Kenneth Sandin

Psychological therapy with work-focus for patients on or at risk of sick leave due to depression and anxiety: Heterogeneity in sick leave and self-reported health

Thesis for the Degree of Philosophiae Doctor

Trondheim, December 2023

Norwegian University of Science and Technology, NTNU
Faculty of Social and Educational Sciences
Department of Psychology

Table of Contents

ABSTRACT

Depression and anxiety are mental disorders that have broad effects on health and well-being. Affected individuals often suffer from a reduced ability to function in daily life, including at work. Compared to non-clinical populations, patients suffering from depression and anxiety have lower rates of employment and higher rates of sick leave, and an increased risk of long-term exclusion from the labour market. The combined cost of these disorders to individuals and society makes them comprehensive public health issues in need of effective interventions.

The last ten to fifteen years have seen efforts to expand and improve treatment for these patients. This includes an increased focus on interventions that also target employment status through aiding return to work for patients on sick leave. These interventions could help reduce individual suffering though alleviating symptoms, but also lower societal expenses associated with lost productivity from sick leave.

A key challenge in developing effective interventions is the heterogeneity of depression, anxiety, and sick leave. A more nuanced understanding of how these patients differ, and how they thus may have different needs, could represent a key step in improving interventions and outcomes. The aim of this PhD thesis was thus to examine the characteristics of patients with depression and anxiety referred for work-focused treatment.

In Paper I, registry data was used to examine trajectories of sick leave among patients before, during, and after work-focused treatment. The registry data were provided by the Norwegian Labour and Welfare Administration. This data gave an objective measure of sick leave with no loss to follow-up and were analysed using latent growth mixture modelling. Clinical and sociodemographic characteristics of the different trajectories were then analysed post-hoc. Three different sick leave trajectories were found. Female gender and age were associated with higher sick leave at baseline, whilst residual depressive symptoms increased risk of continued sick leave after treatment. One group comprised of almost half the patients had little to no sick leave for the entire follow-up period, despite

comparable symptom levels to the other two groups. The findings in this study indicate that work-focused treatment may be improved by tailoring treatment more closely to patient needs.

In Paper II, the validity of the EQ-5D questionnaire in patients on or at risk of sick leave was examined. The EQ-5D is a generic measure of health that can also be used for health economic analyses. The health status recorded by the EQ-5D was compared to disorder-specific measures of depression and anxiety using correlation and regression analyses. This cross-sectional study showed that the patients in the sample reported substantially reduced health status compared to the general population. Furthermore, that the EQ-5D was sensitive to different levels of depression and anxiety severity, and to functional impairment in the form of sick leave. The findings of the study suggest that the EQ-5D shows indications of validity in this patient group.

Paper III examined the longitudinal responsiveness of the EQ-5D to changes in health for patients that completed treatment. As in Paper II, the performance of the EQ-5D was compared to the performance of disorder-specific measures of depression and anxiety symptoms. Effect size and correlation suggested similar magnitude of change for the EQ-5D as the disease-specific measures during treatment. In addition, ROC-analyses indicated that the EQ-5D was able could differentiate patients who had "recovered" during treatment, and those who were "improved" or "unchanged". The findings suggest that the EQ-5D can be a useful instrument when evaluating change during clinical interventions for these patients.

In sum, the thesis offers insights on characteristics of patients on or at risk of sick leave due to depression and anxiety. The registry data analysed in Paper I suggests that there may subgroups of patients with different treatment needs, and that tailoring interventions may improve patient outcomes. Paper II and Paper III showed that a generic measure of self-reported health was able to capture health status and sick leave, as well as recovery during treatment. In addition, the EQ-5D questionnaire can also be used for calculating quality of life, and for health economic evaluations. Use of this questionnaire in future research can thus help further expand our knowledge of the impact of depression, anxiety, and sick leave.

Scientific environment

The work presented in this PhD thesis was carried out as part of the doctoral programme at the Faculty of Social and Education Sciences at the Norwegian University of Science and Technology (NTNU). The research itself was conducted the at Division of Mental Health and Substance Abuse at Diakonhjemmet Hospital in Oslo.

The supervisor for the thesis has been Professor Odin Hjemdal of NTNU and Diakonhjemmet Hospital. The papers included in the thesis are part of a wider research project on mental health and work focused treatment called “The Norwegian studies of psychological treatment and work (NOR-WORK)”.

Financial support for the work in this thesis was provided by a PhD grant from Stiftelsen DAM (previously Extrastiftelsen), with additional funding from Diakonhjemmet Hospital and the Southern and Eastern Regional Health Authority.

Acknowledgements

This PhD project, like all PhD projects, has been a collaborative endeavour. This is especially true as the work presented relies on clinical data collected from hundreds of patients over several years. There are thus a great many people without whose contribution this PhD would not have been possible. Firstly, I want to thank all the patients for their courage in entering treatment, and their generosity in sharing their data with us. Our humble effort to advance the understanding of integrating work in psychotherapy would not have been possible without their participation. Secondly, the staff at Diakonhjemmet Hospital; the clinicians, administrative staff, and research assistants. Their dedication to providing the best treatment possible while also remaining focused on data collection has been and continues to be invaluable. Finally, our clinical supervisors, Marit Hannisdal and Ragne G.H. Gjengedal, for their unwavering commitment to prioritising clinical research, and their patience and understanding with the twists and turns along the way.

A big thanks also to my co-authors. Cooperating with Frederick Anyan and Gemma E. Shields on the papers has been a privilege and a pleasure. Likewise, it has been educational and productive to receive input from Kåre Osnes, Ragne G.H. Gjengedal, Jonas S.R. Leversen. Silje E. Reme, and Marianne T. Bjørndal. In addition to the three papers in the thesis, we also published a study protocol for a randomised controlled trial during the course of this PhD. The co-authors on this protocol also provided feedback that has been helpful for the PhD work as a whole. The ones not already listed above are Lars G. Røv, Torkil Berge, Suzanne Lagerveld, Roland Blonk, Hans M. Nordahl, and Adrian Wells. Thanks also to the editors and reviewers who helped us on the papers in the process to publication. We are also grateful to our external benefactors, Stiftelsen Dam and the Regional Southern and Eastern Health Authority, and to the user organisation Mental Helse.

A particular note of gratitude to my supervisor, Odin Hjemdal. Your insightful comments and mild temper have taught me a great deal about what research is, and how to go about it.

To wrap up, a nod to Sigmund Freud. Although there is precious little mention of superegos or ids in the text that follows, the man had a point regarding the importance of working, while also maintaining

meaningful relationships. On that note, my gratitude and love to 7960 for providing much needed respite from pandemics and lockdowns. Thanks to V.V., G.S., and A.P. for the tunes. Many thanks to my friends for their diverse interests, none of which include work-focused treatment. To my parents, Ann Mari and Lars, for their continued support and enthusiasm over many years of study. And finally, to my wife Lam and my son William for making it all worthwhile.

List of publications

Paper I:

Sandin, K., Anyan, F., Osnes, K., Gjengedal, R.G.H., Risberg Leversen, J. S., Reme, S.E., & Hjemdal, O. (2021). Sick leave and return to work for patients with anxiety and depression: A longitudinal study of trajectories before, during and after work-focused treatment. BMJ Open, 11(9), e046336. https://doi.org/10.1136/bmjopen-2020-046336

Paper II:

Sandin, K., Shields, G. E., Gjengedal, R. G. H., Osnes, K., Bjørndal, M. T., & Hjemdal, O. (2021). Self-Reported Health in Patients on or at Risk of Sick Leave Due to Depression and Anxiety: Validity of the EQ-5D. Frontiers in Psychology, 12, 655151. https://doi.org/10.3389/fpsyg.2021.655151

Paper III:

Sandin, K., Shields, G., Gjengedal, R. G. H., Osnes, K., Bjørndal, M. T., Reme, S.E., & Hjemdal, O. (2023). Responsiveness to change in health status of the EQ-5D in patients treated for depression and anxiety. Health and Quality of Life Outcomes, 21(1), 35. https://doi.org/10.1186/s12955-023-02116-y

List of abbreviations

BAI: Beck's Anxiety Inventory

BDI-II: Beck's Depression Inventory-II

CBT: Cognitive behavioural therapy

CMD: Common mental disorders

EQ-5D Euro-Qol 5 dimensions questionnaire

IAPT: Improving access to psychological therapies

ICER Incremental Cost Effectiveness Ratio

M.I.N.I Mini-International Neuropsychiatric Interview

MCT Metacognitive therapy

NTNU Norwegian University of Science and Technology

OECD Organisation for Economic Co-operation and Development

QALY Quality-adjusted life year

RCT Randomised controlled trial

RTW Return to work

RTW-SE Return to Work Self-efficacy

SHC Subjective Health Complaints

W-CBT Work-focused cognitive behavioural therapy

1. INTRODUCTION

1.1 Impact of common mental disorders on work capacity

Common mental disorders and their impact on individuals and society have increasingly been recognised as major public health concern over the recent years (Lépine & Briley, 2011; World Health Organization, 2013; Yang et al., 2021). Typically, "common mental disorders" is taken to mean depression and anxiety disorders, which are the most prevalent mental illnesses in the population (World Health Organization, 2017). Point prevalence estimates of common mental disorders indicate that around one in five people suffer from one of these disorders at any given time. Over the course of a life, around one in three people will at some point satisfy the criteria for a depression or anxiety disorder (Steel et al., 2014).

These disorders are frequently comorbid (Kalin, 2020), and depression and anxiety have wide-ranging impact on the health and quality of life of the people who are affected (Angermeyer et al., 2002; Mendlowicz & Stein, 2000). People with depression and anxiety experience earlier physical and cognitive decline as they get older, and they also have an increased risk of a multitude of physical health problems, such as cardiovascular disease, stroke, and obesity (Penninx et al., 2013). Common mental disorders thus have a profound impact on health and are leading causes of ill health worldwide. Findings from the Global Burden of Disease study, a large multidisciplinary study examining trends in mortality and disability, indicate that both depression and anxiety are often chronic or recurring disorders resulting in both high levels of disability and excess mortality (Baxter et al., 2014; Ferrari et al., 2013). Recent estimates indicate that in terms of years lived with disability, mental illness may account for as much as a third (32.4 %) of the total global disease burden (Vigo et al., 2016). Further population-based studies have indicated that both depression and anxiety are significant contributors to excess mortality (Gilman et al., 2017; Meier et al., 2016). These findings underline the importance of finding cost-effective interventions that reduce the suffering caused by these disorders (Baxter et al., 2014; Ferrari et al., 2013; Vigo et al., 2016).

One of the key ways in which common mental disorders impact people's health and wellbeing is through the reduced ability to work (OECD, 2015). Mirroring the effect seen in studies of health, studies of welfare and benefit claims show common mental disorders to be a leading cause of disability pension (Ahola et al., 2011), and that they are also the of cause most of the long term sickness absence (Henderson et al., 2011). Even subthreshold symptoms of depression and anxiety, i.e. below the level warranting a clinical diagnosis, have been shown to be associated with a substantial increase in disability (Rai et al., 2010). As for diagnosed disorders, between a 30 - 50 % of new disability claims in the OECD countries are due to mental ill health, and for young people the number may be as high as 70 % (OECD, 2012). The younger a person is when going on disability, the more working years are lost, further underlining the role of common mental health problems in reduced employment and work disability (Knudsen et al., 2012).

Work can give structure and purpose to everyday life, promotes participation in society, and provides independence through income (Waddell et al., 2007). Conversely, unemployment is associated with future physical health problems such as poorer cardiovascular health (Meneton et al., 2015), and an increase in all-cause mortality (Reme, 2020), and even increased risk of suicide (Milner et al., 2013). It is also well known that unemployment is associate with problematic substance use and substance abuse disorders, with all the health problems that these disorders entail, but again, the direction of causality is not always clear (Henkel, 2011). As with general mental health, there may be a negative self-selection in that people who have problems with substance abuse have lower rates of workforce entry. One five-year follow-up study of an unemployed Norwegian sample examined this question in relation to alcohol abuse. The study concluded that negative self-selection into long-term unemployment does occur, but the strongest support was found for the impact of employment on reducing problematic alcohol use (Claussen, 1999).

In general, it is accepted that employment is good for both physical and mental health (OECD, 2019), and review studies confirm its importance for mental health in particular (van der Noordt et al., 2014). One review sought to investigate the association between mental health and work by analysing what is often framed as the "social selection hypothesis" of mental health and employment versus the "social

causation hypothesis". Briefly, the social selection hypothesis assumes that persons with poorer health will be excluded from the labour market, and that this negative selection accounts for the relationship between employment and mental health. The social causation hypothesis assumes the opposite, that the association between mental health and employment is due to improvement of health status by gaining or maintaining employment (Reme, 2020). The study found support for both hypotheses, but the improvement in mental health from gaining employment was larger than the relative deterioration from losing employment. Strongest support was thus found for the social selection hypothesis, that gaining employment improves mental health (Murphy & Athanasou, 1999).

Given the well-established link between employment and health, there has been a growing recognition over the latter years that treatment for depression and anxiety should also include interventions aimed at helping people maintain employment (OECD, 2012). For depression and anxiety disorders, it is the case that they carry symptoms that sometimes may warrant respite from work in the form of sick leave. But sick leave should probably be used sparingly. The evidence suggests that even short term sick leave may decrease future earnings, and increase the chance of long term absence and unemployment, risking the detrimental health outcomes associated with being excluded from the labour market (Hultin, Lindholm, & Möller, 2012; Markussen, 2012). Furthermore, the relationship seems dose dependent. The longer the sick leave spell lasts, the smaller the chances of the person making a full return to work (RTW) (Blank et al., 2008). Effective interventions for depression or anxiety should thus help patients avoid sick leave or assist with RTW to avoid prolonged absence.

1.2 Treatment approaches in depression and anxiety.

1.2.1 Cognitive behavioural therapy

Cognitive behavioural therapy (CBT) is a form of psychological treatment that is widely used for a range of mental health problems, and the method is commonly recommended for treating mood and anxiety disorders (NICE, 2020, 2022). Research has underlined the overall efficacy of CBT, especially when treating mood and anxiety disorders (Hofmann et al., 2012). CBT is a structured, goal-oriented therapeutic approach that is typically short-term and solution-focused. The underlying premise of CBT is that the way in which a person interprets events in their lives has a significant

impact on their wellbeing (Dobson & Dozois, 2021). The point is that it is not necessarily what happens to a person that has the greatest impact on their life, but how that person interprets that event. The main thrust in CBT then is commonly to help a patient adjust their interpretation of events in a process called cognitive restructuring. This involves identifying and challenging unhelpful or maladaptive interpretations of situations, and finding more accurate, helpful, or healthy ways to interpret events.

CBT incorporates both cognitive and behavioural elements in its approach. While cognitive restructuring is a central aspect, CBT also focuses on improving positive coping behaviours in individuals. For instance, in mood disorders, patients often display behavioural symptoms such as isolation and withdrawal, while those with anxiety disorders tend to avoid anxiety-inducing situations. CBT employs strategies to challenge and test negative beliefs and emphasizes the interconnectedness of situations, behaviours, and cognitions. Behavioural interventions typically employ exposure and desensitization techniques, particularly in addressing anxiety disorders. Behavioural experiments and homework assignments are common practice, the goal of which is for individuals to apply what they learn in therapy to real-life situations, which may also include the workplace (Dobson & Dozois, 2021).

The evidence base for CBT is solid, and it is often touted as the “gold-standard” for treating anxiety and depression. However, research over many years have shown that about 40 - 50 % of patients with mood and anxiety disorders recover in this treatment (Anderson & Maloney, 2001; Gyani et al., 2013). Although this recovery rate has formed the rationale for implementing CBT in large national programs such as Increasing access to psychological therapies (IAPT) in the UK (Clark, 2018), a recovery rate of 40 - 50 % still leaves room for improvement. One newer therapy that has aimed to improve on these results is metacognitive therapy.

1.2.2 Metacognitive Therapy

Metacognitive Therapy (MCT) was developed by professor Adrian Wells, and is grounded in the Self-Regulatory Executive Function (S-REF) (Wells & Matthews, 1996). The S-REF model provides a framework for understanding the brain as an information processing organ, and that negative bias in attention contributes to the development and persistence of mental disorders (Wells, 2009). The model states that biased metacognitions foster maladaptive attentional focus and the regulation of thought processes, including rumination, worry, and the use of ineffective coping strategies. Metacognitions are "thoughts about thinking" that are often implicit and unexamined, but that still govern the use of attention when processing information.

In MCT, a key goal is then to uncover and challenge the maladaptive metacognition that contribute to negative attentional bias. In depression, for example, a typical metacognition will take the form of "if I focus on why I am sad, I may discover the reason for my sadness, and this may help me improve". In anxiety, a typical metacognition may be "if I focus on things that can make me anxious, I may be better prepared in the future if something bad should happen". MCT states that rather than increasing patient coping through finding answers or improving preparedness, these metacognitions are more likely to just keep patients focused on distressing stimuli. This is called the cognitive attentional syndrome (CAS). MCT states that the CAS negatively skews attention towards negative thoughts and feelings, which perpetuates rumination and worry, which again are key mechanism in maintaining depression and anxiety (Wells, 2009).

Although there are differences between CBT and MCT, as the names imply, these treatments do have some factors in common. Both assert that the way in which patients process and interpret information is an important and determining factor in mental health. Both are concerned with changing negative bias in cognition, CBT does this by directly challenging the content of negative thoughts while MCT focuses on shifting attention through challenging metacognitions. Both are structured, even schematic, short-term treatment approaches, where concrete everyday situations form the basis for case formulations. Regarding treatment evidence, MCT is still a relatively new treatment approach, but initial results are promising. Norwegian research has shown recovery rates of 65% and 70% for mood

an anxiety disorders (Hagen et al., 2017; Hjemdal et al., 2019; Solem et al., 2019, 2021). In Denmark, one study compared MCT with CBT for depression, and showed that MCT was superior, achieving a 74% recovery rate compared with 52% in CBT (Callesen et al., 2020).

1.2.3 Developing treatment that also aids return to work.

CBT and MCT are both recommended treatments for depression and anxiety (NICE, 2020, 2022). Although the evidence indicates these methods do reduce symptoms and help people recover, the research also shows that effective treatment of CMD symptoms alone does not in itself help people return to work (RTW) (Cullen et al., 2018). Clinical treatment for mental health such as CBT and MCT has traditionally been preoccupied with alleviating clinical symptoms, and evaluations have pointed out that work is not an area of focus, nor a core competency for most clinicians, despite the importance of work for health and wellbeing (OECD, 2013). The growing recognition of this fact has prompted attempts over the last 15 – 20 years to develop interventions aimed at building on recommended symptom-treatment for CMD by integrating work-focused interventions to improve RTW outcomes (Schultz & Gatchel, 2015). This type of work-focused treatment is still in its infancy, and best practice approaches have yet to be established.

Much of the research so far on work-focused treatment has originated from the Netherlands, where occupational physicians and multidisciplinary approaches play a large role (Nigatu et al., 2016). One of the earliest studies was done in 2003 by van der Klink and colleagues. Employees with adjustment disorder received an intervention based on CBT principles aimed at helping them RTW after sick leave (van der Klink, 2003). The main thrust of the intervention was to help patients to improve their coping skills in the workplace, that is, transferring coping skills learned during therapy and to use this actively in their place of work. This was combined with a gradual RTW to help people adapt in their own tempo, not least of all because RTW after sick leave can be a demanding process. These features of transferring coping skills learned from therapy to situations in the work environment and facilitating a gradual RTW are features that are still central aspects of work-focused treatment (Schultz & Gatchel, 2015). The “activating intervention” in van Der Klink and colleague’s study was compared to a “care as usual” condition. The care as usual condition consisted of psychoeducation on

work and stress given by occupation therapists. Patients who received the intervention in the study returned to work earlier than patients in the control group, though their symptoms improved at a slightly slower rate (van der Klink, 2003). A subsequent Dutch study in 2007 found similar results for patients with major depressive disorder. Adding occupational therapy interventions to treatment as usual did not negatively impact treatment of depression but did enhance RTW rates. This study by Schene and colleagues also included a cost-effectiveness analysis that concluded that the intervention had a 75.5 % probability of being more cost-effective than treatment as usual (Schene et al., 2007).

Lagerveld, Blonk, and colleagues were likely the first to compare gold standard CBT with and without work-focused interventions for CMD in a study from 2012. The treatment was named work-focused-CBT, or W-CBT, and the intervention was compared to regular CBT in a sample of patients on sick leave due to CMD diagnoses. Symptoms of CMD decreased in both the intervention and control group, but the intervention group had earlier RTW. Counting reduced loss of working days due to earlier RTW, the findings also seemed to show that the intervention was cost-effective, and that costs associated with the W-CBT intervention were 20 % lower than the standard CBT condition due to the reduced number of sick days in the intervention group (Lagerveld et al., 2012). The W-CBT used in the study by Lagerveld et al. built on the earlier work by van der Klink et al. (van der Klink, 2003), integrating interventions aimed at enhancing coping at work with traditional CBT.

The approach that pairs evidence-based treatment for depression and anxiety and work-focused intervention has since been further adapted to a Norwegian context (Sandin, Gjengedal, et al., 2021). Building on the Dutch researchers' work, a work-focused module has been adapted for a Norwegian setting. Overall, the work-focused intervention still builds on the ideas of early studies like the 2003 study by van der Klink et al. In addition to working with patients to transfer coping skills learned in therapy to the working situation, the approach also includes a structured approach to assessing the individual working situation, psychoeducation on work and mental health, identification and Mapping of RTW Barriers, and, as in the earlier research, planning of gradual RTW (Sandin, Gjengedal, et al., 2021).

This approach has been previously evaluated in an observational study, where these work-focused interventions were combined with CBT. In a wait-list trial by Gjengedal et al., patients on sick had higher RTW rates and higher rates of recovery from depression and anxiety compared to the wait-list condition. Additionally, in the intervention group, 41.2 % of patients made a full return to work versus 26.3 % in the wait-list condition (Gjengedal et al., 2020).

Other studies pairing work-focused interventions with CBT have had similar results for CMD and sick leave in a Norwegian context, while also finding partial support for a positive cost-benefit ratio. Reme et al. conducted a randomised controlled trial evaluating the effectiveness of W-CBT and individual job support (IPS) compared to care as usual for people with CMD struggling with job participation. The study found that a larger proportion of patients in the intervention group managed to increase or maintain their work-participation compared to the control group. This effect was also present at 18-month follow-up. A cost-benefit analysis found that the economic cost of the intervention did not outweigh the benefits for the programme as a whole. The benefits of the programme would outweigh the costs only if it was provided solely to patients on long term benefits (Reme et al., 2015). This study by Reme and colleagues was, to the authors knowledge, the first study in Norway to examine the efficacy of work-focused CBT. Similar results were seen with W-CBT in Germany. Kröger and colleagues compared usual CBT with W-CBT for patients with MDD on measures of symptoms and RTW. The groups showed similar improvement on symptoms as indicated by last pre- to post effect sizes, but more patients RTW in the W-CBT group. This study also contained an analysis of sick leave, and found that the reduced number of sick days associated with W-CBT resulted in a net positive cost-effectiveness gain (Kröger et al., 2015).

1.2.4 Variation in approaches and results.

Overall, these findings point to adding work-focused interventions as a promising avenue of research with potential benefits for both patients and society. However, some studies fail to find that adding work-focused interventions improves RTW. One Dutch study by Hees et al. examined the effect of

occupational therapy on work-status for patients with MDD (Hees et al., 2010). The intervention was given by occupational therapists and built on a "Quality of work" model focused on the characteristics of the work, such as demands and autonomy. The intervention was a mixture of individual and group sessions, and did not contain CBT or other standardised therapies for symptoms of MDD (Hees et al., 2013). The occupational therapy approach was no better than care as usual for aiding recovery from depression or aiding RTW (Hees et al., 2010).

In Sweden in 2017, Salomonsson et al. compared traditional CBT, RTW interventions, and a condition combining the two. They found no difference between the three conditions on RTW. And similar to van der Klink, they found that symptom improvement was similar across the conditions, but that the work-focused interventions delayed improvement in mental health symptoms compared to CBT alone (Salomonsson et al., 2017). A study in Denmark found that a tailored intervention aimed at overcoming barriers for RTW actually delayed RTW compared to case management as usual for patients with mental health problems. Although the description of the RTW intervention indicates similarities with the other studies described, the intervention did not include what the authors describe as "formal psychotherapy", such as CBT This Danish study was similar to the study by Hees and colleagues in that it focused on patients with CMD without offering CBT or other specific treatment aimed at the disorder, focusing instead on work-related interventions. That neither of these studies demonstrated effect may indicate that work interventions alone have limited effect on CMD and return to work, and that Lagerveld et al. (Lagerveld et al., 2012) and similar studies may be more fruitful as they also include treatment aimed at the CMD. As the body of research on interventions aimed at helping people with CMD stay at work or RTW grows, reviews of the literature have appeared. Nigatu's (2016) review and meta-analysis of 16 randomised controlled trials concluded that the interventions did not improve the rates (i.e. number of people) that returned to work, but did reduce sick leave by reducing the number of days until RTW by a modest amount (13.38 days) (Nigatu et al., 2016). The review highlights the variation in interventions, with high focus on multidisciplinary approaches. It concludes that the reviewed studies do not support the efficacy of work-focused interventions for improving RTW in patients with CMD. A more recent review of

psychological treatments for patients on sick leave due to CMD examined 45 studies. It concluded that there is small effect of these interventions on reducing sick leave and symptoms (Hedges g = 0.15, and g = 0.21 respectively) compared to care as usual. There was no difference between CBT or multidisciplinary approaches. This review also highlights the wide variety of interventions, and also the lack of standardised measurements in the field (Salomonsson et al., 2018).

A recent Cochrane review evaluated the evidence on interventions to improve RTW in people with depression. It concluded, similarly to previous results, that the main effect was to be found in the reduction of sickness days, but that amount of people returning to work did not improve at one year follow-up. Work-focused interventions in combination with psychological treatment such as CBT was the approach with the highest likelihood of improving RTW, work ability, and symptoms. Crucially, it also concluded that there is need for further research on which combination of work-focused interventions and clinical treatment work best (Nieuwenhuijsen et al., 2020).

Overall, despite promising results in some studies, there is still considerable uncertainty attached to work-focused interventions for CMD. Although clinical treatment that also includes work-focus seems the most promising avenue, more research is needed before best practice approaches are established. A key challenge in integrating work-focused interventions in psychotherapy is the inherent complexity in sick leave, depression, anxiety, and in the relationship between them. Symptoms of CMD may be the justification for sick leave in the first place, but symptom improvement in itself does not necessarily lead to RTW (Reme, 2020).

The reasons for sick leave are manifold. One comprehensive study of register data in Norway looked at 1.78 million individuals who had a combined total of 3.8 million sick leave spells from 2001 – 2005. The study included almost 400 explanatory variables of the workplace, the employees, and the behavioural patterns of general practitioners who certified the sick leave. The study found support for associations that are well-known from the literature, e.g. that female gender, lower education, and lower socio-economic status all increase the rate of sick leave. However, despite a large dataset of potential explanations to draw from, the vast majority of variance in sick leave behaviour was explained by unobserved individual heterogeneity. It is worth noting that the study, as it used registry

data, did not include clinical factors such as symptom scores, or cognitive factors such as perceived barriers for RTW (Markussen et al., 2011).

Some review studies have looked more closely at the literature of prognostic factors for RTW after sick leave due to CMD. In a scoping review, de Vries et al. determined that higher symptom severity, co-morbidity, female gender, lower education level and low perceived general health were all predictors of sick leave in CMD (de Vries et al., 2018). A Finnish study by Ervasti and colleagues looking at prognostic factors for RTW after depression-related disability found similar results. They found strong support for older age, more severe depression symptoms, somatic comorbidity, and psychiatric comorbidity (i.e. one or more psychiatric diagnoses in addition to the depression diagnosis) delaying RTW. There was, however, significant heterogeneity between the studies included in the review (Ervasti et al., 2017). In addition to personal factors such as demographic characteristics and health issues, work-related factors such as high demand and low control have also been shown to influence RTW for CMD in reviews (Nigatu et al., 2017). Ultimately, the study by de Vries et al. found 53 predictive factors of sick leave and RTW across 21 studies (de Vries et al., 2018).

The challenge of devising effective work-focused treatment takes place against this backdrop of a multitude of factors influencing sick leave and RTW for CMD. For clinicians, effort must necessarily focus on factors that can be addressed in a clinical setting. Although the relationship is not straight forward, symptom severity of depression and anxiety is a prognostic factor for sick leave and RTW (de Vries et al., 2018; Ervasti et al., 2017; Nigatu et al., 2017). Providing gold standard treatment (e.g. CBT) for relieving depression and anxiety symptoms thus seems a reasonable starting point. Pairing this treatment with work-focused interventions to account for barriers for RTW that are not ordinarily addressed in this type of treatment also seems warranted. However, the vast variation in prognostic factors entails that treatment response will likely see large individual variation in any sample of patients on or at risk of sick leave due to CMD. A key challenge in developing effective work-focused treatment is thus to map the heterogeneity of patient characteristics and response in relation to such treatment. A better understanding of this variance could help stratify treatment to patient needs.

Beyond the heterogeneity of patients and their response to treatment, there is the societal perspective to consider. There is a long-standing treatment gap when it comes to mental illness, in particular for common mental disorders (Patel et al., 2010). Although mental ill health, and in particular CMD, account for a large proportion of the global burden of disease, treatment programs for these disorders are underfunded (Chisholm et al., 2016). Some studies have suggested that as much as 50 – 90 % of patients in need of mental health care are left untreated worldwide (Patel et al., 2010). In the UK in 2007, one study estimated that only 1 % of people needing specialised treatment for mental ill health received evidence-based treatment (Layard & Clark, 2015). In Norway, anxiety and depression have been shown to be the largest contributors to permanent disability (Knudsen et al., 2012). Despite this fact, studies have found that as much as a third of people awarded disability pension for these diagnoses in Norway have not previously received specialised treatment for these disorders (Overland et al., 2007). The acknowledgement of the widespread nature of these disorders in combination with limited treatment capacity has been behind initiatives in multiple countries to "scale up mental health" (Layard, 2006). The rationale from a health-economic perspective is that scaling up mental health would pay for itself given that treatment is relatively cheap and effective (Layard & Clark, 2015). The most well-known the initiatives based on this rationale is probably "Increasing Access to Psychological Therapies" (IAPT) in the UK (Clark, 2018), although Norway has its own version modelled on IAPT called "Prompt Mental Health Care" (Knapstad et al., 2018).

A key argument for such scaling up of treatment, is that the treatment would basically pay for itself through reducing disability, increasing quality of life and overall health, and through reducing lost working days (Chisholm et al., 2016). There may be merit to this assertion but there are methodological difficulties to overcome. Although early studies by Schene et al., Lagerveld et al., and Reme et al. did include an evaluation of the health economic benefits of work-focused CBT, no standardised method for such evaluation exists. There is therefore a need to develop methods that allow for the measurement of the economic benefits of such treatment. If standardised measurement of health economic outcomes could be reliably used with this patient group, it would facilitate comparison between studies and even with other patient groups. This could potentially provide a

clinical marker of whether the health economic assertion that scaling up mental health treatment would pay for itself is true.

2. AIMS

Work-focused treatment may be promising for patients suffering from common mental disorders given the risk of sick leave that these disorders carry (OECD, 2015). The development of such treatment is relatively new, and so is the study of its efficacy. As with the advent of any new field of inquiry, there is substantial heterogeneity in the methods and measurements of work-focused treatment (Nigatu et al., 2016). The aim of the present PhD work was thus to explore this heterogeneity in Norwegian patients receiving CBT or metacognitive therapy (MCT) with work-focused interventions for depression and anxiety.

Broadly, two avenues were pursued. Firstly, we wanted to know what registry data on sick leave can tell us about patients receiving work-focused treatment. Analysing registry data, we also wanted to know whether clinical and demographic characteristics of patients were associated with different sick leave behaviour over time. Secondly, we wanted to know if the impact of depression, anxiety, and sick leave on these patients could be recorded using a generic measure of health.

The research questions addressed in this thesis are thus:

1. Can subgroups of patients be identified based on sick leave behaviour before, during, and after work-focused treatment (Paper I)?
2. Do subgroups of patients differ on demographic and clinical characteristics (Paper I)?
3. Does the EQ-5D, a generic measure of health used in health economic analyses, show indications of validity in measuring health status and sick leave in these patients (Paper II)?
4. Can the EQ-5D adequately reflect change in health status during work-focused treatment (Paper III)?

The findings from our first line of inquiry could enhance our understanding of the heterogeneity among these patients. It could give us a better understanding of the relationship between symptoms, work-focused treatment, and sick leave over time. It could also potentially help prepare the ground for more targeted interventions through stratifying treatment according to patient characteristics. Our

second line of inquiry could help increase our understanding of the impact of depression, anxiety, and sick leave on wider health status. If the EQ-5D showed indications of validity for these patients, it could also pave the way for health economic analyses of treatment for patients with common mental disorders, including those receiving work-focused interventions.

3. METHODS

3.1 Context

The work presented in this thesis is based on data collected at the Division of Mental Health and Substance Abusee at Diakonhjemmet Hospital, Oslo. The clinic is part of national specialised mental health care service and offers work-focused treatment for patients suffering primarily from depression and anxiety. The studies used for the PhD work is part of an ongoing research project called "The Norwegian studies of psychological treatment and work (NOR-WORK)". The three articles presented are based on two separate studies in the NOR-WORK project. The first study uses data gathered in an observational study that ran from 2013 – 2016. The second and third articles uses data from a subsequent observational study that ran from 2017 – 2020. Data were primarily gathered through self-report questionnaires, and through patient journals. Additional data on sick leave used in the first article was gathered from national registries.

As the thesis and the papers presented herein to a large extent focus on sick leave, return to work, and work-focused treatment, a word on the Norwegian context is in order. The Norwegian welfare system is relatively generous. Employees receive 100 % compensation for lost pay from day one of sick leave. The first 16 days of sick leave is covered by the employer, after which the state welfare system pays for the employee's sick leave compensation. The employee can in principle remain on sick leave for up to a year. If the employee is still unable to return to work, he or she is then transferred to a separate long-term benefit ("Arbeidsavklaringspenger" or "AAP"), corresponding to 66 % of the employee's original salary. It also bears mentioning that protection of workers is quite strong in Norway, and no employee can be terminated for causes related to sick leave for the first 12 months of sick leave.

As Norway has a relatively high level of sick leave (OECD, 2013), several initiatives have been launched over the years to attempt to address this issue. Most notably, the "Faster return" programme, which was launched in 2007. The programme aimed at incentivising health care providers to develop and provide work-focused treatment. Particular emphasis was given to the patient groups responsible

for the majority of sick leave: common mental disorders and musculoskeletal disorders. The clinic in which the research in the present thesis took place was initially funded by the "Faster return" programme.

3.2 Design and study population

The patients included in the studies underpinning the three papers of this PhD were all referred to work-focused treatment at Diakonhjemmet Hospital by their general practitioners. The inclusion criteria at the clinic were that patients were on or at risk of sick leave due to depression and anxiety, and that they completed treatment. This implies that patients were adults of working age (18 – 70 years). Patients were excluded if they were on long term disability benefits, if they were suffering from cluster A or B personality disorder, or if they were suffering from severe mental illness such as bipolar disorder, schizophrenia, or other psychotic disorders. Patients were also excluded if they were engaging in active substance abuse, or if they were deemed to be at high risk of suicide. Patients were diagnosed after an initial screening session with a clinical psychologist, during which the Mini International Diagnostic Interview (M.I.N.I) was used (Sheehan et al., 1998). Diagnoses was set according to ICD-10 guidelines (WHO, 1992).

