

Fair or square? Experiences of introducing a new method for assessing general work ability in a sickness insurance context

Christian Ståhl, Ida Seing, Björn Gerdle and Jan Sandqvist

The self-archived postprint version of this journal article is available at Linköping University Institutional Repository (DiVA):

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-146118>

N.B.: When citing this work, cite the original publication.

This is an electronic version of an article published in:

Ståhl, C., Seing, I., Gerdle, B., Sandqvist, J., (2017), Fair or square? Experiences of introducing a new method for assessing general work ability in a sickness insurance context, *Disability and Rehabilitation*, , 1-10. <https://doi.org/10.1080/09638288.2017.1401675>

Original publication available at:

<https://doi.org/10.1080/09638288.2017.1401675>

Copyright: Taylor & Francis (STM, Behavioural Science and Public Health Titles)

<http://www.tandf.co.uk/journals/default.asp>



Fair or square? Experiences of introducing a new method for assessing general work ability in a sickness insurance context

Christian Ståhl^{1,2}, Ida Seing¹, Björn Gerdle³, Jan Sandqvist⁴

1: Division of Community Medicine, Department of Medical and Health Sciences, Linköping University, Linköping, Sweden

2: HELIX Competence Centre, Linköping University, Linköping, Sweden

3: Pain and Rehabilitation Centre, Department of Medical and Health Sciences, Linköping University, Linköping, Sweden

4: Department of Social and Welfare Studies, Linköping University, Norrköping, Sweden

Corresponding author:

Christian Ståhl, PhD

Department of Medical and Health Sciences, Linköping University

SE-581 83 Linköping, Sweden

e-mail: christian.stahl@liu.se

phone: +46 13282690

This is an Accepted Manuscript of an article published by Taylor & Francis in Disability & Rehabilitation on [date of publication], available online: [http://www.tandfonline.com/\[Article DOI\]](http://www.tandfonline.com/[Article DOI]).

Abstract

Purpose: To study social validity and perceived fairness of a new method for assessing general work ability in a sickness insurance context. Assessments are based on self-reports, combined with examinations by physicians, and, if needed, occupational therapists, physiotherapists and/or psychologists.

Materials and methods: Interviews with 36 insurance officials, 10 physicians, and 36 sick-listed persons, which were analyzed through a qualitative content analysis.

Results: Insurance officials and physicians considered the method useful and that it facilitated benefit decisions. The experiences of persons who had undergone the assessment differed, where the dialogue with insurance officials seemed to have had an influence on experiences of the assessment and the decisions it led to.

Conclusions: The perceived fairness and social validity of the assessment depended on how it was carried out; organizational conditions and priorities; communication skills; and decision outcomes. Professionals have an important pedagogical task in explaining the purpose and procedure of the assessment in order for the sick-listed to perceive it as fair rather than square, i.e., too standardized and not considering individual conditions. If the assessment could be used also for rehabilitative purposes, it could possibly be perceived as more acceptable also in cases where it leads to denied benefits.

Keywords: work ability; work disability; sickness insurance; return to work; assessment; Sweden

Introduction

To comprehend a process fully, it is often essential to see the results of it.

[1, p. 311]

The legitimacy of a social insurance system is based on its trust in the population, where key aspects are that it operates in a fair way with a high quality of services, and manages to strike a balance between equal treatment and respect for individual needs [2]. In sickness insurance and workers' compensation systems, a central element for upholding the system's legitimacy is to secure that eligibility for benefits is assessed in a reliable way, meeting criteria of both equality and equity, i.e. that similar cases are treated alike, and that cases are given an individual assessment with regard to the characteristics of the situation. In many systems, eligibility decisions are based on assessments of a person's functioning, and their ability to work [3]. Work ability is a complex phenomenon where the research points to increasingly broad and dynamic definitions, emphasizing the relation between individual medical conditions, workplace conditions, and societal conditions [4-8]. Tengland argues that we need two definitions of work ability – *specific* work ability, which is the ability to perform a specific job with certain competence and skill requirements, and *general* work ability, which is the ability to perform a type of job that most people can manage given a short period of training [8].

There are a variety of different methods and instruments that measures work ability, work capacity or functional capacity. In a review [3], it was concluded that the terminology differs, as does the purposes and contexts of the assessments (e.g., assessing eligibility to benefits; a part of a rehabilitation or return to work intervention; related to a specific job, or decontextualized as a general functional assessment), and the dimensions assessed (e.g., specific physical or mental functions, environmental aspects). In the review, it was also concluded that most instruments focus on physical functions, and that few include environmental and occupational factors in the assessment [3]. Some attempts have been made to include such factors in order to make the assessments more contextually relevant [9, 10], which however is easier when assessing specific rather than general work ability.

Assessing general work ability is a common feature of different insurance and social security schemes, where the aim generally is to determine eligibility for benefits at a point where it is

no longer likely that the person will return to the previous job, or because regulations stipulate that this option is no longer available. There are a number of different assessment tools for carrying out general work ability assessments (or similar broad assessments), that may be used for different purposes, e.g., the Work Rehabilitation Questionnaire (WORQ) which is based on the International Classification of Functioning (ICF) [11]. Many available tools however are not sufficiently robust and lacks proper evaluation of their measurement properties [12, 13].

Legitimacy, fairness and social validity

This study is specifically concerned with the perceived fairness of procedures within the Swedish sickness insurance system. Ultimately, this is related to the legitimacy of the system, as a system which operates in ways that are perceived unfair by claimants is due to lead to distrust; such a system is bound to lose its public legitimacy. When designing assessment methods, and when performing such assessments, it is therefore important that they offer a relevant and legitimate rendering of the person's abilities, and that the assessment is communicated in a comprehensible way, in order to be perceived as fair.

The general concept of fairness is used in this paper as whether or not procedures and outcomes are unbiased and whether or not they take into account characteristics of the individual case (i.e., a balance between equal treatment and attention to individual needs [2]). Fairness is, in this conceptualization, related to the perceived justice of outcomes and procedures, which in studies of organizational justice has been divided into several dimensions, most commonly distributive, procedural, and interactional justice [14]. The concepts of fairness and justice are used interchangeably in this paper, and it is specifically *perceived* fairness/justice that is studied, rather than fairness in a general sense. Unless indicated, it is therefore the *perceptions* of fairness/justice that are referred to.

Fairness is also related to the concept of social validity, which comprises the acceptability of the goals, procedures and outcomes of an intervention or procedure [15, 16]. Social validity also makes use of the notion of social importance, which is concerned with whether or not an intervention or procedure concurs with what the society wishes it to fulfil. The literature on social validity hence adds the dimension of the desirability of the goals an intervention is designed to achieve, which is a dimension that the literature on justice seldom includes.

