period B to period C, from period C to period D,), indicating that applying different criteria also lead to different number of switches within the same patient.

Conclusions considering mood switches

Switch rate is a course variable that is mainly used in studies measuring medication effects on the clinical BD course. In general, for the calculation of switches different criteria than DSM criteria are used. However, no consensus seems to exist over what exactly should be considered a mood switch. Very rapid mood changes (any crossing of the midline) might be a good indicator of mood instability and rapid cycling patterns. But, as Leverich et al. (78) already mentioned, mood switches into a more sustained period of mood symptoms might be clinically more relevant, since very brief switches might be considered less problematic since their impact on daily functioning might be minimal.

2.4 Discussion

This review reveals the challenging issues when using course variables derived from the NIMH LCM. These course variables can be used in research, but also when interpreting the rich data provided by this chart in clinical care. The NIMH LCM is first and foremost an important clinical tool for monitoring the bipolar mood course and a valuable resource for both clinician and patient to gain more insight in the longitudinal mood course of the patient and its possible correlates. The current review shows that this tool is also widely used in research.

Both the clinician rated version of the LCM-p and the patient rated version are now well-validated (61, 88). Both version are suitable for long-term mon-

itoring of the mood course and both have their advantages and disadvantages. The benefits of the clinician version is that clinicians can perform the prospective ratings every time they see the patients, and irrespective of whether or not patients have completed their own ratings. Consequently, clinician ratings are more likely to be complete, although probably less reliable when not using the ratings by the patients themselves, especially in cases with the rating of longer intervals. However, rating of the clinician version is time-consuming and therefore costly in both clinical and research settings. With respect to the patient version, a recent validation study (88) shows good validity and reliability of this version, suggesting this version to be a good alternative to the clinician version. However, especially daily self-rating might be rather burdening to patients and may lead to higher drop-out rates than the clinician version (89). Alternatively, electronic versions of the life chart might lead to more compliance (90).

Further, the main issue we address in this review is the fact that the LCM-p is used to measure different course variables which are differently defined. The most commonly used measures are: number of mood episodes, proportion of time impaired, average severity of impairment, and number of mood switches. In most studies a combination of these course variables are used.

The course variables which have the most consistent definitions across the literature are: proportion of time impaired/ill and average severity of illness. Calculating these course variables requires minimal manipulation of the raw LCM-p data and they reflect global impact and severity of the disorder over an observed period of time. A benefit of the calculation of proportion of time ill is that it allows for a distinction between patients that have an equal number of episodes, but very different duration of time in certain mood states. A disadvantage of these variables is that they do not reflect cycling patterns or overall mood instability. Both average severity and proportion of time ill/impaired are therefore often used in combination with other LCM-p variables, such as number of episodes or switch rates.

Number of episodes based on the LCM-p is the most frequently used variable across the reviewed literature. However, at the same time several different definitions are used to calculate the number of episodes. Especially on the duration criteria for episodes no strict consensus exists. Most studies used

duration and severity criteria based on the DSM-III-R or DSM-IV. However, also other criteria are used such as the leapfrog rule and the modified duration criteria (76, 78) of Angst. These two methods use shorter duration criteria for episodes, leading to higher numbers of episodes counted. The relevance of additionally including brief mood episodes is demonstrated by the studies of Angst et al. (84, 91) showing that brief episodes are part of the clinical reality of the course of bipolar disorder and that the occurrence of episodes is not restricted to DSM duration criteria. Consequently, DSM duration criteria might be useful to detect more general cycling patterns and full-duration episodes, with the risk of underestimating number of mood episodes. As such, the shorter duration criteria are suitable for the detection of rapid cycling patterns, with the risk of overestimating the number of episodes. The leapfrog rule might be most suitable for the detection of brief episodes, since it minimalizes the risk for overestimation of number of episodes by setting more strict criteria to what constitutes an euthymic period.

For the calculation of number of switches the same problems arise which apply to the calculation of mood episodes. Again, when very rapid switches are included, mood instability might be overestimated. However when stricter duration criteria are applied, clinically relevant mood changes might be overlooked.

The current review has several (potential) limitations. While effort was made to uncover all relevant publications on the LCM in the English language, some relevant studies may have been missed. Moreover, this review was limited to an evaluation of the prospective NIMH LCM and did not include an overview of the use of the retrospective LCM as well as other charting methods. The strengths of the current review is that it is the first to describe the different interpretation methods of the LCM-p when using it as an course measure which may help to increase comparability across future studies.

In conclusion, our systematic review showed that the LCM-p is not only a clinical useful tool (not discussed as such in this paper), but is also a challenging but invaluable and widely used method in longitudinal research. Strict recommendations on how to use the LPM-p in course research are hard to give. Since there exists strong variety in methods to analyse and report LCM-

p data it is important that researchers report the exact method they applied to their LCM-p data and motivate why this is their method of first choice.