Data for Paper I was gathered in a naturalistic observational study that ran at the clinic from 2013 – 2016. In all, 619 patients were included based on the abovementioned criteria. Data for Paper II and Paper III were gathered in a subsequent study that ran from 2017 – 2020. During this time period, 890 patients were included at baseline. This sample was used in the second paper, which was a cross-sectional study examining patient characteristics at baseline. The third paper used data from the same study as the second paper but included patients who had completed treatment during this period. This longitudinal study looked at the responsiveness to change in the EQ-5D compared to the BDI-II and the BAI over the course of treatment. All patients included in this sample thus had to score above clinical threshold on either the BDI-II (≥ 14) or BAI (≥ 16) at baseline. In all, 416 patients fulfilled the criteria.

Although the data for these three papers were gathered in two separate observational studies, these studies took place in close succession at the same clinic. The samples were therefore similar, as reflected by the characteristics of patients recorded at baseline (i.e., at start of treatment). Average age of patients in the three papers ranged from 36.8 – 37.9 years (Paper 1 = 37.9, Paper 2 = 36.8, Paper 3 = 37.6), proportion of females from 68.5% - 71.9 % (Paper 1 = 68.5, Paper 2 = 69.6, Paper 3 = 71.9), and proportion of patients with higher education ranged from 76.5% - 82.8% (Paper 1 = 76.5, Paper 2 = 79.5, Paper 3 = 82.8). Depression diagnoses were the most common diagnostic group in all three papers, accounting for roughly half the primary diagnoses given (46.1% - 51.9%). The second most common diagnostic group was anxiety disorders, proportion ranging from 36.1% - 38.7 % of diagnoses in the three papers. A smaller part of the sample had diagnoses related to depression or anxiety, such as F41.2 "Mixed anxiety and depression" or F43.2 "Adjustment disorder". Symptom severity as recorded by the BDI-II and the BAI was also similar across the samples. For depression, the mean BDI-II scores ranged from 24.7 - 26.4, indicating moderate symptom severity. For anxiety, the mean BAI scores ranged from 17.5 -18.7, indicating mild levels of anxiety. Finally, proportion of patients on sick leave at start of treatment was also similar in the three papers. Approximately half the patients were on some type of sick leave, ranging from 45.7% – 49.9%.

3.3 Intervention

The treatment at the clinic was provided by a team of primarily psychologists, but also psychiatrists and psychiatric nurses. The treatment did not diverge from the treatment normally offered at the clinic, and this was thus a naturalistic observational study. Patients in the first study (2013 – 2016) which underpins Paper I received either CBT or MCT paired with work-focused intervention. Patients included in the second study (2017 - 2020) underpinning Paper II and III received MCT paired with the same work-focused intervention.

3.3.1 Cognitive behavioural therapy

CBT, as outlined in the introduction, is usually a short-term, problem-oriented approach aimed at challenging negative interpretations of events and situations in a patient's life. The therapist

commonly uses a case formulation to help hypothesise about aspect that give rise to and maintain the patient's mental difficulties. This will usually include a mapping of the patient's current difficulties and symptoms, and how they relate to maladaptive interpretations of events, as well as maladaptive coping behaviours. The therapist will also be interested in negative automatic thoughts that may influence the maintenance of maladaptive coping, and also the patient's strengths and resources (Ruggiero et al., 2021). In the treatment in the present project, patients were encouraged to use examples from the workplace in case formulations, especially if they were on sick leave.

3.3.2 Metacognitive therapy

As with CBT, case formulation was used to conceptualise treatment in MCT in the present project. This included symptom mapping and understanding the patients triggers for maladaptive coping. Patients are then introduced to the MCT model so that therapist and patient together can gain a common understanding of the patients' main issues. By learning to identify and recognise triggers, patients can become aware of which situations and what behaviour leads them to rumination or worry. Through challenging the patients assumption about thinking and through MCT techniques like attention training, postponement of worry and rumination, ban threat monitoring, patients learn to reduce coping strategies that backfire maladaptive coping strategies, and replace these with new, adaptive plans for regulating actions (Wells, 2009). As with CBT, patients were encouraged to use work situations as basis for case formulations if appropriate.

3.3.3 Work-focused interventions

CBT or MCT was paired with work-focused interventions. The guiding principles of the work-focused interventions was a structured checklist designed to ensure the maintenance of a work-centric focus throughout the entire treatment. It was especially important to make sure that work and working life was addressed early on in treatment. The key components of checklist are as follows:

Work Situation Assessment: A comprehensive examination of the patient's working environment, job characteristics, working conditions, interpersonal relationships, and the impact of their symptoms on work functioning.

Psychoeducation on Work and Mental Health: This includes pros and cons of sick leave, relationship between work, activity, and mental health. This included an exploration of the advantages and disadvantages associated with taking sick leave, including the risk of social isolation and withdrawal normally associated with depression and anxiety.

Additionally, the work-focused intervention included a plan for gradual RTW, identification and mapping of barriers for RTW, exploration of proper communication with the workplace regarding the patient's situation, and an exploration of the need for workplace adjustment. These guiding principles collectively served as the methodical framework to ensure that treatment maintained a work-focused approach. A description of the metacognitive therapy and work-focused interventions is also available in a protocol-article published for a forthcoming randomised controlled trial (Sandin, Gjengedal, et al., 2021).

3.4 Data sources

3.4.1 Sociodemographic characteristics

We included sociodemographic characteristics of patients in all three papers, both for descriptive and analytical purposes. Of particular interest were variables that previous research has shown to influence sick leave: age, gender, education, and marital status. Age in years was used as a continuous variable and gender was used as a dichotomous variable (male/female). Education was also used as a dichotomous variable denoting whether a participant had higher education or not. "Higher education" in this sense was operationalised as and degree completed beyond 12 years of school. Finally, marital status was a dichotomous variable (yes /no).

For Paper I, we also included data on whether patients had somatic comorbidity. Patients were asked "Do you have a somatic illness diagnosed by health personnel – Yes / No".

3.4.2 Sick leave

Registry data. Data on sick leave for the observational study used in the first article was collected from national registries. The registry features comprehensive data on sick leave as reported by general practitioners, who are responsible for certifying and reporting sick leave under the Norwegian system. The data spanned 29.5 months, or approximately two and a half years. The average duration of the treatment featured in the study was 5.5 months. We thus had registry data on sick leave from one year prior to one-year post treatment. The data is reported as sick leave dates throughout this period, giving us the start and end date of sick leave periods for each individual patient. We also had data on degree of sick leave for each period e.g., whether a patient was 100 % or 50 % off work. For the study, we operationalised sick leave as a dichotomous variable, patients who were fully working versus all other degrees of sick leave. We set measurement points at 90-day intervals, giving us 11 measurement points in all. The benefit of using registry data was that it gave us detailed and objective knowledge of sick leave behaviour over an extended period of time.

Self-reported sick leave was recorded from questionnaires for articles two and three, as we did not have registry data available for this study population. As with the registry data, self -reported sick leave was encoded as a dichotomous variable, "fully working" versus all other degrees of sick leave.

3.4.3 Clinical measures

We recorded primary diagnoses from patient journals, i.e. the diagnosis for which the patient received treatment. Secondary diagnosis was not recorded. The following self-report questionnaires were used to collect further clinical data:

Depression

The Beck Depression Inventory II (BDI-II). The BDI-II is a self-report questionnaire that measures depression over the last two weeks using 21 items that are answered on a scale from 0 (least severe) to 3 (most severe). The questionnaire thus has a sum range from 0 – 63, and a higher score indicates more severe symptoms. Example items include sadness, where 0 indicates "I do not feel sad" and 3 indicates "I am so sad and unhappy that I can't stand it", and suicidality, where 0 indicates "I don't have any thoughts of killing myself", and 3 "I would kill myself if I had the chance". Guidelines suggest that a BDI-II sum score of 0 – 13 denote minimal depressive symptoms, 14 – 19 indicate mild symptoms, 20 – 28 moderate, and 29 – 63 severe symptoms (Beck et al., 1996).

The BDI-II was used in all three papers, and mean score on the BDI-II indicate moderate depressive symptoms in the study population at baseline. The psychometric properties of the BDI-II are considered to be good for patients with depression, showing an Cronbach's alpha of 0.90 and a test-retest reliability ranging from 0.73 – 0.96 (Wang & Gorenstein, 2013). In the second paper presented in this thesis, the Omega (Peters, 2014) was calculated to be 0.86, indicating good reliability.

Anxiety

The Beck Anxiety Inventory (BAI). The BAI, like the BDI-II, is a 21 item self-report questionnaire that measures severity of anxiety on a sum scale from 0 – 63. Each item is scored on a Likert scale from 0 – 3, and patients are asked to indicate how severe their symptoms have been over the last month. For instance, "Unable to relax" 0 – 3, and "Nervous" 0 – 3, higher scores indicate more severe symptoms. Scoring guidelines suggest that 0 – 15 indicate minimal symptoms of anxiety, 16 – 25 moderate symptoms, and 26 – 63 severe symptoms (Beck & Steer, 1990).

The BAI was used in all three papers, and psychometric properties in patient populations with anxiety have in general been shown to be good, with and Cronbach's alpha of 0.92 and a test-retest reliability of 0.75 (Beck et al., 1988). In our second paper, we calculated the omega to be 0.90, indicating satisfactory reliability.

Subjective health complaints

The subjective health complaints questionnaire (SHC) is a self-report measure containing 29 items. It measures subjective complaints where there are no objective findings. It uses five dimensions: musculoskeletal pain, pseudo-neurology, gastrointestinal problems, allergy, and flu. Patients rate pains and complaints in various body parts on a scale from 0 (no pain or problem) to 3 (severe pain or complaints) during the last 30 days. Total score thus ranges from 0 – 87, where more severe complaints are reflected by higher scores. The aim of the scale is to provide a simple, theory-free measure of the most common complaints seen by general practitioners. The largest proportion of variance in the scale is explained by musculoskeletal complaints (Eriksen et al., 1999).

The SHC was used in Paper II, where we calculated its Omega to 0.82, indicating good reliability.

Self-reported health

Two separate measures of self-reported health were used in this PhD work. For the first article, we used a single-item measure of overall subjective health. Patients were asked to rate their health on a four-point scale: bad, not so good, good, very good. This is a question that is often used in larger health-surveys and has been shown to be good predictor of overall health status as well as sick leave (Schnittker & Bacak, 2014).

In Paper II and Paper III, the focus of interest was the EQ-5D. The EQ-5D is a generic measure of self-reported health (Herdman et al., 2011). The current version measures health on five subscales or dimensions: Mobility, Self-care, Usual activities, Pain / discomfort, and Anxiety / depression. Each of these five dimensions are scored on a scale from 1 – 5, where higher values indicate more severe problems. The level 1 corresponds to “No problems”, whereas level 5 corresponds to “Extreme problems”. Thus, the EQ-5D yields a five-digit profile where “11111” would indicate “No problems” on all dimensions, and “55555” would indicate “Extreme problems” on all dimensions (Devlin et al., 2020).

The EQ-5D profile can be converted, using preference-based weights, into the EQ-5D value. This EQ-5D value records perfect health as 1.000, and death is anchored at 0.000. Health states with negative values are possible, indicating a health state worse than death, but these are uncommon. For reference, a Norwegian survey of the general population found an EQ-5D value of 0.805 (Garratt et al., 2022). In Paper II, we found that the mean EQ-5D value at baseline was 0.631. The EQ-5D value can be used to calculate QALYS, which can be used in cost-benefit analyses. Please also note that it is common to report the EQ-5D utility using three decimal spaces (Devlin et al., 2020).

The EQ-5D also contains a visual analogue scale where respondents are asked to rate their health from 0 – 100. The EQ-5D was used in Paper II and Paper III. It is not recommended practice to calculate the Omega for the EQ-5D as it is not intended to measure a single construct (Konerding, 2013).

RTW-SE

The return-to-work self-efficacy scale (RTW-SE) is an 11-item questionnaire. Respondents reply to questions such as "I will be able to perform my tasks at work" and "I will be able to concentrate at work" on a six-point scale from "Totally agree" to "Totally disagree". The RTW-SE has been shown to predict full return to work after work-focused treatment in a Norwegian sample (Gjengedal et al., 2021). The RTW-SE was used in Paper I, where we calculated the Omega to 0.86, indicating satisfactory reliability.

3.5 Statistics

Missing data in all three papers was generally low. In the first paper, we used registry data for sick leave, ensuring no loss to follow-up, and missing data was thus not an issue. The clinical and demographic data were recorded from self-report questionnaires and patient journals. For all three papers, missing data was consistently less than 5 % on all items. For missing items, missing data were

replaced using weighted means. This is a method developed for handling missing data sets from patients with depression and has shown good precision with this patient group (Gale & Hawley, 2001).

3.5.1 Statistics Paper I

In Paper I we used latent growth mixture modelling (LGMM) to analyse the sick leave registry data. This analysis models unobserved heterogeneity in a sample of repeated measures by identifying patients with similar values across the measures. In our case, this translated to trajectories of sick leave across our 11 measurement points. Fit statistics guided model selection, in particular the Bayesian Information Criterion (BIC), sample-size adjusted BIC (ABIC), and Akaike information criterion indices, Entropy values and the bootstrap likelihood ratio test (BLRT). The model yields a probabilistic class assignment for patients based on sick leave trajectories only.

Once the best fitting model had been selected, we wanted to analyse differences in demographic and clinical characteristics across the classes, or groups. This post-hoc analysis of auxiliary variables is a viable approach where LGMM entropy is ≥ 0.8 (BMJ 34). The original class assignment in the LGMM is probabilistic, while in the post-hoc analyses, the groups are treated as categorical. This may lead to an underestimation of standard errors. It is therefore recommended to choose a more stringent significance level for these post-hoc analyses, and we therefore chose a significance level of $p < 0.01$.

Demographic characteristics were analysed as predictors of class membership using multinomial logistic regression. Clinical characteristics as predictors of class membership were analysed using a one-way analysis of variance (ANOVA) and Tukey's honestly significant difference test.

3.5.2 Statistics Paper II

In the second paper, we presented descriptive statistics showing the EQ-5D scores as per recommendations (Devlin et al., 2020). Descriptive statistics were calculated and shown in relation to clinical and demographic variables, and in relation to the norm scores obtained from a study of the general Norwegian population (Garratt et al., 2022). We then divided the patients in the sample into quartiles based on severity of depression and anxiety symptoms and tested the EQ-5D utility median using Cuzicks test for trends (Cuzick, 1985). The same severity quartiles were then used as the dependant variable in an ordinal logistic regression where the EQ-5D dimensions were the independent variables. This was done to analyse whether the health status reported by the EQ-5D utility and the EQ-5D dimensions deteriorated with increasing depression and anxiety symptoms.

We then performed a Pearson's correlation, correlating the EQ-5D value, the EQ-VAS and the EQ dimensions with the BDI-II, the BAI, and the SHC. Finally, a multiple linear regression model using the EQ-5D value as the dependant variable was run to examine whether and which of the clinical and demographic variables in the study predicted overall health as recorded by the EQ-5D. In the regression model, we wanted to see the unique contribution to variance of each factor, and partial correlation was thus calculated for each variable.

3.5.3 Statistics Paper III

In the third paper, we examined the responsiveness of the EQ-5D to change in health status during work-focused treatment by comparing its performance to the BDI-II and the BAI. We calculated effect size and standardized response mean from start to end of treatment. We also calculated the change scores (Δ) by subtracting the baseline scores from the scores obtained at end of treatment for the BDI-II, the BAI, the EQ-5D value, and the EQ-VAS. We then calculated the correlation of change scores using Spearman's correlation.

According to the clinical guidelines for the BDI-II and the BAI, we classified patients as either "Recovered", "Improved", or "Unchanged" from pre- to post-treatment. We then ran receiver

operating characteristics (ROC) analyses to determine if the EQ-5D value at end of treatment could correctly identify patient outcomes. Finally, we calculated Youden's index (J) to determine optimal cut-off point for identifying "Recovered" patients based on combined highest sensitivity and specificity.

3.6 Ethical considerations

The research underpinning the three articles presented in this thesis was classified as "health service research" according to Norwegian regulations. This entails that the Norwegian Data Protection Agency has mandated that treatment providers are responsible for the proper handling of data. There is, under the Norwegian ethical guidelines, no requirement to submit health service research to approval by an independent ethical committee. The reasoning behind this is that the intervention in such research does not deviate from the ordinary intervention at the clinic. Thus, the participants in the study are not subjected to any treatment or intervention other than what normally follows from being a patient at the clinic in question. Ethical guidelines for the treatment are thus provided by the rules and regulations that apply to any health care services in Norway. The practice of treating "health service research" in this way is not unique to Norway, but the practice varies from country to country. As science is a collaborative, and often international endeavour, this variation in ethical regulation between countries may make it more difficult to properly assess whether ethical guidelines have been followed. It is also the case that the division between "health service research" and "health research" is not necessarily self-evident or without overlap, which may potentially further complicate the matter. A learning point from the present study and a hope for the future, may thus be that we may arrive at more unified standards for ethical assessment of health and health service research. This would make it easier to assess the ethical aspects of research projects.

For the present project, the research presented in this thesis followed the principles of the Helsinki declaration (World Medical Association, 2022). Privacy and confidentiality ensured by the Data Protection officer at Diakonhjemmet hospital. All data collection and security were managed by

Diakonhjemmet Hospital, after approval of data handling was granted by Oslo University Hospital, approval number 2015/15606. To ensure that participation was voluntary, all patients provided written consent before being enrolled in the studies. To ensure that participants had the necessary information about the purpose of the study, the consent form stated clearly what data was collected, and what the data was used for. The consent form also stated that the treatment at the clinic was not contingent on or affected by agreeing to participate in the research. Finally, the consent form stated that consent to participate in the research could be withdrawn by the patient at any time without having to give any reason for withdrawing. An English version of the consent form is provided in the Appendix section. To ensure user input and that vulnerable groups were given a voice in the conceptualisation of the research, user representatives were involved in planning of segments of the research, such as giving feedback design and in on questionnaires used for data collection. A user representative from the user organisation Mental Helse ("Mental Health") was present at research meetings for this purpose.

It is also worth considering the potential upsides of the study when making the overall ethical assessment. As outlined in the introduction, depression, anxiety, and sick leave are all highly prevalent phenomena that impact a wide range of the population. The cost for both for individuals and society is substantial. Best practice interventions to alleviate the suffering captured in cost estimates is not yet established. There is thus an ethical case to be made for the necessity of conducting studies that evaluate the different aspects of services aimed at common mental disorders and sick leave. This necessity must be weighed against the potential burden placed on the individual participant in the study.

Finally, it is crucial that once data has been gathered from patients, this data is put to meaningful use. To ensure that the potential benefits of the project's insights reach the patients, researchers must actively disseminate their findings. This sharing of new knowledge holds the promise of informing further research and hopefully enhancing healthcare services. The present project is an attempt to contribute to this process.

4. RESULTS

4.1 Results from Paper I

Sick leave and return to work for patients with anxiety and depression: A longitudinal study of trajectories before, during, and after work-focused treatment.

In the first paper, we examined the characteristics of patients (N = 619) on or at risk of sick leave due to depression and anxiety who received work-focused treatment. We used registry data on sick leave collected before, during, and after treatment. The observational period before treatment was 12 months, average duration of treatment was 5.5 months, and the observational period after treatment was again 12 months. In all, we thus had registry data on sick leave for these patients spanning 29.5 months, or roughly two and a half years. The registry data gave us an objective measure of sick leave, before, during, and after work-focused treatment. In addition, clinical data and patient characteristics were collected from patient journals and self-report questionnaires at start and end of treatment.

The first objective of the study was to determine if subgroups of patients could be identified based on their sick leave behaviour over the observational period. We used latent-growth mixture modelling, identifying three different subgroups. These three groups showed distinctly different sick leave trajectories during the observational period. The largest group of patients consisting of 47.7% (n = 295) of the total sample had low sick leave throughout the entire period. We labelled this section of the sample as the "Resilient" group. The second group of patients consisted of 31.8 % (n = 197) of the patients in the sample. This group had low sick leave initially, increasing to high sick leave at start of treatment, before the majority of patient in this group returned to work within three months of completing treatment. We labelled this the "Recovery" group. Finally, the third group comprised 20.5 % of the patients (n = 127). These patients had high sick leave throughout the 29.5 months, and we labelled this the "High risk" group.

The second objective of the study was to determine if these subgroups differed on non-modifiable demographic characteristics, and whether they differed on modifiable clinical characteristics. Multinomial logistic regression showed that the relative risk of belonging to the "High risk" group

versus the "Resilient" group increased with higher age (RR = 0.95, $p < 0.001$), female gender (RR = 0.36, $p < 0.001$), not having higher education (RR = 2.13, $p < 0.01$), and having a previously diagnosed somatic illness (RR = 0.50, $p < 0.01$). The three groups showed similar clinical characteristics at the start of treatment, indicating that it would not have been possible to screen for "High risk" patients at the onset of treatment. All groups saw clinical improvement during the treatment period, but the "High risk" had consistently lower effect sizes on all clinical measures. At end of treatment, the "High risk" group also had residual depressive symptoms, with a mean score above clinical threshold.

4.2 Results from Paper II.

Self-reported health in patients on or at risk of sick leave due to depression and anxiety: Validity of the EQ-5D.

In the second paper, we investigated the validity of the EQ-5D, a commonly used measure of self-reported health, in patients on or at risk of sick leave due to depression and anxiety ($N = 890$). In addition to capturing a wider health status than what is recorded by clinical measures of mental health, the EQ-5D can also be used for health economic analyses. This could potentially help inform decision-making on work-focused treatment.

We investigated known-groups and convergent validity. We hypothesised that known-groups validity would be supported if the health status captured by the EQ-5D showed reduced health status compared to general population norms. Validity would be further supported if the EQ-5D could distinguish between different levels of severity of depression and anxiety. Convergent validity would be supported if the EQ-5D showed at least moderate negative correlation with symptom-specific measures of depression and anxiety. Furthermore, we hypothesised that the overall health status captured by the EQ-5D should be explained by variance in severity of depression and anxiety symptoms, and functional impairment operationalised as sick leave.

The EQ-5D value, capturing overall health, yields a health status where "perfect health" is denoted as 1.000, while "dead" is anchored at 0.000. In the study, the mean EQ-5D value for the patients was 0.631, indicating that these patients had reduced health status compared to "perfect health". Furthermore, a Norwegian general population survey yielded an EQ-5D value 0.820 [13], and the patients in the study thus reported worse health than the general population. The same was seen in the EQ-VAS, the visual-analogue scale of the EQ-5D where patients rate their perceived health from worst to best on a scale from 0 – 100. The Norwegian population survey found an EQ-VAS of 79.4, compared to an EQ-VAS 55.7 in our study. The patients in our study also reported more problems on all individual dimensions of the EQ-5D compared to the general population: Mobility, Self-care, Usual activities, Pain/Discomfort, Anxiety/Depression. The patients in the study thus reported worse health on all parts of the EQ-5D compared to the general population norms. The difference between the patient cohort and the population norms were most pronounced on the Anxiety/depression dimension where 96.6 % of patients reported problems, versus 35.4 % of the norm population reported problems on this dimension.

For convergent validity, the EQ-5D value showed a strong negative correlation with the BDI-II (r = -0.52, $p \leq 0.001$) and a moderate negative correlation with the BAI (r = -0.49, $p \leq 0.001$) and the SHC (r = -0.44, $p \leq 0.001$), indicating that higher symptoms were associated with lower self-reported health. SHC, measuring subjective health complaints with a particular emphasis on musculoskeletal pain [14], showed a strong correlation (r = - 0.50, $p \leq 0.001$) with the Pain/discomfort dimension of the EQ-5D. The BDI-II showed moderate correlation with the Usual activities (r = 0.45, $p \leq 0.001$) and Anxiety/depression dimensions (r = 0.46, $p \leq 0.001$), while the BAI showed a moderate correlation (r = 0.42, $p \leq 0.001$) with the Anxiety/depression dimension. Note again that higher EQ-5D value denotes better health, whereas higher EQ-5D dimension replies denote worse health.

Ordinal logistic regression further showed that symptom severity of depression and anxiety was predicted by more severe problems reported on the EQ-5D dimensions. Largest odds ratios were seen for Anxiety/depression (OR = 2.56, $p < 0.001$) and Pain/discomfort (OR = 1.77, $p < 0.001$). Our final regression model showed that the EQ-5D value was significantly predicted by depression and anxiety

symptom severity, being on sick leave, and female gender, predicting 40.1% of the explained variance ($F(8, 876) = 65.24$, $p < 0.000$, $R^2 = 0.40$.). The largest proportion of variance in the model was explained by depression (BDI-II $r = 0.38$, $p < 0.001$) and anxiety (BAI $r = 0.28$, $p < 0.001$), respectively.

4.3 Results from Paper III.

Responsiveness to change in health status of the EQ-5D in patients treated for depression and anxiety.

In Paper III, our aim was to explore the degree to which the self-reported health recorded by the EQ-5D was responsive to change during work-focused treatment for patients with depression and anxiety. Data were collected at baseline and end of intervention for patients who had been referred for work-focused treatment ($N = 416$).

Responsiveness refers to an instruments ability to detect "clinically significant change" over time [15, 16]. Two criteria for what should constitute clinically significant change have been suggested in the literature. Firstly, that the magnitude of change recorded should be statistically reliable, and secondly, that the change recorded renders patients indistinguishable from the normal population on the characteristics of interest [17]. Thus, being able to detect "recovery" in patients is of particular importance for determining an instruments responsiveness. Using these criteria, it is evident that responsiveness will vary according to context and population, which makes it important to examine this property in multiple patient groups. Very few studies have examined the responsiveness of the EQ-5D to depression and anxiety. Determining the responsiveness of the EQ-5D to change in depression and anxiety is therefore imperative to determine its usefulness for these patients [18].

We tested two hypotheses. Our first hypothesis stated that the magnitude of change in the health status recorded by the EQ-5D from start to end of treatment would be similar to changes in symptoms severity recorded by disease-specific measures. Our disease-specific measures with proven validity in the patient group was the BDI-II for depression, the BAI for anxiety. Our second hypothesis stated

that the EQ-5D would correctly identify patients who had recovered according to the established criteria used for the disease-specific measures at the end of treatment.

For the first hypothesis, we calculated the Cohen's *d*, standardised response mean (SRM), and Pearson's correlation for the change scores of the EQ-5D, the Becks Depression Inventory II (BDI-II), and the Beck Anxiety Inventory (BAI). For the second hypothesis, patients were classified as either "Recovered", "Improved", or "Unchanged" according to the BDI-II and the BAI. We then performed ROC analyses to determine if the EQ-5D could correctly classify patient outcomes.

For the first hypothesis, we saw a large magnitude of change across the EQ-5D value (d = 1.07), EQ-VAS (d = 1.25), the BDI-II (d = 1.84), and the BAI (d = 1.39). This is consistent with our hypothesis that we should see similar magnitude of change in these measures from start to end of treatment. We saw moderate correlation between the change scores of the EQ-5D value and the BDI-II (r = -0.48), and the BAI (r = -0.41). The EQ-VAS change score showed moderate correlation with the BDI-II (r = -0.56), but weak correlation the BAI (r = -0.32).

For the second hypothesis, the EQ-5D value could consistently identify "Recovered" patients for the total sample, for primary depression diagnoses, and primary anxiety diagnoses. The EQ-5D value could not distinguish between "Improved" and "Unchanged" patients.

Given that the EQ-5D value could consistently identify "recovered" patients, we also calculated Youden's index to determine which cut-off value provided the highest sensitivity and specificity in distinguishing "recovered" from "non-recovered" patients. For both depression and anxiety, best results (highest combined sensitivity and specificity) were seen using a cut-off value for the EQ-5D value of J = 0.768.

5. DISCUSSION

The aim of the current thesis was to explore the heterogeneity and self-reported health of patients on or at risk of sick leave due to common mental disorders receiving work-focused treatment. Two main topics were explored. Paper I dealt with sick leave and its relationship to demographic and clinical characteristics using registry data and self-report collected during the course of work-focused treatment. Paper II and Paper III dealt with whether the impact of depression, anxiety, and sick leave could be validly recorded using a generic measure of health intended for health economic evaluations.

5.1 Main findings from Paper I

Research question 1: Can subgroups of patients be identified based on sick leave behaviour before, during, and after work-focused treatment?

Our findings showed that there were three distinct subgroups with distinctly different trajectories over the two-and-a-half-year observation period. The first group comprised almost half the patients, 47.7 % (n = 295), and these patients had very low incidence of sick leave for the entire period. As they mostly avoided sick leave despite their symptoms of depression and anxiety, we named these patients the "Resilient" group. The second group made up almost a third of the patients, or 31.8 % (n = 197). These patients had low incidence of sick leave at the start of the observation period, which was one year before entering work-focused treatment. Their sick leave incidence then rose, and this group had the highest rates of sick leave at start of treatment. From start of treatment and until three months after finishing treatment, these patients' sick leave rates declined sharply, and at one-year follow-up, they had similar rates of sick leave as the "Resilient" group. This group of patients was named the "Recovery" group. The last group made up around a fifth of the patient sample, or 20.5 % (n = 127). These patients had fairly high rates of sick leave throughout the observational period. During the treatment period, they had similar rates of sick leave as the "Recovery" group, and although they did see a small decline after end of treatment, their sick leave rate remained high. At one-year follow-up,

their sick leave rate was 51.2 %, compared to 2.0 % for the "Recovery" group, and 3.4 % for the "Resilient group".

The answer to this research question was thus that there was unobserved heterogeneity in sick leave behaviour in the sample that the model classified as three distinct groups. These three groups differed in their sick leave trajectories before, during, and after work-focused treatment. Previous research has shown that there is great variability in determinants of sick leave and prognostic factors in RTW (de Vries et al., 2018; Nigatu et al., 2017). Uncovering unobserved heterogeneity in sick leave trajectories would seem to be in line with this research. It also implies that, in addition to heterogeneous determinants of sick leave, there may be corresponding heterogeneity in determinants treatment outcome.

Research question 2: Do subgroups of patients differ on demographic and clinical characteristics?

For this question, we saw patterns of differences between the groups on both demographic and clinical characteristics.

Firstly, looking at the two groups with high sick leave at the start of treatment ("High risk" and "Recovery"), these groups did not differ on demographic characteristics. Their mean age, gender distribution, levels of education, and rates of somatic comorbidity were not significantly different. Furthermore, at baseline, there was hardly any differences in clinical scores between these two groups. The difference in depression symptoms, anxiety symptoms, and RTW-SE were not significant. The only difference between these two groups at start of treatment on our measures was self-rated health. This implies that for a clinician in this study, it would not be possible to distinguish between members of the two groups at start of treatment. Thus, it would probably not be possible to predict individual sick leave trajectories at start of treatment using the measures included in the study. Research has shown that previous sick leave may be the best predictor of future sick leave (Hultin, Lindholm, Malfert, et al., 2012). In our research, sick leave at baseline would not be a good indicator for clinicians attempting to anticipate future sick leave. Most patients at sick leave at start of treatment

made a full RTW within three months of completing treatment (Sandin, Anyan, et al., 2021). One implication of this is that screening patients at start of treatment with aim of adapting treatment would not be feasible based on our included variables: the groups on sick leave at start of treatment were too similar.

Although they had similar clinical scores at baseline, the difference in clinical measures widened between the "High risk" and "Recovery" groups during the course of treatment. The "Recovery" group did not only have higher rates of RTW but had higher effect sizes for clinical change on all measures. At end of treatment, these two groups were significantly different on all clinical outcomes. In that sense, the "Recovery" group approached the scores of the "Resilient" group. This raises the question: What mediated the difference in change during treatment between the "High risk" and "Recovery" groups given their similar starting points? Demographic characteristics included in the study did not differ between the two groups at baseline. Despite previous research showing that higher age, female gender, comorbid somatic illness, and lower education increases risk of sick leave (de Vries et al., 2018; Nigatu et al., 2017), these variables did not differentiate between full RTW and continued sick leave in our study (Sandin, Anyan, et al., 2021). Which factors that mediate outcome of work-focused treatment for patients already on sick leave thus remains a pertinent question for future research.

We thus saw substantial differences between the "High risk" and "Recovery" group despite their similarities at baseline. There are several factors that may help explain these findings. It could be that characteristics of the "High risk" risk group that were not recorded in the study impeded progress during treatment for these patients. Conversely, there could be similar characteristics associated with the "Recovery" group that helped improve their treatment outcomes. This could be aspects associated with the working situation which tend to influence sick leave, such as degree of autonomy (Boštjančič & Galič, 2020), or harassment and bullying (Aarestad et al., 2020). The research has shown that there are a great many factor that influence sick leave and RTW rates in depression and anxiety, one review found 78 such potential factors (de Vries et al., 2018). It is implausible that a study design can account for all of these factors, and a lack of factors associated with the workplace is a limitation in the papers

presented in the thesis. One thing that is worth noting, however, is that we saw larger magnitude of change with the "Recovery" group on all measures included in the study compared to the "High risk group". This does suggest that improvement in depression, anxiety, RTW-SE and return to work are related. This can be useful when designing future interventions for this group and does suggest that symptom improvement is a necessary component for successful RTW for these patients.

Further support for this notion comes from the fact that the "High risk" group with its continued sick leave was the only group with mean symptom scores above clinical threshold at end of treatment. Depression and anxiety symptom scores of the "Recovery" and "Resilient" groups declined to below clinical threshold at end of treatment. This implies that, in general, the patients in these groups recovered from depression and anxiety. For the "High risk" group, depression symptoms remained above clinical threshold at end of treatment. This indicates that residual depressive symptoms may be a predictor of continued sick leave or relapse after work-focused treatment. This implies that even though it may not be possible to screen patients and adapt treatment at the start of the intervention, it may be possible to do so at the end of treatment. The implication in our study is that patients on sick leave at end of treatment who also have residual depressive symptoms have a low chance of making a full RTW.

The "Recovery" and the "'Resilient" groups were not significantly different on depression, anxiety, or self-rated health at end of treatment. Only RTW-SE remained different between the two at end of treatment. It would thus seem that RTW-SE provides insight on these patient's ability to RTW, in line with previous research (Lokman et al., 2017). In fact, research conducted at the same clinic as the on this thesis originates from has found that likelihood of full RTW increases as the RTW-SE score goes up (Gjengedal et al., 2021). This was also confirmed by the RTW-SE scores of our trajectory groups. The direction of any causality in this relationship remains unresolved. One interpretation would be to assume that specifically targeting RTW-SE in future interventions could help patients RTW. Or, it could be that an increase in RTW-SE is a by-product of effective treatment that successfully targets depression and anxiety symptoms and transfers coping skills to the work situation.

Finally, the majority of patients in the "Resilient" group were not on sick leave throughout the 28.5 months, despite similar clinical severity as the two other groups. It would be of interest to further research to uncover which factors that contributed to this group's resilience to sick leave. Given their similar symptom score to both the other groups at baseline, the implication is that it is possible to stay at work despite relatively high symptom levels. This raises the question of whether there are characteristics of family or working life that contributed to these patients avoiding sick leave. Previous research has shown that this may be the case. The balance between the demands of the work and employee autonomy (Haveraaen et al., 2017), as well as support from supervisors and colleagues are important in this respect. We did not include workplace factors in our study and can therefore not adequately answer if this was a mediating factor in our sample (Cancelliere et al., 2016). We did record education level which could be interpreted as a reasonable proxy for working conditions, and there were significantly more patients with higher education in the "Resilient" group compared to the "High risk" group. In general, it is worth considering that the "High risk" and "Recovery" groups were similar on demographic characteristics, but that these two groups differed from the "Resilient" group. This indicates that female gender, not having higher education, and somatic comorbidity predicted sick leave in this study. But as the "Recovery" group mostly had successful RTW, these demographic factors seem not to have been predictive of the outcome of work-focused treatment. This could lend further support to the interpretation that the impact of the intervention superseded the impact of demographic factors for these patients. It may be the case that symptom reduction and interventions aimed at RTW provided the tools that these patients needed to recover.