In this paper, we consider distributional justice as equivalent to the acceptability of outcomes in the social validity literature, while procedural and interactional justice is considered to be dimensions of the acceptability of procedures. The acceptability of goals is related to the social importance of the intervention. In order for an intervention to be considered having social validity, it needs to account for all these dimensions. We argue that social validity is a relevant measure of the fairness of the procedures within a system, and thereby a contributing factor to whether or not a system is considered legitimate. In this paper, we limit the analysis to perceived fairness with respect to these dimensions, and to the specific groups targeted by the study design. Hence, we do not claim to analyze any other potential determinants of system legitimacy.

In the context of work disability and interactions with insurance systems, it has been shown that people that get their benefits granted generally perceive a higher level of justice, on the distributive, procedural and interactional dimensions, while those who feel they have been forced to return to work too early feel that they have had an unfair assessment [17]. Fairness and social validity have not been extensively studied in relation to general work ability assessments; however, there are examples of assessment methods that have been criticized on these accounts, e.g., the by now notorious Work Capacity Assessment (WCA) that was used to re-assess all claimants in the UK disability benefit system. Scholars and professionals have noted how it has led to detrimental health effects (including suicides) [18], to unemployment rather than employment [19], and to psychological stress that is counterproductive to its stated aims [20]. In a report on how to improve the WCA, it was suggested to increase the contextualization by introducing real-world assessments, and to combine this with standardization through developing a database of requirements in different occupations, as has been done in the Netherlands and the US [21].

The Activity Capacity Assessment (AFU)

In Swedish social insurance legislation, work ability is described on two levels, similar to Tengland's suggestion [8], which is related to how it is assessed at different time points to determine eligibility for sickness benefits [22]. For the first 180 days of sick leave, assessments are made in relation to the workplace where the person is employed, first to the specific work task, and after 90 days to any task the employer can offer (i.e., specific work

ability). This assessment is based on a medical certificate by a GP, and information from the person and/or the employer about the characteristics of the job. After the 180-day mark, assessments are to be made in relation to any job normally available on the labor market (i.e., general work ability).

To manage the complexities of assessing general work ability, the Swedish Social Insurance Agency (SSIA) has developed a standardized method for “activity capacity assessments” (aktivitetsförmågeutredning, AFU). The AFU has been tested on a limited scale since 2011, and is at the time of writing being gradually implemented [23]. The AFU is of international interest since it is an example of a broad all-purpose assessment focusing on general work ability, i.e., not related to a specific job, and designed to fit all types of diagnoses.

The purpose of the AFU is to assess a person’s general work ability, and it is used when the insurance official does not consider the existing medical certificates to be sufficient for making such an assessment. The aim of the AFU is to offer a holistic assessment not only of the person’s functional and activity limitations, but also the remaining capabilities. The AFU is based on a self-assessment by the person on sick leave, comprising four pages of questions about their health condition, activity limitations, work ability (measured by a single-item question from the Work Ability Index, which has been shown to correlate well with the WAI as a whole [24, 25], and a second question about how they assess their chances of working in six months), expectations of return to work and whether or not they can consider another type of job. This is combined with an examination by a specially trained independent physician (i.e. not the treating physician), generally working within the healthcare system. The physician’s examination comprises physical and mental status; physical strength and endurance; balance and coordination; and a neuropsychiatric interview if there is a psychiatric diagnosis, a suspicion of such a diagnosis, or a mental condition. The assessment results in a profile over the person’s physical and mental abilities. If needed, an extended assessment may be carried out by occupational therapists, physiotherapists and/or psychologists. These professionals should use specific pre-determined assessment methods. The process of an AFU assessment is schematically described in figure 1.

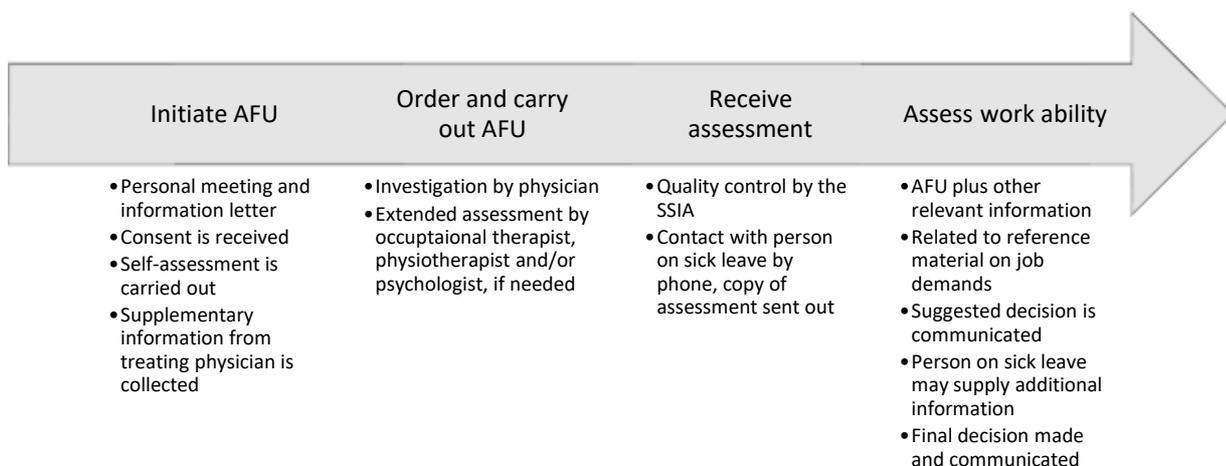


Figure 1: The different steps in an AFU assessment.

The AFU is initiated by the insurance official at the SSIA, who informs the person about the purpose of the assessment, and the procedure. After the assessment by the physician, he/she is to follow up and report whether or not the person agrees with the results. The insurance official may thereafter consult with a specialist in insurance medicine at the SSIA, to determine if the assessment is complete or if it needs to be supplemented with further information. A copy of the assessment is also sent to the person on sick leave. The official uses the results (along with other available information) for assessing general work ability, where they may relate the assessment profile to a database with descriptions of job demands in 40 common occupations. This material has no legal status and is to be used as a reference only. The work ability assessment determines the person's eligibility for continued sickness benefits. When the work ability is not considered to be sufficiently reduced (i.e., less than 25% in relation to any job normally available on the labor market), the official communicates that the SSIA considers not granting further benefits, after which the person may submit responses and/or supplementary information to strengthen their case.

