



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SciVerse ScienceDirect

journal homepage: [www.elsevier.com/locate/jval](http://www.elsevier.com/locate/jval)

## SYSTEMIC REVIEW

## The Use and Performance of Productivity Scales to Evaluate Presenteeism in Mood Disorders

Nicolas Despiégel, MSc<sup>1\*</sup>, Natalya Danchenko, PhD<sup>2</sup>, Clément François, PhD<sup>2</sup>, Benedikte Lensberg, MSc<sup>3</sup>, Michael F. Drummond, PhD<sup>4</sup>

<sup>1</sup>OptumInsight, Nanterre, France; <sup>2</sup>Lundbeck, Issy-les-Moulineaux, France; <sup>3</sup>OptumInsight, Uxbridge, UK; <sup>4</sup>University of York, York, UK

## ABSTRACT

**Objective:** Mood disorders are associated with a high societal cost, mainly due to presenteeism. The objective of this study was to review the use of 10 instruments that rate presenteeism in mood disorders and to provide recommendations regarding the appropriateness of instruments in different study settings. **Methods:** A systematic review of the literature was conducted to identify scales used to measure presenteeism, including the World Health Organization Health and Work Performance Questionnaire, the Lam Employment Absence and Productivity Scale, the Sheehan Disability Scale, the Work Limitation Questionnaire, and Work Productivity and Activity Impairment questionnaire. Study characteristics and major results (by symptom level, by treatment arm, correlation to other scales, and use of monetization) were data extracted. **Results:** Twenty-nine studies were identified. The Sheehan Disability Scale, the Work Limitation Questionnaire, and Health and Work Performance Questionnaire were the most commonly used instruments. The majority (60%) of scales demonstrated higher presenteeism in individuals with mood disorders than in individuals without. The Lam Employment Absence and

Productivity Scale, the Sheehan Disability Scale, and the Work Limitation Questionnaire showed that presenteeism increased with increasing severity of disease. Few studies reported results on presenteeism by treatment, with only small between-treatment differences observed. Good correlations between presenteeism instruments and clinical or quality-of-life scales were reported. Three studies converted results from presenteeism scales into monetary units. **Conclusions:** Limited experiential evidence exists comparing the performance of presenteeism scales in mood disorders. Therefore, recommendations for inclusion of a presenteeism tool must be driven by instrument properties (ease of administration, amenability to monetization) and the study type. Future research should focus on the responsiveness of the instrument and on how mood disorders impact self-reported assessment.

**Keywords:** mood disorders, presenteeism, productivity, self-report.

Copyright © 2012, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

### Introduction

Mood disorders include major depressive disorder in which patients experience one or several depressive episodes or bipolar disorder, characterized by intermittent episodes of mania or hypomania, usually interspersed with depressive episodes. These conditions are associated with a high societal cost, primarily due to productivity losses [1]. Mood disorders lead to higher unemployment, absence, and turnover rates as well as higher at-work performance deficits (or presenteeism). *Presenteeism* is defined as “the decrease in productivity for the much larger group of employees whose health problems have not necessarily led to absenteeism and the decrease in productivity for the disabled group before and after the absence period” [2].

For some disorders, presenteeism is an even greater cause of productivity loss than absenteeism (i.e., migraine, seasonal allergies). Although the cost of presenteeism is not routinely estimated in economic evaluations, it is estimated to account for 54% to 82% of the total lost productivity in employees with mood

disorders, as observed in US workers by using the World Health Organization Health and Work Performance Questionnaire (HPQ) or the Work and Health Interview [3,4]. These values, however, may overestimate the true impact because they do not account for the hiring and training of replacement workers. The relative importance of presenteeism compared with absenteeism in this disease area is likely because individuals with depression or anxiety tend to stay at work and perform suboptimally rather than take sick leave [4,5].

Productivity loss due to absenteeism is often taken into account in economic evaluations that adopt a societal perspective and is measured simply by counting the number of days off work; measuring productivity loss due to presenteeism, on the other hand, is more complex. First, the evaluation of presenteeism requires the estimation of a “normal productive output” for a given individual in a given role, after which the impairment in productive output may be quantified [6]. Second, reduction in symptoms reduces absenteeism, but its impact on presenteeism is more uncertain [7,8]. Third, depression affects productivity

\* Address correspondence to: Nicolas Despiégel, OptumInsight, 41 rue des trois fontanot, 92000 Nanterre, France.

E-mail: [Nicolas.despiegel@optum.com](mailto:Nicolas.despiegel@optum.com).

1098-3015 – see front matter Copyright © 2012, International Society for Pharmacoeconomics and Outcomes Research (ISPOR).

Published by Elsevier Inc.

<http://dx.doi.org/10.1016/j.jval.2012.08.2206>

differently depending on the occupational status, with various aspects of productivity more typically impaired in different professions [9]. For example, in jobs in which individuals have to exercise judgment such as nurses, engineers, social workers, marketing managers, attorneys, and financial analysts, depressed workers may experience limitations on output demands (handle workload, work fast enough, meet simultaneous demand, etc.) and be more absent. In jobs with a high degree of contact such as teachers, customer service managers, sales people, and consultants, workers may have limitations in handling mental (keep mind on work, think clearly, do precise work, handle demanding/stressful work, etc.), interpersonal (speak on the phone, communicate well, maintain contacts, etc.), and physical (lift/carry/move objects, use handled tools/equipment, get to work from parking/bus/train, etc.) demands. Fourth, the impact on presenteeism is directly related to the nature of the depressive symptoms presented by the patient. According to Lerner et al. [9], concentration difficulties and distractibility lead to lower overall productivity, whereas tiredness and sleep disturbance induce higher absence and problems with mental, interpersonal, time (getting to work, work without breaks or rests, adjusting to work pace changes, etc.), and output demands. Finally, while self-report instruments have been developed to evaluate presenteeism, they run the risk of presenting biased and different results compared with objectively measured time lost at work. Stewart et al. [10] demonstrated that self-reported time spent working was higher than that which was obtained when evaluating official workplace data on time absent, time away from desk, and electronic continuous performance data specific to the workplace. Other researchers state that self-reported evaluations might lead to an overestimation of lost productivity due to presenteeism [11] because subjective feelings of high discomfort may cause employees to report lower productivity even if their tasks have been completed successfully. Moreover, depression can influence the self-reporting of productivity because of a loss of concentration, attention, and/or motivation [12] or cognitive deficits [13].

The selection of the appropriate self-report instrument should be based on both its ability to translate health states into at-work productivity estimates and its relevance to the study setting and objectives, that is, whether the objective is to assess the impact of the disease on aspects of work performance or to estimate the economic consequences of the disease. The optimal self-report instrument needs to be sufficiently sensitive to detect the effects of treatment interventions.

Several reviews have previously been published with the aim of evaluating instruments used to assess presenteeism. Prasad et al. [14] published an extensive review of the psychometric properties of productivity instruments, assessing their validity, reliability, responsiveness, generalizability, and ease of administration. Evidence on psychometric properties from this and other studies has been provided to various degrees depending on the nature of the instrument. Prasad et al. concluded that the Work Limitation Questionnaire (WLQ) and the Work Productivity and Activity Impairment (WPAI) questionnaire offer the most significant advantages with good psychometric properties, being ready, and easy to use in various study settings (e.g., clinical studies and employee populations). Lofland et al. [15] and Mattke et al. [16] assessed the psychometric properties of available instruments and the ability of each instrument to provide monetary estimates. The authors of both studies concluded that the main hurdle currently relates to the absence of established and validated methods for monetization of productivity estimates. A more recent review by Brooks et al. [17] discussed several issues related to measurement. The authors emphasized the limitations related to the conversion into economic outcomes and advocated against evaluating productivity at an individual level. Lerner and Henke [12] assessed the impact of depression on lost productivity, reviewing the four most widely used self-report instruments to

assess presenteeism in this disease area and their use in various settings (e.g., population-based studies, workplace studies, and clinical studies). They concluded that the WLQ was the most appropriate instrument for accurately measuring work productivity in individuals with depression.