Following this line of reasoning, it could be that the "High risk" group experienced complicating factors that in sum superseded the effect of the treatment. Whether there are workplace factors at play, family life, or whether there are mediating factors of treatment outcome not included in our study remains a question for future research. A large, population-based study in Norway found that overall, the variance in sick leave predictors remain largely unexplained (Markussen et al., 2011). Of particular interest for future research targeting work-focused treatment would be uncovering mediating factors that are possible to target during the course of an intervention. Still, it is important

to note that most patients in the study sample did RTW within one year of the intervention. These patients were referred for work-focused treatment due to depression and anxiety, and half the patients (50.1 %) were on sick leave at the start of treatment. Within one year of completing the intervention, 87.2 % were fully working with no sick leave. This seems a reasonably good outcome, despite comparison to other studies of work-focused treatment being difficult due to lack of standardised methods for measuring RTW (Salomonsson et al., 2018). Future research will have to determine the impact of the combined psychological therapy and work-focused intervention on the RTW rates, and the sustainability of RTW beyond the 12-month follow-up period included in our study.

5.2 Main findings from Paper II

Research question 3: Can the EQ-5D, a generic measure of health used in health economic analyses, be a valid measurement of health status and sick leave in these patients (Paper II)?

For this paper, we saw that the EQ-5D did show indications of being a valid measure of health status in the patient sample. The patients were all referred and admitted to work-focused treatment for depression and / or anxiety and diagnosed by clinicians in line with national health service guidelines based on the ICD-10 (WHO, 1992). As such, we can be reasonably certain that these patients had a mental health problem, which should entail reduced health status. The patients in the sample reported problems on each of the five EQ-5D dimensions, and the Anxiety / depression dimension had the highest rates of patients reporting problems (96.6 %). The patients in the sample reported more problems on every dimension compared to the general Norwegian population norms (Garratt et al., 2022), and on the EQ-5D value, and the EQ-VAS. These findings indicate known-groups validity, i.e. that a clinical sample of patient diagnoses with a mental health problem reports worse health status compared to the wider non-clinical population.

The patterns of correlations examined showed that our symptoms scores of depression and anxiety had the highest correlation with the “Anxiety / depression” subscale. Furthermore, our measure of subjective health complains which has a particular emphasis on musculoskeletal pain, had the highest

correlation with the "Pain" dimension of the EQ-5D. Finally, linear regression analysis showed that our clinical variables significantly predicted the health status recorded by the EQ-5D value. The largest contribution to the variance in the model was made by the depression and symptom scores. Given the sample it was expected. Sick leave also made a significant contribution to the model. As the patients were receiving work-focused treatment for depression and anxiety, these findings seem to support the convergent validity between the EQ-5D and our clinical measures.

The findings in the study support the validity of the EQ-5D for patients on or at risk of sick leave due to depression and anxiety, indicating that it can be a useful measure of health-related quality of life for these patients. Though demonstrating the validity of an instrument in a patient population is important for assessing its usefulness, validity in a cross-sectional study does not indicate whether the instrument is responsive to change over time. This was the topic of our third paper.

5.3 Main findings from Paper III

Research question 4: Can the EQ-5D adequately reflect change in health status during work-focused treatment?

Being able to adequately reflect change in health status over time is of course of particular importance for evaluating interventions, i.e., whether patients improve, stay the same, or deteriorate during treatment. For the third paper, we investigated whether the EQ-5D could record changes in health status for patients during work-focused treatment.

The change of the EQ-5D scores from start to end of treatment were compared to the change scores of our symptoms scores for depression and anxiety. We saw that the effect size of the EQ-5D change from start to end of treatment was in a similar range as our symptom scores for depression and anxiety. Furthermore, that there was moderate correlation between the change scores from start to end of treatment. This indicates that the EQ-5D had similar magnitude of change in recorded health status during treatment as out "gold standard" symptom measures.

We then classified patients as either “Recovered”, “Improved”, or “Unchanged” according to the symptoms scores. The ROC analyses showed that the EQ-5D value could correctly classify “Recovered” patients, distinguishing these patients from the “Improved” and “Unchanged” patients. These findings show that the EQ-5D did show indications of responsiveness to change during work-focused treatment. Furthermore, that the magnitude of change was on par with the depression and anxiety measures. That the EQ-5D could also correctly identify recovered patients suggests that the instrument may be sensitive to the difference between clinical and non-clinical populations. Lastly, we calculated a cut-off scoring for making this distinction between the clinical and sociodemographic scores. Taken together, the results from Paper II and Paper II indicate that the EQ-5D can be a useful instrument for measuring health in patients on or at risk of sick leave due to depression and anxiety. This supports the feasibility of using the EQ-5D as a basis for health-economic analyses when evaluating work-focused treatment for these patients.

5.4 Strengths and limitations of the work presented in this thesis.

Sampling, inclusion and exclusion, and data used.

Inclusion and exclusion criteria in these studies were dictated by necessity as they followed the criteria for treatment at the clinic where the research took place. There are a number of limitations that follow. As with any clinical research, some patient groups of interest were excluded. Patients suffering from serious mental illness such as psychotic disorders, or substance abuse disorders, were not included in any of the samples, which is regrettable as these disorders are consistently linked to lower work-force participation, including sick leave (Markussen et al., 2011).

It is also important to be aware that the patients included in the samples for paper I and III were “completers” in the sense that these were patients who completed the treatment at the clinic. This selection was done to explore the study data both pre- and post-treatment for those that completed the treatment. This means that patients that dropped out of treatment were not included in the analyses. This is a limitation because it is reasonable to assume that patients who drop out may exhibit other

characteristics than people who choose to complete treatment. For example, any treatment effect may potentially over-estimate changes by including only those that chose to complete treatment, as patients who drop out may because they responded poorly to, or even deteriorated due to the treatment. Although the present thesis does not focus on treatment effect, the inclusion of only completers should still be factored in when interpreting the present findings. Patients not included may possibly also be those that returned to work of their own accord due to factors contributing to sick leave resolving themselves.

As previously mentioned, the studies underlying the thesis were naturalistic observational studies. This upside to this approach is that the research takes place in an ordinary clinic belonging to national health care services, and thus provides a good snapshot of this particular sample at this clinic. But one should be careful to generalise our findings to other settings, particularly those that serve patient groups with other diagnostic profiles. It also a key point that the lack of randomisation and control in this design precludes drawing any conclusions about causality in treatment effect.

The risk of committing type I or type II errors can arise from inappropriately specified statistical models, but also from improper sampling. Type I, or false positives, identify a significant relationship where none exists, while type II errors, or false negative, miss a significant relationship that is actually present. In the present thesis, for example, completer analysis can be seen to increase the risk of a type I error as patients who drop out are more likely to have poorer response to treatment (Dettori, 2011). Although symptom change was not an outcome measure in the present thesis, symptom measures were still used in categorisation of patient groups, raising the question of whether the groups may have looked different with another sample. The counterargument would be that the sample, and groups were quite large, as were the effect sizes, making it less likely that the relationship between clinical characteristics and sick leave trajectories in Paper I would disappear entirely.

An important aspect in the study design is that the lack of control group means that we do not know what would have happened to these patients had they not received the work-focused treatment. We can therefore not say to what degree RTW or symptom improvement was due to spontaneous remission. Relatedly, type III errors occur when a study arrives at the right answer while having asked

the wrong question (Schwartz & Carpenter, 1999). To put it differently, diseases and social problems in research are always multifactorial. As touched upon, there are many reasons for sick leave, and probably as many reasons for returning to work. The relationship between treatment and sick leave found in a study is vulnerable to the influence of what was not measured. In our studies, lacking control groups, we cannot know whether patients returned to work due to treatment, or for a any number of other reasons. It is also the case that effect size is proportional to the influence of the other explanatory variables included in the analyses. If important factors for explaining a phenomenon are left out, the resultant effect size is likely to give an inflated impression of the relationship. Reverse causation can also be a source of type III error, it may for instance not always be straight forward whether reduction in symptoms leads to RTW, or whether the revers may be true. Although we should be mindful of these potential errors when interpreting the results, it should also be mentioned that the present thesis had several strengths. Sample sizes were relatively large, the measures included are widespread and validated measures of symptoms for the primary diagnoses that these patients were receiving treatment for, and the method of intervention was previously researched at same clinic, where patients showed the overall same patterns of symptom distribution and sick leave behaviour (Gjengedal et al., 2020).

Measures and analyses used in the thesis.

The data used in these studies is quantitative in nature, meaning that patient experiences were recorded on numerical scales from self-report questionnaires, or recorded as demographic variables that were then coded numerically. For the measures used in the present thesis, the papers mainly used the BDI-II to measure depression, BAI to measure anxiety, EQ-5D to measure health-related quality of life, and the SHC to measure subjective health complaints. In addition, sick leave from registry data and demographic characteristics were used. All these measures are quantitative, and thus have the advantages and disadvantages of quantitative instruments.

The BDI-II and the BAI are widely used measures of depression and anxiety, potentially facilitating comparison with studies using the same or even similar measures. The EQ-5D is the most commonly used instrument for measuring health-related quality of life (Devlin et al., 2020), and there has been a push from Norwegian health authorities to expand the use of the instrument (Hansen et al., 2020). One important aspect of the instrument is that its scores can be converted to quality-adjusted life years, which may form the basis for health economic evaluations of burden of disease in a specific group (such as in Paper II), or the purported effect of an intervention (such as in Paper III). If the instrument over- or underestimates the burden of disease of impact of an intervention, this can lead to misallocation of scarce health resources.

This is one of the important findings in the present study, that the EQ-5D may be a fruitful avenue to further explore in future research. The present thesis adds to the limited knowledge of the psychometric properties of the five-level version of the EQ-5D in mental health, especially in a Norwegian setting. In the future, studies may use a value set specifically developed for the EQ-5D (Hansen et al., 2020), and it will then be interesting to see if studies corroborate the indications found in the present thesis that the EQ-5D can be capture health status in patients on or at risk of sick leave due to depression and anxiety. So, although the present findings point towards the EQ-5D being fit for purpose in common mental disorders in Norway, it is necessary to replicate the present findings using a Norwegian tariff in future studies when these are available.

Overall, there are several advantages to using quantitative measures, not least of all that it facilitates hypotheses testing using statistical analyses. A quantitative approach may facilitate replication, because one can use the same measures and run the same analyses. Although there is some effort involved from patients and staff in filling out and recording the scores of the questionnaires, the process is less time-consuming than for example conducting in-depth interviews, and facilitates inclusion of a larger sample size, potentially increasing the generalisability of the findings. On the note of filling in questionnaires, it is worth mentioning that overall, the amount of missing data was low in these studies (<5 %), which may indicate that the burden of filling in questionnaires was at

least tolerable for patients. Again, this conclusion must be tempered by the knowledge that analyses were only run on completers.

For the analyses, sick leave in Paper I was operationalised as data recorded from NAV registries. This had the advantage of providing objective records of sick leave, but we did lose nuance and information as the data were dichotomised to facilitate analyses. The sick leave data in Paper I was analysed using Latent Growth Mixture Modelling. This is a statistical technique for longitudinal data analysis. Its advantages are that it effectively models heterogeneity and is especially useful when studying populations with diverse growth trajectories, in this case, sick leave over time. In particular, the methods allow for the discovery of latent subgroups, which can reveal significant variability not captured by conventional growth models (Jung & Wickrama, 2008). The potential implications of identifying latent subgroups may help inform further intervention development and coming research projects.

For the post-hoc analyses in Paper I, we ran ANOVA and multinomial logistic regression. ANOVA, or Analysis of Variance, is a widely used method for measuring the differences in means between two or more groups. Multinomial logistic regression is useful in addressing multiclass classification scenarios, it provides probabilistic output, which gives estimations of probabilities associated with each dependent variable category (Fang, 2017).

ROC analysis is widely used to select appropriate clinically optimal cut-of scores by testing the ability of a scale to discriminate between groups. ROC analysis provides a comprehensive view of a classification model's performance across various thresholds. It plots the trade-off between sensitivity (true positive rate) and specificity (true negative rate) at different decision thresholds. ROC analysis is threshold-independent, meaning it evaluates the model's performance across all possible classification thresholds. This is especially useful when the optimal threshold is not known, such as in the examination of recovery from CMD and sick leave in Paper III. ROC analysis can help in understanding the trade-off between sensitivity and specificity, and depending on the application, one can choose an operating point on the ROC curve that balances these trade-offs according to the specific needs of the problem (Hajian-Tilaki, 2013).

Although these methods are useful, the challenge with both the ROC analyses and the LGMM in the current thesis is the reduction of complex phenomena to dichotomisation. These were the "sick leave groups" in the case of LGMM in Paper I, and the "Recovered", "Improved", and "Unchanged" groups in the ROC analyses in Paper III. As touched on previously, this is perhaps one of the key limitations of the present thesis. In trying to map some of the heterogeneity of sick leave, accommodations were made to facilitate statistical analyses, at the cost of reduced nuance.

Quantitative methods do have their advantages, but there are notable limitations to this approach. One such limitation is the potential lack of depth in capturing the intricacies of human experiences and behaviours when reducing human experiences to numerical scales. Quantitative methods give priority to measurable aspects and do risk overlooking more nuanced but equally important parts of psychological phenomena. Depression, for instance, is a heterogeneous disorder that is notoriously difficult to measure. People affected by depression experience a wide range of symptoms across cognitive, emotional, somatic, and behavioural spectrums, and there is research to suggest that much of what patients and next of kin experience as the most pressing concerns in depression is often not measured by commonly used questionnaires (Chevance et al., 2020).

Additionally, quantitative research may fall short in providing a comprehensive understanding of underlying reasons and context crucial in psychological phenomena. Given the complex relationship between sick leave and common mental disorders, this is a limitation of the work presented in the present thesis. The risk is that complex human phenomena are shoehorned into numerical values and statistical relationships, potentially resulting in the loss of valuable qualitative insights. Specifically, for the present thesis, some nuance was undoubtedly lost due to treating sick leave as a dichotomous measure in Paper I. This was done as a necessity as the data quality of the graded sick leave records were inconsistently coded from the data provider, the National Labour and Welfare Administration, and we could therefore not use graded sick leave as basis for our analyses. Another example comes from operationalising severity of depression and anxiety as sum scores from the BDI-II and BAI. Recent research from the same clinic has shown, that there may be some depression domains that are more associated with sick leave than others (Gjengedal et al., 2022). This casts doubt on whether sum

scores questionnaires for complex phenomena such as depression and anxiety can adequately represent patient experiences in a way that is helpful for designing interventions or making clinical decisions. For Paper I in the present thesis for example, residual depressive symptoms were associated with continued sick leave after treatment. Perhaps a better understanding could have been gained by examining which aspects of depression were the most salient for patients who were still on sick leave.

An alternative approach to the methodological choices made for the present thesis could have been to pursues a qualitative research design or augment the quantitative approach with qualitative tools in a mixed-methods approach. Qualitative methods, such as in-depth interviews, may provide more rich and in-depth data on human experiences, perceptions, and behaviours through open-ended questions. Moreover, qualitative methods give more room for participants to expand on their perspectives, adding depth and authenticity to data representation. The present project could for example have benefited from a more in-depth understanding of why patients were on sick leave, and why they chose to return to work when they did. The present project aimed to map some of the heterogeneity involved in the complex relationship between common mental disorders and sick leave. A mixed methods approach drawing on both quantitative and qualitative methods, could likely have provided further insights on the heterogeneity in the relationship between common mental disorders and sick leave.

5.5 Context and contribution of the thesis

The underlying question guiding much of the research on work-focused treatment, including the present thesis is: To which degree can sick leave from depression and anxiety be alleviated through work-focused treatment? Given the heterogeneity of depression, anxiety, and sick leave, it seems a reasonable assumption that treatment response, even if something like an optimal work-focused treatment were to be found, will not be uniform. This is further underlined by consistent findings that the recommended treatment for depression and anxiety does not itself lead to return to work (Cullen et al., 2018). At the clinical level, the idea driving the present thesis then, has been to begin to map if

there are different patient groups receiving work-focused treatment, and whether there are different sick leave behaviours associated with these patient groups. For future research, we may hope that a more differentiated approach could be possible so that treatment, or other interventions such as work-place adjustments, may be tailored to patients' individual needs.

But sick leave due to depression and anxiety is not just a question for clinical practice, nor just a challenge for individual health and wellbeing, but also one of the biggest challenges facing the modern welfare state. The cost associated with sick leave due to mental health issues is a key part of the rationale for nationwide programmes such IAPT in the UK (Clark, 2018), and Prompt Mental Health Care in Norway, which was explicitly modelled on IAPT (Knapstad et al., 2018). The economic argument underpinning these large-scale programmes justifies not only the scale of investment in the intervention, but also to a significant extent the *content* of the intervention. The literature on IAPT, for instance, makes the argument that one-size fits all CBT is an appropriate response due to its purported recovery rates of 50 %, and the main thrust of the argument is that this approach would be the most *cost-effective* (Clark et al., 2018).

But the findings so far from IAPT are perhaps not encouraging on this point. When looking at recovery rates, public data from IAPT shows that 52.1 % supposedly recover after treatment. However, this number is based on an analysis of patients who had at least five sessions. When the recovery rate is calculated from the original number of referrals, it looks very different, and approximately 22.3 % of original referrals are moved to the "recovered" category. This means that approximately only one in five referrals move on to recovery in IAPT, and this has been a consistent feature since at least 2015 (McInness, 2018). One reason for the low rate of recovery is the attrition rate: around a third of referrals do not enter therapy (defined as having two or more treatment sessions), and of the people that do enter therapy, around half drop out before completing treatment (the figure for 2018 was 45 %) (McInness, 2018). Attrition introduces uncertainty in any study of clinical outcomes. Some patients may drop out due to natural recovery or not fulfilling diagnostic criteria in the first place. However, research shows that patients who drop out have, on average, poorer prognoses than completers. Furthermore, a rule of thumb for dealing with attrition in research

indicates that rates < 5 % introduce little risk of bias, whereas rates > 20 % carry a substantial risk of bias (Dettori, 2011). Given the reported attrition rate of 45 %, it seems reasonable to question if IAPTS reported recovery rate of 51.4 % is a valid representation of the programme.

For Prompt Mental Health Care in Norway, reliable recovery rates from common mental disorders were 58.1 %, which yielded a moderate effect size of $d = 0.61$ compared to the treatment as usual group (Knapstad et al., 2020). Although this may be encouraging, the programme did not improve work participation, which is in line with the previous research: treatment for common mental disorders that does not include specific work-focused interventions does not improve return to work or increase work rates (Cullen et al., 2018). There is room for improvement in the treatment that is offered to patients on or at risk of sick leave due to CMD. Should the findings in Paper I be replicated, future directions for research and treatment may be to differentiate more what is offered to whom.

There is a wide range of reasons for sick leave and the diverse reasons for why patients return to work seen in the literature (de Vries et al., 2018; Nigatu et al., 2017). Although some of these factors seem like targets for work-focused treatment, such as symptom burden, previous history of CMD, and overall health, some do not. Workplace factors such as competency of managers (Stengård et al., 2021), low control over working situation while stressed with high demands, and psychosocial milieu at work are all factors that are known to contribute to sick leave (de Vries et al., 2018), but that are not obvious targets to address in a therapy room primarily providing treatment for CMD symptoms. This is perhaps what is seen with the "High risk" group in Paper I. This group continued to have high degrees of sick leave after work-focused treatment. Perhaps what is needed for future research is to discover whether work-focused treatment is the right approach for these patients, or whether there are other interventions that would be more useful. Interventions at the workplace aimed at making adjustments to working situations or improving management competence through training is one example of an alternative approach (Milligan-Saville et al., 2017). Another example is bullying and harassment, which is a well-known source of absenteeism. In cases of bullying, it is not obvious that it is ethically defensible to provide treatment as the only solution for someone who is the victim of bullying, without addressing workplace issues. The answer may be that work-focused treatment for

anxiety and depression can perhaps be helpful for some, but that it may not be the correct approach for all patients currently included in these programs(Aarestad et al., 2022).

Apart from the discussion on individual needs in interventions, arguments for providing large-scale programmes aimed at CMD and sick leave also rest on an economic foundation. In the case of IAPT, for instance, the argument is that increased work participation would mean that the program would pay for itself (Clark, 2018). The cost of sick leave, not just to individuals, but also to employers, insurance providers, and the welfare system is likely why research on work-focused treatment often has included either economic evaluations of treatment, or measures fit for health-economic analyses (Lagerveld et al., 2012; Reme et al., 2015; Schene et al., 2007). The evaluation of Prompt Mental Health Care continues this tradition by including the EQ-5D to measure health-related quality of life (Knapstad et al., 2020). This is also in line with the policy of Norwegian health authorities where the National Institute of Public Health is investing in developing a Norwegian value set for the EQ-5D, and the instrument is also included in nationwide quality registries for the health services (*Forsiden | Nasjonalt Servicemiljø for Medisinske Kvalitetsregistre*, n.d.; Hansen et al., 2020).

Given the popularity of the EQ-5D in recent years, it is worth pointing out that that Papers II and III in the present thesis are the first papers to examine the psychometric properties of the five-level EQ-5D in mental health in Norway. This is important because the methodology underlying the EQ-5D assumes a societies' willingness to pay for health services, and this willingness, or value placed on different health states will vary by country (Devlin et al., 2020). Although validity and psychometric properties are always subject to and will vary with context, papers II and III provide a starting point for evaluating the use of the EQ-5D on mental health in Norway. Given the widespread and increasing use of the EQ-5D establishing that it accurately represents the health status of the patient groups in which it is used is crucial.

6. CONCLUSION

This doctoral thesis has looked at work-focused treatment for common mental disorders from a health service research perspective. The underlying question in research on working-focused treatment is whether this is a fruitful avenue to pursue for patient on or at risk of sick leave due to common mental disorders, and under which conditions this treatment can be useful. This is still a new field of research, the first international study was published in 2003, the first Norwegian study in 2015. There are many unanswered questions, and much of the research is, like the present thesis, exploratory in nature.

Adding to the growing literature, the three papers presented in this thesis have looked at the overarching challenge in the inherent heterogeneity in depression, anxiety, and sick leave, and the complexity of the relationship between them. Our findings show that in this sample of patients, there were groups with distinctly different sick leave trajectories. This indicates that patient outcomes may be improved by either differentiating treatment according to the individual needs or by considering if alternative interventions could be more helpful. One patient group showed resilience to sick leave despite similar symptom severity as other patient groups on sick leave. This implies that resilience and protective factors may be an important area for future sick leave research.

Although Paper I was an observational study, it is also worth noting that effect for change in clinical measures during treatment were large across all groups, and that the vast majority made successful, sustainable RTW within three months of completing work-focused treatment. We can hope that future research may uncover the role of the work-focused treatment model on depression, anxiety, and RTW.

Our findings from Paper II and Paper III demonstrated that the EQ-5D may be useful in measuring decrements in health status for these patients and may also be sensitive to change in health status during the course of treatment. This gives hope that future development and evaluation of work-focused treatment for common mental disorders may be augmented by establishing good practices for health economic evaluation. The assertion that scaling up treatment by investing in nationwide

programs aimed at alleviating depression and anxiety will pay for itself is widespread in the literature. It would be helpful if future research can build on ours and similar research to put this assertion to the test in a clinical setting using properly validated measures.

REFERENCES

Aarestad, S. H., Einarsen, S. V., Hjemdal, O., Gjengedal, R. G. H., Osnes, K., Sandin, K., Hannisdal, M., Bjørndal, M. T., & Harris, A. (2020). Clinical Characteristics of Patients Seeking Treatment for Common Mental Disorders Presenting With Workplace Bullying Experiences. *Frontiers in Psychology*, *11*. https://www.frontiersin.org/articles/10.3389/fpsyg.2020.583324

Aarestad, S. H., Harris, A., Hjemdal, O., Gjengedal, R. G. H., Osnes, K., Sandin, K., Reme, S. E., Hannisdal, M., & Valvatne Einarsen, S. (2022). Healing the wounds of workplace bullying: Evaluating mental health and workplace participation among victims seeking treatment for common mental disorders. *Work*, *Preprint*(Preprint), 1–13. https://doi.org/10.3233/WOR-210920

Ahola, K., Virtanen, M., Honkonen, T., Isometsä, E., Aromaa, A., & Lönnqvist, J. (2011). Common mental disorders and subsequent work disability: A population-based Health 2000 Study. *Journal of Affective Disorders*, *134*(1–3), 365–372. https://doi.org/10.1016/j.jad.2011.05.028

Anderson, D. A., & Maloney, K. C. (2001). The efficacy of cognitive-behavioral therapy on the core symptoms of bulimia nervosa. *Clinical Psychology Review*, *21*(7), 971–988. https://doi.org/10.1016/s0272-7358(00)00076-3

Angermeyer, M. C., Holzinger, A., Matschinger, H., & Stengler-Wenzke, K. (2002). Depression and quality of life: Results of a follow-up study. *International Journal of Social Psychiatry*, *48*(3), 189–199.

Baxter, A. J., Vos, T., Scott, K. M., Ferrari, A. J., & Whiteford, H. A. (2014). The global burden of anxiety disorders in 2010. *Psychological Medicine*, *44*(11), 2363–2374. https://doi.org/10.1017/S0033291713003243

Beck, A. T., Brown, G., Epstein, N., & Steer, R. A. (1988). *An Inventory for Measuring Clinical Anxiety: Psychometric Properties*. 5.

Beck, A. T., & Steer, R. A. (1990). Manual for the Beck anxiety inventory. *San Antonio, TX: Psychological Corporation*.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the beck depression inventory-II*. San Antonio, TX: Psychological Corporation.

Blank, L., Peters, J., Pickvance, S., Wilford, J., & Macdonald, E. (2008). A systematic review of the factors which predict return to work for people suffering episodes of poor mental health. *Journal of Occupational Rehabilitation*, *18*(1), 27–34. https://doi.org/10.1007/s10926-008-9121-8

Boštjančič, E., & Galič, K. (2020). Returning to Work After Sick Leave – The Role of Work Demands and Resources, Self-Efficacy, and Social Support. *Frontiers in Psychology*, *11*. https://www.frontiersin.org/articles/10.3389/fpsyg.2020.00661

Callesen, P., Reeves, D., Heal, C., & Wells, A. (2020). Metacognitive Therapy versus Cognitive Behaviour Therapy in Adults with Major Depression: A Parallel Single-Blind Randomised Trial. *Scientific Reports*, *10*(1), 7878. https://doi.org/10.1038/s41598-020-64577-1

Cancelliere, C., Donovan, J., Stochkendahl, M. J., Biscardi, M., Ammendolia, C., Myburgh, C., & Cassidy, J. D. (2016). Factors affecting return to work after injury or illness: Best evidence synthesis of systematic reviews. *Chiropractic & Manual Therapies*, *24*(1), 32. https://doi.org/10.1186/s12998-016-0113-z

Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., & Saxena, S. (2016). Scaling-up treatment of depression and anxiety: A global return on investment analysis. *The Lancet Psychiatry*, *3*(5), 415–424. https://doi.org/10.1016/S2215-0366(16)30024-4

Clark, D. M. (2018). Realizing the Mass Public Benefit of Evidence-Based Psychological Therapies: The IAPT Program. *Annual Review of Clinical Psychology*, *14*(1), 159–183. https://doi.org/10.1146/annurev-clinpsy-050817-084833

Clark, D. M., Canvin, L., Green, J., Layard, R., Pilling, S., & Janecka, M. (2018). Transparency about the outcomes of mental health services (IAPT approach): An analysis of public data. *The Lancet*, *391*(10121), 679–686. https://doi.org/10.1016/S0140-6736(17)32133-5

Claussen, B. (1999). Alcohol disorders and re-employment in a 5-year follow-up of long-term unemployed. *Addiction*, *94*(1), 133–138. https://doi.org/10.1046/j.1360-0443.1999.94113310.x

Cullen, K. L., Irvin, E., Collie, A., Clay, F., Gensby, U., Jennings, P. A., Hogg-Johnson, S., Kristman, V., Laberge, M., McKenzie, D., Newnam, S., Palagyi, A., Ruseckaite, R., Sheppard, D. M., Shourie, S., Steenstra, I., Van Eerd, D., & Amick, B. C. (2018). Effectiveness of Workplace Interventions in Return-to-Work for Musculoskeletal, Pain-Related and Mental Health Conditions: An Update of the Evidence and Messages for Practitioners. *Journal of Occupational Rehabilitation*, *28*(1), 1–15. https://doi.org/10.1007/s10926-016-9690-x

Cuzick, J. (1985). A wilcoxon-type test for trend. *Statistics in Medicine*, *4*, 543–547. https://doi.org/10.1002/sim.4780040112

de Vries, H., Fishta, A., Weikert, B., Rodriguez Sanchez, A., & Wegewitz, U. (2018). Determinants of Sickness Absence and Return to Work Among Employees with Common Mental Disorders: A Scoping Review. *Journal of Occupational Rehabilitation*, *28*(3), 393–417. https://doi.org/10.1007/s10926-017-9730-1

Dettori, J. R. (2011). Loss to follow-up. *Evidence-Based Spine-Care Journal*, *2*(1), 7–10. https://doi.org/10.1055/s-0030-1267080

Devlin, N., Parkin, D., & Janssen, B. (2020). *Methods for Analysing and Reporting EQ-5D Data*. Springer International Publishing. https://doi.org/10.1007/978-3-030-47622-9

Dobson, K. S., & Dozois, D. J. A. (2021). *Handbook of Cognitive-Behavioral Therapies: Fourth Edition*. Guilford Press. https://www.guilford.com/books/Handbook-of-Cognitive-Behavioral-Therapies/Dobson-Dozois/9781462547722

Eriksen, H. R., Ihlebæk, C., & Ursin, H. (1999). A scoring system for subjective health complaints (SHC). *Scandinavian Journal of Public Health*, *27*(1), 63–72. https://doi.org/10.1177/14034948990270010401

Ervasti, J., Joensuu, M., Pentti, J., Oksanen, T., Ahola, K., Vahtera, J., Kivimäki, M., & Virtanen, M. (2017). Prognostic factors for return to work after depression-related work disability: A systematic review and meta-analysis. *Journal of Psychiatric Research*, *95*, 28–36. https://doi.org/10.1016/j.jpsychires.2017.07.024

Fang, J.-Q. (2017). *Handbook of Medical Statistics*. WORLD SCIENTIFIC. https://doi.org/10.1142/10259

Ferrari, A. J., Charlson, F. J., Norman, R. E., Patten, S. B., Freedman, G., Murray, C. J. L., Vos, T., & Whiteford, H. A. (2013). Burden of Depressive Disorders by Country, Sex, Age, and Year: Findings from the Global Burden of Disease Study 2010. *PLoS Medicine*, *10*(11), e1001547. https://doi.org/10.1371/journal.pmed.1001547

Forsiden | Nasjonalt servicemiljø for medisinske kvalitetsregistre. (n.d.). Retrieved 18 September 2023, from https://www.kvalitetsregistre.no/

Gale, T., & Hawley, C. (2001). A model for handling missing items on two depression rating scales. *International Clinical Psychopharmacology*, *16*(4), 205–214. https://doi.org/10.1097/00004850-200107000-00004

Garratt, A. M., Hansen, T. M., Augestad, L. A., Rand, K., & Stavem, K. (2022). Norwegian population norms for the EQ-5D-5L: Results from a general population survey. *Quality of Life Research*, *31*(2), 517–526. https://doi.org/10.1007/s11136-021-02938-7

Gilman, S. E., Sucha, E., Kingsbury, M., Horton, N. J., Murphy, J. M., & Colman, I. (2017). Depression and mortality in a longitudinal study: 1952–2011. *Canadian Medical Association Journal*, *189*(42), E1304–E1310. https://doi.org/10.1503/cmaj.170125

Gjengedal, R. G. H., Lagerveld, S. E., Reme, S. E., Osnes, K., Sandin, K., & Hjemdal, O. (2021). The Return-to-Work Self-efficacy Questionnaire (RTW-SE): A Validation Study of Predictive Abilities and Cut-off Values for Patients on Sick Leave Due to Anxiety or Depression. *Journal of Occupational Rehabilitation*, *31*(3), 664–673. https://doi.org/10.1007/s10926-021-09957-8

Gjengedal, R. G. H., Osnes, K., Reme, S. E., Lagerveld, S. E., Johnson, S. U., Lending, H. D., Sandin, K., Bjørndal, M. T., & Hjemdal, O. (2022). Changes in depression domains as predictors of return to work in common mental disorders. *Journal of Affective Disorders*, *308*, 520–527. https://doi.org/10.1016/j.jad.2022.04.080

Gjengedal, R. G. H., Reme, S. E., Osnes, K., Lagerfeld, S. E., Blonk, R. W. B., Sandin, K., Berge, T., & Hjemdal, O. (2020). Work-focused therapy for common mental disorders: A naturalistic study comparing an intervention group with a waitlist control group. *Work*, *66*(3), 657–667. https://doi.org/10.3233/WOR-203208

Gyani, A., Shafran, R., Layard, R., & Clark, D. M. (2013). Enhancing recovery rates: Lessons from year one of IAPT. *Behaviour Research and Therapy*, *51*(9), 597–606. https://doi.org/10.1016/j.brat.2013.06.004

Hagen, R., Hjemdal, O., Solem, S., Kennair, L. E. O., Nordahl, H. M., Fisher, P., & Wells, A. (2017). Metacognitive Therapy for Depression in Adults: A Waiting List Randomized Controlled Trial with Six Months Follow-Up. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00031

Hajian-Tilaki, K. (2013). Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian Journal of Internal Medicine*, *4*(2), 627–635.