Aim

The aim of this study was to analyze the initial experiences from applying the AFU method, from the perspectives of insurance officials, physicians and people on sick leave. The perspective of multiple actors is specifically relevant in order to analyze how the assessment

is used in practice, and whether or not it is perceived as fulfilling basic demands on perceived fairness and social validity.

Methods

The data collection was based on an explorative qualitative approach [26], aiming to capture the experiences from the central actors involved in AFU assessments, i.e., people on sick leave, physicians, and insurance officials. The material analyzed in this article is part of a larger study, also involving quantitative and qualitative case material, which will be reported in separate articles.

A list of 203 persons (122 women and 81 men) on sick leave who had undergone an AFU assessment was retrieved by the SSIA, based on the date of the AFU. The SSIA informed these persons about the study by mail, and asked whether they approved to be contacted by the researchers. 36 persons (24 women and 12 men, mean age of 52) approved and were contacted for interviews. Contact information to insurance officials and physicians were retrieved from a database of 300 AFU case files. 54 insurance officials were contacted, of which 36 consented to participate (30 women and 6 men). 11 physicians were contacted, of which all consented to participate, although one interview was not carried out due to problems with scheduling. See tables 1 and 2 for details on the respondents. The recruitment of respondents was carried out in the pilot phase of the AFU, which affected the availability of professionals with experiences of using the method.

Table 1: Overview of the sick-listed respondents in the study. Where information is missing, the respondent did not consent to sharing documentation.

No	Sex	Birth year	Diagnostic category (ICD-10)	Extended assessment	Decision after the AFU
1	F	1957	M	PT	Disability pension, after supplementing information
2	F	1971	F, M	OT	Disability pension
3	F	1954	M	None	Sickness benefits
4	M	1956	G, M	None	Denied benefits
5	M	1967	F	None	Sickness benefits
6	F	1965	M	OT	Sickness benefits

7	M	1959	M, F	OT	Denied benefits
8	F	1950	M	OT, PT	Denied benefits
9	F	1977	-	-	-
10	F	1974	-	-	-
11	M	1958	M	None	Sickness benefits
12	F	1957	F	PT, OT	Sickness benefits
13	F	1958	I, L	None	Sickness benefits
14	M	1952	J, M	None	Disability pension
15	F	1950	F	None	Disability pension
16	F	1974	M	Psychologist	Sickness benefits after supplementary information
17	M	1954	M	None	Sickness benefits after supplementary information
18	F	1958	F, M	PT, OT, psychologist	Sickness benefits
19	F	1962	M, F	PT, OT, psychologist	Disability pension
20	F	1982	Q, M	None	Sickness benefits
21	M	1950	-	-	-
22	M	1954	I	OT, psychologist	Denied benefits
23	F	1952	M	PT, OT	Denied benefits
24	F	1960	M, H	PT	Disability pension
25	F	1969	F, R	PT, OT	Sickness benefits
26	M	1960	M	None	Sickness benefits
27	F	1979	G	PT, OT, psychologist	Sickness benefits
28	F	1977	-	-	-
29	F	1959	F, R	Psychologist	Denied benefits
30	F	1979	G	None	Denied benefits
31	F	1981	M	None	Denied benefits
32	F	1969	-	-	-
33	M	1967	R, M	OT	Sickness benefits
34	M	1951	M	OT	Denied benefits
35	M	1950	F	OT, psychologist	Denied benefits
36	F	1956	-	-	-

OT=occupational therapist; PT=physiotherapist. Diagnostic categories refer to the diagnostic manual ICD-10.

Table 2: Overview of interviewed insurance officials and physicians.

Insurance officials									Physicians			
No	Sex	Number of AFU ordered	No	Sex	Number of AFU ordered	No	Sex	Number of AFU ordered	No	Sex	Specialty	Share of AFU w. additional professionals (approx.)
1	F	2	13	F	2	25	F	4	1	M	Orthopedics	20%
2	F	2	14	F	6	26	F	1	2	M	Rehabilitation medicine	20%
3	F	1	15	F	2	27	F	2	3	F	-	10%
4	F	1	16	M	10	28	F	7	4	M	Orthopedics	5%
5	F	6	17	F	5	29	F	-	5	F	-	-
6	F	4	18	F	1	30	M	3	6	F	Rehabilitation medicine	30-40%
7	F	5	19	F	8	31	F	4	7	M	Psychiatrist	10%
8	F	5	20	F	5	32	M	7	8	M	Rehabilitation medicine	90%
9	M	4	21	F	5	33	F	7	9	M	Orthopedics	25-35%
10	M	10	22	F	4	34	F	6	10	M	Orthopedics	2-3%
11	M	5	23	F	9	35	F	0				
12	F	7	24	F	-	36	F	2				

People on sick leave were asked about their experiences of the assessment procedure, their contacts with the SSIA and the health care professionals, if they considered that the assessment was a fair representation of their work ability, and if the SSIA made a fair decision. Insurance officials were asked about their experiences of the AFU; how it is used in practice; how it affects decisions; how the information in the assessment is valued; conditions for using it; and use of the reference material on job demands. Physicians were asked about their opinions about the usefulness of the AFU, if any aspects are especially difficult to assess, and in which situations and for what reasons they used extended assessments. All interviews were carried out by phone, and lasted between 10 and 90 minutes, where some interviews with physicians were particularly short. The interviews were recorded using a digital audio recorder, and transcribed verbatim.

The material was analyzed through a manifest content analysis [27], focusing mainly on what respondents explicitly said, rather than performing interpretative analyses. The interview

transcripts were first read repeatedly to familiarize with the material, after which categorizations was made separately for each respondent group, based on identification of meaning units that formed sub-categories and categories. In this phase, the three respondent groups were categorized by different authors. The draft categorizations were cross-checked by another author. This first step sorted the data in descriptive sub-categories, which was later bundled into larger categories. For a description of the categorization in this step, see table 3.

Table 3: Categories and sub-categories in the first step of the analysis.