Because of the importance of productivity losses in mood disorders and the challenges for their estimation, the selection of the appropriate tool to measure presenteeism is essential. Because three of the existing reviews did not relate to mood disorders and the fourth, considering depression, reviewed only a subset of the available instruments, the objective of this article was to review and discuss the scales available for assessing presenteeism related to mood disorders. We extend the work by Lerner et al. [12], by including six additional instruments and more recently published articles of randomized clinical trials (RCTs) and observational studies. We also provide recommendations on the use of productivity scales in the area of mood disorders.

## Methods

### Identifying Instruments That Evaluate Presenteeism

Taking the four most recent reviews detailing rating scales measuring presenteeism, we evaluated all reported scales, all of which were self-report instruments. In addition to the initial pool of 20 scales, a more recent scale not mentioned in any of the four reviews and a disability scale with a presenteeism component were evaluated. Instruments were selected according to the following inclusion/exclusion criteria (Table 1). All scales should be generic or specific to mood disorders, demonstrate (at least partially) good psychometric properties, and be applicable to any type of worker. It should be possible to use these instruments to measure presenteeism in studies such as RCTs or prospective observational studies; nongeneral instruments or techniques (applicable to a specific study design or a specific worker category) were excluded. Generic instruments that were used in a limited number of disease areas other than mood disorders were also excluded. Ten instruments were finally considered. Among the 10 excluded instruments, 7 were specific to a disease area other than mood disorders.

Two categories of instruments were defined: (1) instruments that are amenable to monetization (with specific development regarding this matter or with published experience on conversion into monetary units) and (2) instruments that cannot be used to monetize productivity loss. This article briefly addresses the major characteristics and psychometric properties, as well as the methods for monetization. Full details regarding the major domains of the scales, how each instrument assesses presenteeism, major properties, possibility for conversion into monetary units, and use reported in the literature are provided in Table 2 [10,18–29].

### Health & Labour Questionnaire/Short Form Health & Labour Questionnaire

The Health & Labour Questionnaire (HLQ) gathers data regarding reduction in work performance due to illness [28,29]. It consists of four modules that assess (1) workplace absenteeism, (2) workplace presenteeism, (3) unpaid work, and (4) impediments to paid and unpaid work. Workplace presenteeism is measured as the number of additional hours that should have been worked to compensate for production losses due to illness at work. Individuals also provide responses to questions designed to determine specific productivity problems (e.g., concentration difficulties) related to presenteeism. A short-form version of the HLQ (SF-HLQ) has also been developed and comprises three modules: absenteeism from paid work, production losses without absenteeism from paid work, and hindrance in the performance of paid and unpaid work [19].

**Table 1 – Selection of presenteeism instruments.**

Instruments	Selected in the review	Generic	Specific to mood disorder	Specific to other disease areas	Generic but tested in a limited number of other disease areas	Not applicable to any employee	Not applicable to studies in mood disorders
ALWQ				Angina			
EWPS			X				
HAQ/HAQ-II				Rheumatoid arthritis			
HLQ/SF-HLQ		X					
HPQ	X	X					
HRPQ-D					Parkinson's disease, infectious mononucleosis		
HWQ	X	X					
LEAPS			X				
MIDAS				Migraine			
MWPLQ				Migraine			
Osterhaus technique				Migraine			
SDS	X	X					
SPS	X	X					
Unnamed hepatitis instruments				Hepatitis			
WHI	X	X					
WIS				Rheumatoid arthritis and traumatic brain injury			
WLQ	X	X					
The WPAI questionnaire	X	X					
WPI		X				Call centers	
WPSI							Used to inform on the relative importance of health conditions that affect productivity at work for a large group of employees

EWPS, Endicott Work Productivity Scale; HLQ/SF-HLQ, Health & Labour Questionnaire/Short-form Health and Labour Questionnaire; HPQ, Health and Work Performance Questionnaire; HWQ, Health and Work Questionnaire; LEAPS, Lam Employment Absence and Productivity Scale; MIDAS, Migraine Disability Assessment; MWPLQ, Migraine Work and Productivity Loss Questionnaire; SDS, the Sheehan Disability Scale; SPS, Stanford Presenteeism Scale; WHI, Work and Health Interview; WLQ, Work Limitations Questionnaire; WPAI, Work Productivity and Activity Impairment; WPI, Worker Productivity Index; WPSI, Work Productivity Short Inventory.

**Table 2 – Summary of productivity measurement instruments.**

Instrument	Assessments	Assessment and scoring: presenteeism	Recall period and estimated time to complete	Psychometric properties				Option for monetization
				Content validity	Construct validity	Reliability	Responsiveness	
<i>Instruments amenable to monetization</i>								
Health & Labour Questionnaire (HLQ)	Workplace absenteeism, workplace presenteeism, unpaid work, impediments to work and unpaid work	[Employed] level of impediment while working on a 0–10 scale	Recall: 2 wk Number of questions: 10 items (short-form version) Time to complete: Not reported		Convergent discriminant			Specified
Health and Work Performance Questionnaire (HPQ)	Presenteeism, absenteeism, accident/injuries	Absolute presenteeism: actual performance on a 1–100 scale  Relative presenteeism: ratio of actual performance to the performance of most workers at the same job, as reported by the respondent	Recall: 7 d or 4 wk Number of questions: 13 items (7-d RCT version) Time to complete: Not reported	Yes Cognitive debriefing	Convergent discriminant	Internal		Specified. Absolute measures should be used because relative score of presenteeism can be >1
Health and Work Questionnaire (HWQ)	Productivity, impatience/irritability, concentration, work satisfaction, satisfaction with supervisor, personal life satisfaction	Efficiency, quality, and amount of work completed over the recall period on a 0–10 scale  Indication of how a supervisor and a coworker would respond to the same questions about the respondents' productivity	Recall: 1 wk Number of questions: 27 questions Time to complete: Not reported		Convergent	Internal		HWQ responses could be interpreted as a percentage reduction in normal productivity, though there is no example of monetization of this scale in the recent literature
Work and Health Interview (WHI)	Absenteeism, presenteeism, time spent caring for ill family members, salary from paid employment	Concentration loss while unwell, repeated a job while unwell, worked more slowly while unwell, felt fatigue while unwell. % of effectiveness while ill (from categorical responses on presenteeism)	Recall: 2 wk Number of questions: 25 questions Time to complete: Not reported		Convergent			Specified. But defined from the conversion of categorical response into an effectiveness score
Work Limitations Questionnaire (WLQ)	Time management, physical, mental, interpersonal, output	Average of the four domains to productivity-related questions and converting the average to an interval scale to achieve a 0–100 scale of productivity	Recall: 2 wk Number of questions: 25 questions Time to complete: Not reported	Yes Content developed including depressed patients Cognitive debriefing	Convergent discriminant	Internal		Not explicitly specified. Would be defined from the conversion of categorical response into an effectiveness score
Work Productivity and Activity Impairment (WPAI) questionnaire	Absenteeism, presenteeism, usual activities impairment	Self-reported work performance is assessed with a 0–100 visual analogue scale	Recall: 7 d Number of questions: Six questions Time to complete: Not reported	Cognitive debriefing	Convergent discriminant	Test-retest	Yes	Specified
Lam Employment Absence and	Absenteeism, productivity impairment	Total impairment score on a 0–28 scale (level of energy, concentration,	Recall: 2 wk Number of questions: 10 items		Convergent discriminant	Internal		Yes, using a subscore including the three Likert items