Hansen, T. M., Helland, Y., Augestad, L. A., Rand, K., Stavem, K., & Garratt, A. (2020). Elicitation of Norwegian EQ-5D-5L values for hypothetical and experience-based health states based on the EuroQol Valuation Technology (EQ-VT) protocol. *BMJ Open*, *10*(6), e034683. https://doi.org/10.1136/bmjopen-2019-034683

Haveraaen, L. A., Skarpaas, L. S., & Aas, R. W. (2017). Job demands and decision control predicted return to work: The rapid-RTW cohort study. *BMC Public Health*, *17*(1), 154. https://doi.org/10.1186/s12889-016-3942-8

Hees, H. L., de Vries, G., Koeter, M. W. J., & Schene, A. H. (2013). Adjuvant occupational therapy improves long-term depression recovery and return-to-work in good health in sick-listed employees with major depression: Results of a randomised controlled trial. *Occupational and Environmental Medicine*, *70*(4), 252–260. https://doi.org/10.1136/oemed-2012-100789

Hees, H. L., Koeter, M. W., de Vries, G., Ooteman, W., & Schene, A. H. (2010). Effectiveness of adjuvant occupational therapy in employees with depression: Design of a randomized controlled trial. *BMC Public Health*, *10*, 558. https://doi.org/10.1186/1471-2458-10-558

Henderson, M., Harvey, S. B., Overland, S., Mykletun, A., & Hotopf, M. (2011). Work and common psychiatric disorders. *Journal of the Royal Society of Medicine*, *104*(5), 198–207. https://doi.org/10.1258/jrsm.2011.100231

Henkel, D. (2011). Unemployment and Substance Use: A Review of the Literature (1990-2010). *Current Drug Abuse Reviewse*, *4*(1), 4–27. https://doi.org/10.2174/1874473711104010004

Herdman, M., Gudex, C., Lloyd, A., Janssen, Mf., Kind, P., Parkin, D., Bonsel, G., & Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research*, *20*(10), 1727–1736. https://doi.org/10.1007/s11136-011-9903-x

Hjemdal, O., Solem, S., Hagen, R., Kennair, L. E. O., Nordahl, H. M., & Wells, A. (2019). A Randomized Controlled Trial of Metacognitive Therapy for Depression: Analysis of 1-Year Follow-Up. *Frontiers in Psychology*, *10*, 1842. https://doi.org/10.3389/fpsyg.2019.01842

Hofmann, S. G., Asnaani, A., Vonk, I. J. J., Sawyer, A. T., & Fang, A. (2012). The Efficacy of Cognitive Behavioral Therapy: A Review of Meta-analyses. *Cognitive Therapy and Research*, *36*(5), 427–440. https://doi.org/10.1007/s10608-012-9476-1

Hultin, H., Lindholm, C., Malfert, M., & Möller, J. (2012). Short-term sick leave and future risk of sickness absence and unemployment—The impact of health status. *BMC Public Health*, *12*(1), 861. https://doi.org/10.1186/1471-2458-12-861

Hultin, H., Lindholm, C., & Möller, J. (2012). Is There an Association between Long-Term Sick Leave and Disability Pension and Unemployment beyond the Effect of Health Status? – A Cohort Study. *PLoS ONE*, *7*(4), e35614. https://doi.org/10.1371/journal.pone.0035614

Jung, T., & Wickrama, K. A. S. (2008). An Introduction to Latent Class Growth Analysis and Growth Mixture Modeling: Latent Trajectory Classes. *Social and Personality Psychology Compass*, *2*(1), 302–317. https://doi.org/10.1111/j.1751-9004.2007.00054.x

Kalin, N. H. (2020). The Critical Relationship Between Anxiety and Depression. *The American Journal of Psychiatry*, *177*(5), 365–367. https://doi.org/10.1176/appi.ajp.2020.20030305

Knapstad, M., Lervik, L. V., Sæther, S. M. M., Aarø, L. E., & Smith, O. R. F. (2020). Effectiveness of Prompt Mental Health Care, the Norwegian Version of Improving Access to Psychological Therapies: A Randomized Controlled Trial. *Psychotherapy and Psychosomatics*, *89*(2), 90–105. https://doi.org/10.1159/000504453

Knapstad, M., Nordgreen, T., & Smith, O. R. F. (2018). Prompt mental health care, the Norwegian version of IAPT: Clinical outcomes and predictors of change in a multicenter cohort study. *BMC Psychiatry*, *18*(1), 260. https://doi.org/10.1186/s12888-018-1838-0

Knudsen, A. K., Øverland, S., Hotopf, M., & Mykletun, A. (2012). Lost Working Years Due to Mental Disorders: An Analysis of the Norwegian Disability Pension Registry. *PLoS ONE*, *7*(8), e42567. https://doi.org/10.1371/journal.pone.0042567

Konerding, U. (2013). What does Cronbach's alpha tell us about the EQ-5D? A methodological commentary to "Psychometric properties of the EuroQol Five-Dimensional Questionnaire (EQ-5D-3L) in caregivers of autistic children". *Quality of Life Research*, *22*(10), 2939–2940. https://doi.org/10.1007/s11136-013-0430-9

Kröger, C., Bode, K., Wunsch, E.-M., Kliem, S., Grocholewski, A., & Finger, F. (2015). Work-related treatment for major depressive disorder and incapacity to work: Preliminary findings of a controlled, matched study. *Journal of Occupational Health Psychology*, *20*(2), 248–258. https://doi.org/10.1037/a0038341

Lagerveld, S. E., Blonk, R. W. B., Brenninkmeijer, V., Wijngaards-de Meij, L., & Schaufeli, W. B. (2012). Work-focused treatment of common mental disorders and return to work: A comparative outcome study. *Journal of Occupational Health Psychology*, *17*(2), 220–234. https://doi.org/10.1037/a0027049

Layard, R. (2006). The depression report: A new deal for depression and anxiety disorders. *Centre for Economic Performance, London School of Economics and Political Science*, 16.

Layard, R., & Clark, D. M. (2015). Why More Psychological Therapy Would Cost Nothing. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01713

Lépine, J.-P., & Briley, M. (2011). The increasing burden of depression. *Neuropsychiatric Disease and Treatment*, *7*(Suppl 1), 3–7. https://doi.org/10.2147/NDT.S19617

Lokman, S., Volker, D., Zijlstra-Vlasveld, M. C., Brouwers, E. P., Boon, B., Beekman, A. T., & Smit, F. (2017). Return-to-work intervention versus usual care for sick-listed employees: Health-economic investment appraisal alongside a cluster randomised trial. *Open Access*, 11.

Markussen, S. (2012). The individual cost of sick leave. *Journal of Population Economics*, *25*(4), 1287–1306. https://doi.org/10.1007/s00148-011-0390-8

Markussen, S., Røed, K., Røgeberg, O. J., & Gaure, S. (2011). The anatomy of absenteeism. *Journal of Health Economics*, *30*(2), 277–292. https://doi.org/10.1016/j.jhealeco.2010.12.003

McInness, B. (2018). *IAPT 2018: Why do less than 1 in 5 referrals reach recovery?* Therapy Meets Numbers. https://therapymeetsnumbers.com/iapt-2018-why-do-less-than-1-in-5-referrals-reach-recovery/

Meier, S. M., Mattheisen, M., Mors, O., Mortensen, P. B., Laursen, T. M., & Penninx, B. W. (2016). Increased mortality among people with anxiety disorders: Total population study. *British Journal of Psychiatry*, *209*(3), 216–221. https://doi.org/10.1192/bjp.bp.115.171975

Mendlowicz, M. V., & Stein, M. B. (2000). Quality of life in individuals with anxiety disorders. *American Journal of Psychiatry*, *157*(5), 669–682.

Meneton, P., Kesse-Guyot, E., Méjean, C., Fezeu, L., Galan, P., Hercberg, S., & Ménard, J. (2015). Unemployment is associated with high cardiovascular event rate and increased all-cause mortality in middle-aged socially privileged individuals. *International Archives of Occupational and Environmental Health*, *88*(6), 707–716. https://doi.org/10.1007/s00420-014-0997-7

Milligan-Saville, J. S., Tan, L., Gayed, A., Barnes, C., Madan, I., Dobson, M., Bryant, R. A., Christensen, H., Mykletun, A., & Harvey, S. B. (2017). Workplace mental health training for managers and its effect on sick leave in employees: A cluster randomised controlled trial. *The Lancet Psychiatry*, *4*(11), 850–858. https://doi.org/10.1016/S2215-0366(17)30372-3

Milner, A., Page, A., & LaMontagne, A. D. (2013). Long-Term Unemployment and Suicide: A Systematic Review and Meta-Analysis. *PLOS ONE*, *8*(1), e51333. https://doi.org/10.1371/journal.pone.0051333

Murphy, G. C., & Athanasou, J. A. (1999). The effect of unemployment on mental health. *Journal of Occupational and Organizational Psychology*, *72*(1), 83–99. https://doi.org/10.1348/096317999166518

NICE. (2020). *Overview | Generalised anxiety disorder and panic disorder in adults: Management | Guidance | NICE*. NICE. https://www.nice.org.uk/guidance/cg113

NICE. (2022). *Overview | Depression in adults: Recognition and management | Guidance | NICE*. NICE. https://www.nice.org.uk/guidance/cg90

Nieuwenhuijsen, K., Verbeek, J. H., Neumeyer-Gromen, A., Verhoeven, A. C., Bültmann, U., & Faber, B. (2020). Interventions to improve return to work in depressed people. *Cochrane Database of Systematic Reviews*, *2020*(12). https://doi.org/10.1002/14651858.CD006237.pub4

Nigatu, Y. T., Liu, Y., Uppal, M., McKinney, S., Gillis, K., Rao, S., & Wang, J. (2017). Prognostic factors for return to work of employees with common mental disorders: A meta-analysis of cohort studies. *Social Psychiatry and Psychiatric Epidemiology*, *52*(10), 1205–1215. https://doi.org/10.1007/s00127-017-1402-0

Nigatu, Y. T., Liu, Y., Uppal, M., McKinney, S., Rao, S., Gillis, K., & Wang, J. (2016). Interventions for enhancing return to work in individuals with a common mental illness: Systematic review and meta-analysis of randomized controlled trials. *Psychological Medicine*, *46*(16), 3263–3274. https://doi.org/10.1017/S0033291716002269

OECD. (2012). *Sick on the Job?: Myths and Realities about Mental Health and Work*. OECD. https://doi.org/10.1787/9789264124523-en

OECD. (2013). *Mental Health and Work: Norway*. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/employment/mental-health-and-work-norway_9789264178984-en

OECD. (2015). *Fit Mind, Fit Job: From Evidence to Practice in Mental Health and Work | en | OECD*. https://www.oecd.org/els/fit-mind-fit-job-9789264228283-en.htm

OECD. (2019). *Health at a Glance 2019: OECD Indicators*. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-2019_4dd50c09-en

Overland, S., Glozier, N., Krokstad, S., & Mykletun, A. (2007). Undertreatment before the award of a disability pension for mental illness: The HUNT Study. *Psychiatric Services (Washington, D.C.)*, *58*(11), 1479–1482. https://doi.org/10.1176/ps.2007.58.11.1479

Patel, V., Maj, M., Flisher, A. J., De Silva, M. J., Koschorke, M., Prince, M., WPA Zonal and Member Society Representatives, Tempier, R., Riba, M., Sanchez, M., Campodonico, F. D., Risco, L., Gask, L., Wahlberg, H., Roca, M., Lecic-Tosevski, D., Soghoyan, A., Moussaoui,

D., Baddoura, C., … Richardson, G. (2010). Reducing the treatment gap for mental disorders: A WPA survey. *World Psychiatry*, *9*(3), 169–176. https://doi.org/10.1002/j.2051-5545.2010.tb00305.x

Penninx, B. W., Milaneschi, Y., Lamers, F., & Vogelzangs, N. (2013). Understanding the somatic consequences of depression: Biological mechanisms and the role of depression symptom profile. *BMC Medicine*, *11*(1), 129. https://doi.org/10.1186/1741-7015-11-129

Peters, G.-J. Y. (2014). The alpha and the omega of scale reliability and validity: Why and how to abandon Cronbachâ€™s alpha and the route towards more comprehensive assessment of scale quality. *European Health Psychologist*, *16*(2), Article 2.

Rai, D., Skapinakis, P., Wiles, N., Lewis, G., & Araya, R. (2010). Common mental disorders, subthreshold symptoms and disability: Longitudinal study. *British Journal of Psychiatry*, *197*(5), 411–412. https://doi.org/10.1192/bjp.bp.110.079244

Reme, S. E. (2020). Common mental disorders and work. In *Handbook of Disability, Work and Health.* Springer Nature.

Reme, S. E., Grasdal, A. L., Løvvik, C., Lie, S. A., & Øverland, S. (2015). Work-focused cognitive–behavioural therapy and individual job support to increase work participation in common mental disorders: A randomised controlled multicentre trial. *Occupational and Environmental Medicine*, *72*(10), 745–752. https://doi.org/10.1136/oemed-2014-102700

Ruggiero, G. M., Caselli, G., & Sassaroli, S. (Eds.). (2021). *CBT Case Formulation as Therapeutic Process*. Springer International Publishing. https://doi.org/10.1007/978-3-030-63587-9

Salomonsson, S., Hedman-Lagerlöf, E., & Öst, L.-G. (2018). Sickness absence: A systematic review and meta-analysis of psychological treatments for individuals on sick leave due to common mental disorders. *Psychological Medicine*, *48*(12), 1954–1965. https://doi.org/10.1017/S0033291718000065

Salomonsson, S., Santoft, F., Lindsäter, E., Ejeby, K., Ljótsson, B., Öst, L.-G., Ingvar, M., Lekander, M., & Hedman-Lagerlöf, E. (2017). Cognitive–behavioural therapy and return-to-work intervention for patients on sick leave due to common mental disorders: A randomised

controlled trial. *Occupational and Environmental Medicine*, *74*(12), 905–912. https://doi.org/10.1136/oemed-2017-104342

Sandin, K., Anyan, F., Osnes, K., Gunnarsdatter Hole Gjengedal, R., Risberg Leversen, J. S., Endresen Reme, S., & Hjemdal, O. (2021). Sick leave and return to work for patients with anxiety and depression: A longitudinal study of trajectories before, during and after work-focused treatment. *BMJ Open*, *11*(9), e046336. https://doi.org/10.1136/bmjopen-2020-046336

Sandin, K., Gjengedal, R. G. H., Osnes, K., Hannisdal, M., Berge, T., Leversen, J. S. R., Røv, L. G., Reme, S. E., Lagerveld, S., Blonk, R., Nordahl, H. M., Shields, G., Wells, A., & Hjemdal, O. (2021). Metacognitive therapy and work-focused interventions for patients on sick leave due to anxiety and depression: Study protocol for a randomised controlled wait-list trial. *Trials*, *22*(1), 854. https://doi.org/10.1186/s13063-021-05822-4

Schene, A. H., Koeter, M. W. J., Kikkert, M. J., Swinkels, J. A., & McCRONE, P. (2007). Adjuvant occupational therapy for work-related major depression works: Randomized trial including economic evaluation. *Psychological Medicine*, *37*(03), 351. https://doi.org/10.1017/S0033291706009366

Schnittker, J., & Bacak, V. (2014). The Increasing Predictive Validity of Self-Rated Health. *PLoS ONE*, *9*(1), e84933. https://doi.org/10.1371/journal.pone.0084933

Schultz, I. Z., & Gatchel, R. J. (Eds.). (2015). *Handbook of Return to Work: From Research to Practice* (1st ed. 2016 edition). Springer.

Schwartz, S., & Carpenter, K. M. (1999). The right answer for the wrong question: Consequences of type III error for public health research. *American Journal of Public Health*, *89*(8), 1175–1180. https://doi.org/10.2105/AJPH.89.8.1175

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry*, *59 Suppl 20*, 22-33;quiz 34-57.

Solem, S., Kennair, L. E. O., Hagen, R., Havnen, A., Nordahl, H. M., Wells, A., & Hjemdal, O. (2019). Metacognitive Therapy for Depression: A 3-Year Follow-Up Study Assessing

Recovery, Relapse, Work Force Participation, and Quality of Life. *Frontiers in Psychology*, *10*, 2908. https://doi.org/10.3389/fpsyg.2019.02908

Solem, S., Wells, A., Kennair, L. E. O., Hagen, R., Nordahl, H., & Hjemdal, O. (2021). Metacognitive therapy versus cognitive–behavioral therapy in adults with generalized anxiety disorder: A 9-year follow-up study. *Brain and Behavior*, *11*(10). https://doi.org/10.1002/brb3.2358

Steel, Z., Marnane, C., Iranpour, C., Chey, T., Jackson, J. W., Patel, V., & Silove, D. (2014). The global prevalence of common mental disorders: A systematic review and meta-analysis 1980–2013. *International Journal of Epidemiology*, *43*(2), 476–493. https://doi.org/10.1093/ije/dyu038

Stengård, J., Peristera, P., Johansson, G., & Nyberg, A. (2021). The role of managerial leadership in sickness absence in health and social care: Antecedent or moderator in the association between psychosocial working conditions and register-based sickness absence? A longitudinal study based on a swedish cohort. *BMC Public Health*, *21*(1), 2215. https://doi.org/10.1186/s12889-021-12236-z

van der Klink, J. J. L. (2003). Reducing long term sickness absence by an activating intervention in adjustment disorders: A cluster randomised controlled design. *Occupational and Environmental Medicine*, *60*(6), 429–437. https://doi.org/10.1136/oem.60.6.429

van der Noordt, M., IJzelenberg, H., Droomers, M., & Proper, K. I. (2014). Health effects of employment: A systematic review of prospective studies. *Occupational and Environmental Medicine*, *71*(10), 730–736. https://doi.org/10.1136/oemed-2013-101891

Vigo, D., Thornicroft, G., & Atun, R. (2016). Estimating the true global burden of mental illness. *The Lancet Psychiatry*, *3*(2), 171–178. https://doi.org/10.1016/S2215-0366(15)00505-2

Waddell, G., Burton, A. K., Great Britain, & Department for Work and Pensions. (2007). *Is work good for your health and well-being?* TSO. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/214326/hwwb-is-work-good-for-you.pdf

Wang, Y.-P., & Gorenstein, C. (2013). Psychometric properties of the Beck Depression Inventory-II: A comprehensive review. *Revista Brasileira de Psiquiatria*, *35*(4), 416–431. https://doi.org/10.1590/1516-4446-2012-1048

Wells, A. (2009). *Metacognitive Therapy for Anxiety and Depression* (1st edition). The Guilford Press.

Wells, A., & Matthews, G. (1996). Modelling cognition in emotional disorder: The S-REF model. *Behaviour Research and Therapy*, *34*(11), 881–888. https://doi.org/10.1016/S0005-7967(96)00050-2

WHO. (1992). *The International Classification of Diseases-10 (ICD-10).* World Health Organisation.

World Health Organization. (2013). *Investing in mental health: Evidence for action*. WHO: Geneva.

World Health Organization. (2017). *Depression and other common mental disorders: Global health estimates* (WHO/MSD/MER/2017.2). World Health Organization. https://apps.who.int/iris/handle/10665/254610

World Medical Association. (2022). *Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects*. https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/

Yang, X., Fang, Y., Chen, H., Zhang, T., Yin, X., Man, J., Yang, L., & Lu, M. (2021). Global, regional and national burden of anxiety disorders from 1990 to 2019: Results from the Global Burden of Disease Study 2019. *Epidemiology and Psychiatric Sciences*, *30*, e36. https://doi.org/10.1017/S2045796021000275

BMJ Open

Sick leave and return to work for patients with anxiety and depression: a longitudinal study of trajectories before, during and after work-focused treatment

Kenneth Sandin ,[1,2] Frederick Anyan,[2] Kåre Osnes,[1] Ragne Gunnarsdatter Hole Gjengedal,[1,2] Jonas Sigurd Risberg Leversen,[1] Silje Endresen Reme,[3] Odin Hjemdal[1,2]

To cite: Sandin K, Anyan F, Osnes K, *et al*. Sick leave and return to work for patients with anxiety and depression: a longitudinal study of trajectories before, during and after work-focused treatment. *BMJ Open* 2021;**11**:e046336. doi:10.1136/bmjopen-2020-046336

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-046336).

Received 26 October 2020
Accepted 13 September 2021

[1]Division of Mental Health and Substance Abuse, Diakonhjemmet Hospital, Oslo, Norway
[2]Department of Psychology, Norwegian University of Science and Technology, Trondheim, Norway
[3]Department of Psychology, University of Oslo, Oslo, Norway

Correspondence to
Kenneth Sandin;
kenneth.sandin@diakonsyk.no

ABSTRACT

Objectives Sick leave due to anxiety and depression is a heterogeneous process constituting a pressing public health issue. This longitudinal study aimed to identify sick leave trajectories among patients before, during and after work-focused treatment, in all 29.5 months. We then aimed to determine the background and clinical characteristics of these trajectory groups.

Methods Background and clinical data were collected by patient self-report (N=619) in an observational study in a specialised mental healthcare clinic. Sick leave was recorded from national registry data. A latent growth mixture model identified trajectories. Multinomial logistic regression determined differences in background characteristics while a one-way analysis of variance (ANOVA) identified clinical differences.

Results We identified three trajectories: The 'Resilient' group (47.7%) had low sick leave throughout the period. The two other groups ('Recovery', 31.8% and 'High risk', 20.5%) had similar pretreatment trajectories: lower sick leave one year prior which increased to high sick leave at the start of treatment. After treatment, the 'Recovery' group made an almost full return to work while the 'High risk' group remained at high sick leave. The two groups with high sick leave had more women and higher age compared with the 'Resilient' group. All groups had similar clinical scores at the start of treatment, but the 'High risk' groups had residual depressive symptoms at the end of treatment. Effect sizes for anxiety and depression were moderate or large for all groups, (Cohen's *d*=0.74–1.81), and 87.2% of the total sample were fully working one year after treatment.

Conclusion We found three subgroups with distinctly different trajectories. Female gender and higher age were associated with high sick leave at the start of treatment, while residual depressive symptoms at the end of treatment predicted continued sick leave. The study points to the possibility of improving patient outcomes in the future by stratifying and tailoring treatment to patient characteristics.

Strengths and limitations of this study

- This is the first study that uses registry data spanning pretreatment and post-treatment to examine sick leave trajectories among patients with common mental disorders.
- Latent growth modelling over 2.5 years gave detailed knowledge of sick leave behaviour.
- The study contributes to a better understanding of sick leave over time with implications for treatment and research.
- The study was not a randomised controlled trial and lacks a control group, precise impact of the intervention is yet to be determined.

INTRODUCTION

Sick leave due to common mental disorders (CMDs) such as anxiety and depression is a pressing public health issue. Globally, one in five people fulfil diagnostic criteria for a CMD at any given time, and the lifetime prevalence is estimated to almost one-third of the population (29.2%).[1] People suffering from mental health problems commonly report that their daily activities are impeded, and mental illness leads to low employment rates and reduced productivity.[2 3] Measured in 'years lived with disability', mental ill health is the largest contributor to burden of disease worldwide.[4] The economic cost of mental disorders in Europe alone is estimated to be €600 billion per year, the majority of which comes from reduced employment and lost productivity.[5]

Several policy initiatives have been launched to address the rising cost of CMD. Most well known is perhaps Improving Access to Psychological Therapies in the UK. The programme was launched to alleviate the cost of mental illness, including lost productivity, by scaling up access to psychological therapy.[6] Similarly, the Faster Return (FR) programme was initiated in Norway in 2007. As Norway has the highest rates of sick leave globally, the

target population for the programme were patients 'on or at risk of sick leave'.[7] The data in the present study were collected in a clinic originally funded by the FR programme.

Offering effective treatment for CMD symptoms that also reduces sick leave is not straightforward. Reviews indicate that psychotherapy alone has little impact on sick leave, but that pairing it with work-focused interventions can help patients return to work (RTW).[8] There is still room for improvement. Some studies struggle to find effect on work status, while others find that the effect is unevenly distributed among patient groups.[9–11] This is not surprising, as sick leave is a heterogeneous process. Personal characteristics, features of the work, workplace and health issues including symptom severity all contribute to sick leave.[12 13] Whether this heterogeneity of prognostic factors contain more homogeneous subgroups of patients with similar risk profiles is currently not clear. If so, identifying these groups could be an important next step towards more effective treatment.

Developing more effective treatment is also likely to depend on a better understanding of how RTW is sustained over time. First, previous sick leave predicts future absence, regardless of health status, raising the question of whether treatment outcome may in part be determined by sick leave history.[14] Second, longitudinal studies evaluating how patients fare after interventions show that they often struggle to increase and maintain work participation.[15 16] Cross-sectional measurements of sick leave (eg, at the end of treatment) may therefore be of limited value for understanding patient outcomes and needs. A more thorough understanding of sick leave behaviour may depend on including data before, during and after an intervention.

In the present study we therefore wanted to examine trajectories of sick leave before, during and after work-focused treatment. We then examined if known risk factors from the literature differed between the various trajectories. We included both non-modifiable background characteristics and modifiable clinical characteristics: gender, age, education level, somatic comorbidity, CMD symptom severity, self-rated subjective health and RTW self-efficacy (RTW-SE). All factors that have consistently been found to predict sick leave and RTW.[13 17]

This study is unique in being the first to combine clinical data with longitudinal registry based sick leave data covering both preintervention and postintervention periods for patients with CMD. Furthermore, registry data gave us an objective measure of sick leave with no loss to follow-up. The observational period stretched from one year prior to treatment to one year after end of treatment (24 months). Average duration of treatment was 5.5 months, making the total observation period 29.5 months, or approximately 2.5 years.

Our primary objective was: (1) to determine if subgroups of patients could be identified based on their sick leave trajectories before, during and after they received work-focused treatment in a specialised healthcare clinic.

Our secondary objectives were to (1) examine if these groups differed on non-modifiable background characteristics: age, gender, education and somatic comorbidity, and (2) examine if these groups differed on modifiable clinical characteristics: CMD symptom severity, self-rated subjective health and RTW-SE.

METHODS

Study context

Data were collected in an outpatient clinic at Diakonhjemmet Hospital in Oslo, Norway. The clinic is part of the specialised healthcare service and the observational study ran from 2013 to 2016. Cognitive behavioural therapy (CBT) and metacognitive therapy (MCT) are recommended for treating anxiety and depression.[18 19] Our treatment consisted of short-term CBT or MCT based on diagnose-specific manuals,[20 21] which was paired with work-focused interventions. The work-focus consists of clinicians addressing the patients work status from the onset of therapy, mapping resources and barriers for maintaining work status, or in the case of sick leave, returning to work. A gradual RTW plan is developed in cooperation with the patient, and interventions tailored to address challenging issues, for example, role playing to help patients gain confidence communicating their needs in the workplace. The work-focused interventions are based on Dutch research,[22] and have been further developed and adapted to a Norwegian context. They have been described in more detail in a previous publication.[23] The treatment was provided by 20 therapists who were clinical psychologists, psychiatrists or psychiatric nurses.

The Norwegian welfare system ensures that patients receive 100% coverage of lost income through sick pay from day one and up to one year. Employers cover payments for the first 16 days, and the Norwegian Labour and Welfare Administration cover the remaining part of the year. Patients may then be eligible for long-term benefits that cover roughly 66% of their original income.[24]

Procedure and participants

Participants were referred for treatment by their general practitioners (GPs), who are responsible for certifying sick leave. To be eligible for inclusion, patients had to be adults of working age (18–67 years), either on or at risk of sick leave due to anxiety and/or depression. All patients in the study were either on sick leave or deemed 'at risk' of sick leave by their GPs. Patients on long-term benefits such as disability pension at baseline were excluded from the study. Patients were also excluded if they were suffering from psychosis (ie, bipolar disorder, schizophrenia or other psychotic disorders), engaged in active substance abuse, suffered from cluster A or B personality disorder or considered high suicide risk. Patients received oral and written information about the study and signed a consent form before being enrolled. In all, 619 patients were recruited.

All patients were screened at an initial assessment session by a clinical psychologist and diagnosed according to the International Classification of Diseases-10.[25] After the screening session, patients waited an average of 42.71 days (*SD*=29.4) before starting treatment, after which patients were given weekly sessions of work-focused treatment. Average number of sessions including screening was 10.71 (*SD*=3.24, duration *M*=121.92 days, *SD*=57.92). Average duration of entire intervention including screening was thus 5.5 months. Background variables and clinical scores were collected from self-report questionnaires during treatment at the clinic. Registry data on sick leave was collected from national registries.

ETHICAL CONSIDERATIONS

This observational study qualifies as health service research and was therefore pre-approved by the Norwegian Data Protection Authority. Patients were informed that they could withdraw consent without providing any explanation. The study was conducted according to the principles of the Helsinki Declaration. User representatives were involved in the planning of the current study, including design and data collection. The primary user representatives involved represented the user interest organisation Mental Health.

MEASURES

Registry data

Sick leave was collected from national registries containing comprehensive records as reported by medical doctors. The data covered each patient's sick leave spanning 29.5 months: One year prior to inclusion, the duration of treatment, and one year after the end of treatment. Sick leave for each individual is given as a start and end date of episode. We had information on degree of sick leave, for example, 50% or 100%. However, for the purpose of this study, we operationalised sick leave as a dichotomous variable: Fully working with no sick leave of any kind vs all degrees of sick leave. We did this to give us a conservative measure of RTW, and to facilitate comparisons with other studies.

For the latent growth mixture model (LGMM), we set measurement points at 90-day intervals. Thus, we had four measurements prior to treatment: at 360 days, 270 days, 180 days and 90 days. The same intervals were applied after the end of treatment. The treatment period was included in the model as three measurement points: Screening session, start of treatment and end of treatment. Thus, the final model contained 11 measurement points (figure 1). Each measurement point reflects the sick leave of the patient population at the time. Each patient was given a value of either 0 ('Fully working') or 1 ('On sick leave') for each of the measurement points based on model estimates. The sick leave of 100 patients assigned in this manner, where 46 were working and 54 were on sick leave would thus be '54 %'.

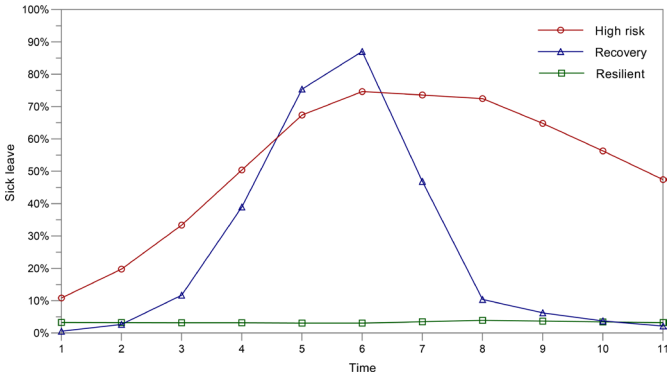


Figure 1 The three trajectories: 'High risk' 20.5% (n=127), 'Recovery' 31.8% (n=197), 'Resilient'47.7% (n=295). Time point 1–4 denotes the year prior to intervention at 90-day intervals. Time point 5, 6 and 7 denotes assessment, start of treatment and end of treatment. Time point 8–11 denote the year after intervention at 90-day intervals.

Background characteristics

Background characteristics were recorded as self-report at assessment and at end of treatment. Age, gender and education have all been shown to impact sick leave in previous studies.[13 26] In addition, comorbid somatic diagnosis have been shown to negatively impact work functioning over time in patients with CMD in a previous longitudinal study.[17] The patients answered the question 'Do you have any somatic illness diagnosed by a health professional?'. Somatic diagnoses were diverse, with musculoskeletal (24.2%), cardiovascular (12.7%) and obesity (12.1%), making up the main diagnostic categories. Both 'Somatic diagnosis' and 'Higher education' were included as dichotomous variables (yes/no). 'Higher education' was defined as any completed degree beyond upper secondary school, that is, the first 12 years of school.

Clinical characteristics

Anxiety and depression were measured with the Beck Depression Inventory-II (BDI-II) and the Beck Anxiety Inventory (BAI). Both questionnaires contain 21 items that are scored from 0 to 3, giving a total score of 0–63. Higher score indicates higher symptom severity. Both the BDI-II and the BAI have demonstrated good reliability and validity.[27 28] All patients had either a score ≥14 on the BDI-II or ≥15 on the BAI at baseline in keeping with criteria for clinically significant symptoms of depression or anxiety.

Self-rated health (SRH) was measured by a single item rating of total subjective health. Respondents answered the question 'How would you rate your health at the moment?' on a 4-point scale: 'Bad', 'Not so good', 'Good', 'Very good'. SRH is commonly used in public health surveys, and is a robust predictor of general health status including all-cause mortality.[29]

The RTW-SE questionnaire is an 11-item self-report questionnaire shown to predict RTW, including in a Norwegian sample.[23] Patients are asked how they would

deal with overcoming obstacles in RTW by responding to statements like 'I will be able to cope with potential problems at work' or 'I will be able to manage set-backs' using a 6-point Likert-scale. Total score is recorded as the mean of all responses, and higher score indicates higher self-efficacy.[30]

There was no missing sick leave data in the study due to the source being national registries. The same was true for the background characteristics and diagnoses, as they were recorded from patient journals. The exception was 'Higher education' with 3.1% missing values. The clinical data collected from patient self-report was also mostly complete, and more than 95% of patient questionnaires had ≤1 missing item. In these cases, missing items were replaced by weighted means.

Statistical analyses

The registry data was analysed with LGMM using Mplus V.7.4.[31] This method helps model unobserved heterogeneity in a population by identifying subgroups with similar individual growth trajectories across repeated measures. We favoured piecewise growth modelling, thus setting the start of treatment as the status factor reference point. Once the best fitting latent growth curve model had been established, we estimated the LGMM. To determine the appropriate class solution, we examined fit statistics with classification accuracy so that average probability of belonging to the most likely class should be high, and the average probability of belonging to the other classes should be low.[32] Particular attention is given to the Bayesian Information Criterion (BIC), sample-size adjusted BIC (ABIC), and Akaike information criterion indices, Entropy values and the bootstrap likelihood ratio test (BLRT). We sought a model with lower values for the criterion indices, higher entropy values and significant p values for the BLRT.[33] Fit indices in combination with theoretical interpretability guided the final model selection.

Once the best fitting model had been established, we wanted to know if auxiliary variables like background and clinical characteristics differed across the groups. Only sick leave was included in the LGMM, and we analysed auxiliary variables as predictors of class membership post hoc. This is a viable approach where LGMM model entropy is ≥0.8.[34] However, in the post hoc analyses, class membership is necessarily treated as categorical, despite assignment being probabilistic. Standard errors inherent in the parameter may thus be underestimated. It is therefore recommended to choose a more stringent significance level than p<0.05 for these analyses.[34] Significance level for all post hoc analyses in this study was thus set to p<0.01.

Background characteristics (age, gender, education level and previously diagnosed somatic illness) were selected based on the literature and analysed as predictors using multinomial logistic regression. First, we used the 'High risk' group as the reference category. We then ran the same analyses with the 'Resilient' group as reference category, to also obtain comparisons between the 'Recovery' and 'Resilient' groups. The clinical measures are reported as pre to post sum scores. Effect size (ES) was calculated using Cohen's *d*. Between-group differences on clinical measures pre and post were analysed using a one-way analysis of variance(ANOVA) and Tukey's honestly significant difference test.[35]

Patient and public involvement

User representatives and patients were involved in the planning of the study, including giving feedback on data collection, the content of work-focused interventions and self-report questionnaires prior to the start of the study. Findings from the study will be disseminated in cooperation with Stiftelsen Dam, a not-for-profit trust, and Mental Health, a user interest organisation.

RESULTS

Characteristics of patients

For the total sample, average age at inclusion was 37.9 years (*SD*=10.5). There were more women (68.5%). More than three quarters (76.5%) had some form of higher education. The most prevalent primary diagnoses were depressive disorders (47.8%, n=296), anxiety disorders (36.7%, n=227) and mixed anxiety and depression (12.9%, n=80). The remaining 2.6% (n=16) had other primary diagnoses, such as eating or sleeping disorders (table 1).

Trajectories of sick leave

The LGMM identified a three-piece unconditional linear growth trajectory that showed adequate fit according to the root mean square error of approximation (RMSEA), the comparative fit index (CFI), and the Tucker-Lewis index (TLI): (χ^2=170.869, *df*=43, p<0.001; RMSEA=0.069, (90% CI=0.059 to 0.080); CFI=0.973; TLI=0.966) (figure 1). Model fit indices of all unconditional LGMM models under comparison are shown in table 2.