Respondent group	Categories	Sub-categories
<i>Insurance officials</i>	Administrative and legal aspects	Number of AFU ordered
		Selection to AFU
		Compatibility with rules and procedures
	Usefulness	Relation to other available methods
		Use of reference material
		Use of extended assessments
		Easier/harder to make decisions
		Amount of information in AFU
	Interpretation and outcomes	Interpretation of AFU results
		Outcomes from AFU
	Perceived fairness	Representation of work ability in AFU
		Communication with sick-listed
Participation aspects		
<i>Physicians</i>	Usefulness	Practical aspects of performing AFU
		Aspects difficult to assess
		Relation to other available methods
	Use of and views on extended assessments	Use of extended assessments
		Professionals consulted in extended assessments
	Usefulness of extended assessments	
<i>Sick-listed persons</i>	Assessment procedure	Purpose of the assessment
		Experiences of contacts with the SSIA
		Professionals involved in AFU
		Experiences of contacts with assessing professionals
		Clarity of assessment procedure
		Self-assessment of work ability
		Participation aspects
	Decision-making procedure	Experiences of contacts with the SSIA
		Clarity of decision-making
		Outcomes from AFU
	Perceived fairness	Correspondence between AFU and self-assessment
		Perceived fairness of assessment
		Perceived fairness of decision

In the next step, the results were compared between respondent groups, where the material was related to the key concepts of the literature on social validity [16] and perceived justice [17]. These concepts were used to answer the question whether or not the AFU could be considered to be a fair and socially valid assessment tool. Social validity was related to the perceived acceptability of goals, procedures and outcomes, where acceptability of procedures and outcomes was considered equivalent to procedural and distributional justice, respectively. Interactional justice was considered as referring to contacts with both assessment professionals and insurance officials in the assessment and decision-making procedures. The acceptability of goals was related to how the respondents described their general perceptions of and attitudes to the AFU, i.e., the purpose and use of the method. Social importance, as described in the social validity literature, was considered as a broader concept containing all aspects described above.

Results have been validated through presentation for and discussions with representatives from the SSIA, and physicians carrying out AFU assessments.

Ethical considerations

All respondents gave informed consent for participation. People on sick leave gave written consent, while professionals gave oral consent. The study, including the procedures for receiving informed consent, was approved by the Regional Ethics Review Board in Linköping, Sweden (dnr. 2014/15-31).

Results

In this section, results will be presented based on the categorization of responses from each group (insurance officials, physicians, and people on sick leave); thereafter, a section in which the different actors' perspectives on the perceived fairness and social validity of the AFU is presented.

Insurance officials: the AFU as fair and useful

As for the administrative and legal aspects of using the AFU, the insurance officials working with the AFU received “information rather than training” on how to use the assessment. At the time of the interviews, the officials had carried out relatively few assessments (ranging

from none to about ten), and commonly had not done so for a long time; the assessments were primarily carried out around the time it was introduced in the offices. Reasons given for this limited use was that the work load was too high, and that they currently prioritized paying out benefits over carrying out time-consuming assessments. Officials also noted how the management's priorities tended to shift rapidly within the SSIA, and that the AFU was not currently on the agenda. As a consequence, officials either chose not to send people to assessments, or used other procedures instead of the AFU.

The officials had discretion to choose who to send for assessments, where they (according to instruction) primarily determined this based on the length of the case, and if there were uncertainties or lack of information. Overall, the officials were of the opinion that the criteria should be based on the need for an assessment, rather than that they had passed a specific time point. Some chose to do an AFU when the person ought to be “declared healthy” (i.e., able to work); others primarily in cases where return to work was considered unlikely but more information was needed to make such a decision; others when a person could be considered eligible for permanent disability pension. A recurrent comment was that the AFU was used when the official had a gut feeling that something was “wrong”, where the assessment tended to confirm this feeling.

Many times, when I've sent people on assessments, I have had a gut feeling, and I think I often get that feeling verified. For instance, a person with such serious memory problems that it qualifies for disability pension, or that a person needs a change of jobs. (Insurance official 7)

The officials considered the AFU to be a useful assessment method that facilitates for officials to make relevant and correct decisions. It was also perceived as facilitating dialogue with external actors, such as the Public Employment Service, for people who were assessed as able to work in another job. One positive aspect of the AFU that the officials emphasized was the participation from the sick-listed person, through the self-assessment and dialogue with the physician, other healthcare professionals (in some cases), and the insurance official. It was thought to make the assessment and decision-making process more transparent and thereby to facilitate for the person on sick leave to understand and accept the decisions.

...it forces me to talk to the person who I'm sending to this, and to reason about what an assessment is, what kind of questions will be asked, what it's used for and what it can lead to. [...] It can lead to benefits being denied, or it could lead to this or that. So, you will have had those discussions.

(Insurance official 5)

The officials considered the information they received from an AFU to be relevant, nuanced and useful, albeit extensive. Sometimes, the officials wished that the assessment could be done over a longer time period, to capture fluctuations in work ability. They also wished to be able to choose if an extended assessment should be carried out or not, a decision currently made by the physician in charge of the assessment. This was considered a weakness compared to other available methods, where team assessments were used in all cases.

The officials had different views on the usefulness of the reference material of demands in typical jobs. Some had not used it at all, some were positive, and some were negative. All agreed that it was only to be used as a reference, and that it could not be a basis for a decision to deny benefits. The officials with a positive view considered it as something to use in a dialogue with the person on sick leave, while those with a negative view thought it did not give an adequate description of the labor market and hence could not be used in communications with the person on sick leave.

The interpretation of the assessment was partly complicated by the amount of information in the assessments, and the officials frequently used insurance medicine specialists at the SSIA to help with this.

It doesn't always provide a clear answer. Sometimes it's hard to get a coherent picture of what comes in [...] Then I'll need to discuss it with a group and a specialist. (Insurance official 22)

The assessments did not give much guidance on what conclusions to draw regarding the work ability of the person, which was either implied “between the lines” or were open for interpretation. The officials wanted more support and guidelines for making such conclusions. The assessments varied in terms of quality, and in some cases the numerical assessments and the text descriptions of the person's abilities did not seem to match very well, which according to the respondents made it difficult to know how to use the information. It was

considered positive that the AFU offered an evaluation that was independent from the physician who certified the sick leave.

The officials did not think that the assessment led to a typical outcome – this depends on the type of case and the assessment results. Most commonly, it led to the person returning to the job they were sick-listed from, continued sick leave, disability pension, or denied benefits if the person was assessed to be able to work in another job. In either case, the AFU facilitated for the official to make and argue for the decision.

As for the fairness of the AFU, the respondents considered it to give a fair assessment of the person's work ability, and that the person on sick leave generally agreed with it, although there could be discord if the assessment led up to a decision where benefits were denied. In some cases, officials reported, people on sick leave could consider the assessment to be correct and fair, but that the decision was wrong. This emphasizes the need for dialogue about the results to make sure the decisions are considered legitimate.