Table 2 – (continued)

Instrument	Assessments	Assessment and scoring: presenteeism	Recall period and estimated time to complete	Psychometric properties				Option for monetization
				Content validity	Construct validity	Reliability	Responsiveness	
Productivity Scale (LEAPS)	related to mental health-related problems	anxiety). This scale could be interpreted as a presenteeism measure						related exclusively to problems at work
<i>Instruments not amenable to monetization</i> Endicott Work Productivity Scale (EWPS)	Behaviors, subjective feelings or attitudes	Frequency of productive behaviors during the previous 1 wk using a five-point Likert scale. A sum of scores is computed, ranging from 0 (best score) to 100 (worst score)	<u>Time to complete:</u> 3–5 <u>Recall:</u> 1 wk <u>Number of questions:</u> 25 items <u>Time to complete:</u> Not reported		Convergent discriminant	Internal test retest		No. No obvious link between the score and % productivity, which prevents obvious monetization
Sheehan Disability Scale (SDS)	Work/school impairment, social life impairment, family life impairment, number of days unable to attend work or school because of symptoms, number of days underproductive due to symptoms	Rate (0–10 scale) on how symptoms have disrupted work or school work Number of days “underproductive” due to symptoms	<u>Recall:</u> 1 wk, 1 mo, or 3 mo <u>Number of questions:</u> Five items <u>Time to complete:</u> Not reported		Convergent discriminant	Internal test-retest	Yes	No. The question on how symptoms have disrupted work or schoolwork includes absenteeism The number of days underproductive has no link with the previous assessment
Stanford Presenteeism Scale (SPS)	Stress, focus, energy at work	Stress, focus, energy at work, no global score combining these attributes	<u>Recall:</u> 1 mo <u>Number of questions:</u> Six items <u>Time to complete:</u> Not reported		Convergent discriminant	Internal		No

RCT, randomized controlled trial.

### World Health Organization Health and Work Performance Questionnaire

The HPQ is a self-report instrument designed to estimate the consequence of health problems on job performance, sickness absence, and work-related accidents and injuries [20]. Presenteeism as measured by the HPQ is a measure of actual performance in relation to possible performance. Respondents are asked to rate normal and impaired performance in their job on a scale of 1 to 10, understood as “percentage effectiveness,” for the purposes of scoring and monetizing presenteeism. Absolute presenteeism has a lower bound of 0 (or 0% effectiveness during hours worked) and an upper bound of 100 (or perfect effectiveness during hours worked). Relative presenteeism is measured as the ratio of actual performance to the performance of most workers in the same job. The instrument’s developers recommend restricting the measure of relative presenteeism from 0.25 to 2.0, where 0.25 is the worst relative performance (25% or less of other workers’ performance) and 2.0 is the best performance (200% or more of other workers’ performance).

### Health and Work Questionnaire

The Health and Work Questionnaire is a multidimensional measure of productivity. Its domains measure productivity, impatience/irritability, concentration, work satisfaction, satisfaction with supervisor, and personal life satisfaction [26]. Presenteeism measures the efficiency, quality, and amount of work completed over the recall period on a 0 to 10 scale. Respondents are also asked to give an indication of how a supervisor and coworker would respond to the same questions as to the respondent’s productivity.

### Work and Health Interview

The Work and Health Interview is a questionnaire that provides a measurement of productive time lost because of work absence and reduced performance at work [10,27]. Presenteeism is measured as the number of days in the previous 2 weeks the respondent worked while ill. Four questions require respondents to report how often, on average, they (1) lost concentration, (2) repeated a job, (3) worked more slowly, and (4) felt fatigued on days they attended work while feeling unwell. Response options are categorical (“all of the time,” “most of the time,” “half of the time,” “some of the time,” and “none of the time”). In Stewart et al. [10] these categorical responses were converted to percentages for monetization (“all of the time” = 100%, “most of the time” = 75%, through “none of the time” = 0%).

### Work Limitation Questionnaire

The WLQ is a 25-item questionnaire designed to measure the impact of chronic disease on work performance [23]. The 25 items of the WLQ are grouped into four scales: (1) time management, (2) physical, (3) mental-interpersonal, and (4) output. The scale scores represent the amount of time in the past 2 weeks that an individual was limited on the job. To create a score for each scale, the questionnaire’s categorical responses are converted to an interval scale. The average scores for items within each scale are summed, divided by the total number of scale items, and then multiplied by a factor of 25 (WLQ “Scale Score”):  

$$\text{WLQ Scale Score} = 25(\text{average item score} - 1)$$

The scale scores calculated from the four scale scores range from “0” (least limited) to “100” (most limited). A score of “50,” for example, indicates that the individual was limited in performing work demands during 50% of the reporting period. Another score, the WLQ “Productivity Loss Index,” has been developed [30] as the weighted sum of scale scores and retransformed into a WLQ “Productivity Loss Score,” which corresponds to “the estimated

percent difference in an employee’s at-work productivity compared to employees who do not have health-related work limitations.” This represents a percent reduction in output in the past 2 weeks compared with the output of a healthy (not limited) employee. The weights used are empirically based statistical estimates relating self-report scores to objectively measured productivity.

### Work Productivity and Activity Impairment

The WPAI questionnaire consists of six questions that ask respondents to identify the number of hours missed from work and usual activities, as well as the degree to which work or regular daily activities were limited over the past 7 days [24,25]. The questionnaire yields four scores: (1) percentage of work time missed because of ill health, (2) percentage impairment while working due to ill health, (3) percentage activity impairment due to ill health, and (4) an overall percentage work impairment score due to health problems. The four scores are expressed as “impairment percentages,” where higher numbers reflect greater impairment and decreased productivity.

### Lam Employment Absence and Productivity Scale

LEAPS is a 10-item self-report questionnaire developed to assess work functioning and impairment in a clinically depressed population [21]. It includes a component related to absenteeism (based on two items inquiring as to the number of scheduled work hours and work hours missed) and another related to presenteeism. The latter component consists of seven items rated on a five-point Likert scale (“none of the time” to “all of the time”), asking directly as to the most common problems encountered at work (“making more mistakes,” “doing poor quality of work”) and the clinical symptoms most associated with work impairment (levels of energy, concentration, and anxiety). The total score on the seven Likert items therefore yields a global impression of impairment due to mental health-related problems.

### Endicott Work Productivity Scale

The EWPS is a 25-item questionnaire created to assess impairment in employees with depression [18]. This instrument was designed specifically to capture information on productivity loss during clinical trials, measuring both absenteeism and presenteeism. The frequency of productive behaviors during the previous week is captured by using a five-point Likert scale (anchored from 0 = “Never” to 4 = “Almost always”), and a sum of the scores is computed, ranging from 0 (best score) to 100 (worst score). The instrument lacks any obvious means to convert employees’ scores directly into monetary units, limiting its function to the relative ranking of employees according to reported productivity impairment.