Determining the optimal class solution was carried out incrementally until model fit ceased to improve in a 5-class solution with non-significant results for the likelihood ratio tests. Despite a lower BIC in the 4-class model than the 3-class model, classes 1 and 3 in the 4-class model were not distinct from each other as both classes followed similar trajectory. Thus, guided by theoretical interpretability, the class profile plot based on the estimated posterior probabilities and the best performing BIC, we favoured a 3-class solution whose BIC was lower than a 2-class solution.[33] The 3-class solution provided a narrower class assignment probability for most likely latent class membership than a 4-class solution, showing a higher degree of precision and reliability of the classification. Compared with the 4-class solution, the 3-class solution also showed easy-to-interpret conditional response probabilities than a 4-class solution,

Table 1 Patient characteristics and multinomial logistic regression at assessment

	Characteristics			Multinomial logistic regression					
	High risk	Recovery	Resilient	1 vs 2		1 vs 3		2 vs 3	
	n=127	n=197	n=295	RR	95% CI	RR	95% CI	RR	95% CI
Age, mean±SD (y)	41.5±11.3	38.4±10.1	36.0±10.0	0.97	0.95 to 0.99	0.95*	0.93 to 0.97	1.03†	1.02 to 1.04
Female gender, n (%)	99 (78.0)	145 (73.6)	180 (61.0)	0.69	0.39 to 1.2	0.36*	0.21 to 0.60	1.92†	1.28 to 2.90
Higher education, n (%)	82 (67.2)	146 (76.4)	231 (80.5)	1.60	0.96 to 2.7	2.13†	1.3 to 3.5	0.75	0.48 to 1.19
Somatic diagnosis, n (%)	61 (48.0)	69 (35.0)	90 (30.5)	0.60	0.37 to 0.95	0.50†	0.32 to 0.78	1.20	0.81 to 1.78

1: 'High risk', 2: 'Recovery', 3: 'Resilient'.
*P<0.001 level.
†P<0.01 level.
RR, Relative risk.

providing a more reasonable representation of the data and more parsimonious model.

For the first trajectory, sick leave at one year pre-treatment was 10.8%, at screening 67.4% and at one year post-treatment 47.5%. Patients belonging to this trajectory class were labelled the 'High risk' group (n=127; 20.5%). For the second trajectory, sick leave at one year pre-treatment was 0.6%, at screening it was 87.1%, and at one year post-treatment it was 2.2%. Patients belonging to this trajectory class were labelled the 'Recovery' group (n=197; 31.8%). For the third trajectory, sick leave was low throughout the period, estimated sick leave was consistently below 4%. Patients belonging to this trajectory were labelled the 'Resilient' group (n=295; 47.7%). At one year post treatment, 87.2% (n=540) of the total study population were fully working. Patients fully working one year post-treatment per trajectory were as follows: 'High risk' group, 48.8% (n=62); 'Recovery' group, 98.0% (n=193); and for the 'Resilient' group 96.6% (n=285).

Background and clinical characteristics of the groups

Proportion of depressive disorder was largest in the 'Recovery' group (55.3%) compared with the 'High risk' (48.8%) and 'Resilient' groups 3 (42.4%). There was no significant difference in waiting time or number of sessions between the groups, one-way ANOVA $F(109, 509)=1.16$, p=0.14 and $F(19, 599)=0.98$, p=0.47, respectively. Half the patients (49.9%) were on sick leave at the start of treatment, decreasing to 12.8% one year after the intervention. Multinomial logistic regression showed that there were no significant differences between the 'High risk' and 'Recovery' groups. The following covariates significantly predicted the log-odds of being in the 'High risk' group: higher age, female gender, not having higher education and having a previously diagnosed somatic illness. The results of the multinomial logistic regression are presented in table 1.

Figure 2 shows clinical scores pretreatment and post-treatment. Within-group (ES) from pretreatment to post-treatment were moderate to large on all measures. The lowest ES were consistently associated with the 'High risk' group (d=0.54–0.84), while the 'Recovery' group had the highest ES on all measures (d=1.10–1.81).

Table 3 shows a one-way ANOVA at pretreatment and post-treatment including Tukey's honestly significant difference (HSD) for between-group differences on clinical measures. Largest pretreatment differences were found on SRH and RTW-SE, $F(2,593)=38.17$, p=0.00, and $F(2,596)=81.41$, p=0.00, respectively. For between-group differences, the post hoc Tukey HSD showed that the 'High risk' and 'Recovery' groups were

Table 2 Model fit indices for latent growth mixture model (N=619)

	AIC	BIC	ABIC	Entropy	LMR LR test p value	ALMR LR p value	Classification accuracy	BLRT p value
1-Class	6335.842	6353.561	6340.862	–			–	–
2-Class	13 276.114	13 382.427	13 306.231	0.850	0.0000	0.0000	0.962 to 0.965	***
3-Class	5228.056	5290.072	5245.625	0.828	0.0424	0.0457	0.913 to 0.939	***
4-Class	5124.340	5208.504	5148.182	0.828	0.0007	0.0008	0.853 to 0.948	***
5-Class	5052.982	5159.301	5083.105	0.852	0.0771	0.0818	0.832 to 0.971	***

***Denotes significant level at p<0.001.
ABIC, sample size adjusted BIC; AIC, Akaike information criterion; ALMR LR, Lo-Mendell-Rubin Adjusted Lrt Test; BIC, Bayesian information criterion; BLRT, bootstrap likelihood ratio test; LMR LR, Vuong-Lo-Mendell-Rubin Likelihood Ratio Test.

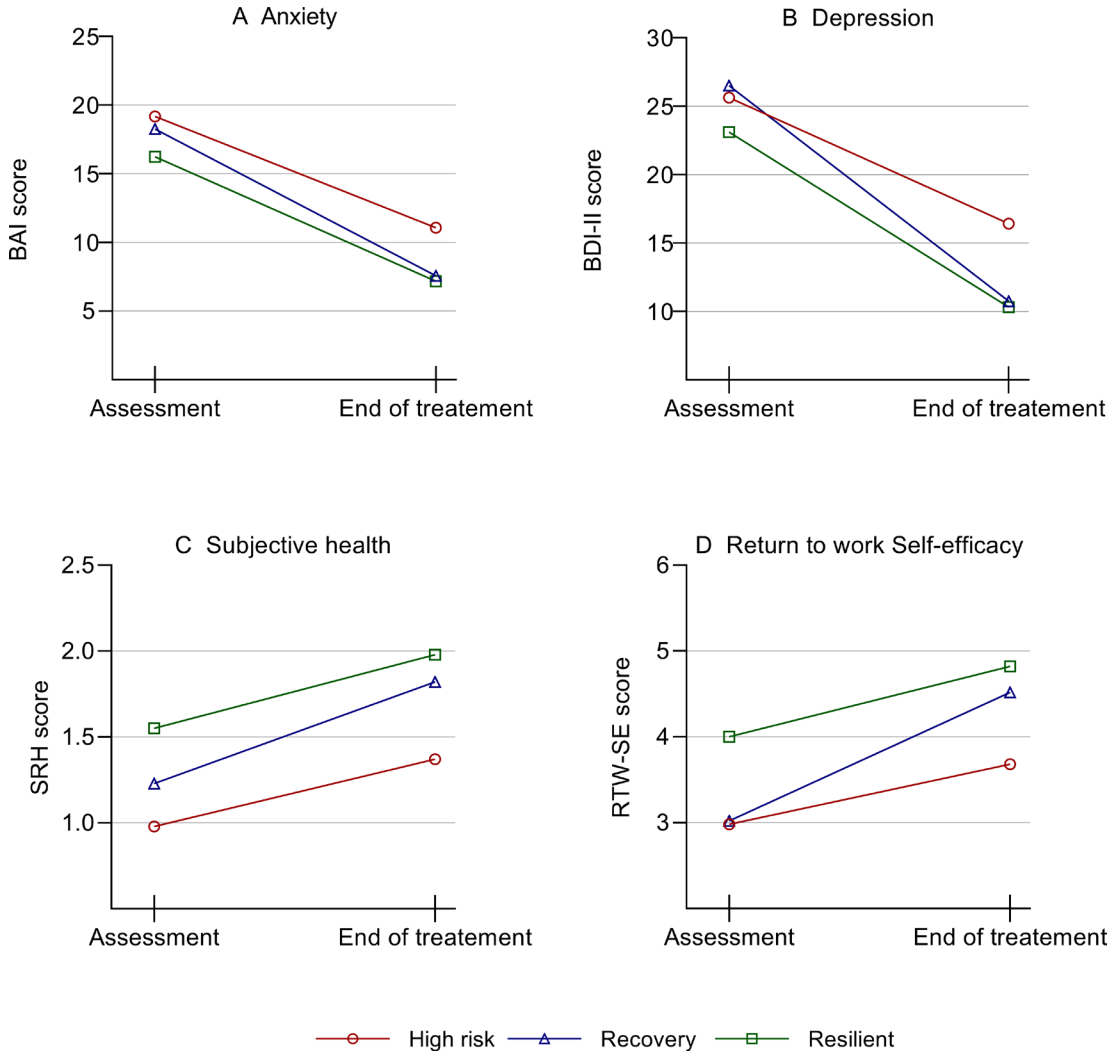


Figure 2 Effect size (ES) by Cohen's d. Anxiety: 'High risk': d=0.74, 'Recovery': d=1.81, 'Resilient' d=1.04. Depression: 'High risk': d=0.84, 'Recovery': d=1.71, 'Resilient' d=1.41. Subjective health: 'High risk': d=0.54, 'Recovery': d=1.82, 'Resilient' d=0.70. RTW-SE: 'High risk': d=0.66, 'Recovery': d=1.58, 'Resilient' d=0.98. Cohen's d>05 indicates moderate ES, >0.8 indicates large ES. BAI, Beck Anxiety Inventory; BDI-II, Beck Depression Inventory-II; RTW-SE, return to work self-efficacy; SRH, self-rated health.

only different on SRH at baseline. Post-treatment, the differences between the 'High risk' and 'Recovery' groups had increased markedly (table 3, figure 2).

DISCUSSION

Our main aim was to determine if trajectories of sick leave could be detected in a population of patients with CMD who received work-focused treatment. We identified three groups with distinctly different sick leave trajectories. Almost half the patients ('Resilient', 47.7%) had low sick leave throughout the 29.5 months, whereas the two remaining groups had either a near total RTW following treatment ('Recovery', 31.8%), or sustained high sick leave ('High risk', 20.5%). Our model showed good fit indices and was based solely on objective sick leave data from national registers spanning 2.5 years. We thus regard these groups to be robust classifications of sick leave behaviour in this patient cohort.

These subgroups and their characteristics have several implications. The 'Resilient' group largely avoided sick leave altogether despite relatively high symptom levels, on par with the two other groups. This suggests that there may be an element of prophylactic effect to the treatment. However, this is not a controlled study and we therefore do not know if their sick leave rates would have been higher without the intervention. We also do not know to which degree the 'Resilient' group experienced problems at work. Previous research would suggest at least some degree of presenteeism.[2] Both are potential questions for future research.

At the start of treatment, 92.4% of the patients in the 'Recovery' group were on sick leave. This number steadily decreased for the remainder of the

Table 3 Group differences on clinical measures at pretreatment and post-treatment (N=619)

	ANOVA		Tukey HSD		
	F	η^2	1 vs 2	2 vs 3	1 vs 3
Assessment					
BDI-II	8.21*	0.025	0.79	3.86*	2.49
BAI	4.38	0.014	0.74	2.12	2.67
SRH	38.17*	0.38	0.43†	5.38*	8.43*
RTW-SE	81.41*	0.21	0.36	11.01*	9.90*
End of treatment					
BDI-II	19.96*	0.028	5.28*	0.49	6.07*
BAI	11.30*	0.011	3.87*	0.54	4.61*
SRH	43.00*	0.41	6.51*	2.62	9.25*
RTW-SE	73.02*	0.14	8.30*	3.62*	12.07*

One-way ANOVA and Tukey honestly significant difference (HSD) test. 1: 'High risk', 2: 'Recovery', 3: 'Resilient'.
*P<0.001 level.
†P<0.01 level.
ANOVA, Analysis of variance; BAI, Beck Anxiety Inventory; BDI-II, Beck Depression Inventory-II; RTW-SE, return to work self-efficacy; SRH, self-rated health.

study's observational period. Within three months of completing treatment, 14.7% remained on sick leave, and at 12 months after the end of treatment, only 2% of the patients in this group remained on sick leave. A recent review defined sustainable RTW as no new sick leave within three months of initial RTW.[36] The majority of patients who went on sick leave in this group can thus be said to have made a highly sustainable RTW.

Perhaps our most interesting findings are the contrasts and similarities between this 'Recovery' group and the 'High risk' group. These groups had similar initial sick leave trajectories, culminating in high sick leave at the start of treatment. During the treatment, however, their sick leave trajectories diverged sharply. The high rates of patients fully working in the 'Recovery' group was contrasted by the continued high sick leave of the 'High risk' group: 47.5% of patients in this group were still on sick leave 12 months after the end of treatment. In this respect, their sick leave prognosis at the start of treatment could hardly have been more different (Time 6 in figure 1). This was evident at the end of treatment but became even more pronounced three months after treatment had ended (Time 8 in figure 1). In fact, minimal changes in sick leave occurred after this time point. This implies that the probability of RTW is drastically reduced if it has not occurred within three months after ended treatment.

In the second aim of our study, we asked if the background or clinical characteristics at baseline could help distinguish between the groups. Differences in risk factors could help clinicians identify patients with poor sick leave prognosis and potentially tailor and adapt interventions accordingly. Significant differences were found on all background characteristics between the 'High risk' and 'Resilient' groups. Higher age, female gender, lower education and somatic comorbidity were all associated with higher risk of sick leave in our study. This is in line with previous research.[13] However, no significant differences were found between the 'High risk' and 'Recovery' groups on background characteristics. Furthermore, clinical differences between the groups on anxiety and depression at baseline were small. Thus, it would likely not be possible for clinicians to predict sick leave prognosis at baseline in our study.

Despite similarities at baseline, clinical scores diverged for the 'High risk' and 'Recovery' groups during treatment, mirroring their sick leave trajectories. Although all three groups showed substantial clinical improvement, greatest change were seen in the 'Recovery' group and the smallest changes in the 'High risk' group. This was true of all clinical measures. Two scores are of particular interest. Firstly, the BDI-II showed that the depression score for the 'High risk' group remained above clinical threshold at the end of treatment, in line with previous research linking depression symptoms severity to delayed RTW.[12] Secondly, the RTW-SE score of the 'High risk' group remained below the threshold that previous research has indicated is required for high probability of RTW.[30] This implies that clinicians should be wary of residual depressive symptoms and low RTW-SE scores at the end of treatment.

A further implication is that these patients may need more follow-up. Clinicians could schedule future sessions or assess whether there are other services more appropriate for helping the patient recover. It is also possible that there are factors explaining the continued sick leave of the 'High risk' patients that were not recorded in the present study. Examples may include personal circumstance or characteristics of the work or workplace. As residual depressive symptoms predicted future sick leave, it is also worth noting that the wide-ranging impact of depression on a patient's life may not always be adequately recorded by standard clinical instruments.[37] Future research investigating longitudinal sick leave trajectories in this patient group could thus benefit by a more thorough recording of workplace factors, and perhaps by supplementing quantitative measures with a qualitative approach to get a deeper understanding of the causes of continued sick leave.

Finally, it is worth mentioning that overall, the patient outcomes were generally favourable. In addition to clinical improvements in all groups, of the total sample, 87.2% of patients were fully working with no sick leave 12 months after the end of treatment. This is a marked increase compared with the start of treatment, where 50.1% of the total sample were fully working. Previous estimates indicate that a 5% increase in both the ability to work and productivity at work would mean a threefold return on investment in CMD treatment.[38] The increase in work

participation seen in the present study are well above these estimates.

Strengths and limitations

The longitudinal design of the study is a strength as sick leave and work status has been shown to fluctuate over time.[39] Other studies have used latent growth modeling to analyse longitudinal sick leave data after interventions.[15 16] But to the best of our knowledge, ours' is the first study to also include data from before the intervention in classification of sick leave trajectories, thereby providing a much broader perspective on the issue. The sick leave data in question was objective data collected from registries with no loss to follow-up. In sum, this gave the study a robust classification of sick leave behaviour. The sample was also relatively large for a clinical study and was carried out in a national health service clinic, increasing ecological validity. Finally, a recent article using similar methodology emphasised the importance of developing better interventions grounded in a more thorough understanding of individual variation. Several factors that were highlighted as important areas of future research are included in the present study, including information on previous sick leave, comorbidity, psychological variables including self-efficacy, whether persons received an intervention or not and their specific diagnosis.[40]

Still, the observational design means that this study does not have a control group. For the future, a randomised controlled trial would be helpful to better understand the impact of the intervention on sick leave. As for the background and clinical data used in the post hoc analyses, no data were collected on workplace conditions. We do not know to which degree workplace issues affected outcomes. Future research on the topic would benefit from including workplace factors known to influence the RTW process, such as job sector, size of workplace and support from colleagues and supervisors.[8 40] We also do not have clinical follow-up data for the year following treatment, and do not know if CMD relapse occurred, which could potentially explain the lower RTW rates in the 'High risk' group.'

Our findings should be interpreted with some caution. Overall, the variability in response to work-focused treatment has been demonstrated in numerous studies and is an obvious object of interest.[11] However, the exploratory nature of our study and the inherent differences in welfare structures across countries indicates that one should be careful to generalise. Growth mixture modelling describes classes based on estimated posterior probabilities, which is then used to assign individuals to their most likely class membership. Growth mixture modelling is, therefore, exploratory, and as such more research is needed in replicating the classes that were identified in this study.

Conclusion and future research

The present study is part of a growing body of literature that uses latent growth modelling to analyse longitudinal sick leave data from patients suffering from anxiety and depression. The aim of these studies is to better understand the heterogeneity of these patients, with the hope of developing more tailored, and thus more effective, interventions. Our study found three groups with distinctly different sick leave trajectories. These trajectories were in turn associated with differences in background and clinical characteristics.

The two groups with high sick leave at the start of treatment had more women and higher age than the 'Resilient' group. However, gender and age did not differ between the two groups with high sick leave. Moreover, clinical symptoms of anxiety and depression at the start of treatment were similar across all three groups. Thus, neither background nor clinical characteristics could be used to predict sick leave prognosis at the start of treatment in this study. At the end of treatment, the 'High risk' group showed consistently poorer response than the two other groups on all clinical measures and also had residual depressive symptoms. A question for future research is thus to which degree long-term sick leave prognosis is linked to clinical outcomes at the end of treatment.

Should future research uncover similar groups with similar patterns, all groups identified in this study are of interest, for different reasons. The 'High risk' group can shed light on which factors maintain high sick leave despite treatment, and whether other interventions may be of more use. The 'Recovery' group may reveal factors that influence rapid RTW. The 'Resilient' group mostly avoided sick leave, despite symptoms levels comparable to the other groups. Future investigation of this group can help determine which factors act as protection against sick leave in the first place.

The findings in the present study point to the possibility of improving outcomes for patients on or at risk of sick leave due to anxiety and depression through stratifying and tailoring treatment.

Twitter Kenneth Sandin @kennethsandin

Acknowledgements We would like to thank patients, next of kin and user representatives for their participation and valuable feedback. We would also like to thank therapists and administrative staff at Diakonhjemmet Hospital for their participation and technical assistance.

Contributors KS led the writing of the manuscript and is the principle author of the funding application. FA contributed to data analyses and writing. KO was responsible for management and structuring of data. RGHG is the clinic leader and contributed to study design and intervention development. JSRL contributed to intervention development and writing. SER contributed to design and writing. OH is the project manager and contributed to design, analyses and writing. All authors contributed to, read and approved the final manuscript.

Funding The trial was mainly sponsored by Diakonhjemmet Hospital with additional funding from Southern and Eastern Norway Regional Health Authority, and Stiftelsen Dam (previously Extrastiftelsen), a not-for-profit trust that funds health research and innovation projects.

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Consent obtained directly from patient(s)

Ethics approval The study was approved by the Norwegian Data Protection Authority as health service research. The Norwegian Data Protection Authority has in such cases preapproved the use of data for research, and responsibility for secure and proper data management rests with the treatment provider, which in this case was Diakonhjemmet Hospital. There is therefore no file number or approval number or any similar reference for this type of research in Norway. This arrangement stipulates that written consent for the use of data for research must be obtained from each individual patient, which was done.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Additional data can be accessed on request by email (kenneth.sandin@diakonsyk.no).

ORCID iD
Kenneth Sandin http://orcid.org/0000-0002-1673-9606

REFERENCES

1 Steel Z, Marnane C, Iranpour C, *et al*. The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. *Int J Epidemiol* 2014;43:476–93.
2 OECD. *Mental Health and Work. Switzerland - Mental health and work*. Paris: OECD Publishing, 2014.
3 OECD. *Health at a glance 2019: OECD indicators*. Paris: OECD Publishing, 2019.
4 Vigo D, Thornicroft G, Atun R. Estimating the true global burden of mental illness. *Lancet Psychiatry* 2016;3:171–8.
5 OECD/EU health at a glance: Europe 2018: state of health in the EU cycle 2018.
6 Clark DM. Realizing the mass public benefit of evidence-based psychological therapies: the IAPT program. *Annu Rev Clin Psychol* 2018;14:159–83.
7 OECD. *Mental health and work*. Norway, 2013.
8 Cullen KL, Irvin E, Collie A, *et al*. Effectiveness of workplace interventions in return-to-work for musculoskeletal, pain-related and mental health conditions: an update of the evidence and messages for practitioners. *J Occup Rehabil* 2018;28:1–15.
9 Salomonsson S, Santoft F, Lindsäter E, *et al*. Cognitive–behavioural therapy and return-to-work intervention for patients on sick leave due to common mental disorders: a randomised controlled trial. *Occup Environ Med* 2017;74:905–12.
10 Reme SE, Grasdal AL, Løvvik C, *et al*. Work-focused cognitive–behavioural therapy and individual job support to increase work participation in common mental disorders: a randomised controlled multicentre trial. *Occup Environ Med* 2015;72:745–52.
11 Nigatu YT, Liu Y, Uppal M, *et al*. Interventions for enhancing return to work in individuals with a common mental illness: systematic review and meta-analysis of randomized controlled trials. *Psychol Med* 2016;46:3263–74.
12 Ervasti J, Joensuu M, Pentti J, *et al*. Prognostic factors for return to work after depression-related work disability: a systematic review and meta-analysis. *J Psychiatr Res* 2017;95:28–36.
13 de Vries H, Fishta A, Weikert B, *et al*. Determinants of sickness absence and return to work among employees with common mental disorders: a scoping review. *J Occup Rehabil* 2018;28:393–417.
14 Hultin H, Lindholm C, Malfert M, *et al*. Short-term sick leave and future risk of sickness absence and unemployment - the impact of health status. *BMC Public Health* 2012;12.
15 Hellström L, Madsen T, Nordentoft M, *et al*. Trajectories of return to work among people on sick leave with mood or anxiety disorders: secondary analysis from a randomized controlled trial. *J Occup Rehabil* 2018;28:666–77.
16 Arends I, Almansa J, Stansfeld SA, *et al*. One-year trajectories of mental health and work outcomes post return to work in patients with common mental disorders. *J Affect Disord* 2019;257:263–70.
17 Ubalde-Lopez M, Arends I, Almansa J, *et al*. Beyond return to work: the effect of multimorbidity on work functioning trajectories after sick leave due to common mental disorders. *J Occup Rehabil* 2017;27:210–7.
18 NICE. Depression in adults: recognition and management, 2009. Available: https://www.nice.org.uk/guidance/cg90
19 NICE. *Generalised anxiety disorder in adults, in national clinical guideline number 113*. London, 2019.
20 Dobson KS, Dozois DJA. *Handbook of cognitive behavioral therapies*. 4th edn. Guilford Press, 2019.
21 Wells A. *Metacognitite therapy for anxiety and depression*. New York, NY: Guildford Press, 2011.
22 Lagerveld SE, Blonk RWB, Brenninkmeijer V, *et al*. Work-focused treatment of common mental disorders and return to work: a comparative outcome study. *J Occup Health Psychol* 2012;17:220–34.
23 Gjengedal RGH, Reme SE, Osnes K, *et al*. Work-focused therapy for common mental disorders: a naturalistic study comparing an intervention group with a waitlist control group. *Work* 2021;66:657–67.
24 Norwegian Labour and Welfare Administration. Absence due to illness and sick-pay, 2018. Available: https://www.altinn.no/en/start-and-run-business/working-conditions/print-sickness-and-leaves-of-absence/absence-due-to-illness-and-sick-pay/ [Accessed 28 Jan 2020].
25 WHO. *The International classification of diseases-10 (ICD-10)*. Geneva: World Health Organisation, 1992.
26 Nigatu YT, Liu Y, Uppal M, *et al*. Prognostic factors for return to work of employees with common mental disorders: a meta-analysis of cohort studies. *Soc Psychiatry Psychiatr Epidemiol* 2017;52:1205–15.
27 Beck AT, Steer RA, Brown GK. *Manual for beck depression inventory II (BDI-II)*. San Antonio: Psychology Corporation, 1996.
28 Beck AT, Epstein N, Brown G, *et al*. An inventory for measuring clinical anxiety: psychometric properties. *J Consult Clin Psychol* 1988;56:893–7.
29 Schnittker J, Bacak V. The increasing predictive validity of self-rated health. *PLoS One* 2014;9:e84933.
30 Gjengedal RGH, Lagerveld SE, Reme SE, *et al*. The return-to-work self-efficacy questionnaire (RTW-SE): a validation study of predictive abilities and cut-off values for patients on sick leave due to anxiety or depression. *J Occup Rehabil* 2021;31:664–73.
31 Muthén LK, Muthén BO. *Mplus user's guide*. Seventh ed. Los Angeles, CA: Muthén & Muthén, 1998-2012. https://www.statmodel.com/download/usersguide/Mplus%20user%20guide%20Ver_7_r6_web.pdf
32 Heck RH, Thomas SL. *An introduction to multilevel modeling techniques: MLM and SEM approaches using mPLUS*. Routledge, 2015.
33 Muthén B. Statistical and substantive checking in growth mixture modeling: comment on Bauer and Curran (2003). *Psychol Methods* 2003;8:369–77.
34 Clark SL, Muthén B. Relating latent class analysis results to variables not included in the analysis, 2012. Available: https://www.statmodel.com/download/relatinglca.pdf [Accessed 8 Jan 2020].
35 Cohen J. *Statistical power analyses for the social sciences*. Hillsdale, NJ: Lawrence Erlbauni Associates, 1988.
36 Etuknwa A, Daniels K, Eib C. Sustainable return to work: a systematic review focusing on personal and social factors. *J Occup Rehabil* 2019;29:679–700.
37 Chevance A, Ravaud P, Tomlinson A, *et al*. Identifying outcomes for depression that matter to patients, informal caregivers, and health-care professionals: qualitative content analysis of a large international online survey. *Lancet Psychiatry* 2020;7:692–702.
38 Chisholm D, Sweeny K, Sheehan P, *et al*. Scaling-up treatment of depression and anxiety: a global return on investment analysis. *Lancet Psychiatry* 2016;3:415–24.
39 Øyeflaten I, Lie SA, Ihlebæk CM, *et al*. Multiple transitions in sick leave, disability benefits, and return to work. - A 4-year follow-up of patients participating in a work-related rehabilitation program. *BMC Public Health* 2012;12:748.
40 Spronken M, Brouwers EPM, Vermunt JK, *et al*. Identifying return to work trajectories among employees on sick leave due to mental health problems using latent class transition analysis. *BMJ Open* 2020;10:e032016.

frontiers in Psychology

ORIGINAL RESEARCH
published: 28 October 2021
doi: 10.3389/fpsyg.2021.655151

Self-Reported Health in Patients on or at Risk of Sick Leave Due to Depression and Anxiety: Validity of the EQ-5D

Kenneth Sandin[1,2], Gemma E. Shields[3], Ragne G. H. Gjengedal[1,2], Kåre Osnes[1], Marianne Tranberg Bjørndal[1,2] and Odin Hjemdal[1,2]*

[1]Division of Mental Health and Substance Abuse, Diakonhjemmet Hospital, Oslo, Norway, [2]Department of Psychology, Norwegian University of Science and Technology, Trondheim, Norway, [3]Manchester Centre for Health Economics, Division of Psychology and Mental Health, Faculty of Biology, Medicine and Health, School of Health Sciences, University of Manchester, Manchester, United Kingdom

Objectives: The EQ-5D is a generic, self-report measure of health that is increasingly used in clinical settings, including mental health. The EQ-5D captures health using five dimensions: Mobility, Self-care, Usual activities, Pain/discomfort, and Anxiety/Depression. The validity of the EQ-5D is previously unexplored in patients on or at risk of sick leave due to depression and anxiety. The study's aim was to examine its validity in this group of patients.

Methods: Baseline data were collected from self-report questionnaires in an observational study ($N = 890$) at a Norwegian outpatient-clinic. Participants were adults on or at risk of sick leave due to depression and anxiety who were referred for treatment by general practitioners. The crosswalk methodology was applied to estimate the EQ-5D value. Validity was assessed by comparing responses on the EQ-5D with the Beck Depression Inventory-II (BDI-II), the Beck Anxiety Inventory (BAI), and Subjective Health Complaints (SHC). An ordinal regression model was used to assess known-groups validity. Convergent validity was assessed using Pearson's correlation coefficient, and a multivariate regression model that included sociodemographic characteristics.

Results: The mean EQ-5D value was 0.631, indicating reduced health status compared to "full health" anchored at 1.0, and patients reported moderate levels of depression and anxiety. Ordinal regression indicated that the EQ-5D could discriminate between different levels of symptom severity for depression and anxiety. The EQ-5D value showed significant correlation with the clinical measures; $r = -0.52$ for the BDI-II, $r = -0.49$ for the BAI, and $r = -0.44$ for SHC. The multivariate regression showed that the clinical variables significantly predicted the EQ-5D value, explaining 40.1% of the variance. Depression and anxiety scores were the largest determinants of EQ-5D value, respectively, whilst sick leave, subjective health complaints, and gender made moderate contributions.

Conclusion: The EQ-5D showed indication of validity in patients on or at risk of sick leave due to depression and anxiety in the present study. The EQ-5D value was sensitive to both symptom severity and functional impairment in the form of sick leave. The findings support the EQ-5D as a feasible and relevant measure of health status in these patients.

Keywords: depression, anxiety, sick leave, EQ-5D, validity, self-reported health

OPEN ACCESS

Edited by:
Patrícia M. Pascoal, Universidade Lusófona, Portugal

Reviewed by:
Richard Xu, Hong Kong Polytechnic University, Hong Kong SAR, China
Carolina Da Motta, Center for Research in Neuropsychology and Cognitive and Behavioral Intervention (CINEICC), Portugal

****Correspondence:***
Kenneth Sandin
kenneth.sandin@diakonsyk.no

Specialty section:
This article was submitted to Psychology for Clinical Settings, a section of the journal Frontiers in Psychology

Received: *18 January 2021*
Accepted: *30 September 2021*
Published: *28 October 2021*

Citation:
Sandin K, Shields GE, Gjengedal RGH, Osnes K, Bjørndal MT and Hjemdal O (2021) Self-Reported Health in Patients on or at Risk of Sick Leave Due to Depression and Anxiety: Validity of the EQ-5D. Front. Psychol. 12:655151. doi: 10.3389/fpsyg.2021.655151

INTRODUCTION

Common mental disorders such as depression and anxiety are frequently comorbid, and affect a fifth of the working population at any given time (Lamers et al., 2011; OECD, 2015a). Functional impairment is a key feature of these disorders, which may partially be related to typical symptoms like withdrawal and isolation (OECD, 2015b). Globally, mental illness is a leading cause of disease burden, estimated to account for 32.4% of all years lived with disability (Vigo et al., 2016). Across the EU region mental ill health costs in excess of € 600 billion per year (4.4% of GDP), and the majority of the cost comes from lost productivity through sick leave and disability (OECD/EU, 2018). Employment rates among people with depression and anxiety are 10–15% lower than for the general population (Norstrom et al., 2019). Loss of employment leads to worse health, including an increase in all-cause mortality (Voss et al., 2004), highlighting the impact of these disorders on wider health status.

The cost of mental health problems for individuals and society has led to calls for increased funding for mental health care (Chisholm et al., 2016). But any increase in investment in mental health must be weighed against potential gains of investing in other areas of health. This inherent dilemma of health care prioritisation has led to a growing interest in instruments that can help compare disease burden across patient groups (Drummond et al., 2015). Generic measures of health can help facilitate such comparisons, for instance through generating quality-adjusted life-years (QALYs) used in cost-effectiveness analyses (Devlin et al., 2020).

The most widely used generic measure of patient-reported health is the EQ-5D (Devlin et al., 2020). The instrument was initially developed by an interdisciplinary group with the aim of measuring and valuing health states (Devlin and Brooks, 2017). Expert reviews of existing literature and empirical testing resulted in the publication of a self-report questionnaire that recorded health across five dimensions: Mobility, Self-care, Usual activities, Pain/discomfort, and Anxiety/Depression. These five dimensions were rated on a three-level severity scale from "No problems" to "Moderate problems" to "Extreme problems" (EuroQol, 1990). The EQ-5D has since seen increasing use in clinical research, and its use in appraising health care interventions is recommended by bodies such as the National Institute of Health and Care Excellence in the United Kingdom and the National Institute of Public Health in Norway (NICE, 2018; NIPH, 2019).

Substantial use of the three-level version of the EQ-5D has since led to concerns that the instrument has limited range in capturing variation in health. Studies on both general and clinical populations showed that health problems were not adequately measured, for instance through pronounced ceiling effects (Herdman et al., 2011). This was also the case for mental health populations: reasonable validity was seen in depression, whilst for anxiety disorders the results were more mixed (Sonntag et al., 2013; Brazier et al., 2014). Given the variable performance of the three-level EQ-5D across multiple patient groups, a new version of the EQ-5D, containing five levels of severity, was developed to improve the instruments measurement characteristics (Herdman et al., 2011). Evidence on the validity of the new five-level version is so far limited, and there is thus a need for studies investigating its validity across different patient groups (Mulhern et al., 2014), including mental health patients (Brazier et al., 2014).

To be a valid measure of self-reported health for patients with depression and anxiety, the EQ-5D would need to adequately reflect the wide impact that these disorders have on health. In addition to symptom severity, reduced functioning is a key feature of these disorders (Chevance et al., 2020). This is supported by the high prevalence of sick leave and disability seen among people with depression and anxiety (Norstrom et al., 2019). For this reason, increasing attention is given to work status and sick leave in studies of interventions for depression and anxiety (Cullen et al., 2018; Salomonsson et al., 2018). There is now broad agreement on the importance of helping these patients avoid sick leave, and that success of interventions should also be measured in terms of maintaining employment or returning to work (OECD, 2012). Sensitivity to functional impairment such as sick leave would thus support the validity of the EQ-5D for this patient group, and its usefulness for evaluating interventions.

Research on the previous three-level version of the EQ-5D showed some indication of ability to capture functional impairment in depression and anxiety. One study found that patients with depression in primary care had substantially lower health status as recorded by the EQ-5D. Furthermore, patients in the sample who were on sick leave reported a 10% lower EQ-5D value compared to those who were not on sick leave (Sobocki et al., 2007). A study that used a random sample of 43,589 individuals from the general Swedish population found that sick leave was associated with more problems reported on the three-level EQ-5D (Eriksson et al., 2008). Another Swedish study showed that lower EQ-5D scores predicted an increase in sick leave in patients with musculoskeletal complaints (Stigmar et al., 2013). In Norway, a randomised controlled trial found significantly reduced health status in patients with common mental disorders and work-impairment (Reme et al., 2015).

In addition to symptom severity and reduced functioning, overall health status may also be affected by sociodemographic factors such as age, gender, marital status, and level of education (Braveman and Gottlieb, 2014). These sociodemographic factors have also been shown to be associated with sick leave (Mastekaasa and Melsom, 2014; de Vries et al., 2018). The degree to which these factors impact the health status of patients with depression and anxiety could thus also help shed light on the instrument's validity.