...they often feel that it was good that they did this assessment [...], that they were part of the decision, even if it means that "oh, shit, I probably need to report myself healthy", they are in on it. (Insurance official 7)

Some officials felt that the AFU focused too much on matching against other jobs, since this was considered difficult for people on sick leave. Most of the people undergoing the AFU, according to the officials, also assessed their work ability in relation to their regular job, which may explain why some did not consider the assessment to be fair, since it is related to any type of job. Hence, the officials noted that the feedback to the person was crucial, and that this could be improved to secure that decisions are made in dialogue, and thus considered acceptable.

Physicians: the AFU as complex but with potential

The physicians generally appreciated the AFU as an assessment tool since it gives much and relevant information. It did however take some time before they could master the method.

It took time to conquer the assessment form [...]. There's a long starting process where you need to carry out a lot of assessments. I've done three to

four hundred by now, and I still feel that I'm developing for each one.

(Physician 8)

Aspects that the physicians found specifically difficult to assess were activity limitations and endurance in relation to different activities. Activity limitations was considered difficult to assess using only the physician's method, where occupational therapists and psychologists were seen as useful. Endurance was difficult to assess given the time available for the assessment. The assessment of these aspects was primarily based on self-reports by the person being assessed, or a heuristic assessment by the physician. A general difficulty with self-reports, according to one physician, is that people tend to adapt to their limitations which makes them avoid certain activities; if they did not take this into account when reporting their abilities, the self-report would lead to false conclusions that they could manage activities which they could not. Further, mental capacities were considered difficult to assess. The design of the descriptors was also criticized, where a single figure was supposed to indicate a variety of aspects.

There were also some practical obstacles to performing the AFU, including long travel distances. This could, for instance, result in that an extended assessment was not carried out, since it would imply multiple long-distance trips in a short period for the sick-listed person. Another practical issue was the irregularity of when the SSIA ordered an AFU, which made it difficult to prepare for them. These aspects were all related to the AFU being a pilot method, which meant that there were a limited number of assessment clinics, and a limited number of assessments being ordered.

A common opinion was that extended assessments was not regularly required; most of the physicians claimed to use it in a range from a few percent to around a third of the cases (see table 2). Some physicians noted that the use of extended assessments was limited by the time required to perform them. One physician stood out by claiming to use it in almost all cases. Extended assessments were however generally considered helpful, since it validated the physician's assessment and gave additional insights. The physicians emphasized how extended assessments were helpful in more complex and longer sick leave cases, where a team assessment was considered useful; many of these persons have comorbidity where multiple diagnoses demand a broader assessment. Some physicians also pointed out how

extended assessments could add to a more objective and legally secure assessment, e.g., through using validated assessment methods that can clarify the activity limitations.

It is important that the person is given the opportunity to show the capabilities in an actual activity. (Physician 6)

The information from the extended assessments was considered helpful for physicians to document limitations in their communication with the SSIA, where occupational therapists and psychologists were considered to offer valuable contributions.

Sick-listed persons: the AFU as an administrative tool

The people on sick leave considered the main purpose of the AFU to be to give the SSIA material for granting or denying benefits. Hence, the assessment was considered having an administrative purpose, rather than a rehabilitative one. The AFU was presented to them as an opportunity or a privilege, since it would give them a certificate that they cannot do any type of job on the labor market, which they could use in further contacts with the Public Employment Service. Most interviewees had been informed about the AFU through mail, and in some cases through personal meetings.

“We happen to have this project we have gotten money for, from the state, so maybe we should use it. It’s a work ability assessment.” And I replied that it sounds interesting, because I am open to trying different things. I won’t be difficult. (Sick-listed 24)

In some cases, the officials explicitly described in their communication how the current medical certificates were not considered enough to assess their work ability, where the person on sick leave tended to experience the invitation as coercive, questioning their right to benefits. In their meetings with the SSIA, these people felt like a burden or a problem. Some of these people were very critical, describing the SSIA and their regulations in harsh terms.

What can I say? I was met in a professional but ultimately inhumane way. And I know there is no use blaming the people who work with it, because they didn’t decide upon the rules or laws. But it’s terribly inhumane, all contacts with the SSIA, really. (Sick-listed 7)

The people on sick leave varied in their experiences of the assessment procedure. Some had positive encounters with physicians and other healthcare professionals, and perceived them as

empathetic, committed, accommodating and listening to them. In some cases, the assessment resulted in new diagnoses and had the consequence that the SSIA approved continued benefits, after previously having questioned their eligibility. These people were very positive to the assessment.

I talked with [a psychiatrist] for a couple of hours and got my diagnosis, post-traumatic stress, and she made sure I didn't have to do this job anymore [...]. She was a professional, I should have met her years ago. But since I wasn't on sick leave before, I didn't. So that experience was just fabulous. (Sick-listed 15)

Others were more critical to the assessment, especially that it was performed in an isolated clinical setting that did not represent a realistic work situation. These respondents mentioned several aspects that the assessment did not capture, such as the social dimension of work, the stress of having to perform different tasks simultaneously, or endurance over a longer period of time. Some respondents described the AFU as “thin” and including too few aspects, while others considered it too extensive for the time allotted to the assessment, leading to physical and mental strain. Respondents had varying experiences of whether or not the physician included and related the assessment to previous medical certificates.

A recurrent theme in the interviews was that the respondents had problems understanding the relevance of the different parts of the assessment, and in some cases the overall purpose. The different questions and tests appeared strange to some, and their usefulness was difficult to grasp. Some respondents also reacted to what they perceived as standardized tests and questions that they could not see the relevance of in relation to their condition, e.g., questions about daily routines and hygiene. It was difficult to understand how this was related to their work ability.

The questions are for a large group of people, so to speak, and didn't have anything to do with me or my condition [...]. I have a problem, so why don't we focus on that, instead of the whole body? (Sick-listed 6)

This perspective was contrasted by other respondents who lacked a holistic perspective in the AFU, and that the assessment failed to take into account aspects of their health that had relevance for their sick leave case, and hence experienced the assessment to be insulting and unprofessional. Both of these perspectives can be said to relate to the design of the

questionnaire in the AFU, where the standardized questions and procedures were either considered too broad (if the person had a specific problem) or too specific (if the person wanted a more in-depth assessment of their individual situation to not miss important aspects).

Commonly, the persons on sick-leave related their work ability to their previous job, rather than to any job on the labor market. They directed critique towards the AFU for not taking the work conditions in their job into consideration, and that their work ability could have been better assessed at their workplace. Some of these respondents considered the physicians performing the assessment to be “in league” with the SSIA, with the purpose to get them off benefits.