### Sheehan Disability Scale

Although not specifically designed to measure presenteeism, the Sheehan Disability Scale (SDS) assesses functional impairment in three areas: work/school, social, and family experiences. For each domain, a 10-point visual analogue scale with labels is used to quantify the level of functional impairment [22,31,32]. The three components of the questionnaire are then summed into a single score that ranges from 0 (unimpaired) to 30 (highly impaired). Two additional questions are used as part of the assessment: one asks respondents to estimate the number of days they were less than fully productive at work. It should be noted that there is no direct link between this question and the assessment of the functional impairment of activities at work/school. In other words, the SDS could potentially provide spurious results if



the 0 to 10 “level of impairment scale” were interpreted as a “% effectiveness” scale for the number of days of impairment reported. As a result, no monetization seems possible based on the SDS evaluation.

#### Stanford Presenteeism Scale

The Stanford Presenteeism Scale 6 (SPS-6) is a six-item questionnaire [33], with each question linked to a Likert five-item response scale ranging from “strongly disagree with the statement” to “strongly agree with the statement.” No clear indication is provided as to how the ordinal responses from individuals in the SPS could be monetized. The scale is intended to rank individuals according to impairment at work.

#### Instrument Validations

Prasad et al. [14], Lofland et al. [15], Mattke et al. [16], and Brooks et al. [17] have published extensive reviews on the reliability and validity of all available instruments measuring presenteeism, with the exception of the LEAPS scale. This scale, which was developed after these reviews were published, has since been formally validated [21] and is addressed in this article. In these reviews, information regarding content validity has not been extensively addressed. Some aspects related to the inclusion of patients with mood disorders during the scale development and the use of cognitive debriefing are presented further in this article. The summary of psychometric properties by instrument is presented in Table 2.

The face validity of each instrument was evaluated by checking the scale format and whether the recall period is appropriate for this population. The 10 instruments present different response formats, Likert scales being the most frequent. Some instruments include questions leading to complex calculations. For example, the HPQ asks the “Number of hours in the past 4 weeks” (B6) and includes detailed examples to guide the respondent. The recall period ranges from 1 week to 1 month, but more recent instruments tend to use a 2-week recall period. The main challenge is to mitigate the recall bias without adding too many additional assessments and increasing the burden of the study [17,34]. SDS is primarily a functionality scale and not an instrument to estimate presenteeism or productivity loss *per se*.

Three instruments (HPQ, WLQ, and the WPAI questionnaire) reported details on the scale development beyond the fact that item generation was based on a review of the literature. In particular, reported cognitive debriefing was part of the scale development that addressed one of the identified causes of bias in self-reports in this population. This supports content validity for these three instruments. In addition to the depression-specific instruments (EWPS and LEAPS), only WLQ included patients with depression as part of the scale development.

Construct validity is divided into convergent validity (instruments measuring the same concepts should correlate with one another) and discriminant validity (instruments measuring different concepts should not correlate with one another). Construct validity has been established for all instruments categorized as demonstrating either convergent validity or discriminant validity. Reliability, which pertains to internal consistency (of items within a domain), as well as interrater and test-retest reliability, has been formally demonstrated for the Health and Work Questionnaire, LEAPS, WLQ, and EWPS. Because the focus for the SDS is toward the work domain (comprising only one question), it is not possible to assess the internal consistency of this instrument. Responsiveness (the ability to assess clinically important change over time) has been demonstrated for three instruments: the WPAI questionnaire, WLQ, and SDS.

Because most of these instruments were not initially developed to assess mood disorders, it is important to assess their suitability for use in this population to account for the specific

impact of depression on presenteeism. Only four instruments have been partially or fully validated for use in this disease area: psychometric properties have been investigated in major depressive disorders for LEAPS [21], WLQ [5], and EWPS [18], while validity has been fully assessed for the SDS in both major depressive disorder [32] and bipolar disorder [35].

#### Monetization

The translation of scores derived from presenteeism instruments to monetary units has been discussed extensively in Brooks et al. [17]. This translation can be divided into two parts: the conversion from a score into a quantifiable measure, such as lost time or percentage of productivity lost, and the translation into monetary units. Translations of presenteeism scales into monetary units are detailed by Lofland et al. [15] and Brooks et al. [17] for HLQ, HPQ, WLQ, the WPAI questionnaire, and the Worker Productivity Index (WPI). Monetization of lost productivity due to presenteeism is not feasible in instruments producing a score that does not represent (directly) impaired productivity (e.g., EWPS, SDS, and SPS-6).

The HPQ appears to be the instrument best suited for conversion into monetary units. This is due to the clear distinction that is made by this instrument between absence for any reason and absence due to ill health, enabling the quantification of the actual lost productivity that was caused by the illness. Moreover, full guidance on monetization and missing data handling is provided. A major weakness in this instrument, however, is that some questions can be very complex, especially in the context of mood disorders.

While not considered in the review by Brooks et al. [17], LEAPS allows for the generation of costs associated with presenteeism, including the Likert items related exclusively to problems at work (three items). These subscores have the potential for monetization, for which the development and testing of a specific methodology is under way by the scale developers. Further assessment of the psychometric properties for these subscores in major depressive disorder by comparing it against the HPQ is also planned.

In theory, the translation from WLQ into monetary units could be achieved by using the WLQ “Scale Score” or the WLQ “Productivity Loss Score.” The former is restricted to one domain. While the feasibility of monetization has not been explicitly mentioned by the designers of the latter instrument, it was purposely created to express the percentage productivity reduction in nonhealthy individuals compared with healthy workers and is better suited to generating monetary estimates of presenteeism because all the domains are weighted according to their impact on objective estimates. This is the only conversion algorithm derived by using the relation to self-report measures to objective data.

Valuation of presenteeism is usually handled by applying the human capital approach according to the formulae provided by instrument developers or other research teams although this does not account for the impact on other workers’ productivity. The multiplier methods proposed by Pauly et al. [11] could be applied to monetize this impact. The impact of using multipliers to estimate the cost of presenteeism has been assessed in other disease areas [36].

#### Review of the Use and Performance of Productivity Scales in Mood Disorders

The use and performance of these instruments for assessing productivity in mood disorders in RCTs and observational studies was systematically reviewed by using PubMed. All articles published before March 30, 2011, were selected by using keywords, and this search was completed through a manual search. The following characteristics for the identified studies were extracted:

**Table 3 – Summary of information available in studies using productivity instruments in mood disorders.**

Reference	Title	Type of study	Productivity scale	Other scale	Result by presence of disorder	Result by severity of disorder	Result by evolution of disorder	Result by treatment arm*	Correlation to other scales
Adler et al. [7]	Job performance deficits due to depression	OS	WLQ	PHQ-9 (depression), SF-36	WLQ (output, time, mental) × Depression		WLQ (output, time, mental) × Improvement		
Ansseau et al. [37]	Objective: remission of depression in primary care The Oreon Study	OS	SDS	HDRS7 (depression)	SDS work × Remission				
Arbuckle et al. [35]	The psychometric validation of the SDS in patients with bipolar disorder	OS	SDS, the WPAI questionnaire	BFSQ (bipolar), EQ-5D, SF-36		SDS work × Bipolar severity			SDS work × (SDS family; SDS social; the WPAI questionnaire overall work impairment; BFSQ; MCS, EQ-5D)
Burton et al. [13]	The association of medical conditions and presenteeism	OS	WLQ		WLQ (output, time, mental) × Depression				
Calabrese et al. [38]	Impact of bipolar disorder on a US community sample	OS	SDS	SAS-SR, MDQ	SDS work × Bipolar				
Calabrese et al. [39]	Impact of depressive symptoms compared with manic symptoms in bipolar disorder: results of a US community-based sample	OS	SDS	SAS-SR, MDQ	SDS work × Depressive / Manic symptoms				
Endicott and Nee [18]	Assessment measures for clinical studies	RCT (post hoc)	EWPS	SCL-90, HAMD (depression), CGI	EWPS × Depression				EWPS × (HAMD; CGI)
Esposito et al. [40]	Mood and anxiety disorders, the association with presenteeism in employed members of a general population sample	OS	SPS	MINI (mood disorder)	SPS × Mood disorders/ Anxiety / Control				
Greco et al. [41]	The outcome of physical symptoms with treatment of depression	RCT	WLQ (output, time, mental)	PCP					
Holden et al. [42]	Psychological distress is associated with a range of high-priority health conditions affecting working Australians	OS	HPQ <sup>†</sup>						
Kamat et al. [43]	Prevalence and humanistic impact of potential misdiagnosis of bipolar disorder among patients with major depressive	Retro-spective claims database + survey data	SDS	SF-12, MDQ (Bipolar disorder)	SDS work × Bipolar				