The sensitivity of the five-level version to depression, anxiety, and functional impairment in the form of sick leave has yet to be investigated. Therefore, the aim of the present study was to help address this gap by exploring the construct validity of the EQ-5D for patients on or at risk of sick leave due to depression and/or anxiety. Construct validity is the degree to which an instrument measures the intended construct (Piedmont, 2014). Two types of construct validity were examined:

known-groups and convergent validity. Known-groups validity indicates that an instrument should be able to discriminate between groups known to differ on the variable of interest (Davidson, 2014). Convergent validity indicates that two instruments that measure related constructs should be highly correlated (Chin and Yao, 2014). To assess the validity of the EQ-5D on these counts, the associations with condition-specific measures of depression and anxiety were assessed.

The current study investigated the following hypotheses: that known-groups validity was supported by (1a) patients on or at risk of sick leave due to depression and/or anxiety reporting reduced health status on the EQ-5D compared to the general population norms, and (1b) that the EQ-5D was able to distinguish between patient groups with different levels of depression and anxiety severity. Additionally, that convergent validity was supported by (2a) the EQ-5D showing significant negative correlations with symptom-specific measures, and (2b) health status recorded by the EQ-5D was significantly explained severity of depression and anxiety symptoms, and by sick leave.

MATERIALS AND METHODS

Study Context and Participant Characteristics

Data were collected in a naturalistic observational study at an outpatient clinic at Diakonhjemmet Hospital in Oslo, Norway. The clinic is part of the national specialised mental health care services. This observational study is part of the project "The Norwegian studies of psychological treatment and work (NOR-WORK)." The treatment at the clinic consists of either Metacognitive therapy (MCT) or Cognitive behavioural therapy (CBT), paired with work-focused interventions. The work-focused interventions are aimed at either helping patients remain at work, or in the case of sick leave, return to work (Gjengedal et al., 2020).

The patients who participated in the study were initially referred by their general practitioners for treatment of depression and/or anxiety. At the clinic, patients are initially screened by clinical psychologist for treatment eligibility according to clinical and diagnostic criteria, including by use of the Mini-International Neuropsychiatric Interview (MINI; Lecrubier et al., 1997). As the clinic offers work-focused treatment, the target population consists of patients on or at risk of sick leave due to depression and/or anxiety. That the patients conform to these criteria is firstly assessed through the referral done by the general practitioner, which is evaluated by a clinical psychologist. A second clinician then sees the patient for an assessment session, determining in cooperation with the patient that the patient has clinically relevant symptoms of depression and anxiety, and is experiencing work-related difficulties that could benefit from work-focused treatment. Patients thus had to be adults of working age (age 18–70 years) to participate in the study. Patients were not included in the study if they were suffering from severe mental illness such as bipolar disorder or other psychotic disorders, if they were considered to be at high risk of suicide, or if they were engaging in active substance abuse, or suffered from cluster A or B personality disorder. All patients gave written, informed consent before participation in the study. Data were collected from May 2017 through December 2019, and 890 patients fulfilled the inclusion criteria and consented to participate in the study.

Ethical Considerations

The study is classified as health service research under Norwegian regulation. The Norwegian Data Protection Agency has designated that treatment providers (i.e., hospitals) are responsible for proper data management in such cases. As the information being collected is part of ongoing provision of health care, no further approval is needed beyond consent from the individual patient. Written consent was obtained from all participants. Data collection and security in the present study was managed by Diakonhjemmet Hospital, and approval of data handling was granted by Oslo University Hospital, approval number 2015/15606. The study was carried out in accordance with the principles of the Helsinki declaration.

Measures

Clinical and sociodemographic data were collected from patient journals and from self-report questionnaires filled in by patients at the clinic.

EQ-5D

The EQ-5D questionnaire measures health status using five dimensions (Mobility, Self-care, Usual activities, Pain/discomfort, and Anxiety/depression). Designed to improve upon the three-level version, The EQ-5D-5L scores each dimension on five levels of severity ranging from 1 = "No problems" to 5 = "Extreme problems" (Herdman et al., 2011). For example, on the Anxiety/depression dimension, patients report their responses from 1 ("I am not anxious or depressed") to 5 ("I am extremely anxious or depressed"). The responses on the five dimensions yield the EQ-5D profile, e.g., "11,111" in the case of "No problems" on all dimensions, or "55,555" in the case of "Extreme problems" on all dimensions. There are 3125 (5^5) possible EQ-5D profiles in the five-level version (Devlin et al., 2020).

These health profiles can in turn be converted into a single EQ-5D value using preference based weights. Value sets (or tariffs) are available to support the calculation of the EQ-5D values (Devlin et al., 2020). A study is underway to acquire a value set for Norway, but this is not yet available (NIPH, 2019; Moen Hansen et al., 2020). In such cases it is recommended to use a crosswalk (or mapping) system (NICE, 2019), and this crosswalk system was used in a recent study obtaining Norwegian EQ-5D population norms (Garratt et al., 2021). The same crosswalk methodology was used in the present study when calculating the EQ-5D value. Although negative values are possible, the EQ-5D value ordinarily ranges from 0, which represents death, to 1 which represents full health. A score of 1.000 (i.e., full health) corresponds to a health profile of "11,111,"

i.e., reporting "No problems" across all dimensions. Healthy populations typically report EQ-5D values close to 1; for instance, the study obtaining data from the Norwegian general population found a mean value of 0.805 in a postal survey (Garratt et al., 2021). Note also that when reporting the EQ-5D values it is common to use three decimals (Devlin et al., 2020).

In addition to the EQ-5D profile and the EQ-5D value, the EQ-5D also contains a visual-analogue scale of health, the EQ visual analogue scale (VAS). On the EQ VAS, patients indicate their subjective health state on a visual barometer from a minimum of 0 = worst imaginable health, to a maximum of 100 = best imaginable health (Herdman et al., 2011).

Anxiety

The Beck Anxiety Inventory (BAI) is a self-report measure of anxiety severity over the last week. Examples of items in the BAI are "Heart pounding or racing" and feeling "Nervous." The BAI has 21 such items where these symptoms of anxiety are scored on a scale of severity ranging from 0 to 3, giving total score ranging from 0 to 63. Higher scores indicate more severe symptoms. Recommended scoring of the BAI suggests that 0–15 indicate minimal symptoms, 16–25 moderate symptoms, and 26–63 severe symptoms. In literature reviews, the BAI has shown high internal consistency with an alpha of 0.92 and a test-retest reliability of 0.75 (Beck et al., 1988). In the current study, we report the Omega as this may be a more precise measurement (Peters, 2014). The Omega of the BAI in this study was 0.90.

Depression

The Beck Depression Inventory-II (BDI-II) is a 21 item self-report measure of depression symptom severity over the last 2 weeks. As with the BAI, the BDI has 21 items that are scored on a severity scale ranging from 0 to 3, giving a score range of 0–63. Higher score indicates more severe symptoms (Beck et al., 1996). As an example, the first item asks patients to rate their sadness from 0 ("I do not feel sad") to 3 ("I am so sad or unhappy that I cannot stand it"). A BDI-II score of 0–13 indicates minimal symptoms, 14–19 mild symptoms, 20–28 moderate symptoms, and 29–63 severe symptoms. A review of the literature indicates that the BDI-II is psychometrically sound with internal consistency showing an alpha around 0.90, and a test-retest reliability ranging from 0.73 to 0.96 (Wang and Gorenstein, 2013). In the current study, we found the Omega to be 0.86.

Subjective Health Complaints

The subjective health complaint (SHC) is a self-report questionnaire that contains 29 items measuring subjective health complaints along five factors: musculoskeletal pain, pseudo-neurology, gastrointestinal problems, allergy, and flu. For example, patients are asked to rate pain in arms, leg, or lower back. The aim of the SHC is to provide a simple measure of the most common complaints seen by general practitioners while "avoiding diagnoses and theoretical bias." The severity of complaints on each item is rated on a four point Likert-scale from 0 (no complaints) to 3 (severe complaints) during the last 30 days. The total score of the scale thus ranges from 0 to 87 where higher score indicates worse complaints. Factor analysis of the questionnaire has shown that the greatest proportion of variance of scores is explained by musculoskeletal pain (Eriksen et al., 1999). This measure of subjective health complaints was included as depression and anxiety both have well-known comorbidity with musculoskeletal pain (Bair et al., 2003; Asmundson and Katz, 2009). In the current study, the Omega for the SHC was 0.82.

Sick Leave

Sick leave in the present study was collected from patients *via* self-report questionnaires. For the purpose of the study, we encoded sick leave as a dichotomous variable where patients who were fully working with no social benefits of any kind were coded as "0," and patients on sick leave were coded as "1." We did not collect data on degree of sick leave (e.g., whether a patient was on 100 or 50% sick leave).

Sociodemographic Variables

We included age, gender, cohabitation, and level of education in the analyses to measure relevant sociodemographic aspects of health. Cohabitation was dichotomised as living with partner or living alone. Education level was included as a dichotomous variable, those without higher education were coded as 0, and those with higher education were coded as 1. "Higher education" in this regard refers to any completed degree beyond upper secondary school, i.e., the first 12 years of schooling.

Statistical Analyses

All analyses were carried out using STATA 16.1 (StataCorp, 2019). Assessment of missing data found low incidences throughout the measures. The BDI-II, the BAI, and the EQ-5D, <2% on all items. The SHC had <5% missing on all items. Little's MCAR test was not significant for our dependant variable, the EQ-5D value (χ^2 19.69, DF = 13, $p = 0.103$). This indicates that these values were missing completely at random. Little's MCAR test was significant for the BAI (χ^2 1113.19, DF = 1,040, $p = 0.006$), the BDI-II (χ^2 704.38, DF = 628, $p = 0.018$), and SHC (χ^2 1918.09, DF = 1,566, $p < 0.001$), indicating that these variables were not missing completely at random. Further exploration of missing patterns in the BAI, the BDI-II, and the SHC showed that missing data were explained by the covariate "education," i.e., patients with higher education were more likely to return complete forms. Guidance on EQ-5D data states that general methods used for handling missing data also apply to the EQ-5D (Devlin et al., 2020). Recent guidelines indicate that, as a rule of thumb, it may be a valid approach to ignore missing data if missingness is below 5% (Jakobsen et al., 2017). Although this was the case in the present study, we chose to replace missing data on individual items by weighted means. This method was developed for handling missing data in patients with depression and has shown good precision when used with this patient population (Gale and Hawley, 2001). Data were tested for normality and the clinical variables were found to be within the acceptable range for use of parametric tests as skewness and kurtosis were within −1 to +1 on all measures (Hair et al., 2017).

We defined floor effect for the EQ-5D as proportion of patients reporting "No problems" on all dimensions (i.e., an EQ-5D profile of "11,111"). We defined ceiling effect of the EQ-5D as reporting "Extreme problems" on all dimensions (i.e., an EQ-5D health profile of "55,555"). For the BDI-II, the BAI, and the SHC, floor and ceiling were defined as patients reporting either the lowest or highest possible sum score, that is 0 or 63 for the BAI and the BD-II, and 0 or 87 for the SHC.

It is recommended to present EQ-5D scores with descriptive statistics before presenting any further findings (Devlin et al., 2020). Therefore, we report the proportion of patients that indicated each level of severity for each dimension of the EQ-5D. We also present the mean EQ-5D values and EQ VAS scores by groups based on clinical and sociodemographic characteristics. We then compared the proportion of patients reporting "No problems" to patients reporting any other levels of severity (Devlin et al., 2020). Using the recently published Norwegian population norms (Garratt et al., 2021), we explore known-groups validity by comparing the patients in our study and participants in the general population study who reported "no problems" vs. all other levels of severity. For known-group validity within the sample, we divided the patients into quartiles based on severity of depression and anxiety symptoms as recorded by the BDI-II and BAI scores. Test of Cuzick (1985) for trends, which is a Wilcoxon rank-sum type test for three groups or more, was used to examine if the EQ-5D utility could distinguish between the groups. For the EQ-5D dimensions, we performed an ordinal logistic regression. The severity groups divided by quartiles was used as the dependant variable, and the EQ-5D dimensions were used as predictor variables. The model was tested for multicollinearity. No predictor variable had a variance inflation factor (VIF) higher than 1.38, indicating that multicollinearity was not a problem.

We then explored convergent validity by analysing to which degree the EQ-5D correlated with clinical measures of anxiety, depression, and subjective health complaints (De Vet et al., 2015). The tests were carried out using Pearson's correlation coefficient, a common approach when exploring EQ-5D validity in different patient groups (Byford, 2013; Mulhern et al., 2014). Correlations with the clinical measures were analysed for the EQ-5D values, the EQ VAS, and for all five dimensions. Absolute values larger than +/− 0.50 are considered strong correlations, values between 0.30 and 0.49 moderate, and values beneath 0.30 are considered weak correlations (Fleiss, 1982).

Convergent validity was further explored using a multiple linear regression model. Analyses of multicollinearity were carried out for the explanatory variables in the regression model. No explanatory variable had a VIF higher than 1.58, indicating that multicollinearity was not an issue. The regression model explored the relationship between the EQ-5D values, clinical variables, and sociodemographic variables. We were thus interested in the unique variance contribution of each explanatory variable. Partial correlation was thus calculated for each variable to determine its unique contribution to variance.

RESULTS

Participant Characteristics

Table 1 shows characteristics of patients. The average age was 36.8 years, and there were more females than males (69.6%).

TABLE 1 | Characteristics of patients ($N = 890$).

	n	%	Mean	SD	Median	IQR
Gender						
Female	619	69.55				
Male	271	30.45				
Age, years			36.83	10.45	35	28–45
18–30	313	35.17				
31–40	272	30.56				
41–50	189	21.24				
51–60	102	11.46				
61–70	14	1.57				
Cohabiting/married	535	60.45				
Education						
Primary/Secondary	179	20.48				
Higher education ≤4 yrs	324	37.07				
Higher education >4 yrs	371	42.45				
Employment status						
Sick leave	405	45.66				
Fully working	482	54.34				
Health status						
Anxiety (BAI)			18.74	10.12	18	11–26
Depression (BDI-II)			26.09	8.99	26	20–31
Subjective health (SHC)			23.03	10.17	22	16–29
EQ-5D value			0.631	0.187	0.696	0.501–0.750
EQ VAS			55.7	17.7	60	40–70

BAI, the beck anxiety inventory; BDI-II, the beck depression inventory-II; SHC, subjective health complaints; and IQR, interquartile range.

The majority were currently living with a partner, either as cohabiting or married (60.5%). On average the patients had a high level of education, there were 79.5% who had some form of higher education, whilst 20.5% had primary or secondary education. Almost half the patients were on some form of sick leave (45.7%), whilst the rest (54.3%) were fully working with no form of social benefits. Scores on the BDI-II and the BAI indicated moderate levels of depression and anxiety. The most common primary diagnosis was F41.1 Generalised Anxiety Disorder (16.1%), followed by F32.1 Moderate Depressive Disorder (12.6%). Depression disorders accounted for 46.07% and anxiety disorders accounted for 36.07% of the diagnoses in the sample. The most prevalent diagnoses that were not strictly an anxiety or depression disorder were still diagnoses associated with these disorders: F43.2 Adjustment disorder (7.6%), and F41.2 Mixed anxiety and depressed mood (4.4%). Secondary diagnoses were not recorded in the study. The mean EQ-5D value was 0.631 indicating that these patients perceived their health status as reduced compared to "full health" anchored at 1.0 on this measure (Devlin et al., 2020). The mean score of the EQ-5D VAS was 55.7. Floor and ceiling effects were negligible for all self-report questionnaires. There were 10 patients (1.1%) who reported "No problems" on all EQ-5D dimensions, indicating a ceiling effect, no patients responded "Extreme problems" on all dimensions. No patients reported scores indicating a ceiling effect on the BDI-II, the BAI, or the SHC. Three patients (0.3%) reported scores indicating a floor effect on the BDI-II, one patient (0.1%) on the SHC.

The proportion of participant responses across domains and by level is reported in **Table 2**. More than two thirds of the patients (68.9%) reported "moderate" to "extreme" problems on the Anxiety/depression dimension of the EQ-5D. No participants reported the highest level of severity on the Mobility or Self-care dimensions. These two dimensions also had the largest number of patients reporting "No problems," which was 75.7 and 84.8%, respectively.

The mean EQ-5D values and EQ VAS scores by sociodemographic and clinical characteristics are presented in **Table 3**.

Known-Groups Validity

All participants in the study reported that they had at least some problems on EQ-5D. The highest proportion of problems was seen in the Anxiety/depression, where 96.6% of patients reported some level of problems. A minority of patients reported problems on the Mobility and Self-care; 24.3 and 15.2%, respectively, whilst a large majority of patients in the sample reported some level of problems on Usual activities and Pain/discomfort (78.3 and 76.2%, respectively). **Figure 1** shows a comparison between the proportion of patients in the present study and the proportion of participants in the study obtaining data for the general population norms (Garratt et al., 2021) reporting "No problems" on the sub-scales. The patient cohort reported more problems on all dimensions compared to the norm population.

Test of Cuzick (1985) for trends showed that there was significant difference between the EQ-5D utility scores when patients were divided into quartiles based on severity of depression and anxiety symptoms, $Z = -16.58$, $p = <0.001$. As severity of symptoms increased, health as recorded by the EQ-5D utility decreased (**Table 4**). Similarly, the ordinal logistic regression showed that EQ-5D dimensions significantly predicted the symptom severity groups. All dimensions were significant predictors, while the largest contribution was made by the Anxiety/depression dimension (**Table 5**).

Convergent Validity

Pearson's correlation coefficient showed that the EQ-5D values had a significantly strong negative correlation with the BDI-II depression score, and a moderate negative correlation with the BAI anxiety score and the SHC score. This indicates that for all clinical scales, higher symptom severity correlated with worse reported health status on the EQ-5D value.

For the EQ-5D dimensions, the BDI-II showed a significant moderate correlation with Usual activities and the Anxiety/depression dimension. The BAI showed a significant moderate correlation with Anxiety/depression, whilst SHC showed a significant strong correlation with the Pain/discomfort. Note that lower scores on each dimension indicate better health, i.e., a "1" indicates "no problems." The moderate to strong correlations in these results thus indicate that lower clinical scores signifying better health were associated with better health reported across the EQ-5D dimensions. Pearsons's correlation coefficients are presented in **Table 6**.

The multivariate regression model for convergent validity showed that higher levels of depression and anxiety symptoms, more subjective health complaints, being on sick leave, and being female, all significantly predicted lower EQ-5D value,

TABLE 2 | Distribution of all recorded EQ-5D responses in the patient sample ($N = 890$).

Severity	Mobility		Self-care		Usual activities		Pain/discomfort		Anxiety/depression	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%
1	674	75.7	755	84.8	193	21.7	212	23.8	30	3.4
2	146	16.4	101	11.4	340	38.2	375	42.1	234	26.3
3	43	4.8	17	1.9	237	26.6	229	25.7	363	40.8
4	12	1.4	4	0.5	99	11.1	51	5.7	230	25.8
5					9	1.0	8	0.9	20	2.3

Severity of problems: 1 No problems; 2 Some problems; 3 Moderate problems; 4 Severe problems; and 5 Extreme problems.

TABLE 3 | Mean EQ-5D value and EQ VAS by sociodemographic and clinical characteristics ($N = 890$).

	EQ-5D value	EQ VAS
	Mean (SD)	Mean (SD)
Gender		
Female	0.635 (0.181)	55.5 (17.4)
Male	0.623 (0.201)	56.0 (18.3)
Age		
18–30	0.622 (0.187)	65.9 (18.6)
31–40	0.623 (0.197)	64.6 (18.8)
41–50	0.667 (0.168)	62.9 (19.1)
51–60	0.618 (0.209)	66.2 (19.7)
61–70	0.562 (0.190)	55.7 (16.5)
Cohabitation		
Living with partner	0.633 (0.182)	55.8 (17.6)
Living alone	0.616 (0.189)	55.3 (17.8)
Education		
Primary/secondary	0.586 (0.191)	51.8 (17.5)
Higher education	0.643 (0.184)	56.5 (17.6)
Employment		
Sick leave	0.608 (0.193)	52.0 (18.4)
Fully working	0.651 (0.179)	58.7 (16.4)
Primary diagnosis		
Depression	0.624 (0.184)	53.3 (18.1)
Anxiety	0.625 (0.190)	57.6 (16.9)

Variation in health status by patient characteristics as recorded by the EQ-5D value and the EQ Visual analogue scale (VAS).

i.e., worse health status, $F(8, 876) = 65.24$, $p < 0.000$, $R^2 = 0.401$. We examined the partial correlation for the variables that were significant predictors in the model: For gender it was $r = 0.13$, $p < 0.001$; for BDI-II $r = 0.38$, $p < 0.001$; for BAI $r = 0.28$, $p < 0.001$; for SHC $r = 0.13$, $p < 0.001$; and for sick leave $r = 0.09$, $p < 0.001$. The largest proportion of the variance in the model was thus explained by depression and anxiety, respectively. Results from the regression model is presented in **Table 7**.

DISCUSSION

Our aim was to investigate the validity of the EQ-5D in patients on or at risk of sick leave due to depression and anxiety by examining the health status reported by the EQ-5D. Patients in the study reported poorer health status on the EQ-5D than the normal population. Known-groups validity was supported by both the EQ-5D utility value and the dimensions being able to discriminate between patient groups based on severity of depression and anxiety symptoms. Convergent validity was supported by the EQ-5D showing strong correlations with the BDI-II, and moderate correlation with the BAI and the SHC. Finally, the clinical measures in the study significantly predicted overall health as recorded by the EQ-5D value.

In the current study, all dimensions of the EQ-5D had patients who reported at least some degree of problems. As would be expected in a sample of patients with depression and anxiety diagnoses, highest incidence of problems was reported on the Anxiety/depression dimension. A total of 96.6% of patients reported problems of varying severity on this dimension. The majority of patients also reported problems on the Usual activities and Pain/discomfort dimensions, 78.3 and 76.2%, respectively. This is in line with previous research, which has shown that both functional impairment and pain are prevalent in depression and anxiety (de Heer et al., 2014; McKnight et al., 2016; Hammer-Helmich et al., 2018). A majority of patients reported "No problems" on the Mobility and Self-care dimensions, 75.7 and 84.8%, respectively. We would suggest that this is consistent with the clinical characteristics of the sample. The patients reported moderate levels of depression and/or anxiety, which would not usually entail difficulties with mobility or washing and dressing. Overall, patients in the study reported more problems across all dimensions compared to the respondents in the study that collected the Norwegian EQ-5D norm data (Garratt et al., 2021).

Our findings show that patients experienced reduced health status with a mean EQ-5D value of 0.631 ($SD = 0.187$) and a mean EQ VAS score of 55.6 ($SD = 17.7$). The EQ-5D value was reduced compared to the "full health" anchoring at 1.0, and also compared to the Norwegian study obtaining population norms which found a mean EQ-5D value of 0.805 and a mean EQ VAS of 77.9 in their postal survey (Garratt et al., 2021). A previous study of Norwegian patients with common mental disorders used the three-level version of the EQ-5D, and reported a mean EQ VAS of 65.6 (Reme et al., 2015). The present study seems to add to this finding and indicates that the EQ-5D as expected reports reduced health status in patients with depression and anxiety when compared to a non-clinical population.

When the patients in the study were divided into quartiles based on severity of depression and anxiety symptoms, and the EQ-5D value reported significantly poorer health with increasing symptom severity. Similarly, the ordinal regression model showed that problems reported on all EQ-5D dimensions increased with symptom severity. The largest contribution to the model was made by the Anxiety/depression dimension, which seems to support validity.

The EQ-5D value showed moderate correlations with the measures of anxiety and subjective health complaints, and strong correlation with the depression measure. The patients in the current study had all been referred to specialised care for treatment of depression and anxiety, and we would thus want to see significant correlations with condition-specific measures to support the validity in this patient group. For the five dimensions of the EQ-5D, the BDI-II and the BAI showed moderate correlations with the Anxiety/depression dimension. The BDI-II also showed a moderate correlation with Usual activities, whilst the BAI only had a weak correlation with this dimension. This reflects previous research which indicates that depression has a clear link to functional impairment, whilst the link to anxiety is more ambiguous (McKnight et al., 2016; Hammer-Helmich et al., 2018). For musculoskeletal complaints, the SHC showed a strong correlation with the Pain/discomfort dimension. That the BDI-II and the BAI both showed the strongest correlation with the Anxiety/depression

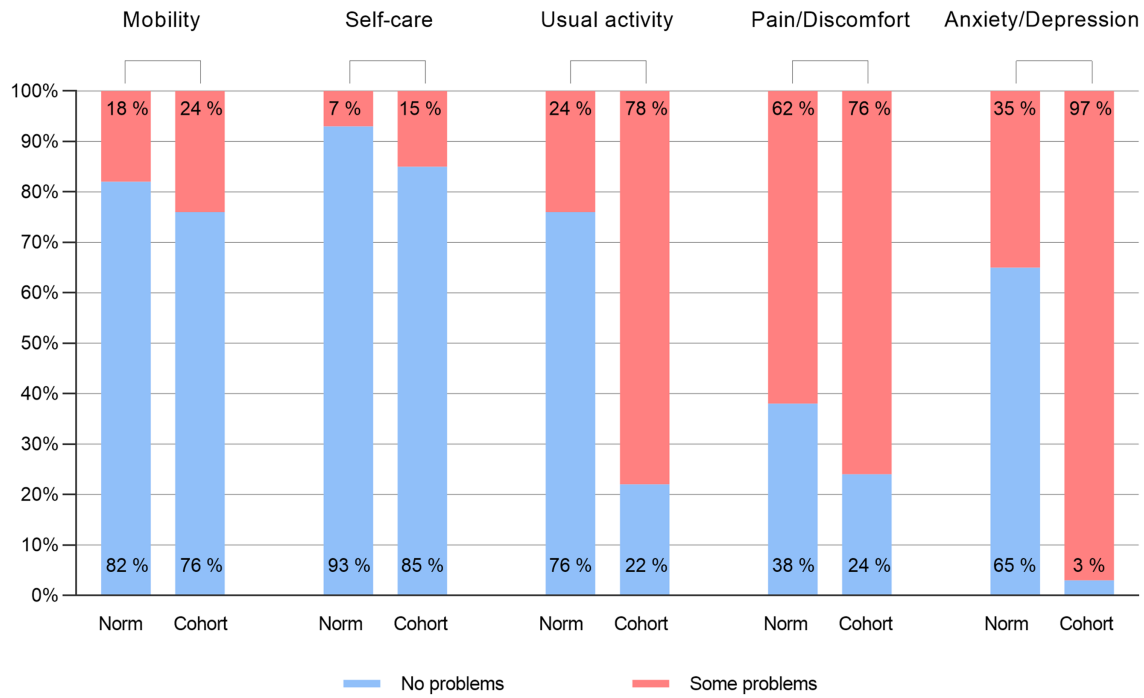


FIGURE 1 | Comparing proportion of EQ-5D "No problems" responses to the Norwegian general population. The "Cohort" columns represent the proportion of patients in the current study who reported "No problems" on each dimension of the EQ-SD compared to the patients who reported any other level of severity. The "Norm" columns represent the same proportion from the study that collected Norwegian population norms; respondents who reported "No problems "on the various dimensions compared to all other levels of severity (Garrat et al., 2021). Percentages are rounded to the nearest integer.

TABLE 4 | Severity of depression and anxiety symptoms by quartiles ($N = 890$).

Severity quartile	*n*	BDI-II	BAI	EQ-5D utility
		Median	Median	Median
1	224	17	9	0.767
2	230	24	13	0.721
3	214	28	20	0.689
4	222	35	29	0.476

BAI, the beck anxiety inventory; BDI-II, the beck depression inventory-II.

dimensions, whilst the SHC showed the strongest correlation with the Pain/discomfort dimensions is consistent with discriminant validity as the dimensions provide a differentiated pattern of correlations. The pattern of correlations between the condition-specific measures and the relevant EQ-5D dimensions thus seems to further support convergent validity as the highest correlations are found between conceptually related dimensions and conditions-specific measures.

The final regression analysis indicated that a substantial part of the EQ-5D value was explained by the condition-specific measures in the study. The only significant socioeconomic variable in the study was gender. This finding is consistent with previous research which has shown a gender gap in self-reported health, where women report poorer health than men (Boerma et al., 2016). Women also have generally higher rates of sick leave than men across developed countries, including Norway, where the present study was conducted (Mastekaasa and Melsom, 2014). Although several explanations have been offered, such as a potential extra burden on women as caretakers in the home, the reasons for this gender gap is still poorly understood (Ostby et al., 2018).

That age, education level and cohabitation did not influence health as recorded by the EQ-5D is perhaps more unexpected. Previous research has shown that these factor tend to influence health status (Braveman and Gottlieb, 2014). This is also true when considering health as recorded by the EQ-5D, where age in particular has been shown to influence self-reported health (Stavem et al., 2018). It may be that the sample was too heterogenous to detect differences in the current study. The patients were quite young with a mean age of 36.8 years, and most had higher education. Perhaps a more diverse selection of patients would produce different results on this count.

The BDI-II, which measures depression, was the largest predictor in the regression model, followed by anxiety measured by the BAI. Furthermore, SHC and sick leave also made significant contributions, indicating that the EQ-5D value was sensitive to musculoskeletal pain and functional impairment. The second regression model explained 40.1% of the variance of the EQ-5D value. The explanatory variables of the model represent a fairly broad clinical evaluation of patients with depression and anxiety. These variables in turn explained a reasonable proportion of the variance of the EQ-5D value. Furthermore, the largest contributors to explained variance were instruments measuring the severity of these patients' primary diagnoses. The results of the regression analyses thus

TABLE 5 | Ordinal logistic regression predicting severity of depression and anxiety symptoms ($N = 890$).

EQ-5D dimension	OR	SE	*z*	*p*	95% CI	Wald	Prob. X^2	Pseudo R^2
Mobility	1.45	0.170	3.17	0.002	1.15–1.83	381.60	0.000	0.158
Self-care	1.41	0.240	2.03	0.042	1.01–1.97			
Activity	1.49	0.117	5.06	<0.001	1.28–1.73			
Pain	1.77	0.143	7.05	<0.001	1.51–2.07			
Anxiety/ depression	2.56	0.227	10.58	<0.001	2.15–3.04			

Severity of depression and anxiety symptoms by quartiles is the dependant variable, and severity of problems reported on the EQ-5D dimensions are the predictor variables.

TABLE 6 | Pearson's correlation coefficient between the EQ-5D and clinical measures ($N = 890$).

	BAI	BDI-II	SHC
EQ-5D value	−0.49	−0.52	−0.44
EQ VAS	−0.27	−0.46	−0.31
EQ-5D dimensions			
Mobility	0.30	0.23	0.23
Self-care	0.15	0.30	0.18
Usual activities	0.24	0.45	0.25
Pain/discomfort	0.38	0.33	0.50
Anxiety/depression	0.42	0.46	0.29

All correlations significant at $p \leq 0.001$. Correlations below 0.3 are considered weak, above 0.4 moderate, above 0.5 are considered strong (Fleiss, 1982). Note that for the EQ-5D value and VAS, higher scores indicate better health. For the dimensions, lower scores indicate better health. BAI, the beck anxiety inventory; BDI-II, the beck depression inventory-II; and SHC, subjective health complaints.

suggest that the variation in the EQ-5D value may be a valid proxy for overall health status as it is associated with the variations in severity of the symptoms reported in this patient group.

Finally, it is worth mentioning that the ceiling and floor effects of the EQ-5D were negligible in the study. This indicates that the EQ-5D seems to have had adequate range in capturing health status for these patients. It is a particular interesting aspect as of the first version of the EQ-5D had difficulties with floor and ceiling effects, including for mental health (Herdman et al., 2011). There were also few missing items, less than 2% on all dimensions. This further suggests that the EQ-5D may be a feasible instrument for these patients.

Implications

The current study suggests that the five-level version of the EQ-5D may be a useful generic measure for evaluating health status in patients on or at risk of sick leave due to depression and anxiety. Including the instrument when assessing burden of disease in these patients may thus facilitate comparison with other patient groups.

Furthermore, functional impairment has emerged as a key component of depression and anxiety. This is especially true of depression, where it also increases risk of relapse (Hardeveld et al., 2010). This functional impairment often manifests as sick leave and work disability, incurring high costs for both individual patients and wider society (OECD, 2015b). This has led to calls for including broader measures of function in evaluating the impact of depression and anxiety on patients (Hardeveld et al., 2010; Chevance et al., 2020). The present study indicates that the EQ-5D may be a valid option to provide a broader measure of health for these patients.

There is also considerable interest in calculating the cost associated with depression and anxiety, and the potential benefits associated with treatment. Multiple studies suggest that better access to treatment would pay for itself, which is one of the key arguments underpinning the UK's Increasing Access to Psychological Therapies (IAPT) programme (Layard and Clark, 2015). These arguments are often based on broad estimates of increased productivity due to beneficial treatment outcomes (Chisholm et al., 2016). The EQ-5D values may help inform such estimates by providing data from clinical trials supporting cost-effectiveness analyses using QALYs. The calculation of QALYs does however depend on adequately measuring health status over time. Future studies should assess this ability of the EQ-5D in mental health.

Strengths and Limitations

This is the first study to investigate the validity of the five-level version of the EQ-5D in a large patient cohort on or at risk of sick leave due to depression and anxiety. The study had a large sample size, and patients were screened and diagnosed in a specialised mental health service clinic, providing high-quality measures of clinical characteristics.

In lack of Norwegian tariffs, the recommendation is to use of the EQ-5D UK value set. Recent research demonstrates that choice of value set can have a significant impact on EQ-5D values produced (Camacho et al., 2018). It is therefore necessary to replicate the present findings using a Norwegian tariff in future studies when these are available.

However, the health profile recorded from the EQ-5D questionnaire would remain the same and therefore many of the conclusions of the study are fixed. The study included a varied, but limited, range of clinical measures and sociodemographic. Further research could explore the correlation between the EQ-5D and other types of measures, such as capability measures, and wider determinants of health. The relationship between type of sick leave, and the role of the welfare system is worth considering. The current study did not include information on degree or

TABLE 7 | Regression analysis predicting the EQ-5D value ($N = 890$).

	Coef.	*SE*	*T*	*P*	95% CI	Beta coef.	*F*	R^2	Adj. R^2
Age	0.0004	0.0005	0.77	0.439	−0.0006 to 0.0014	0.0216	73.34	0.406	0.401
Gender	−0.0423	0.0112	−3.78	<0.001	−0.0643 to −0.0203	−0.1047			
Cohabitation	0.0086	0.0102	0.84	0.399	−0.0114 to 0.0287	0.0227			
Education	0.0209	0.0128	1.73	0.085	−0.0030 to 0.0472	0.0481			
BDI-II	−0.0076	0.0006	−11.96	<0.001	−0.0087 to −0.0062	−0.3621			
BAI	−0.0052	0.0006	−8.40	<0.001	−0.0064 to −0.0040	−0.2825			
SHC	−0.0024	0.0006	−3.81	<0.001	−0.0037 to −0.0012	−0.1339			
Sick leave	−0.0274	0.0101	−2.76	0.006	−0.0478 to −0.0080	−0.0749			

EQ-5D value is the dependant variable and demographic characteristics, BDI-II, BAI, SHC and sick leave are predictors. BAI, the beck anxiety inventory; BDI-II, the beck depression inventory-II; and SHC, subjective health complaints.

duration of sick leave. It is also worth mentioning that the Norwegian welfare system is relatively generous compared to many other countries (Andreß and Heien, 2001). Employees receive compensation equivalent to 100% of their salary from the first day of sick leave. This is covered by employers for the first 16 days, and then subsequently by the state welfare system for up to a year. It is possible that the relationship between health status recorded by the EQ-5D, and sick leave could vary by country, given the substantial variation between national welfare systems and conditions of sick leave.