The decision-making process was described differently depending on the outcome. Those who had their benefits denied after the AFU tended to feel like “closed cases”, where the official was not perceived to be interested in discussing the case further when the decision had been made. Some respondents perceived that they were not able to communicate with the insurance officials about the interpretation of the AFU and what conclusions and decisions to draw from it. Others described positive experiences where the AFU led to them being granted benefits. In these cases, no concerns were raised about the contacts with the SSIA.

Hence, the relationship between the self-assessment and the decisions made by the SSIA determined the retrospective perceptions of whether or not the assessment and the decision-making was fair. There were however examples also of respondents with denied benefits who could accept the decision since the procedure and communication with the insurance officials were perceived as positive.

She's positive and forward-looking [...]. It's not her fault that they have these rules, so she shouldn't get the critique, it's the people who decided on the rules. (Sick-listed 6)

Others considered the assessment to be fair, but the interpretation of it, and hence the decision, as unfair. For instance, one person mentioned how the physician and the insurance official made different interpretations of the figures in the assessment form (given on a scale from zero to four):

The physician said, “yes, I see that you have rather severe limitations, so I’ll give you a two. And the SSIA said that “a two, that’s nothing”. (Sick-listed 16)

Is the AFU a fair and socially valid assessment?

The social validity of the AFU, i.e., the acceptability of goals, procedures and outcomes, was perceived differently between actors. The goal of the AFU was generally perceived by people on sick leave as serving the SSIA’s needs for providing material for decisions on eligibility for benefits; it was hence perceived primarily as an administrative tool. Insurance officials had a slightly broader view, where they also emphasized the usefulness of the assessment in further contacts with the Public Employment Service. Physicians did not speak much of the goals of the AFU, but some noted that an extended assessment may increase the legal security of work ability assessments, and that it was a professional assessment method. The social importance of the goals, hence, goes back to the sickness insurance regulations, where the AFU may be considered by policymakers to be useful for deciding on who should and who should not receive benefits; the acceptability of these goals for those on sick leave, however, seems to be highly dependent on the outcome of the assessment. The goal and outcome dimensions of social validity are hence closely interwoven, where the perceived distributive justice will influence whether or not it is considered socially important and acceptable.

The acceptability of procedures refers to the perceived fairness of the assessment, especially relating to the procedural and interactional dimensions of justice. The results show that the three respondent groups had slightly different perceptions of what a fair assessment means. While the insurance officials considered the self-report to vouch for a more participatory and hence more fair assessment, one physician claimed that there were problems with self-reports, since they could lead to an overestimation of abilities. The people on sick leave, on the other hand, did not generally seem to consider their self-report as particularly important for their perception of fairness – few even mentioned it. The insurance officials and the people on sick leave did however agree that the dialogue about the assessment and the decisions made from it was essential, although the perspectives on this differed. While officials considered this primarily as a way to increase acceptance for the decision and make the person on sick leave understand why benefit claims get denied, people on sick leave seemed rather to consider dialogue important in order to secure a correct and reasonable interpretation of the findings in

the assessment, as in the case with different interpretations of figures in the AFU (described above).

Finally, the “fair or square” question also relates to the design of the AFU, and whether or not it is sensitive enough to offer a relevant assessment of the person under review, including the characteristics of the individual situation, or if it is too standardized and only gives a rudimentary picture of the person’s abilities. The perceptions of this varied greatly among the people on sick leave: some were very positive (especially those who received continued benefits), some very critical (especially those who were denied benefits). Most insurance officials and physicians were positive to the AFU, although admitting that the assessment was a simplification with certain limitations, and that it, even so, took time to master it. In the end, the perceived fairness of the assessment process was dependent on the design of the AFU, the outcome of the assessment, and the dialogue related both to the procedure and to the decision-making, which indicates that its practical implementation will influence whether or not it is considered to be socially valid.

Discussion

The results showed a complex and varied picture of the AFU and how it was considered in terms of fairness and social validity. Some of the issues brought up by the insurance officials and physicians were related to the pilot phase of the AFU (e.g., lacking routines, too few assessment clinics, lacking experience in carrying out assessments). Still, social validity is a relevant concept to use for assessing a process perspective on interventions [16], where the pilot phase examined here may give relevant information about how the assessment was perceived in order to develop or fine-tune it before a full-scale implementation. Social validity and fairness are however not concepts that are easily dichotomized with cut-off points of what can be considered “enough”, but are better thought of as a series of continua [16].

Some aspects mentioned by the insurance officials are relevant to the method’s usefulness and prospects for full-scale implementation. Here, the high case-load of the officials and the priorities of the management prevented the use extensive assessments such as the AFU, since they are administratively burdensome and result in much information that needs to be interpreted. Other aspects raised as problematic were related to the design of the AFU, e.g., the procedures for making extended assessments; another concern was that the assessment

was chosen more based on the timing in the sick leave case (around a specific time limit), rather than based on the need for the assessment. These concerns refer to the acceptability of procedures, which are dependent on organizational and societal factors, such as management priorities and regulations. When the work piles up, it may be easier for the insurance officials to perform a simple assessment based on existing information, which however may risk that important aspects are not considered, as exemplified in the cases where the AFU resulted in new diagnoses and thereby eligibility for benefits.

The social importance of the AFU is perhaps its most intricate dimension, since it deals with the legitimacy of the method as such. It was introduced to facilitate assessments where sickness insurance regulations stipulate that a general work ability assessment [8] is to be carried out, for which there were no available methods. If an assessment marks a person who is obviously feeling sick or work disabled as “fit”, “healthy”, or “able to work”, it will need a great bit of argument to be considered acceptable for the person in question. This is even more so in cases where the person has one or several medical certificates that says he/she is too sick or disabled to work. The legal argument – that work ability is assessed in relation to other jobs at a certain time point and that previous certificates did not assess this – may not sit well with the person if they consider themselves as able to return to their regular job given some more time for rehabilitation. In cases where this dialogue did not take place, our results indicate that the person on sick leave experienced the system as inhumane; they expressed much distrust and questioned the legitimacy of both the assessment and the system as a whole. Concurrently, other people expressed how the assessment was acceptable or even helpful – especially in cases where it helped them getting prolonged benefits. Even where this was not the case, there were examples where people were denied benefits, but could accept this through a relevant dialogue with the official from the SSIA. This indicates that the fairness, and hence the social validity, of the assessment is at least partly dependent on how it is carried out in practice and on the communication skills of the professionals involved. This observation is supported by previous studies, where the interactions with the insurance agency ranged from fragmented to coherent, with subsequent consequences for recovery [28]. The potentially detrimental effects of interactions with insurance systems have been studied also in other contexts [29], and it has been concluded that systems that manages to limit adversarial interactions and that provides substantive protection is better suited to preserve claimants’ dignity [30].