Table 3 – (continued)

Reference	Title	Type of study	Productivity scale	Other scale	Result by presence of disorder	Result by severity of disorder	Result by evolution of disorder	Result by treatment arm*	Correlation to other scales
Kennedy et al. [44]	disorder in a commercially insured population Work, social, and family disabilities of subjects with anxiety and depression	RCT (post hoc)	SDS		SDS work × Depression/Anxiety				
Kessler et al. [3]	Prevalence and effects of mood disorders on work performance in a nationally representative sample of US workers	OS	HPQ	QIDS-SR (depression)	HPQ × (Depression; Bipolar disorders)				
Kessler et al. [45]	Comparative and interactive effects of depression relative to other health problems on work performance in the workforce of a large employer	Retro-spective claims database + survey data	HPQ		HPQ × (Depression; Bipolar disorders)				
Knott et al. [46]	Effect of inadequate response to treatment in patients with depression	OS	The WPAI questionnaire	SF-8	The WPAI questionnaire × responder / partial responder / non-responder				
Lam et al. [21]	A new clinical rating scale for work absence and productivity: validation in patients with major depressive disorder	OS	LEAPS, HPQ, SDS work	QIDS-SR (depression)		LEAPS × depression severity			LEAPS × (SDS work; HPQ Global Work Performance)
Lee [47]	Loss of productivity due to depression among Korean employees	OS	SPS		SPS × Depression				
Lerner et al. [9]	The clinical and occupational correlates of work productivity loss among employed patients with depression	OS	WLQ	PHQ-9 (depression), SF-12	WLQ (output, time, mental) × Depression	WLQ (output, time, mental, physical) × Depression severity			
Lerner et al. [30]	Unemployment, job retention, and productivity loss among employees with depression	OS	WLQ	PHQ-9 (depression)	WLQ (physical, output, time, mental) × Depression				
Lerner et al. [48]	Work performance of employees with depression: the impact of work stressors	OS	WLQ	PHQ-9 (depression), SF-12	WLQ × Depression				
Lin et al. [49]	Can depression treatment in primary	RCT	SDS					SDS work*	

**Table 3 – (continued)**

Reference	Title	Type of study	Productivity scale	Other scale	Result by presence of disorder	Result by severity of disorder	Result by evolution of disorder	Result by treatment arm*	Correlation to other scales
McMorris et al. [50]	care reduce disability? A stepped care approach Workplace productivity, employment issues, and resource utilization in patients with bipolar I disorder	OS	EWPS	SCL-20 (depression), SF-36	EWPS × Bipolar I disorder				
Olley et al. [51]	Persistence of psychiatric disorders in a cohort of HIV/AIDS patients in South Africa: a 6-mo follow-up study	OS	SDS		SDS work × Depression				
Sanderson et al. [5]	Which presenteeism measures are more sensitive to depression and anxiety?	OS	WLQ, SPS	PHQ-9 (depression)		WLQ (output, time, mental) × Depression severity	WLQ × Remission; SPS × Remission		
Sheehan and Sheehan [32]	Assessing treatment effects in clinical trials with the discan metric of the SDS	Review	SDS		SDS work × Depression			SDS work*	
Snedecor et al. [52]	Economic outcomes of eszopiclone treatment in insomnia and comorbid major depressive disorder	RCT	WLQ	HDRS (depression), SF-36				WLQ	
Soares et al. [53]	Assessing the efficacy of desvenlafaxine for improving functioning and well-being outcome measures in patients with major depressive disorder: a pooled analysis of nine double-blind, placebo-controlled, 8-wk clinical trials	Pooled RCTs	SDS	WHO-5 (psychological well-being), MADRS (depression), HDRS (depression)				SDS work*	
Stamouli et al. [54]	Escitalopram in clinical practice in Greece: treatment response and tolerability in depressed patients	Open label surveillance study	SDS	MADRS (depression), CGI	SDS work × Improvement				
Stewart et al. [4]	Cost of lost productive work time among US workers with depression	OS	WHI						
Wang et al. [8]	Chronic medical conditions and work performance in the health and work performance	OS	HPQ						

**Table 3 – (continued)**

Reference	Title	Type of study	Productivity scale	Other scale	Result by presence of disorder	Result by severity of disorder	Result by evolution of disorder	Result by treatment arm*	Correlation to other scales
Wang et al. [55]	questionnaire calibration surveys Telephone screening, outreach, and care management for depressed workers and impact on clinical and work productivity outcomes: a randomized controlled trial	RCT	HPQ	QIDS-SR (depression)					

AIDS, acquired immunodeficiency syndrome; BFSQ, Bipolar Functional Status Questionnaire; CGI, Clinical Global impression; EQ-5D, EuroQol five-dimensional questionnaire; EWPS, Endicott Work Productivity Scale; HDRS, Hamilton Depression Rating Scale; HPQ, Health and Work Performance Questionnaire; LEAPS, Lam Employment Absence and Productivity Scale; MADRS, Montgomery-Åsberg Depression Rating Scale; MDQ, Mood Disorder Questionnaire; MINI, Mini Neuropsychiatric Diagnostic Interview; OS, observational study; PCP, PHQ-9, Patient Health Questionnaire 9; QIDS-SR, Quick Inventory of Depressive Symptomatology Self-Report; RCT, randomized clinical trial; SAS-SR, Social Adjustment Scale Self-Report; SCL-20, symptom checklist-20; SDS, the Sheehan Disability Scale; SF-12, Medical Outcome Study Short Form 12 questions; SF-8, Medical Outcome Study Short Form 8 questions; SF-36, short form 36 health survey; SPS, Stanford Presenteeism Scale; WHI, Work and Health Interview; WHO-5, well-being index; WLQ, Work Limitations Questionnaire; WPAI, Work Productivity and Activity Impairment.

\* Indicates that significant differences were shown between treatment groups on the productivity scale.

† HPQ is used to identify self-reported health status and in particular the presence of psychological distress.

study type, objective, study design, productivity scale(s) included, other rating scales (depressive symptoms, health-related quality of life), assessment dates, and results. Four types of results were considered in each study: (1) results by treatment arms, (2) results by health state (symptom severity level), (3) correlation to other scales, and (4) use of monetization and methods applied. Studies on the SDS were restricted to those that actually reported results on the work domain.