The current study also included more women than men. Although this may raise questions of generalisability, the gender distribution reflects the prevalence patterns of mental disorders (Boyd et al., 2015). We also used a binary approach to gender, and we thus do not know whether the study may have included non-binary participants. Finally, the clinical validity explored in the present study is an important psychometric property of an instrument, but it is not the same as clinical responsiveness (Payakachat et al., 2015). Future research should examine to which degree the EQ-5D is responsive to change in health status in mental health patients, for instance in the shape of recovery from depression and anxiety.

CONCLUSION

In the present study, the EQ-5D showed evidence of construct validity in patients on or at risk of sick leave due to depression and anxiety. The EQ-5D value was sensitive to both clinical symptoms and to functional impairment in the form of sick leave. The findings thus support the validity of the EQ-5D as a measure of health status for these patients. These results need to be replicated in other samples and different sociodemographic settings. However, the current findings suggests that the EQ-5D is a feasible instrument when evaluating health status of patients of patients with depression and anxiety.

DATA AVAILABILITY STATEMENT

The data used in the article is not readily available as the participants have not consented to distribution beyond the use in the study. Requests to access the datasets should be directed to Kenneth Sandin, kenneth.sandin@diakonsyk.no.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

KS led the writing of the manuscript and is the principal author of the funding application. GS contributed to conceptualisation, data analyses, and writing. RG is the clinic leader and contributed to conceptualisation and writing. KO was responsible for management and structuring of data. MB contributed to data structuring and writing. OH is the project manager and contributed to design, analyses, and writing. All authors contributed to the article and approved the submitted version.

FUNDING

The study was mainly sponsored by Diakonhjemmet Hospital with additional funding from Southern and Eastern Norway Regional Health Authority, and Stiftelsen Dam (previously Extrastiftelsen), a not for profit trust that funds health research and innovation projects.

ACKNOWLEDGMENTS

We would like to thank patients, next of kin, and user representatives for their participation and valuable feedback. We would also like to thank therapists and administrative staff at Diakonhjemmet Hospital for their participation and technical assistance. Finally, the reviewers and editor for their valuable contribution in improving the manuscript.

REFERENCES

Andreß, H. J., and Heien, T. (2001). Four worlds of welfare state attitudes? A comparison of Germany, Norway, and the United States. *Eur. Sociol. Rev.* 17, 337–356. doi: 10.1093/esr/17.4.337

Asmundson, G. J., and Katz, J. (2009). Understanding the co-occurrence of anxiety disorders and chronic pain: state-of-the-art. *Depress. Anxiety* 26, 888–901. doi: 10.1002/da.20600

Bair, M., Robinsons, R. L., Katon, W., and Kroenke, K. (2003). Depression and pain comorbidity: a literature review. *Arch. Intern. Med.* 163, 2433–2445. doi: 10.1001/archinte.163.20.2433

Beck, A. T., Epstein, N., Brown, G. K., and Steer, R. A. (1988). An inventory for measuring clinical anxiety: psychometric properties. *J. Consult. Clin. Psychol.* 56, 893–897. doi: 10.1037/0022-006X.56.6.893

Beck, A. T., Steer, R. A., and Brown, G. K. (1996). *Manual for Beck Depression Inventory II (BDI-II)*. San Antonio: Psychology Corporation.

Boerma, T., Hosseinpoor, A. R., Verdes, E., and Chatterji, S. (2016). A global assessment of the gender gap in self-reported health with survey data from 59 countries. *BMC Public Health* 16:675. doi: 10.1186/s12889-016-3352-y

Boyd, A., Van de Velde, S., Vilagut, G., de Graaf, R., O'Neill, S., Florescu, S., et al. (2015). Gender differences in mental disorders and suicidality in Europe: results from a large cross-sectional population-based study. *J. Affect. Disord.* 1, 245–254. doi: 10.1016/j.jad.2014.11.002

Braveman, P., and Gottlieb, L. (2014). The social determinants of health: it's time to consider the causes of the causes. *Public Health Rep.* 129, 19–31. doi: 10.1177/00333549141291S206

Brazier, J., Connell, J., Papaioannou, D., Mukuria, C., Mulhern, B., Peasgood, T., et al. (2014). A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. *Health Technol. Assess.* 18, 1–188. doi: 10.3310/hta18340

Byford, S. (2013). The validity and responsiveness of the EQ-5D measure of health-related quality of life in an adolescent population with persistent major depression. *J. Ment. Health* 22, 101–110. doi: 10.3109/09638237.2013.779366

Camacho, E. M., Shields, G., Lovell, K., Coventry, P. A., Morrison, A. P., and Davies, L. M. (2018). A (five-)level playing field for mental health conditions?: exploratory analysis of EQ-5D-5L-derived utility values. *Qual. Life Res.* 27, 717–724. doi: 10.1007/s11136-017-1768-1

Chevance, A., Ravaud, P., Tomlinson, A., Le Berre, C., Teufer, B., Touboul, S., et al. (2020). Identifying outcomes for depression that matter to patients, informal caregivers, and health-care professionals: qualitative content analysis of a large international online survey. *Lancet Psychiatry* 7, 692–702. doi: 10.1016/S2215-0366(20)30191-7

Chin, C. L., and Yao, G. (2014). "Convergent Validity," in *Encyclopedia of Quality of Life and Well-Being Research*. ed. A. Michalos (Dordrecht: Springer)

Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., et al. (2016). Scaling-up treatment of depression and anxiety: a global return on investment analysis. *Lancet Psychiatry* 3, 415–424. doi: 10.1016/S2215-0366(16)30024-4

Cullen, K., Irvin, E., Collie, A., Clay, F., Gensby, U., Jennings, P., et al. (2018). Effectiveness of workplace interventions in return-to-work for musculoskeletal, pain-related and mental health conditions: an update of the evidence and messages for practitioners. *J. Occup. Rehabil.* 28, 1–15. doi: 10.1007/s10926-016-9690-x

Cuzick, J. (1985). A wilcoxon-type test for trends. *Stat. Med.* 4, 87–90. doi: 10.1002/sim.4780040112

Davidson, M. (2014). "Known-group validity," in *Encyclopedia of Quality of Life and Well-Being Research*. ed. A. Michalos (Dordrecht: Springer)

de Heer, E. W., Gerrits, M. M. J. G., Beekman, A. T. F., Dekker, J., van Marwijk, H. W. J., de Waal, M. W. M., et al. (2014). The association of depression and anxiety with pain: a study from NESDA. *PLoS One* 9:e106907. doi: 10.1371/journal.pone.0106907

De Vet, H. C. W., Terwee, C. B., Mokkink, L. B., and Knol, D. L. (2015). *Measurement in Medicine. 5th Edn*. Cambridge: Cambridge University Press.

de Vries, H., Fishta, A., Weikert, B., Rodriguez Sanchez, A., and Wegewitz, U. (2018). Determinants of sickness absence and return to work among employees with common mental disorders: a scoping review. *J. Occup. Rehabil.* 28, 393–417. doi: 10.1007/s10926-017-9730-1

Devlin, N. J., and Brooks, R. (2017). EQ-5D and the EuroQol group: past, present and future. *Appl. Health Econ. Health Policy* 15, 127–137. doi: 10.1007/s40258-017-0310-5

Devlin, N. J., Parkin, D., and Janssen, B. (2020). *Methods for Analysing and Reporting EQ-5D Data*. Springer International Publishing.

Drummond, M. F., Schulper, M. J., Claxton, K., Stoddart, G. L., and Torrance, G. W. (2015). *Methods for the Economic Evaluation of Health Care Programmes. 4th Edn*. Oxford: Oxford University Press.

Eriksen, H. R., Ihlebaek, C., and Ursin, H. A. (1999). A scoring system for subjective health complaints (SHC). *Scand. J. Public Health* 27, 63–72. doi: 10.1177/14034948990270010401

Eriksson, H. G., von Celsing, A. S., Wahlstrom, R., Janson, L., Zander, V., and Wallman, T. (2008). Sickness absence and self-reported health a population-based study of 43,600 individuals in central Sweden. *BMC Public Health* 8:426. doi: 10.1186/1471-2458-8-426

EuroQol (1990). A new facility for the measurement of health-related quality of life. *Health Policy* 16, 199–208. doi: 10.1016/0168-8510(90)90421-9

Fleiss, J. L. (1982). *Statistical Methods for Rates and Proportions*. New York: Wiley & Sons.

Gale, T., and Hawley, C. (2001). A model for handling missing items on two depression rating scales. *Int. Clin. Psychopharmacol.* 16, 205–214. doi: 10.1097/00004850-200107000-00004

Garratt, A. M., Hansen, T. M., Augestad, L. A., Rand, K., and Stavem, K. (2021). Norwegian population norms for the EQ-5D-5L: results from a general population survey. *Qual. Life Res.* 1–10. doi: 10.1007/s11136-021-02938-7

Gjengedal, R. G. H., Reme, S., Osnes, K., Lagerveld, S., Blonk, R., Sandin, K., et al. (2020). Work-focused treatment for common mental disorders: an observational study comparing an intervention group with a waiting list control group. *Work* 6, 657–667. doi: 10.3233/WOR-203208

Hair, J. F., Hult, G. T. M., Ringle, C. M., and Sarstedt, M. (2017). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). 2nd Edn*. Thousand Oaks, CA: Sage.

Hammer-Helmich, L., Haro, J. M., Jonsson, B., Melac, A. T., Di Nicola, S., Chollet, J., et al. (2018). Functional impairment in patients with major depressive disorder: the 2-year PERFORM study. *Neuropsychiatr. Dis. Treat.* 14, 239–249. doi: 10.2147/NDT.S146098

Hardeveld, F., Spijker, J., De Graaf, R., Nolen, W. A., and Beekman, A. T. (2010). Prevalence and predictors of recurrence of major depressive disorder in the adult population. *Acta Psychiatr. Scand.* 122, 184–191. doi: 10.1111/j.1600-0447.2009.01519.x

Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., et al. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual. Life Res.* 20, 1727–1736. doi: 10.1007/s11136-011-9903-x

Jakobsen, J. C., Gluud, C., Wetterslev, J., and Winkel, P. (2017). When and how should multiple imputation be used for handling missing data in randomised clinical trials—a practical guide with flowcharts. *BMC Med. Res. Methodol.* 17:162. doi: 10.1186/s12874-017-0442-1

Lamers, F., van Oppen, P., Comijs, H. C., Smit, J. H., Spinhoven, P., van Balkom, A. J. L. M., et al. (2011). Comorbitiy patterns of anxiety and depressive disorders in a large cohort study: the Netherlands study of depression and anxity (NESDA). *J. Clin. Psychiatry* 72, 341–348. doi: 10.4088/JCP.10m06176blu

Layard, R., and Clark, D. M. (2015). Why more psychological therapy would cost nothing. *Front. Psychol.* 6:1713. doi: 10.3389/fpsyg.2015.01713

Lecrubier, T., Sheehan, D. V., Weiller, E., Amorim, P., Bonora, I., Sheehan, K. H., et al. (1997). The mini international neuropsychiatric interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *Eur. Psychiatry* 12, 224–231. doi: 10.1016/S0924-9338(97)83296-8

Mastekaasa, A., and Melsom, A. M. (2014). Occupational segregation and gender differences in sickness absence: evidence from 17 European countries. *Eur. Sociol. Rev.* 30, 582–594. doi: 10.1093/esr/jcu059

McKnight, P. E., Monfort, S. S., Kashdan, T. B., Blalock, D. V., and Calton, J. M. (2016). Anxiety symptoms and functional impairment: a systematic review of the correlation between the two measures. *Clin. Psychol. Rev.* 45, 115–130. doi: 10.1016/j.cpr.2015.10.005

Moen Hansen, T., Helland, Y., Augestad, L. A., Rand, K., Stavem, K., and Garratt, A. (2020). Elicitation of Norwegian EQ-5D-5L values for hypothetical

and experience-based health states based on the EuroQol valuation technology (EQ-VT) protocol. *BMJ Open* 10:e034683. doi: 10.1136/bmjopen-2019-034683

Mulhern, B., Mukuria, C., Barkham, M., Knapp, M., Byford, S., Soeteman, D., et al. (2014). Using generic preference-based measures in mental health: psychometric validity of the EQ-5D and SF-6D. *Br. J. Psychiatry* 205, 236–243. doi: 10.1192/bjp.bp.112.122283

NICE (2018). *The National Institute of Health and Care Excellence—Guide to the Processes of Technology Appraisal NICE Process and Methods Guides*. London: NICE.

NICE (2019). The National Institute of Health and Care Excellence—Position statement on use of the EQ-5D-5L value set for England. Available at: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/technology-appraisal-guidance/eq-5d-5l (Accessed Jan 10, 2020).

NIPH (2019). National Institue of Public Health—Fremskaffing av EQ-5D vekter og normative data for helseøkonomiske evalueringer—prosjektbeskrivelse. (English: Acquisition of EQ-5D weights and normative data for health economic evaluations—project description). Available at: https://www.fhi.no/cristin-prosjekter/aktiv/fremskaffing-av-eq-5d-vekter-og-normative-data-for-helseokonomiske-evalueri/ (Accessed September 27, 2020).

Norstrom, F., Waenerlund, A. K., Lindholm, L., Nygren, R., Sahlen, K. G., and Brydsten, A. (2019). Does unemployment contribute to poorer health-related quality of life among Swedish adults? *BMC Public Health* 19:457. doi: 10.1186/s12889-019-6825-y

OECD (2012). *Sick on the Job? Myths and Realities About Mental Health and Work*. Paris: OECD Publishing.

OECD (2015a). *Fit Mind, Fit Job: From Evidence to Practice in Mental Health and Work*. Paris: OECD Publishing.

OECD (2015b). *Mental Health and Work: Fit Mind, Fit Job: From Evidence to Practice in Mental Health and Work*. Paris: OECD Publishing.

OECD/EU (2018). *Health at a Glance: Europe 2018: State of Health in the EU Cycle*. Paris: OECD Publishing.

Ostby, K. A., Mykletun, A., and Nilsen, W. (2018). Explaining the gender gap in sickness absence. *Occup. Med.* 68, 320–326. doi: 10.1093/occmed/kqy062

Payakachat, N., Ali, M. M., and Tilford, J. M. (2015). Can the EQ-5D detect meaningful change? A systematic review. *Pharmacoeconomics* 33, 1137–1154. doi: 10.1007/s40273-015-0295-6

Peters, G. J. Y. (2014). The alpha and the omega of scale reliability and validity. Eur. Health Psychol. 16. [Preprint]. doi: 10.31234/osf.io/h47fv

Piedmont, R. L. (2014). "Construct validity," in *Encyclopedia of Quality of Life and Well-Being Research*. ed. A. Michalos (Dordecht: Springer)

Reme, S. E., Grasdal, A. L., Lovvik, C., Lie, S. A., and Overland, S. (2015). Work-focused cognitive–behavioural therapy and individual job support to increase work participation in common mental disorders: a randomised controlled multicentre trial. *Occup. Environ. Med.* 72, 745–752. doi: 10.1136/oemed-2014-102700

Salomonsson, S., Hedman-Lagerlof, E., and Ost, L. G. (2018). Sickness absence: a systematic review and meta-analysis of psychological treatments for individuals on sick leave due to common mental disorders. *Psychol. Med.* 48, 1954–1965. doi: 10.1017/S0033291718000065

Sobocki, P., Ekman, M., Ågren, H., Krakau, I., Runeson, B., Mårtensson, B., et al. (2007). Health-related quality of life measured with EQ-5D in patients treated for depression in primary care. *Value Health* 10, 153–160. doi: 10.1111/j.1524-4733.2006.00162.x

Sonntag, M., Konnopka, A., Leichsenring, F., Salzer, S., Beutel, M. E., Herpertz, S., et al. (2013). Reliability, validity and responsiveness of the EQ-5D in assessing and valuing health status in patients with social phobia. *Health Qual. Life Outcomes* 11:215. doi: 10.1186/1477-7525-11-215

StataCorp. (2019). *Stata Statistical Software: Release 16*. College Station, TX.: StataCorp LLC.

Stavem, K., Augestad, L. A., Kristiansen, I. S., and Rand, K. (2018). General population norms for the EQ-5D-3 L in Norway: comparison of postal and web surveys. *Health Qual. Life Outcomes* 16:204. doi: 10.1186/s12955-018-1029-1

Stigmar, K. G. E., Petersson, I. F., Jöud, A., and Grahn, B. E. M. (2013). Promoting work ability in a structured national rehabilitation program in patients with musculoskeletal disorders: outcomes and predictors in a prospective cohort study. *BMC Musculoskelet. Disord.* 14:57. doi: 10.1186/1471-2474-14-57

Vigo, D., Thornicroft, G., and Atun, R. (2016). Estimating the true global burden of mental illness. *Lancet Psychiatry* 3, 171–178. doi: 10.1016/S2215-0366(15)00505-2

Voss, M., Nylen, L., Floderus, B., Diderichsen, F., and Terry, P. D. (2004). Unemployment and early cause-specific mortality: a study based on the Swedish twin registry. *Am. J. Public Health* 94, 2155–2161. doi: 10.2105/AJPH.94.12.2155

Wang, Y. P., and Gorenstein, C. (2013). Psychometric properties of the beck depression inventory-II: a comprehensive review. *Braz. J. Psychiatry* 35, 416–431. doi: 10.1590/1516-4446-2012-1048

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Health and Quality of Life Outcomes

RESEARCH Open Access

Responsiveness to change in health status of the EQ-5D in patients treated for depression and anxiety

Kenneth Sandin[1,2*], Gemma Shields[3], Ragne G.H. Gjengedal[1,2], Kåre Osnes[1], Marianne T. Bjørndal[1,2], Silje E. Reme[1,4] and Odin Hjemdal[1,2]

Abstract

Background The EQ-5D is a commonly used generic measure of health but evidence on its responsiveness to change in mental health is limited. This study aimed to explore the responsiveness of the five-level version of the instrument, the EQ-5D-5 L, in patients receiving treatment for depression and anxiety.

Methods Patient data ($N = 416$) were collected at baseline and at end of treatment in an observational study in a Norwegian outpatient clinic. Patients were adults of working age (18–69 years) and received protocol-based metacognitive or cognitive therapy for depression or anxiety according to diagnosis. Responsiveness in the EQ-5D was compared to change in the Beck Depression Inventory-II (BDI-II) and the Beck Anxiety Inventory (BAI). Effect sizes (Cohen's *d*), Standardised response mean (SRM), and Pearson's correlation were calculated. Patients were classified as "Recovered", "Improved", or "Unchanged" during treatment using the BDI-II and the BAI. ROC analyses determined whether the EQ-5D could correctly classify patient outcomes.

Results Effect sizes were large for the BAI, the BDI-II, the EQ-5D value and the EQ VAS, ranging from $d = 1.07$ to $d = 1.84$. SRM were also large (0.93-1.67). Pearson's correlation showed strong agreement between change scores of the EQ-5D value and the BDI-II (r_s -0.54) and moderate between the EQ-5D value and the BAI (r_s -0.43). The EQ-5D consistently identified "Recovered" patients versus "Improved" or "Unchanged" in the ROC analyses with AUROC ranging from 0.72 to 0.84.

Conclusion The EQ-5D showed good agreement with self-reported symptom change in depression and anxiety, and correctly identified recovered patients. These findings indicate that the EQ-5D may be appropriately responsive to change in patients with depression and anxiety disorders, although replication in other clinical samples is needed.

Plain English Summary

The EQ-5D is a questionnaire that people fill in to report their subjective health. It is often used in clinics or hospitals to better understand how patients are affected by their illnesses, and if their health improves after treatment. For this information to be trustworthy, we need to verify how accurately the EQ-5D measures health for the particular patients we want to use it with. This is often done by comparing EQ-5D scores with scores from

*Correspondence:
Kenneth Sandin
ksandin@gmail.com

Full list of author information is available at the end of the article

other questionnaires. For example, if we want to use the EQ-5D with a group of patients with depression, we compare the scores of the EQ-5D with scores from questionnaires that are commonly used to measure depression symptoms.

In this study, we compared the scores of the EQ-5D with scores from questionnaires measuring symptoms of depression and anxiety. Their performances were similar, and the EQ-5D scores could also correctly identify which patients had recovered during treatment. This implies that the EQ-5D can be a useful tool for understanding the impact of depression and anxiety and can help in decision-making regarding these patients.

Keywords Self-rated health, Depression, Anxiety, EQ-5D-5L, Responsiveness

Introduction

One out of every two people will experience a mental health problem during their lifetime and mental ill health is a leading cause of global disease burden [1]. Between 2010 and 2030, mental illness is projected to cost $ 16.1 trillion worldwide, putting it on par with cardiovascular disease [2]. Depression and anxiety disorders account for 40.5% and 14.6% of the disability-adjusted life-years that are due to mental illness, making them the most costly mental health problems [3]. This substantial burden may still be underestimated [4], in part because of the wide ranging effects these disorders have on health and functioning [5].

At the recommendation of decision-making bodies such as the National institute of Health and Care excellence (NICE), generic measures are increasingly used to capture health status [6]. Mental disorders like depression and anxiety have broad, negative impact on quality of life and wellbeing that may not be adequately reflected by condition-specific measures [5, 7]. Generic measures may thus be a valuable supplement to measures of primary symptoms as they capture a broader measure of health. These instruments can also be used to compare burden of disease and impact of interventions between different patient groups, such as in cost-benefit analyses, making them useful tools for decision-makers, researchers, and clinicians [8]. To adequately fill this role, it must be demonstrated that the generic measure in question can accurately capture health status in the relevant patient population.

One of the most commonly used generic measures of health-related quality of life is the EQ-5D [8]. The EQ-5D records health status across five dimensions: Mobility, Self-care, Usual activities, Pain / discomfort, and Anxiety / depression [9]. The previous version of the EQ-5D, the EQ-5D-3L, used three levels of severity and showed good psychometric properties in depression, but mixed results in anxiety disorders [10]. A recent review evaluated the properties of the newer five-level version of the EQ-5D, the EQ-5D-5L, across multiple patient groups [11]. These and other studies of patients with mental health problems have shown moderate to good correlation between condition-specific measures and the EQ-5D-5L in cross-sectional designs [11–15].

These studies did not include data on responsiveness [11]. Responsiveness is often defined as an instruments ability to detect clinically significant change over time [10, 16]. Two criteria have been suggested for defining what constitutes "clinically significant change": That the magnitude of change be statistically reliable, and that patients end up in a clinical range that renders them indistinguishable from the normal population, i.e. they have recovered [17]. Responsiveness according to these criteria is not a fixed parameter, but will likely vary according to populations and context [18]. This makes it necessary to investigate responsiveness across multiple patient groups. One study did find reasonable validity and moderate responsiveness in anxiety on the EQ-5D-3L [19]. But only a few studies have examined this aspect of the five-level EQ-5D-5L in depression and anxiety [11].

One study found that using only the Anxiety/depression dimension of the EQ-5D-5L did not adequately capture responsiveness in anxiety and depression for patients treated in a general internal medicine ward [20]. Another study found that the EQ-5D-5L could adequately screen for depression and anxiety by distinguishing between severity levels in patients with type 2 diabetes. This was true of both the Anxiety / depression dimension and the EQ-5D value [21]. However, this was a cross-sectional design, and the ability of the EQ-5D-5 L to detect change in severity over time in patients with depression and anxiety is not established, and was specifically targeted by a review of the literature as a future research priority [11]. Investigating this aspect of the EQ-5D-5 L is imperative in establishing whether it is a valid tool for capturing the health status of these patients.

To our knowledge, ours is the first study to examine the responsiveness of the EQ-5D-5L in patients treated for depression and anxiety as their primary diagnoses. In line with recommendations and methodology used in previous studies, we explored responsiveness of the EQ-5D-5L by comparing change from start to end of intervention with change in condition-specific measures [17, 20, 21] The aim of the study was thus to test the following hypotheses: (1) that the EQ-5D-5L shows similar range in effect size and an at least moderate correlation with change scores in condition-specific measures, and

(2) that the EQ-5D-5L can identify patients classified as "Recovered" by condition-specific measures at end of treatment.

Methods

Study context

Data were collected in a naturalistic observational study that ran from May 2017 – March 2020 at the Department of Mental Health and Substance Abuse, Diakonhjemmet Hospital in Oslo, Norway. The clinic is part of the national health service, and the study is part of the project "The Norwegian studies of psychological treatments and work (NOR-WORK)". Patients are referred by their general practitioners for treatment of depression and anxiety. Patients at the clinic are generally of working age, and previous research has shown that on average, half the patients are on sick leave due to depression or anxiety at baseline [22]. They are then screened by a clinical psychologist using anamnestic information, the Beck Depression Inventory-II (BDI-II), the Beck Anxiety Inventory (BAI), and the MINI-International Neuropsychiatric Interview [23–25]. Patients are diagnosed during the screening in accordance with the International Classification of Diseases 10 (ICD-10) [26]. Inclusion criteria for the present study were that the patient was an adult of working age (18–70 years) with clinically significant levels of depression and anxiety operationalised as follows: Patients with a primary depression diagnosis had to have a minimum score of 14 on the BDI-II, and patients with a primary anxiety diagnosis had to have a minimum score of 16 on the Beck Anxiety Inventory BAI. In addition to primary depression or anxiety diagnoses, patients with adjustment disorder and mixed anxiety and depression were included in the study. Adjustment disorder is sometimes referred to as "situational depression", underlining its close relationship with depressive disorders [26]. Similarly, patients with a mixed anxiety and depressive disorder were included as the diagnosis is comprised of symptoms of anxiety and depression.

Exclusion criteria were severe mental illness such as bipolar disorder, high risk of suicide, engaging in active substance abuse, or suffering from cluster A or B personality disorder. Patients scoring below clinical thresholds for depression and anxiety on the BDI and BAI at baseline were excluded from the study. All patients who signed a written consent form and completed treatment, including filling in questionnaires at baseline and at end of treatment, were included (N=416). The current study thus focused on patients who completed treatment.

Patients received either Metacognitive therapy (MCT) or Cognitive behavioural therapy (CBT) according to diagnose-specific manuals [27, 28], and average duration of treatment was 10.11 sessions (SD 3.93). Previous research has shown that half the patients are on sick leave when referred, and treatment thus also includes interventions aimed at helping patients return to work [29].

Instruments

Clinical and sociodemographic data were collected at baseline and end of treatment from patient journals and from self-report questionnaires.

The EQ-5D-5L: The EQ-5D-5L questionnaire firstly asks respondents to rate their current health on five dimensions: Mobility, Self-Care, Usual activities, Pain / discomfort, and Anxiety / depression on a severity scale from 1 ("No problems") to 5 ("Severe problems"). The combined severity ratings give an EQ-5D profile, e.g. "11111" in the case of "No problem" on all five dimensions. This health profile can be converted to the EQ-5D value using preference-based weights. A value of 0.00 indicates death and 1.00 indicates perfect health. The EQ-5D value can be used to calculate quality-adjusted life-years (QALYs), i.e. a score of 1.00 for one year equals one QALY. The preference-based weights used to convert responses to EQ-5D values are often referred to as "value sets". A study is underway, but there is currently no Norwegian value set [30]. This study used the crosswalk system recommended by NICE for converting EQ-5D profiles to EQ-5D values [31, 32]. For the EQ-5D value, healthy people generally report scores close to 1.0. In a recent survey of the Norwegian general population, the mean EQ-5D value in a postal survey was 0.848 [33].

The second part of the EQ-5D-5L asks patients to rate their health on a 20 cm visual analogue scale (VAS) where the bottom ("0") indicates worst imaginable health, and the top ("100") indicates best imaginable health. Although it is related to the EQ-5D profile and the value scores, it does not measure the same construct. For instance, the EQ VAS score has been shown to decline with age even for people whose EQ-5D profile show no problems ("11111") [8].

The Beck Depression Inventory-II (BDI-II) is a 21-item questionnaire measuring severity of symptoms over the last two weeks on a scale from 0 to 3, giving a total sum score of 0–63. Examples include feeling sad and change in appetite or sleep. Suggested scoring indicates that 0–13 reflects minimal symptoms, 14–19 mild, 20–28, moderate, and 29–63 severe symptoms [24]. The BDI-II has been found to be psychometrically sound in depression[31], Chronbach's α in the current study was 0.86.

The Beck Anxiety Inventory (BAI) is a self-report measure of anxiety severity over the last week. As with the BDI-II, anxiety symptoms (e.g. "Heart pounding or racing" or feeling "nervous") are scored on a severity range from 0 to 3, giving a total sum score of 0–63. Suggested scoring indicates that 0–15 reflects mild symptoms, 16–25 moderate, and 26–63 severe symptoms. The BAI

has demonstrated good psychometric properties [34], Chronbach's α in the current study was 0.90.

Statistical analyses

Descriptive statistics on age, gender, education level and diagnosis were compiled at baseline. Distribution of scores on the EQ-5D dimensions were calculated in percentages at baseline and at end of treatment and analysed using a non-parametric test of trends developed by Cuzick. The test is similar to the Wilcoxon rank-sum test [35]. Mean scores and standard deviations at baseline and end of treatment, including change (Δ) during treatment, were calculated for the BAI, the BDI-II, the EQ-5D values, and the EQ VAS. Effect sizes (ES) were calculated from baseline to end of treatment using Cohen's *d*. Values < 0.5 are considered small, ≥ 0.5 < 0.8 moderate, and ≥ 0.8 large [36]. We also calculated the standardised response mean (SRM), defined as the mean change in score from baseline to end of treatment divided by the standard deviation of change in scores [37]. For the SRM it is suggested that magnitude of change is dependent on correlation between scores at baseline and end of treatment. For example, SRM > 0.8 can be interpreted as large if this correlation < 0.5, moderate if correlation > 0.5 [38]. Agreement between the change scores on the four measures were also analysed with Pearson's correlation. Pearson's correlation < 0.40 are considered weak, 0.40–0.49 moderate, and > 0.50 are considered strong [39].

Using the BAI and the BDI-II, the patients were then classified according to treatment response. With a minimum score of 14 on the BDI-II for depression patients and 16 on the BAI for anxiety patients at baseline, based on scoring norms for the BDI-II and BAI, patients were classified thus: "Deteriorated" if their scores increased by 9 points or more from baseline to end of treatment, "Unchanged" if the change was less than 9 points in either direction, and "Improved" if the scores decreased by 9 points or more but score at the end of treatment was still above the clinical threshold, . Finally, patients were classified as "Recovered" if their score decreased by 9 points or more and their final score was below clinical threshold (i.e. 14 for the BDI-II and 16 for the BAI) [18, 40, 41].

We ran ROC curve analyses to determine how well the EQ-5D value scores could correctly classify patients according to the clinical criteria of the BDI-II and the BAI: Recovered versus Improved, Recovered versus Unchanged, and Improved versus Unchanged. Analyses of BDI were run to calculate the area under the curve (AUROC) using the entire sample for patients that had a BDI score of at least 14 at baseline, and for all patients who had a BAI score of at least 16, regardless of primary diagnoses. Then, using primary diagnosis as recorded from the medical journals, we then calculated the AUROC for BDI-II for only the patients with depression as primary diagnosis and BDI-II baseline scores of at least 14. Lastly, we calculated the AUROC for BAI for the patients with anxiety as their primary diagnosis and a BAI baseline score of at least 16. The EQ-5D value at end of treatment was used as classifier, when computing the AUROC. AUROC was interpreted as < 0.50 useless test, 0.51–0.69 poor test, 0.7–0.79 fair test, 0.8–0.89 good test, 0.9–0.99 excellent test, 1.0 perfect test [40]. We calculated the sample size needed for the groups included in the ROC analyses. We set the Alpha level to 0.05 and the Beta level to 0.20, area under curve was set to 0.7 and value of null hypothesis was set to 0.5. The ratio of positive to negative cases was set according to the characteristics of the sample. We also computed cut-off values for recovery using Youden's index (J), which displays which values have the highest combined sensitivity and specificity [42].

Generally accepted methods for handling missing data are applicable to the EQ-5D-5L [8]. Missing data on individual items in the current study were replaced by weighted means, a method developed for treating missing data in depression cohorts [43]. All analyses were carried out using STATA 16 [44].

Ethical considerations

All patients included in the study gave written, informed consent to participate. The study is classified as health service research under Norwegian regulation. The Norwegian Data Protection Agency has in such cases designated that treatment providers (i.e. hospitals) are responsible for proper data management. Data collection and security in the present study was managed by Diakonhjemmet Hospital, and approval of data handling was granted by Oslo University Hospital, approval number 2015/15606. The study was carried out in accordance with the principles of the Helsinki declaration.

Results

Characteristics of included patients (N = 416) at baseline are shown in Table 1. Average age of patients was 37.7 years, the youngest was 18 and the oldest 65 years at start of treatment. Females made up 71.9% of the patient sample, which is in line with the gender disparity seen in prevalence studies of depression and anxiety [45]. More than 80% of the sample had some form of higher education. The study only recorded primary diagnosis from the patient's medical journal, but comorbidity was not recorded. The majority of patients had either a primary depression or anxiety diagnosis, the remaining patients were diagnosed with either mixed anxiety / depression, or adjustment disorder. The most prevalent single diagnoses were F32 Major depressive disorder, single episode (n = 114, 26.8%), F 33 Major depressive disorder,

Table 1 Demographic characteristics and diagnoses of patients at baseline (*N* = 416)

		Mean	SD	*n*	%
Age, years		37.66	10.65		
Gender					
	Female			299	71.88
	Male			117	28.13
Education					
	Primary / Secondary			70	17.16
	Higher education ≤ 4 yrs			151	37.01
	Higher education > 4 yrs			187	45.83
Primary diagnosis					
	Depression disorder			216	51.92
	Anxiety disorder			161	38.70
	Mixed anxiety / depression			24	6.77
	Adjustment disorder			15	3.61

recurrent (*n* = 97, 22.8%), and F 41.1 Generalised anxiety disorder (*n* = 86, 20.2%). Missing data in the study was typically low, > 5% on individual items for all measures.

Change in depression, anxiety and the EQ-5D-5 L during treatment

Of the 216 patients with depression diagnoses, 146 (67.59%) were "Recovered" at end of treatment, 31 (14.35%) were "Improved", and 39 were (18.05%) were "Unchanged". Of the 161 patients with anxiety disorder diagnoses, 109 (67.70%) were "Recovered" at end of treatment, 14 (8.69%) were "Improved", and 38 were (23.60%) were "Unchanged". Overall, two patients in the sample were "Deteriorated" on the BAI at end of treatment, both were diagnosed with adjustment disorder. Four patients were "Deteriorated" on the BDI-II, three of which were diagnosed with adjustment disorder, and one with anxiety disorder. No patients with anxiety diagnoses were "Deteriorated" on the BAI at end of treatment, and no patients with depression diagnoses were "Deteriorated" on the BDI-II at end of treatment.

Table 2 shows the distribution of scores on the EQ-5D dimensions at baseline, and after end of treatment. All dimensions had at least some patients reporting problems at baseline. Cuzick's non-parametric test of trends showed that all dimensions saw significant improvement from baseline to end of treatment [33]. The symptom scores reported on the BDI-II and the BAI at baseline in Table 3 indicate moderate levels of depression and anxiety. Patients saw a marked improvement in symptoms over the observation period. Cohen's *d* was >0.8 on all measures from baseline to end of treatment. Similarly, all SRM showed values > 0.8 on all instruments. Correlation between baseline scores and scores at end of treatment were <0.5 on the BDI-II (r_s = 0.39), EQ-5D value (r_s = 0.34), and the EQ-VAS (r_s = 0.31), but >0.5 on the BAI (r_s = 0.51). This indicates that the SRM was large for the BDI-II, EQ-5D value, and the EQ VAS, whilst moderate for the BAI.