The regulation that requires general work ability assessments is based on a broad societal discourse of reforming insurance systems into focusing more on activation, which in a work disability context is translated into focusing on abilities rather than disabilities. This discourse is today prevalent in many jurisdictions, and can be seen in policy, legislation and scientific studies [31]. For instance, the UK system has replaced its previous sick note with a “fit note” [32], which corresponds to the idea of work as healthy and as a central aspect of a person’s rehabilitation process [31, 33]. The same underlying ideas build up the Swedish focus on work ability rather than disability – one purpose of the work ability assessment is to scrutinize if there are remaining abilities that can be used in other types of work. However, one obvious risk with these approaches is that they are solely used to deny benefits, and that the assessments are not related to any types of rehabilitation measures to actually build upon the remaining abilities [31]. This was also reflected in the results in this study, where people assessed as able to work had their cases closed, after which nothing more happened – the contacts with the SSIA simply ended.

Assessing capacities to determine eligibility for benefits is a common feature in compensation systems for work disability, where the “ability paradigm” [31] often has a backside of suspicion and distrust in relation to claimants’ efforts in trying to acquire and manage work [34]. This suspicion of the side of insurance systems goes back to the idea of moral hazard, i.e., that the behavior of people is affected by the availability of insurance, in this case related to income replacement. While this has some relevance (e.g., studies within the Swedish system have shown that the availability and level of benefits were related to sick-listing patterns [35, 36]), there is also a discussion that the idea of moral hazard is essentially a moral and ideological concept born from the economic disciplines [37, 38], and that it has spurred a discourse of abuse [34] leading to expectations of dishonesty and malingering. The literature however suggest that work remain a meaningful goal for people with work disability [39], and workers who collect benefits “exert considerable effort to [return to work] even years following their injury” [34, p. 111]. The notions of mistrust in claimants’ motivation to work hence seem unjustified; rather, disabled workers make efforts that go unrecognized by the system actors [34]. The desire to return to work is related to workers wanting to maintain their identity and to feel as contributing to society, which has been reported in other studies [e.g., 40].

The social importance of the AFU, hence, is related to both the regulations that it is based on (whether or not it is considered appropriate to base benefits on general work ability assessments), and, if these are considered acceptable, how the assessment is presented, carried out, and used. Swedish policy on work disability is much inspired by the “ability paradigm”, which makes the goal of the AFU acceptable from a policy perspective. For it to be considered having social validity, it also needs to be implemented in ways that are considered acceptable to clients, so that they feel that their cases have been managed in a fair way.

Methodological considerations

The qualitative design and the inclusion of multiple respondent groups served to increase the trustworthiness of the results. The trustworthiness of the analysis was increased by a feedback round where results were presented and discussed with the SSIA and physicians carrying out AFU assessments. The analysis was carried out by several researchers independently and jointly, to increase its reliability. The study has some limitations which are generic to qualitative studies, where the results cannot be generalized to a larger population; on the other hand, the study has managed to capture a variety of experiences, ranging from very positive to very negative, which indicates that the results are not biased to reflect only the experiences of a certain group. The study was carried out in the pilot phase of the AFU, and it is possible that both procedures and perceptions of its usefulness will change when it is implemented on a larger scale. Recruitment of people on sick leave was carried out through letters mailed by the SSIA, which may have affected the sample, with a potential bias leaning either towards those not satisfied with the process or outcome, or towards those who considered the AFU as especially helpful. The range of experiences expressed in the data however suggests that bias, if prevalent, did not favor one specific group of respondents. Another potential bias was that the respondents had already had their decision from the SSIA at the time of the interview, meaning that their retrospective accounts of the acceptability of the procedure may have been influenced by the outcome; on the other hand, it could be argued that knowing the outcome may be relevant in order to assess the procedure properly.

Conclusions

The social validity of the AFU, that is, its social importance and acceptability concerning goals, procedures and outcomes, was much dependent on how it was used. If primarily used

to assess eligibility, and if not coupled with sufficient dialogue and participation in influencing and interpreting the results, it may be considered unacceptable for those who are assessed as able to work but who feel that they cannot. They may perceive the assessment as unfair, relating to the distributive, procedural and interactional dimensions of justice. The social importance of the AFU will also differ depending on which actor we consider. From the point of view of system representatives, i.e., legislators, policymakers and authority representatives, the AFU may be seen as socially important as it is currently used. If we include the perspectives of the people being assessed, attention also needs to be placed on its potential for fulfilling rehabilitation goals. If not, the assessment may be perceived simply as a tool used by a suspicious regime to withdraw benefits. In sum, the AFU seems to have potential for increasing its level of social validity if carried out with sensitivity to how the process and its results are communicated.

Declaration of interest

The authors report no conflicts of interest. The study was funded by a research grant from the SSIA; however, the funding agency had no role in study design, data analysis, data interpretation, writing of the report, or the decision to submit for publication. The authors had full access to all the data in the study and had the final responsibility for the decision to submit for publication.

Authors' contributions

CS, BG and JS were involved in study conception and study design. CS, IS and JS carried out data analyses. CS drafted the manuscript. All authors discussed the results, commented on the manuscript, and approved the final version.

References

- [1] Lessing D. The Sirian Experiments. The report by Ambien II, of the Five. London: Flamingo; 1994.
- [2] Cooper T. Big Questions in Administrative Ethics: A Need for Focused, Collaborative Effort. *Public Administration Review*. 2004;64(4):395-407.

