

Måleegenskaper ved den norske versjonen av Beck Depression Inventory 2. utgave (BDI II)

Johan Siqveland¹(johan.siqveland@ahus.no)

Hege Kornør¹² (hege.kornor@r-bup.no)

PsykTestBarn 2011, 1:5

ISSN 1893-9910

Mottatt dato 16. mai 2011

Publisert dato 15. august 2011

Artikkelens DOI 10.21337/0005

Som alle artikler i PsykTestBarn, kan denne fagfellevurderte artikkelen lastes ned, skrives ut og distribueres fritt for alle slags formål på følgende vilkår: korrekt referanse skal oppgis (se under), ingen kommersiell bruk, og ingen bearbeidelse av tekst eller innhold.

Denne artikkelen skal siteres på følgende måte:

Siqveland, J. & Kornør, H. (2011). Måleegenskaper ved den norske versjonen av Beck Depression Inventory 2. utgave (BDI II). *PsykTestBarn*, 1:5. doi: 10.21337/0005

¹ Nasjonalt kunnskapssenter for helsetjenesten

² RBUP Øst og sør

Sammendrag

Beskrivelse. Beck Depression Inventory II (BDI II) er en selv vurderingsskala med 21 ledd som måler grad av depresjon hos ungdom over 13 år og voksne. BDI II ble utgitt i 1996 og er en revidert utgave av det opprinnelige Beck Depression Inventory fra 1961. Norsk utgave kom i 2005 og rettighetshaver for den norske versjonen er Pearson Assessment.

Hvert ledd i skalaen beskriver et symptom på depresjon, og respondenten blir bedt om å velge hvilket av fire utsagn som best beskriver i hvilken grad han/hun har opplevd det aktuelle symptomet de siste to ukene. Skjemaet fylles ut av deltageren selv og tar mindre enn 30 minutter å fylle ut. Skårene fra kartleggingen tolkes av lege eller psykolog

Litteratursøk. Søket resulterte i 221 referanser, hvorav 67 abstracts virket relevante og ble bestilt i fulltekst. Vi inkluderte én artikkel, én hovedoppgave, samt manualen til BDI II. Utvalgene besto utelukkende av voksne.

Psykometri. Indre konsistens og test-retest reliabilitet var adekvat, validitet sammenlignet med andre selvrapporteringskjema var også adekvat. Validitet sammenlignet med diagnostisk intervju var ikke undersøkt.

Konklusjon: Den norske versjonen av BDI II har gode grunnleggende psykometriske egenskaper i voksenpopulasjonen. Måleegenskapene for ungdomspopulasjonen er ikke dokumentert. Hvor egnet skårer fra BDI II er til å skille personer med en behandlingstrengende depresjon fra de som ikke har det, samt skille mellom ulike alvorlighetsgrader av depresjon, er ikke undersøkt. De viktigste egenskapene for klinisk anvendelse av BDI II er dermed ikke undersøkt for den norske utgaven og tolkning av skårer må gjøres med forsiktighet.

Abstract

Description. The Beck Depression Inventory II (BDI-II) is a 21-item multiple-choice self-report instrument developed to measure depression in persons aged 13 and above. The BDI-II was published in 1996, and is a revised version of the Beck Depression Inventory (BDI), published in 1961. The Norwegian version of the BDI-II was published in 2005; copyright holder Pearson Assessments.

Each item in the BDI-II describes a symptom of depression. The participants are instructed to rate the severity of this symptom over the past two weeks on a 4-point Likert-type scale from 0 (absence of symptom) to 3 (severe manifestation of symptom). The BDI-II has a self-rating format and takes less than 30 minutes to complete. The scores are to be interpreted by a clinical psychologist or a medical doctor.

Literature search. Our literature search identified 221 references; and 67 articles were read in full text. We included one article, one master thesis, and the manual in our evaluation. All included materials described the psychometric properties in adult populations.

Psychometrics. Internal consistency reliability, test-retest reliability and external validity, as compared to other self-report forms, were all reported to be adequate. The agreement between the BDI-II score and the scores of various diagnostic interviews has not been investigated in any of the included studies.

Conclusion. The Norwegian version of the BDI-II has demonstrated adequate psychometric properties in an adult population. Its psychometric properties in adolescent samples have not been documented. It has not been reported, whether the instrument is suitable for diagnosing depression severity. Thus, the Norwegian BDI-II provides no evidence for the clinical usefulness of the instrument for assessing or diagnosing depression, and scores need to be interpreted with caution.

Innledning

Beck Depression Inventory 2. utgave (BDI II) er en selvvurderingsskala med 21 ledd som måler grad av depresjon hos ungdom over 13 år og voksne. BDI II ble utgitt i 1996 og er en revidert utgave av det opprinnelige Beck Depression Inventory fra 1961. Hvert ledd i skalaen beskriver et symptom på depresjon, og respondenten blir bedt om å velge hvilket av fire utsagn som best beskriver i hvilken grad han/hun har opplevd det aktuelle symptomet de siste to ukene. Beskrivelsen kan skåres på en skala fra 0 til 3, hvor 0 betegner fravær av symptomet (Beck, Steer & Brown, 1996). Alle leddene summeres og sumskåren kan kategoriseres i fire kategorier; ingen, mild, moderat og alvorlig depresjon. Cutoff skårene for disse kategoriene er basert på amerikanske normer. Den engelskspråklige versjonen av BDI II har god indre konsistens, test-restest reliabilitet og god divergent og konvergent validitet (Beck, Steer & Brown, 1996). BDI II ble oversatt til norsk gjennom en tilbakeoversettelsesprosedyre. BDI II finnes på norsk i papirutgave, og Pearsson Assessment har rettighetene til den norske utgaven.