## Results

Only 31 studies discussing the use of productivity instruments in individuals with mood disorders were retrieved, the majority of which (74%) were observational studies (prospective cohorts, retrospective claims databases complemented with a survey to assess productivity loss or workplace studies). Twelve studies considered the SDS and explicitly reported and discussed results on the work domain. The WLQ and the HPQ were used in five studies. The other productivity instruments appeared in no more than two studies each. Among the four categories of results considered, the majority (68%) of articles reported results on presenteeism instruments by health state or presence/absence of symptoms of mood disorders. An overview of the available information in each of the selected studies is presented in Table 3 [3–5,7–9,13,18,21,30,32,35,37–55].

Overall, results from the retrieved studies showed presenteeism scales to be sufficiently sensitive to demonstrate consistently higher at-work productivity loss in the presence of symptoms of mood disorders, and in depression in particular, compared with the performance of employees with no symptoms. The HPQ, WLQ, SDS, SPS [40], and EWPS [18] clearly showed that employees with mood disorders presented higher productivity loss than did employees without mood disorders. Significantly more impairment (approximately 35% productivity reduction in output, time, and mental-interpersonal demand) was observed by using the WLQ in employees with depressive symptoms [12,30,48]. Excess productivity loss was estimated on the basis of the HPQ in a nationally representative survey of mental disorders in the US population. This amounted to 18.2 workdays per year for employees with major depressive disorders and 35.3 days for employees with bipolar disorders [3].

In addition, studies using several of these instruments have demonstrated that at-work productivity loss increases with increasing severity of symptoms. Impairment as measured by LEAPS and WLQ time, WLQ mental-interpersonal, and WLQ output scales increases with the severity of depression [5,21,30]. Arbuckle et al. [35] studied the psychometric properties of the SDS in bipolar disorder and demonstrated that work impairment increases with increasing severity.

Improvement in symptoms or improvement in health condition and improvement in productivity were evaluated for the WLQ, SPS-6, and SDS in four studies. According to Sanderson et al. [5], remission of symptoms was associated with improvements on WLQ *output demands* domain (10.9-point improvement at 6 months;  $P = 0.006$ ) and on the SPS-6 (2.2-point improvement at 6 months;  $P = 0.033$ ) in a community sample. No significant improvements were measured on the other WLQ domains [7]. Another study, designed to estimate remission rates in primary care [37], demonstrated that failure to achieve remission of symptoms was associated with higher impairment on the SDS work domain (5 vs. 1.6;  $P < 0.001$ ).

Three articles reported results of presenteeism scales by treatment group. A pooled analysis of RCTs in major depressive disorders comparing desvenlafaxine with placebo [53] reported a significant difference on the SDS work domain for all doses of desvenlafaxine versus placebo (between  $-0.6$  and  $-0.8$  compared with placebo) together with lower disability assessed by SDS total

score. Wang et al. [55] used the HPQ to compare care management and telephone therapy versus usual care in 604 employees identified as having significant depression and who were covered by a managed behavioral health plan. The HPQ score in terms of effective hours worked was significantly higher in the intervention group at 6 and 12 months (3 hours more per week at 6 months and 3.3 hours per week at 12 months). Conversely, there was no difference in on-the-job work performance between the two groups at 6 or 12 months. In a pharmacoeconomic analysis conducted alongside a randomized clinical trial comparing eszopiclone + fluoxetine versus fluoxetine + placebo in patients with comorbid insomnia and major depressive disorders, no difference between the two treatment groups was reported on the WLQ [52].

Few studies reported data on how presenteeism instruments correlate to symptomatic rating scales and health-related quality of life. As part of the validation of LEAPS, correlations between LEAPS and the SDS work and the HPQ global work performance domains were reported. The results showed a high correlation between the LEAPS score and both the HPQ ( $r = 0.79$ ;  $P < 0.01$ ) and the SDS work domain (0.63;  $P < 0.01$ ). In the study by Arbuckle et al. [35] reviewing the psychometric properties of the SDS in patients with bipolar disorder, the SDS work domain showed a high correlation to SDS family ( $r = 0.71$ ;  $P < 0.0001$ ) and SDS social ( $r = 0.76$ ;  $P < 0.0001$ ) domains and to the WPAI questionnaire overall work impairment ( $r = 0.55$ ;  $P < 0.0001$ ), the mental component score from the SF-36 ( $r = -0.43$ ;  $P < 0.0001$ ), and the EuroQoL-5 dimension (EQ-5D questionnaire;  $r = -0.46$ ;  $P < 0.0001$ ). In patients seeking treatment for a major depressive episode in an outpatient setting and among nonpatients in the community, fairly high correlations were reported between the EWPS and the Hamilton Rating Scale for Depression (between 0.27 and 0.61) [18].

Four studies translated presenteeism into monetary units. Stewart et al. [27] estimated the costs of depression to the US workforce by calculating lost productive time by using the Work and Health Interview and applying employee wages to obtain costs of productivity loss. This article concluded that depression among US workers was associated with an excess cost of \$31 billion per year in 2002. At-work lost productivity accounted for 81% of the total cost. Kessler et al. [3] evaluated the impact of major depressive disorder and bipolar disorder on lost productivity, and on presenteeism in particular. Presenteeism estimates derived from the HPQ were converted into lost-day equivalents and converted into a salary metric by applying wages increased by 25% to include fringe benefits. Major depressive disorders and bipolar disorders, respectively, resulted in lost time due to presenteeism of 18.2 and 35.3 days per individual per year, which translate to \$2691 and \$5184 per individual per year. On the basis of the WLQ index, hourly wage, and average number of hours worked per week, Snedecor et al. [52] estimated the change from baseline in costs associated with presenteeism and added these to total medical costs and absenteeism costs. The savings over 8 weeks associated with lower presenteeism for eszopiclone + fluoxetine compared with placebo + fluoxetine (\$82) partially offset the additional acquisition costs of eszopiclone (\$209). It should be noted, however, that while differences were observed on the WLQ between the two treatment strategies, these were not statistically significant, and while no minimally important difference in the level of impairment has been established for the WLQ, the reported difference in this trial ( $< 1\%$  of lost productivity while at work) is very small. Further information regarding the minimal important difference on the WLQ productivity loss score is needed to provide definite conclusions.

## Discussion/Conclusions

Several instruments are available to researchers aimed at measuring productivity loss due to presenteeism, all of which are

applicable to mood disorders. Limited evidence exists, however, on their actual use in this disease area, with the exception of the SDS, a scale that cannot be fully classified as a productivity instrument. Although psychometric properties have been verified for most instruments in the general population, little evidence has been generated in mood disorders and only four instruments are at least partially validated in mood disorders. Furthermore, even in a general context, responsiveness is not well documented for most of the scales.

Most of the available instruments are able to show differences in productivity loss between individuals with or without symptoms and allow researchers to detect improvement in productivity associated with health improvements. Nonetheless, the ability to detect differences between treatments or interventions in a randomized setting has not been demonstrated for any of the instruments. While this may be due, in part, to a lack of sensitivity of the instruments, it is probably due to design characteristics such as too short a time horizon and the lack of statistical power to conclude that there was a significant difference between the interventions. This lack of statistical power is expected in clinical trials because these are not initially powered to detect a difference in presenteeism but only on the primary end point, which often is a measure of clinical efficacy. One of the studies retrieved in this review was a pooled analysis, confirming the benefit of performing the analysis on presenteeism instruments with an increase in the statistical power [53]. Furthermore, the low number of studies that report results on presenteeism instruments suggests that these productivity estimates are often not evaluated. This would confirm that it is rare to detect between-treatment differences in presenteeism by using these instruments, and suggests that much larger sample sizes are required to demonstrate a significant difference by using presenteeism instruments compared with symptomatic rating scales.