- [3] Cronin S, Curran J, Iantorno J, Murphy K, Shaw L, Boutcher N, et al. Work capacity assessment and return to work: A scoping review. *Work*. 2013;44(1):37-55.
- [4] Ilmarinen J. Work ability - A comprehensive concept for occupational health research and prevention. *Scandinavian Journal of Work, Environment and Health*. 2009;35(1):1-5.
- [5] Ilmarinen J, Tuomi K, Seitsamo J. New dimensions of work ability. *International Congress Series* 2005. p. 3-7.
- [6] Lederer V, Loisel P, Rivard M, Champagne F. Exploring the Diversity of Conceptualizations of Work (Dis)ability: A Scoping Review of Published Definitions. *Journal of Occupational Rehabilitation*. 2014;24(2):242-67.
- [7] Nordenfelt L. *The Concept of Work Ability*. Brussels: P.I.E. Peter Lang; 2008.
- [8] Tengland PA. The concept of work ability. *Journal of Occupational Rehabilitation*. 2011;21(2):275-85.
- [9] Tengland PA. A qualitative approach to assessing work ability. *Work*. 2013;44(4):393-404.
- [10] Solli HM, Da Silva AB, Egeland J. Usefulness of an ability-based health model in work ability assessments provided by psychiatrists and psychology specialists writing social security certificates. *Disability and Rehabilitation*. 2015;37(7):571-8.
- [11] Finger ME, Escorpizo R, Bostan C, De Bie R. Work Rehabilitation Questionnaire (WORQ): Development and Preliminary Psychometric Evidence of an ICF-Based Questionnaire for Vocational Rehabilitation. *Journal of Occupational Rehabilitation*. 2014;24(3):498-510.
- [12] Abma FI, van der Klink JJJ, Terwee CB, Amick BCI, U. Bl. Evaluation of the measurement properties of self-reported health-related work-functioning instruments among workers with common mental disorders. *Scandinavian Journal of Work, Environment & Health*. 2012;38(1):5-18.
- [13] Noben CYG, Evers SMAA, Nijhuis FJ, de Rijk A. Quality appraisal of generic self-reported instruments measuring health-related productivity changes: a systematic review. *BMC Public Health*. 2014;14(115).
- [14] Cohen-Charash Y, Spector PE. The role of justice in organizations: a meta analysis. *Organizational Behavior and Human Decision Processes*. 2001;86(2):278-321.
- [15] Wolf MM. Social validity: the case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*. 1978;11(2):203-14.
- [16] Foster SL, Mash EJ. Assessing social validity in clinical treatment research: Issues and procedures. *Journal of Consulting and Clinical Psychology*. 1999;67(3):308-19.
- [17] Franche R-L, Severin CN, Lee H, Hogg-Johnson S, Hepburn CG, Vidmar M, et al. Perceived Justice of Compensation Process for Return-to-Work: Development and Validation of a Scale. *Psychological Injury and Law*. 2009;2(3):225-37.
- [18] Barr B, Taylor-Robinson D, Stuckler D, Loopstra R, Reeves A, Whitehead M. 'First, do no harm': Are disability assessments associated with adverse trends in mental health? A longitudinal ecological study. *Journal of Epidemiology and Community Health*. 2016;70(4):339-45.
- [19] Barr B, Taylor-Robinson D, Stuckler D, Loopstra R, Reeves A, Wickham S, et al. Fit-for-work or fit-for-unemployment? Does the reassessment of disability benefit claimants using a tougher work capability assessment help people into work? *Journal of Epidemiology and Community Health*. 2015;70(5):452-8.
- [20] McGurk R. Pain of work capacity assessment. *Psychologist*. 2013;26(11):781.
- [21] Baumberg B, Warren J, Garthwaite K, Bamba C. *Rethinking the Work Capacity Assessment*. London: Demos, 2015.

- [22] Ståhl C, Svensson T, Ekberg K. Swedish rehabilitation professionals' perspectives on work ability assessments in a changing sickness insurance system. *Disability and Rehabilitation*. 2011;33(15-16):1373-82.
- [23] Inspektionen för socialförsäkringen. Aktivitetsförmågeutredningar (AFU). Stockholm: Inspektionen för socialförsäkringen, 2015.
- [24] El Fassi M, Bocquet V, Majery N, Lair ML, Couffignal S, Mairiaux P. Work ability assessment in a worker population: Comparison and determinants of Work Ability Index and Work Ability score. *BMC Public Health*. 2013;13(1).
- [25] Ahlstrom L, Grimby-Ekman A, Hagberg M, Dellve L. The work ability index and single-item question: Associations with sick leave, symptoms, and health - A prospective study of women on long-term sick leave. *Scandinavian Journal of Work, Environment and Health*. 2010;36(5):404-12.
- [26] Patton MQ. *Qualitative research and evaluation methods*. 3 ed. London: Sage; 2002.
- [27] Hsieh H-F, Shannon S. Three Approaches to Qualitative Content Analysis. *Qualitative Health Research*. 2005;15(9):1277-88.
- [28] Hubertsson J, Petersson IF, Arvidsson B, Thorstensson CA. Sickness absence in musculoskeletal disorders - Patients' experiences of interactions with the Social Insurance Agency and health care. A qualitative study. *BMC Public Health*. 2011;11.
- [29] Kilgour E, Kosny A, McKenzie D, Collie A. Interactions Between Injured Workers and Insurers in Workers' Compensation Systems: A Systematic Review of Qualitative Research Literature. *Journal of Occupational Rehabilitation*. 2015;25(1):160-81.
- [30] Lippel K. Preserving Workers' Dignity in Workers' Compensation Systems: An International Perspective. *American Journal of Industrial Medicine*. 2012;55(6):519-36.
- [31] MacEachen E, Kosny A, Ferrier S, Lippel K, Neilson C, Franche R-L, et al. The 'Ability' Paradigm in Vocational Rehabilitation: Challenges in an Ontario Injured Worker Retraining Program. *Journal of Occupational Rehabilitation*. 2012;22(1):105-17.
- [32] Verbeek J, Madan I. From sick notes to fit notes. *BMJ*. 2009;339.
- [33] Waddell G, Burton AK. *Is work good for your health and well-being?* Norwich: TSO; 2006.
- [34] Saunders SL, MacEachen E, Nedelec B. Understanding and building upon effort to return to work for people with long-term disability and job loss. *Work*. 2015;52(1):103-14.
- [35] Johansson P, Palme M. Moral hazard and sickness insurance. *Journal of Public Economics*. 2005;89(9-10):1879-90.
- [36] Hall C, Hartman L. Moral hazard among the sick and unemployed: Evidence from a Swedish social insurance reform. *Empirical Economics*. 2010;39(1):27-50.
- [37] Dembe A, Boden L. Moral hazard: A question of morality? *New Solutions*. 2000;10(3):257-79.
- [38] Stone D. Behind the Jargon: Moral hazard. *Journal of Health Politics, Policy and Law*. 2011;36(5):887-96.
- [39] Saunders SL, Nedelec B. What work means to people with work disability: A scoping review. *Journal of Occupational Rehabilitation*. 2014;24(1):100-10.
- [40] Ståhl C, Stiwne EE. Narratives of Sick Leave, Return to Work and Job Mobility for People with Common Mental Disorders in Sweden. *Journal of Occupational Rehabilitation*. 2013.