Correlation of change scores

Pearson's rank order correlations are shown in Table 4. Note that the BAI and the BDI-II indicate worse health status with higher scores, whereas the reverse is true for the EQ-5D value and the EQ VAS. The EQ-5D value showed strong correlations with both the BDI-II, the EQ VAS, and moderate correlations with the BAI. The EQ VAS showed strong correlation with the BDI-II, but weak correlation with the BAI.

ROC curve analysis

For the total sample, the ROC curve analysis showed that the EQ-5D value consistently distinguished between "Recovered" and "Improved" or "Unchanged" patients according the BDI-II or BAI, AUROC ranging from 0.72 to 0.84 (Table 5). The AUC did not adequately distinguish between "Improved" and "Unchanged" on either measure, AUROC ranged from 0.49 to 0.61.

< Table 5 APPROXIMATELY HERE>

The same pattern repeated when patients scores were analysed according to diagnoses. For patients with depression, the AUC was excellent when distinguishing between "Recovered" and "Unchanged" (0.81) and acceptable distinguishing "Recovered" from "Improved" (0.78), but ineffective separating "Improved" and "Unchanged" (0.52). For patients with anxiety, the AUC showed excellent classification for "Recovered" versus "Unchanged" (0.83). Our analyses of "Recovered" versus "Improved" and "Improved" versus "Unchanged" did not have appropriate statistical power and can thus not be regarded as significant findings. Youden's index indicated that an EQ-5D value of 0.768 had the highest combined sensitivity and specificity when identifying recovered

Table 2 Distribution of EQ-5D dimensions as reported by patients ($N=416$)

		Baseline		End of treatment		*p* value
		n	%	*n*	%	
Mobility						
	No problems	323	78.59	364	88.35	< 0.001
	Slight problems	62	15.09	39	9.47	
	Moderate problems	20	4.87	9	2.18	
	Severe problems	6	1.46	.	.	
	Unable to walk about	.	.	.	.	
Self-care						
	No problems	350	85.37	398	96.60	< 0.001
	Slight problems	52	12.68	11	2.67	
	Moderate problems	6	1.46	2	0.49	
	Severe problems	2	0.49	1	0.49	
	Unable to wash or dress	.	.	.	.	
Usual activities						
	No problems	83	20.19	258	62.93	< 0.001
	Slight problems	161	39.17	111	27.07	
	Moderate problems	113	27.49	35	8.54	
	Severe problems	50	12.17	6	1.46	
	Unable to do usual activities	4	0.97	.	.	
Pain / discomfort						
	No problems	93	22.63	206	50.00	< 0.001
	Slight problems	177	43.07	151	36.65	
	Moderate problems	113	27.49	45	10.92	
	Severe problems	24	5.84	7	1.70	
	Extreme pain or discomfort	4	0.97	3	0.73	
Anxiety / depression						
	No problems	13	3.17	160	38.83	< 0.001
	Slight problems	106	25.85	187	44.39	
	Moderate problems	179	43.66	53	12.86	
	Severe problems	105	25.61	11	2.67	
	Extremely anxious or depressed	7	1.71	1	0.24	

Note. Proportion of patients reporting the various levels of severity on the EQ-5D dimensions at baseline, and at end of treatment. The *P*-values denote significance of a non-parametric test of trends across ordered groups, developed by Cuzick, similar to the Wilcoxon rank-sum test [35]

Table 3 Instrument scores at baseline and end of treatment with ES and SRM ($N=416$)

	Baseline		End of treatment		*d*	SRM
	Mean	*SD*	*Mean*	*SD*		
Anxiety (BAI)	18.48	10.27	6.49	6.47	1.39	1.35
Depression (BDI-II)	26.42	8.77	10.13	8.84	1.84	1.67
EQ-5D value	0.630	0.189	0.816	0.153	1.07	0.93
EQ VAS	54.58	17.13	74.71	14.80	1.25	1.06

Note. Abbreviations: BAI, the Beck Anxiety Inventory; BDI-II, the Beck Depression Inventory-II; ES, Effect Size (reported in Cohen's *d*); SRM, Standardised response mean

patients in the total sample. The value was the same for both depression and anxiety (Table 6).

Discussion

Our aim was to explore the responsiveness of the EQ-5D-5L in patients receiving treatment for depression and anxiety. This was done by comparing change in the EQ-5D-5L to change in the disorder-specific measures BDI-II and BAI. We hypothesised that the EQ-5D-5L should show magnitude of change as the BDI-II and BAI during treatment. The ES was large ($d>0.8$) for all measures, ranging from Cohen's d 1.07–1.84. For the SRM, which accounts for variability in treatment response by dividing change scores by the standard deviation of change scores, the BDI-II, the EQ-5D value and the EQ VAS all showed large magnitude of change. The BAI showed moderate magnitude of change on the SRM when accounting for its higher correlation between baseline and end of treatment

Table 4 Pearson's correlation of change scores ($N = 416$)

	BAI Δ	BDI-II Δ	EQ-5D value Δ
BAI Δ			
BDI-II Δ	0.49		
EQ-5D value Δ	-0.43	-0.54	
EQ VAS Δ	-0.34	-0.58	0.55

Note. Change scores calculated by subtracting the score at end of treatment from the score at baseline. Abbreviations: BAI, the Beck Anxiety Inventory; BDI-II, the Beck Depression Inventory-II. Pearson's correlation < 0.40 are considered weak, 0.40–0.49 moderate, and > 0.50 are considered strong [37]

scores. Furthermore, the EQ-5D-5L change scores showed strong correlation with the BDI-II, and moderate correlation with the BAI. The hypothesis that the EQ-5D-5 L should show similar magnitude of change as the condition-specific measures thus seems confirmed.

We then examined if the EQ-5D value could correctly classify patients deemed as "Recovered" according to the condition-specific measures. Results from the ROC analyses indicate that this was the case: AUROC were from fair to good when distinguishing "Recovered" patients from "Improved" or "Unchanged". This was true for the total sample (AUROC 0.72–0.82), for patients with depression (AUROC 0.75 and 0.80), and for patients with anxiety when distinguishing "Recovered" patients from "Unchanged" patients (AUROC 0.83). In a similarly consistent pattern, the EQ-5D-5L showed poor ability to distinguish between "Improved" and "Unchanged" patients for the total sample, for depression, and for anxiety, (AUROC 0.52–0.64). The ability of the EQ-5D-5L to consistently identify recovered patients indicates that our second hypothesis was confirmed. We also calculated Youden's index, as this may be informative for clinicians and serve as a reference for future research. For recovery from both depression and anxiety in the total sample, cut-off point as defined by highest combined sensitivity and specificity was an EQ-5D value ≥ 0.768 at end of treatment.

Data on the responsiveness of the five-level version of the EQ-5D-5L in mental health is limited, though cross-sectional measures have indicated moderate to good correlation with condition-specific measures [11]. Comparing to the three-level version, one study found moderate responsiveness to anxiety disorders. Similar to the present study, patients were classified as having either "more", "constant", or "less anxiety" according to the BAI. T-tests showed significant differences in change scores for the EQ-5D value and the EQ VAS. However, that study found that the SRM were moderate to small, and ES were large for the EQ-5D value only when patients were deteriorated [19].

Reviews of the literature on the three-level version have indicated reasonable responsiveness in depression and anxiety [10], suggesting that the five-level version may have similar properties. One recent study compared the responsiveness of the three-level and five-level versions of Anxiety / depression dimension for mental health patients. Although the five-level version was found to be more responsive, both showed limited ability to capture changes in mental health [20]. The Anxiety / depression dimension did show significant change from baseline to end of treatment in the present study. Future research may determine how useful it is as a measure on its own.

A previous cross-sectional study did find that the EQ-5D value could screen for depression and anxiety in patients with type 2 diabetes [21]. In the present study, the EQ-5D value showed similar performance in a longitudinal design in patients with depression and anxiety as primary diagnoses. That the EQ-5D value may perform better than the Anxiety / depression dimension alone is perhaps reasonable, as it may better capture the wide-ranging impact of depression and anxiety on health and quality of life [4, 5].

The EQ-5D-5L is increasingly used when evaluating health status in surveys and clinical trials [8], and decision-making bodies recommend its use in evaluating health technologies [6, 46]. Demonstrating its validity in diverse patient groups is therefore essential for sound decision-making when allocating healthcare resources. In this study, the EQ-5D-5L showed good responsiveness to change for patients with depression and anxiety. This suggests that the EQ-5D-5L can be a valid and useful tool for evaluating impact of disease and benefit of treatment for these patients, for instance through estimating QALYs. It also suggests that the EQ-5D-5L can useful when evaluating interventions for patients with depression and anxiety.

Strengths and limitations

The main strength of the study is adding to a limited evidence-base concerning the responsiveness of the five-level version of the EQ-5D in patients with depression and anxiety. The study included a fairly large clinical sample who were assessed and diagnosed by clinical psychologists before entering treatment. We can thus be reasonably certain of the clinical characteristics of the sample. The study took part in a national health service clinic, suggesting that these patients are somewhat representative of clinical populations with depression and anxiety in Norway. The patients saw substantial treatment gains as reflected by the large ES and SRM, which gave an opportunity for evaluating the ability of the EQ-5D-5L to identify recovered patients.

Several limitations to the study have to be considered. The study only included patients who completed

Table 5 Area under the receiver operating characteristic curve (AUROC) using non-parametric ROC analyses ($N = 416$)

		BDI-II				BAI			
		n	C-statistic	SE	95% CI	n	C-statistic	SE	95% CI
Overall									
	Recovered vs. Unchanged	329	0.81	0.032	(0.75–0.88)	215	0.82	0.039	(0.75–0.90)
	Recovered vs. Improved	314	0.78	0.035	(0.71–0.85)	210	0.72	0.061	(0.60–0.84)
	Improved vs. Unchanged	121	0.52	0.052	(0.48–0.68)	49	0.64	0.079	(0.49–0.80)
Depression									
	Recovered vs. Unchanged	184	0.80	0.045	(0.71–0.88)				
	Recovered vs. Improved	176	0.75	0.047	(0.65–0.84)				
	Improved vs. Unchanged	68	0.61	0.070	(0.47–0.75)				
Anxiety									
	Recovered vs. Unchanged					147	0.83		(0.71–0.95)
	Recovered vs. Improved					123	0.84		(0.70–0.97)
	Improved vs. Unchanged					52	0.53		(0.29–0.77)

Note. Area under the curve (AUROC) reported for receiver operating analyses. The BAI and the BDI-II were used as classifiers of whether patients were "Recovered", "Improved", or "Unchanged" at end of treatment. The "Overall" analyses were run using the BAI and the BDI-II for the whole sample regardless of primary diagnosis, the "Depression" analyses were run using the BDI-II for only the patients with depression as primary diagnosis, and the "Anxiety" using the BAI for only the patients with anxiety as primary diagnosis. Abbreviations: BAI, the Beck Anxiety Inventory; BDI-II, the Beck Depression Inventory-II.

Table 6 The central range of operating characteristics of the EQ-5D value post-treatment for identifying recovered versus non-recovered patients ($N=416$)

	EQ-5D cut-off	Sensitivity	Specificity	*J*	Correctly classified
Depression					
	0.740	84.72%	62.69%	0.47	77.73
	0.750	81.25%	62.69%	0.44	75.36
	0.768	**80.56%**	**64.18%**	**0.45**	**75.36**
	0.819	74.31%	67.16%	0.41	72.04
	0.827	74.31%	68.66%	0.43	72.51
Anxiety					
	0.740	84.75%	80.00%	0.65	83.76
	0.750	83.70%	80.00%	0.64	82.91
	0.768	**83.70%**	**84.00%**	**0.68**	**83.76**
	0.819	63.04%	84.00%	0.47	67.52
	0.827	60.87%	84.00%	0.45	65.81

Note. *J*=Youden's index. Depression and Anxiety recovery measured and classified using the BDI-II and the BAI, respectively

treatment, and treatment gains were large. The study could therefore not evaluate the ability of the EQ-5D-5L to detect smaller changes, that still may be of importance to patients. A related limitation is that the large rate of recovered patients in the study meant that "Unchanged" patients formed a small subgroup. The findings concerning the unchanged patients should be treated with caution. We also lack adequate data to determine if the EQ-5D-5L would be equally responsive to deterioration as improvement during treatment. The study also lacked data on comorbidity.

The current study uses the UK value set for converting to EQ-5D value scores, as there is currently no Norwegian value set available. Choice of value sets has shown to influence the estimation of QALYs, which suggests that it would be useful to replicate the present findings when a Norwegian value set is available [14].

As new measures of health status become available, such as the Recovering Quality of Life (ReQoL), it will be important to compare and contrast these against the EQ-5D-5L to judge which instrument is best suited for patients with depression and anxiety [47]. There is evidence that a wide range of outcomes that are important to patients with mental health problems are not adequately captured by commonly used instruments [5, 7]. Further research is needed to assess whether the EQ-5D-5L could reflect key changes in a wider range of outcomes [5], or if other instruments or bolt-on dimensions may be better for capturing psycho-social factors of importance to patients [48].

Conclusion

The findings in this study suggest that the EQ-5D-5L may be responsive to change in health status for patients receiving treatment for depression and anxiety. The EQ-5D-5L showed similar magnitude of change as the condition-specific measures and was also able to consistently identify patients who had recovered from depression and anxiety. Responsiveness of the EQ-5D-5L is likely sensitive to context, and these findings should be replicated in other samples. Still, these findings suggest that the EQ-5D-5L may be a useful tool for evaluating outcomes of treatment for patients with depression and anxiety.

Acknowledgements

The authors would like to thank the patients, next of kin, and user representatives for their participation in the study, and for their valuable feedback. We would also like to thank the staff at the Division of Mental Health and Substance Abuse at Diakonhjemmet Hospital.

Authors' contribution

All authors contributed to the conceptualisation and design of the study. KS led the writing of the manuscript and is the principal author of the funding application. GS contributed to conceptualisation, analyses, and writing. RGHG, KO, MTB, and SER contributed to design of study and intervention, data management, and writing. OH is the project manager and contributed to design, analyses and writing. All authors read and approved the final manuscript.

Funding

Open access funding provided by Norwegian University of Science and Technology

Data Availability

material.

The patient data used for this study is not readily available as the patients have not given consent for use or distribution beyond the research at Diakonhjemmet Hospital. Inquiries can be directed to kenneth.sandin@diakonsyk.no.

Code Availability

Any inquiry can be directed to kenneth.sandin@diakonsyk.no.

Declarations

Conflict of interest

None declared.

Ethics approval

This study was approved as a health service study by the Norwegian Data Protection Authority. The data used in the present study is part of ongoing routine data collection, and no further approval is needed beyond consent from the individual patient. Data collection and security in the present study

was managed by Diakonhjemmet Hospital, and approval of data handling was granted by Oslo University Hospital, approval number 2015/15606.

Consent for publication

All participants gave written, informed consent for use of data for research and publication.

Author details

[1]Division of Mental Health and Substance Abuse, Diakonhjemmet Hospital, Postboks 23 Vinderen, Oslo 0319, Norway
[2]Department of Psychology, Norwegian University of Science and Technology, Trondheim NO-7491, Norway
[3]Manchester Centre for Health Economics, Division of Psychology and Mental Health, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, England
[4]Department of Psychology, University of Oslo, Oslo, Norway

Received: 10 October 2022 / Accepted: 28 March 2023
Published online: 15 April 2023

References

1. OECD, Health at a OECD. ; 2019 [cited 2022 Mar 20]. Available from: https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-2019_4dd50c09-en
2. Bloom DE, Cafiero E, Jané-Llopis E, Abrahams-Gessel S, Bloom LR, Fathima S et al. The Global Economic Burden of Noncommunicable Diseases [Internet]. Program on the Global Demography of Aging; 2012 [cited 2022 Mar 31]. Available from: https://ideas.repec.org/p/gdm/wpaper/8712.html
3. Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, et al. Global burden of disease attributable to mental and substance use disorders: findings from the global burden of Disease Study 2010. Lancet. 2013;382:1575–86.
4. Vigo D, Thornicroft G, Atun R. Estimating the true global burden of mental illness. The Lancet Psychiatry. 2016;3:171–8.
5. Chevance A, Ravaud P, Tomlinson A, Berre CL, Teufer B, Touboul S, et al. Identifying outcomes for depression that matter to patients, informal caregivers, and health-care professionals: qualitative content analysis of a large international online survey. The Lancet Psychiatry. 2020;7:692–702.
6. NICE NI for H and CE. Technology appraisal processes | Technology appraisal guidance | NICE guidance | Our programmes | What we do | About [Internet]. NICE [cited 2022 Mar 31];Available from: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance/process
7. Cuijpers P. Targets and outcomes of psychotherapies for mental disorders: an overview. World Psychiatry. 2019;18:276–85.
8. Devlin N, Parkin D, Janssen B. Methods for Analysing and Reporting EQ-5D Data [Internet]. Cham: Springer International Publishing; 2020 [cited 2022 Mar 20]. Available from: http://link.springer.com/https://doi.org/10.1007/978-3-030-47622-9
9. Herdman M, Gudex C, Lloyd A, Janssen Mf, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20:1727–36.
10. Brazier J, Connell J, Papaioannou D, Mukuria C, Mulhern B, Peasgood T et al. A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. Health Technol Assess 2014;18:vii–viii, xiii–xxv, 1–188.
11. Feng YS, Kohlmann T, Janssen MF, Buchholz I. Psychometric properties of the EQ-5D-5L: a systematic review of the literature. Qual Life Res. 2021;30:647–73.
12. Mihalopoulos C, Chen G, Iezzi A, Khan MA, Richardson J. Assessing outcomes for cost-utility analysis in depression: comparison of five multi-attribute utility instruments with two depression-specific outcome measures. Br J Psychiatry. 2014;205:390–7.
13. Engel L, Chen G, Richardson J, Mihalopoulos C. The impact of depression on health-related quality of life and wellbeing: identifying important dimensions and assessing their inclusion in multi-attribute utility instruments. Qual Life Res. 2018;27:2873–84.
14. Camacho EM, Shields G, Lovell K, Coventry PA, Morrison AP, Davies LM. A (five-)level playing field for mental health conditions?: exploratory analysis of EQ-5D-5L-derived utility values. Qual Life Res. 2018;27:717–24.
15. Sandin K, Shields GE, Gjengedal RGH, Osnes K, Bjørndal MT, Hjemdal O. Self-reported health in patients on or at risk of sick leave due to depression and anxiety: validity of the EQ-5D. Front. Psychol. 2021;12:655151.
16. Streiner DL, Norman GR, Cairney J. Health Measurement Scales: a practical guide to their development and use. 5th ed. Oxford University Press; 2014.
17. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol. 2008;61:102–9.
18. Jacobson NS, Roberts LJ, Berns SB, McGlinchey JB. Methods for defining and determining the clinical significance of treatment effects: description, application, and alternatives. J Consult Clin Psychol. 1999;67:300–7.
19. König HH, Born A, Günther O, Matschinger H, Heinrich S, Riedel-Heller SG, et al. Validity and responsiveness of the EQ-5D in assessing and valuing health status in patients with anxiety disorders. Health Qual Life Outcomes. 2010;8:47.
20. Crick K, Al Sayah F, Ohinmaa A, Johnson JA. Responsiveness of the anxiety/depression dimension of the 3- and 5-level versions of the EQ-5D in assessing mental health. Qual Life Res. 2018;27:1625–33.
21. Al Sayah F, Ohinmaa A, Johnson JA. Screening for anxiety and depressive symptoms in type 2 diabetes using patient-reported outcome measures: comparative performance of the EQ-5D-5L and SF-12v2. MDM Policy & Practice. 2018;3:2381468318799361.
22. Sandin K, Anyan F, Osnes K, Gunnarsdatter Hole Gjengedal R, Risberg Leversen JS, Endresen Reme S, et al. Sick leave and return to work for patients with anxiety and depression: a longitudinal study of trajectories before, during and after work-focused treatment. BMJ Open. 2021;11:e046336.
23. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59 Suppl 20:22–33;quiz 34–57.
24. Beck AT, Steer RA, Brown GK. Manual for the beck depression inventory-II. 1996
25. Beck AT, Steer RA. Manual for the Beck anxiety inventory. San Antonio, TX: Psychological Corporation; 1990.
26. WHO. The International classification of Diseases-10 (ICD-10). Geneva: World Health Organisation; 1992.
27. Dobson KS, Dozois DJA. Handbook of Cognitive-Behavioral Therapies: Fourth Edition [Internet]. Guilford Press [cited 2023 Jan 11];Available from: https://www.guilford.com/books/Handbook-of-Cognitive-Behavioral-Therapies/Dobson-Dozois/9781462547722
28. Wells A. Metacognitive Therapy for Anxiety and Depression. 1st edition. New York, NY: The Guilford Press; 2009.
29. Gjengedal RGH, Reme SE, Osnes K, Lagerfeld SE, Blonk RWB, Sandin K, et al. Work-focused therapy for common mental disorders: a naturalistic study comparing an intervention group with a waitlist control group. WOR. 2020;66:657–67.
30. Hansen TM, Helland Y, Augestad LA, Rand K, Stavem K, Garratt A. Elicitation of norwegian EQ-5D-5L values for hypothetical and experience-based health states based on the EuroQol Valuation Technology (EQ-VT) protocol. BMJ Open. 2020;10:e034683.
31. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in Health. 2012;15:708–15.
32. NICE. Position statement on use of the EQ-5D. -5L value set for England (updated October 2019) | Technology appraisal guidance | NICE guidance | Our programmes | What we do | About [Internet]. NICE [cited 2023 Jan 11];Available from: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/technology-appraisal-guidance/eq-5d-5l
33. Stavem K, Augestad LA, Kristiansen IS, Rand K. General population norms for the EQ-5D-3L in Norway: comparison of postal and web surveys. Health Qual Life Outcomes. 2018;16:204.
34. Beck AT, Brown G, Epstein N, Steer RA. An Inventory for Measuring Clinical Anxiety: Psychometric Properties.:5.
35. Cuzick J. A wilcoxon-type test for trend. Stat Med. 1985;4:543–7.
36. Cohen J. Statistical Power Analysis for the Behavioral Sciences [Internet]. Routledge; 2013 [cited 2023 Jan 11]. Available from: https://www.taylorfrancis.com/books/mono/10.4324/9780203771587/statistical-power-analysis-behavioral-sciences-jacob-cohen

37. Liang MH. Longitudinal construct validity: establishment of clinical meaning in patient Evaluative Instruments. Med Care. 2000;38:II–84.
38. Middel B, Van Sonderen E. Statistical significant change versus relevant or important change in (quasi) experimental design: some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. Int J Integr Care [Internet] 2002 [cited 2022 Mar 20];2. Available from: http://www.ijic.org/article/https://doi.org/10.5334/ijic.65/
39. Fleiss JL, Levin B, Paik MC. Statistical methods for Rates and Proportions. John Wiley & Sons; 2013.
40. Seggar LB, Lambert MJ, Hansen NB. Assessing clinical significance: application to the beck depression inventory. Behav Ther. 2002;33:253–69.
41. Gillis MM, Haaga DAF, Ford GT. Normative values for the Beck anxiety inventory, fear questionnaire, Penn state worry questionnaire, and social phobia and anxiety inventory. Psychol Assess. 1995;7:450–5.
42. Youden WJ. Index for rating diagnostic tests. Cancer. 1950;3:32–5.
43. Gale T, Hawley C. A model for handling missing items on two depression rating scales. Int Clin Psychopharmacol. 2001;16:205–14.
44. StataCorp. Stata Statistical Software: Release 16. 2019
45. Albert PR. Why is depression more prevalent in women? J Psychiatry Neurosci. 2015;40:219–21.
46. National Institute of Public Health. Fremskaffing av EQ-5D vekter og normative data for helseøkonomiske evalueringer - prosjektbeskrivelse. (English: Acquisition of EQ-5D weights and normative data for health economic evaluations – project description) [Internet]. 2019;Available from: https://www.fhi.no/cristin-prosjekter/aktiv/fremskaffing-av-eq-5d-vekter-og-normative-data-for-helseokonomiske-evalueri/
47. Keetharuth AD, Brazier J, Connell J, Bjorner JB, Carlton J, Taylor Buck E, et al. Recovering quality of life (ReQoL): a new generic self-reported outcome measure for use with people experiencing mental health difficulties. Br J Psychiatry. 2018;212:42–9.
48. Chen G, Olsen JA. Filling the psycho-social gap in the EQ-5D: the empirical support for four bolt-on dimensions. Qual Life Res. 2020;29:3119–29.

Publisher's Note

APPENDIX

Questionnaires

Beck's Anxiety Inventory

Indicate to what degree you've been affected by these symptoms the last week, including today (abbreviated).

	Not at all	Mildly (Did not bother me too much)	Moderately (It was very unpleasant, but I could stand it)	Severely. (I could barely stand it)
1. Numbness or tingling	◯	◯	◯	◯
2. Feeling hot	◯	◯	◯	◯
3. Wobbliness in legs	◯	◯	◯	◯
4. Unable to relax	◯	◯	◯	◯
5. Fear of the worst happening	◯	◯	◯	◯
6. Dizzy or light-headed	◯	◯	◯	◯
7. Heart pounding or racing	◯	◯	◯	◯
8. Unsteady	◯	◯	◯	◯
9. Terrified	◯	◯	◯	◯
10. Nervous	◯	◯	◯	◯
11. Feelings of choking	◯	◯	◯	◯
12. Hands trembling	◯	◯	◯	◯
13. Shaky	◯	◯	◯	◯
14. Fear of losing control	◯	◯	◯	◯
15. Difficulty breathing	◯	◯	◯	◯
16. Fear of dying	◯	◯	◯	◯
17. Scared	◯	◯	◯	◯
18. Indigestion or discomfort in abdomen	◯	◯	◯	◯
19. Faint	◯	◯	◯	◯
20. Face flushed	◯	◯	◯	◯
21. Sweating (not due to heat)	◯	◯	◯	◯

Beck's depression inventory-II

Indicate how you've been feeling the last two weeks, including today (translator's abbreviated instruction).

1. Sadness
   - 0 I do not feel sad
   - 1 I feel sad much of the time
   - 2 I am sad all the time
   - 3 I'm so sad or unhappy that I can't stand it

2. Pessimism
   - 0 I am not particularly discouraged about my future
   - 1 I feel more discouraged about my future than I used to be
   - 2 I do not expect things to work out for me
   - 3 I feel my future is hopeless and will only get worse

3. Past failure
   - 0 I do not feel like a failure
   - 1 I feel I have failed more than I should have
   - 2 As I look back, I see a lot of failures
   - 3 I feel I am a total failure as a person

4. Loss of pleasure
   - 0 I get as much pleasure as I ever did from the things I enjoy
   - 1 I don't enjoy things as much as I used to
   - 2 I get very little pleasure from the things I used to enjoy
   - 3 I can't get any pleasure from the things I used to enjoy

5. Guilty feelings
   - 0 I don't feel particularly guilty
   - 1 I feel guilty over many things I have done or should have done
   - 2 I feel quite guilty most of the time
   - 3 I feel guilty all the time

6. Punishment feelings
   - 0 I don't feel I am being punished
   - 1 I feel I may be punished
   - 2 I expect to be punished
   - 3 I feel I am being punished

7. Self-dislike
   - 0 I feel the same about myself as ever
   - 1 I have lost confidence in myself
   - 2 I am disappointed in myself
   - 3 I dislike myself

8. Self-criticalness
   0 I dont criticise or blame myself more than usual
   1 I am more critical of myself than I used to be
   2 I critisice myself for all of my faults
   3 I blame myself for everything bad that happens

9. Suicidal thoughts or wishes
   0 I don't have any thoughts of killing myself
   1 I have thoughts of killing myself, but I would not carry them out
   2 I would like to kill myself
   3 I blame myself for everything bad that happens

10. Crying
   0 I don't cry any more than I used to
   1 I cry more than I used to
   2 I cry over every little thing
   3 I feel like crying, but I can't

11. Agitation
   0 I am no more restless or wound up than usual
   1 I feel more restless or wound up than usual
   2 I am so restless or agitated that it's hard to sit still
   3 I am so restless or agitated that I have to keep moving or doing something

12. Loss of interest
   0 I have not lost interest in other people or activities
   1 I have less interest in other people or things than before
   2 I have lost most of my interest in other people or things
   3 It's hard to get interested in anything

13. Indecisiveness
   0 I make decisions about as well as before
   1 I find it more difficult to make decisions than usual
   2 I have much greater difficulty in making decisions than I used to
   3 I have trouble making any decisions

14. Worthlessness
   0 I do not feel I am worthless
   1 I don't consider myself as worthwhile and useful as I used to
   2 I feel more worthless as compared to other people
   3 I feel utterly worthless

15. Loss of energy
   0 I have as much energy as ever
   1 I have less energy than I used to have
   2 I don't have enough energy to do very much
   3 I don't have enough energy to do anything

16. Changes in sleeping pattern
    - 0 I have not experienced any changes in my sleeping pattern
    - 1a I sleep somewhat more than usual
    - 1b I sleep somewhat less than usual
    - 2a I sleep a lot more than usual
    - 2b I sleep a lot less than usual
    - 3a I sleep most of the day
    - 3b I wake up 1-2 hours early and can't get back to sleep

17. Irritability
    - 0 I am no more irritable than usual
    - 1 I am more irritable than usual
    - 2 I am much more irritable than usual
    - 3 I am irritable all the time

18. Changes in appetite
    - 0 I have not experienced any changes in my appetite
    - 1a My appetite is somewhat less than usual
    - 1b My appetite is somewhat greater than usual
    - 2a My appetite is much less than before
    - 2b My appetite is much greater than before
    - 3a I have no appetite at all
    - 3b I crave food all the time

19. Concentration difficulty
    - 0 I can concentrate as well as ever
    - 1 I can't concentrate as well as usual
    - 2 It's hard to keep my mind on anything for very long
    - 3 I find I can't concentrate on anything

20. Tiredness or fatigue
    - 0 I am no more tired or fatigued than usual
    - 1 I get more tired or fatigued more easily than usual
    - 2 I am too tired or fatigued to do a lot of the things I used to do
    - 3 I am too tired or fatigued to do most of the things I used to do

21. Loss of interest in sex
    - 0 I have not noticed any recent changes in my interest in sex
    - 1 I am less interested in sex than I used to be
    - 2 I am much less interested in sex now
    - 3 I have lost interest in sex completely

EQ–5D–5L

Under each heading, please tick the **ONE** box that best describes your health **TODAY**

MOBILITY

I have no problems in walking about ☐
I have slight problems in walking about ☐
I have moderate problems in walking about ☐
I have severe problems in walking about ☐
I am unable to walk about ☐

SELF-CARE

I have no problems washing or dressing myself ☐
I have slight problems washing or dressing myself ☐
I have moderate problems washing or dressing myself ☐
I have severe problems washing or dressing myself ☐
I am unable to washing or dress myself ☐

USUAL ACITIVITIES

I have no problems doing my usual activities ☐
I have slight problems doing my usual activities ☐
I have moderate problems doing my usual activities ☐
I have severe problems doing my usual activities ☐
I am unable to do my usual activities ☐

PAIN / DISCOMFORT

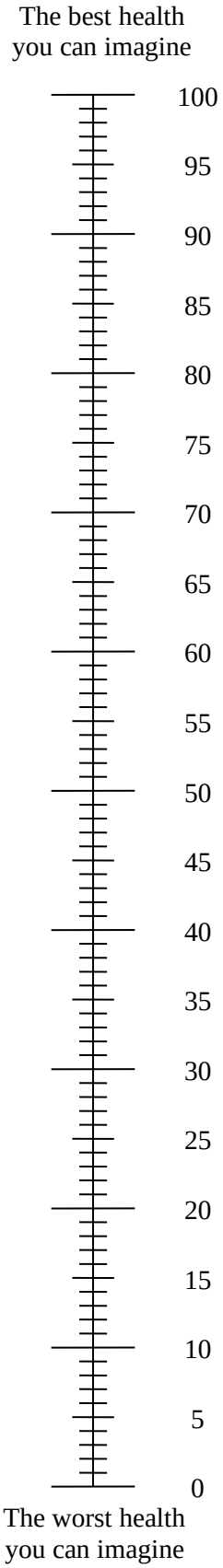
I have no pain or discomfort ☐
I have slight pain or discomfort ☐
I have moderate pain or discomfort ☐
I have severe pain or discomfort ☐
I extreme pain or discomfort ☐

ANXIETY / DEPRESSION

I am not anxious or depressed ☐
I am slightly anxious or depressed ☐
I am moderately anxious or depressed ☐
I am severely anxious or depressed ☐
I am extremely anxious or depressed ☐

- We would like to know how good or bad your health is TODAY
- This scale is numbered from 0 to 100.
- 100 means the best health you can imagine.
  0 means the worst health you can imagine.
- Mark an X on the scale to indicate how your health is TODAY.
- Now, please write the number you marked on the scale in the box below.

YOUR HEALTH TODAY = ☐



Return to work self-efficacy (RTW-SE)

The statements below concern your expectations for work. Imagine that you will be working full time at your current job from tomorrow, with your current mental and physical health. Please fill in this form regardless of whether you are currently working or on sick leave.

	Completely disagree					Completely agree
1. I will be able to cope with setbacks	1	2	3	4	5	6
2. I won't be able to handle potential problems at work	1	2	3	4	5	6
3. I won't be able to complete my work tasks due to my emotional state	1	2	3	4	5	6
4. I will be able to set personal boundaries at work	1	2	3	4	5	6
5. I will be able to perform my tasks at work	1	2	3	4	5	6
6. I will be able to deal with emotionally demanding situations	1	2	3	4	5	6
7. I will have no energy left to do anything else	1	2	3	4	5	6
8. I will be able to concentrate at work	1	2	3	4	5	6
9. I will be able to cope with work pressure	1	2	3	4	5	6
10. I will be able to handle potential problems at work	1	2	3	4	5	6
11. I can motivate myself to perform my job	1	2	3	4	5	6
12. I can deal with the physical demands of my work	1	2	3	4	5	6

Subjective health complaints (SHC)

Here are some common health complaints
(indicate what best fits you)

		Not at all	A little	Some	Serious
1.	Cold, flu	0	1	2	3
2.	Coughing.....	0	1	2	3
3.	Shoulder pain.....	0	1	2	3
4.	Headache	0	1	2	3
5.	Neck pain	0	1	2	3
6.	Upper back pain.....	0	1	2	3
7.	Lower back pain	0	1	2	3
8.	Arm pain.....	0	1	2	3
9.	Shoulder pain.....	0	1	2	3
10.	Migraine	0	1	2	3
11.	Extra heartbeats	0	1	2	3
12.	Chest pain.....	0	1	2	3
13.	Breathing difficulties	0	1	2	3
14.	Leg pain during physical activities	0	1	2	3
15.	Acid reflux, «heartburn»	0	1	2	3
16.	Stomach discomfort	0	1	2	3
17.	Ulcer / non-ulcer dyspepsia.....	0	1	2	3
18.	Stomach pain	0	1	2	3
19.	Gas discomfort.....	0	1	2	3
20.	Diarrhoea.....	0	1	2	3
21.	Constipation.....	0	1	2	3
22.	Eczema	0	1	2	3
23.	Allergies	0	1	2	3
24.	Hot flashes.....	0	1	2	3
25.	Sleep problems	0	1	2	3
26.	Tiredness	0	1	2	3
27.	Dizzines.....	0	1	2	3
28.	Anxiety.....	0	1	2	3
29.	Sadness/depression	0	1	2	3