Few studies have used productivity instruments to estimate the cost of presenteeism, despite the fact that presenteeism is known to be one of the major cost drivers in mood disorders [4] and that monetization is theoretically possible for most of the selected instruments. Similar to the results reported in the review by Lerner and Henke [12], the instruments were most often used in an observational setting such as the workplace, or in a clinical setting, and very little information is published regarding the use of these instruments in an RCT setting. The lack of productivity estimates in clinical trials could be due to a number of factors. First, it is possible that some relevant studies have not been retrieved because the search strategy may have failed to identify studies in which the use of a presenteeism scale is not explicitly mentioned in the abstract. Second, results from the measurement of lost productivity may not necessarily have been reported in the published studies, particularly if the results were negative or uninteresting. Third, the lack of productivity estimates from clinical trials could be due to a failure to publish findings from trials in which the primary objective was to measure changes in productivity if this objective was not met. Further investigations using clinical trial registries are needed to track clinical trials that have assessed lost productivity in an attempt to identify additional published results for further analysis. Finally, the RCT context may not be fully suited for the use of productivity loss instruments because of the limited number of patients, short time horizon, and heterogeneity caused by the participation of patients from different countries.

Recommendations as to which productivity scale is best suited to estimate the cost of presenteeism in mood disorders should be based primarily on instrument properties because of the limited experiential evidence. The HPQ may be the best option for measuring and valuing productivity because this instrument establishes an objective measure of productivity in a workplace context. However, this questionnaire lacks the ability to capture

different levels of presenteeism on different days, and its complexity may prevent its use in RCTs, even with the version adapted to this setting. Its use and efficiency in clinical trials should be further assessed to allow more definite recommendations on the usefulness of this instrument in the area of mood disorders.

Other tools such as the WLQ, the WPAI questionnaire, LEAPS, or SF-HLQ seem to be more adapted to the clinical trial context because they are shorter and more straightforward to complete. WLQ presents important evidence in mood disorders and presents high validity in this disease area although some ambiguity remains related to the monetization scoring. Little evidence exists on the use of the WPAI questionnaire in mood disorders, although it has been used in numerous disease areas, including anxiety disorders [56]. This same limitation should be addressed for SF-HLQ. Nonetheless, the use of cognitive debriefing during the development of the WPAI questionnaire is a piece of evidence that supports its use in patients with mood disorders, whereas such information was not retrieved for the SF-HLQ. LEAPS has the advantage of having been developed specifically for depression and of being validated for this condition, but its amenability for monetization is still under investigation. Despite its wide use in mood disorders and its good properties, the SDS cannot be considered a productivity scale. Other scales such as the EWPS are validated in this specific disease area, but they are of limited use for estimating the cost of presenteeism because conversion into monetary unit is not possible.

To be able to make recommendations as to the use of specific instruments in clinical trials and observational studies in mood disorders, further empirical data are needed. Further research is also needed in the following three areas: (1) Full documentation as to the responsiveness will give researchers important data for sample size calculations for designing RCTs or observational studies; (2) The relevance of the use of these instruments in clinical trials should be further investigated, and guidance should be provided regarding appropriate study design (study duration, population to include, single- or multicountry study) as well as advice on the selection of the most appropriate scale; and (3) Although all the instruments measuring presenteeism were able to detect an excess in productivity loss in patients with symptoms of mood disorders compared with patients without symptoms, the impact of mood disorder on self-rating evaluations is not well documented and, at this time, it is not possible to determine all the potential biases in reported estimates. Lerner and Henke [12] concluded that self-report estimates from the WLQ reported by depressed workers were as accurate as the ones reported by nondepressed workers. Additional evidence is needed regarding the comparison between self-report assessments and objective data for the other presenteeism questionnaires and the validity of the conversion algorithm to generate monetary estimates. This would ensure that unbiased estimates of productivity loss in mood disorder could be generated and reinforce the credibility of economic evaluations in this patient population conducted from a societal perspective.

Source of financial support: This study was funded and fully sponsored by H. Lundbeck A/S.

## REFERENCES

- [1] Collins JJ, Baase CM, Sharda CE, et al. The assessment of chronic health conditions on work performance, absence, and total economic impact for employers. *J Occup Environ Med* 2005;47:547–57.
- [2] Burton WN, Conti DJ, Chen CY, et al. The role of health risk factors and disease on worker productivity. *J Occup Environ Med* 1999;41:863–77.
- [3] Kessler RC, Akiskal HS, Ames M, et al. Prevalence and effects of mood disorders on work performance in a nationally representative sample of U.S. workers. *Am J Psychiatry* 2006;163:1561–8.
- [4] Stewart WF, Ricci JA, Chee E, et al. Cost of lost productive work time among US workers with depression. *JAMA* 2003;289:3135–44.

- [5] Sanderson K, Tilse E, Nicholson J, et al. Which presenteeism measures are more sensitive to depression and anxiety? *J Affect Disord* 2007;101:65–74.
- [6] Schultz AB, Chen CY, Edington DW. The cost and impact of health conditions on presenteeism to employers: a review of the literature. *Pharmacoeconomics* 2009;27:365–78.
- [7] Adler DA, McLaughlin TJ, Rogers WH, et al. Job performance deficits due to depression. *Am J Psychiatry* 2006;163:1569–76.
- [8] Wang PS, Beck A, Berglund P, et al. Chronic medical conditions and work performance in the Health and Work Performance Questionnaire calibration surveys. *J Occup Environ Med* 2003;45:1303–11.
- [9] Lerner D, Adler DA, Chang H, et al. The clinical and occupational correlates of work productivity loss among employed patients with depression. *J Occup Environ Med* 2004;46(6, Suppl.):S46–55.
- [10] Stewart WF, Ricci JA, Leotta C, Chee E. Validation of the work and health interview. *Pharmacoeconomics* 2004;22:1127–40.
- [11] Pauly MV, Nicholson S, Polsky D, et al. Valuing reductions in on-the-job illness: “presenteeism” from managerial and economic perspectives. *Health Econ* 2008;17:469–85.
- [12] Lerner D, Henke RM. What does research tell us about depression, job performance, and work productivity? *J Occup Environ Med* 2008;50:401–10.
- [13] Burton WN, Pransky G, Conti DJ, et al. The association of medical conditions and presenteeism. *J Occup Environ Med* 2004;46(6, Suppl.):S38–45.
- [14] Prasad M, Wahlqvist P, Shikar R, Shih YC. A review of self-report instruments measuring health-related work productivity: a patient-reported outcomes perspective. *Pharmacoeconomics* 2004;22:225–44.
- [15] Lofland JH, Pizzi L, Frick KD. A review of health-related workplace productivity loss instruments. *Pharmacoeconomics* 2004;22:165–84.
- [16] Mattke S, Balakrishnan A, Bergamo G, Newberry SJ. A review of methods to measure health-related productivity loss. *Am J Manag Care* 2007;13:211–7.
- [17] Brooks A, Hagen SE, Sathyanarayanan S, et al. Presenteeism: critical issues. *J Occup Environ Med* 2010;52:1055–67.
- [18] Endicott J, Nee J. Endicott Work Productivity Scale (EWPS): a new measure to assess treatment effects. *Psychopharmacol Bull* 1997;33:13–6.
- [19] Hakkaart-van Roijen L., Bouwmans C.A.M. Manual Short Form-Health and Labour Questionnaire (SF-HLQ). 2010. Available from: [http://www.bmg.eur.nl/fileadmin/ASSETS/bmg/english/iMTA/Publications/Manuals\\_\\_Questionnaires/Manual\\_SF-HLQ\\_2010.pdf](http://www.bmg.eur.nl/fileadmin/ASSETS/bmg/english/iMTA/Publications/Manuals__Questionnaires/Manual_SF-HLQ_2010.pdf). [Accessed December 23, 2011].
- [20] Kessler RC, Barber C, Beck A, et al. The World Health Organization Health and Work Performance Questionnaire (HPQ). *J Occup Environ Med* 2003;45:156–74.
- [21] Lam RW, Michalak EE, Yatham LN. A new clinical rating scale for work absence and productivity: validation in patients with major depressive disorder. *BMC Psychiatry* 2009;9:78.
- [22] Leon AC, Olfson M, Portera L, et al. Assessing psychiatric impairment in primary care with the Sheehan Disability Scale. *Int J Psychiatry Med* 1997;27:93–105.
- [23] Lerner D, Amick BC III, Rogers WH, et al. The Work Limitations Questionnaire. *Med Care* 2001;39:72–85.
- [24] Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics* 1993;4:353–65.
- [25] Reilly M.C. Development of the Work Productivity and Activity Impairment (WPAI) questionnaire. 2008. Available from: <http://www.reillyassociates.net/WPAI-Development.doc>. [Accessed December 23, 2011].
- [26] Shikar R, Halpern MT, Rentz AM, Khan ZM. Development of the Health and Work Questionnaire (HWQ): an instrument for assessing workplace productivity in relation to worker health. *Work* 2004;22:219–29.
- [27] Stewart WF, Ricci JA, Chee E, Morganstein D. Lost productive work time costs from health conditions in the United States: results from the American Productivity Audit. *J Occup Environ Med* 2003;45:1234–46.
- [28] van Roijen L, Essink-Bot ML, Koopmanschap MA, et al. Labor and health status in economic evaluation of health care: The Health and Labor Questionnaire. *Int J Technol Assess Health Care* 1996;12:405–15.
- [29] van RL, Essink-Bot ML, Koopmanschap MA, et al. Societal perspective on the burden of migraine in The Netherlands. *Pharmacoeconomics* 1995;7:170–9.
- [30] Lerner D, Adler DA, Chang H, et al. Unemployment, job retention, and productivity loss among employees with depression. *Psychiatr Serv* 2004;55:1371–8.
- [31] Sheehan DV. *The Anxiety Disease*. New York: Charles Scribner and Sons, 1983.
- [32] Sheehan KH, Sheehan DV. Assessing treatment effects in clinical trials with the discan metric of the Sheehan Disability Scale. *Int Clin Psychopharmacol* 2008;23:70–83.
- [33] Koopman C, Pelletier KR, Murray JF, et al. Stanford Presenteeism Scale: health status and employee productivity. *J Occup Environ Med* 2002;44:14–20.
- [34] Koopmanschap M, Burdorf A, Jacob K, et al. Measuring productivity changes in economic evaluation: setting the research agenda. *Pharmacoeconomics* 2005;23:47–54.
- [35] Arbuckle R, Frye MA, Brecher M, et al. The psychometric validation of the Sheehan Disability Scale (SDS) in patients with bipolar disorder. *Psychiatry Res* 2009;165:163–74.
- [36] Sogaard R, Sørensen J, Linde L, Hetland ML. The significance of presenteeism for the value of lost production: the case of rheumatoid arthritis. *ClinicoEconomics Outcomes Res* 2010;2:105–12.
- [37] Anseau M, Demyttenaere K, Heyrman J, et al. Objective: remission of depression in primary care The Oron Study. *Eur Neuropsychopharmacol* 2009;19:169–76.
- [38] Calabrese JR, Hirschfeld RM, Reed M, et al. Impact of bipolar disorder on a U.S. community sample. *J Clin Psychiatry* 2003;64:425–32.
- [39] Calabrese JR, Hirschfeld RM, Frye MA, Reed ML. Impact of depressive symptoms compared with manic symptoms in bipolar disorder: results of a U.S. community-based sample. *J Clin Psychiatry* 2004;65:1499–504.
- [40] Esposito E, Wang JL, Williams JV, Patten SB. Mood and anxiety disorders, the association with presenteeism in employed members of a general population sample. *Epidemiol Psychiatr Soc* 2007;16:231–7.
- [41] Greco T, Eckert G, Kroenke K. The outcome of physical symptoms with treatment of depression. *J Gen Intern Med* 2004;19:813–8.
- [42] Holden L, Scuffham P, Hilton M, et al. Psychological distress is associated with a range of high-priority health conditions affecting working Australians. *Aust N Z J Public Health* 2010;34:304–10.
- [43] Kamat SA, Rajagopalan K, Pethick N, et al. Prevalence and humanistic impact of potential misdiagnosis of bipolar disorder among patients with major depressive disorder in a commercially insured population. *J Manag Care Pharm* 2008;14:631–42.
- [44] Kennedy BL, Lin Y, Schwab JJ. Work, social, and family disabilities of subjects with anxiety and depression. *South Med J* 2002;95:1424–7.
- [45] Kessler RC, Merikangas KR, Wang PS. The prevalence and correlates of workplace depression in the national comorbidity survey replication. *J Occup Environ Med* 2008;50:381–90.
- [46] Knoth RL, Bolge SC, Kim E, Tran QV. Effect of inadequate response to treatment in patients with depression. *Am J Manag Care* 2010;16:e188–96.
- [47] Lee YM. Loss of productivity due to depression among Korean employees. *J Occup Health* 2010;52:389–94.
- [48] Lerner D, Adler DA, Rogers WH, et al. Work performance of employees with depression: the impact of work stressors. *Am J Health Promot* 2010;24:205–13.
- [49] Lin EH, VonKorff M, Russo J, et al. Can depression treatment in primary care reduce disability? A stepped care approach. *Arch Fam Med* 2000;9:1052–8.
- [50] McMorris BJ, Downs KE, Panish JM, Dirani R. Workplace productivity, employment issues, and resource utilization in patients with bipolar I disorder. *J Med Econ* 2010;13:23–32.
- [51] Olley BO, Seedat S, Stein DJ. Persistence of psychiatric disorders in a cohort of HIV/AIDS patients in South Africa: a 6-month follow-up study. *J Psychosom Res* 2006;61:479–84.
- [52] Snedecor SJ, Botteman MF, Schaefer K, et al. Economic outcomes of eszopiclone treatment in insomnia and comorbid major depressive disorder. *J Ment Health Policy Econ* 2010;13:27–35.
- [53] Soares CN, Kornstein SG, Thase ME, et al. Assessing the efficacy of desvenlafaxine for improving functioning and well-being outcome measures in patients with major depressive disorder: a pooled analysis of 9 double-blind, placebo-controlled, 8-week clinical trials. *J Clin Psychiatry* 2009;70:1365–71.
- [54] Stamouli SS, Yfantis A, Lamboussis E, et al. Escitalopram in clinical practice in Greece: treatment response and tolerability in depressed patients. *Expert Opin Pharmacother* 2009;10:937–45.
- [55] Wang PS, Simon GE, Avorn J, et al. Telephone screening, outreach, and care management for depressed workers and impact on clinical and work productivity outcomes: a randomized controlled trial. *JAMA* 2007;298:1401–11.
- [56] Erickson SR, Guthrie S, Vanetten-Lee M, et al. Severity of anxiety and work-related outcomes of patients with anxiety disorders. *Depress Anxiety* 2009;26:1165–71.