Kompetansekrav for bruk av BDI II er lege, psykolog eller psykoterapeut. Skalaen kan administreres av andre yrkesgrupper, men tolkning av skårer må gjøres av en person som fyller kompetansekravet.

Metode

Vi søkte etter dokumentasjon på testenes psykometriske egenskaper i databasene Norart, BIBSYS, SveMed+, Cochrane Library, PsycINFO, Medline og Embase. Søkedato: 16.03.2010. Søkestrategien er tilgjengelig på <http://www.r-bup.no/CMS/ptb.nsf/pages/bdi2>. Vi kontaktet også førsteforfattere av inkluderte artikler, samt

oversetterne av BDI II, for å identifisere dokumentasjon som eventuelt ikke blir fanget opp av det systematiske søket.

Vi inkluderte alle publikasjoner av studier som har undersøkt og rapportert minst ett av følgende i skandinaviske utvalg:

- normdata for testen
- reliabilitet: indre konsistens, test-retest, interrater og endringssensitivitet
- validitet: samsvar med liknende testskårer, samsvar med referanse-standard eller annet kriterium, og/eller faktorstruktur

I tillegg, og kun for norske versjoner av BDI II, inkluderte vi publikasjoner som rapporterte gjennomsnittsskårer og/eller forekomster for henholdsvis generelle populasjoner og kliniske undergrupper.

To forskere gikk gjennom sammendragene til alle identifiserte publikasjoner, uavhengig av hverandre, etter at dubletter var fjernet. Alle publikasjoner som kunne virke relevante ble bestilt inn i fulltekst, og prosessen over ble gjentatt for fulltekstrapportene.

To forskere vurderte normering, validitet og reliabilitet, uavhengig av hverandre, ved hjelp av en tilpasset versjon av "Test review form and notes for reviewers" (EFPA 2008).

Resultater

Litteratursøk

Søket resulterte i 221 referanser, hvorav 67 abstracts virket relevante og ble bestilt i fulltekst. Kun én av disse (Hynninen, Pallesen, & Nordhus, 2007) fylte inklusjonskriteriene. Blant de 66

Tabell 1. Total BDI gjennomsnittskåre (Mean) og standardavvik (SD) i norske populasjoner

Referanse	Utvalg	n	Mean	SD
Aasen 2001	Normalpopulasjon			
	Kvinner	512	8,9	7,7
	Menn	363	7,3	7,1
Aasen 2001	Studentpopulasjon	303	7,1	6,0

ekskluderte publikasjonene var den første utgaven av BDI brukt i 49, 14 hadde ikke-norske utvalg, én brukte ikke BDI i det hele tatt, én var en dublett og én lyktes vi ikke med å få tak i. Vi kontaktet forfatterne av den inkluderte publikasjonen, andre fagfolk og Pearson Assessment med forespørsler om kjennskap til ytterligere relevant dokumentasjon. Dette førte til inklusjon av en hovedoppgave (Aasen, 2001) og den norske testmanualen.

Hovedoppgaven (Aasen, 2001) inneholdt en tverrsnittstudie av et tilfeldig utvalg fra den generelle populasjonen (N=875) samt en longitudinell studie av en gruppe studenter (N=303). Fra disse undersøkelsene ble det rapportert gjennomsnittsskåre, standardavvik, indre konsistens og test-retest reliabilitet, faktorstruktur og konvergent/divergent validitet med andre selvrappørterte symptomer.

Vi søkte også etter dokumentasjon av psykometriske egenskaper ved den svenske og

danske BDI-II utgaven. Vi fant én relevant publikasjon fra et svensk klinisk utvalg på 87 personer som søkte behandling for depresjon (Holländare, Andersson, & Engström, 2010). Publikasjonen rapporterte indre konsistens.

Hynninen og kolleger (2007) rapporterte resultater fra en tverrsnittstudie med 58 KOLS-pasienter med symptomer på moderat til alvorlig depresjon. Artikkelen oppga gjennomsnittsskåre for hele utvalget og fordelt på kjønn og på KOLS-diagnosens alvorlighetsgrad, samt korrelasjoner med mål på angst. Gitt at dette utvalget var så lite og selektert har vi valgt å ikke gjengi gjennomsnittsskårene her.

Middelverdier

Aasen fanget opp en statistisk signifikant kjønnsforskjell i den generelle populasjonen, men ikke i studentpopulasjonen (Tabell 1). Standardavvik på størrelse med gjennomsnittsskårene angir stor spredning i skårene.

Normer

Aasen rapporterte forekomster av forskjellige grader av depresjon, basert på amerikanske terskelverdier (Mild: 14-19, Moderat: 20-28, Alvorlig: 29-63). Det så ut til at depresjonsforekomstene lå noe lavere i studentpopulasjonen enn i den generelle populasjonen, men dette var ikke testet statistisk (Tabell 2).

Validitet

Konvergent validitet ble undersøkt blant studenter (N = 303) ved å korrelere BDI II med ZungSelf-Report DepressionScale (Zung SDS; Zung, 1965) ($r = 0,71$) og depresjonsskalaen i Revised Symptom Checklist 90 (SCL 90-R; Derogatis, 1983) ($r = 0,84$). Divergent validitet ble undersøkt ved å korrelere BDI II Beck Anxiety Inventory (BAI; Beck & Steer, 1993) ($r = 0,67$) og psykotisismeskalaen i SCL 90-R ($r = 0,61$). Sammenhengen mellom BDI II og andre selvrapporeringsmål på depresjon var her signifikant sterkere enn sammenhengen med selvrapporterte psykotiske symptomer, men ikke signifikant sterkere enn egenrapporterte angstsymptomer (Aasen, 2001).

Som et mål på begrepsvaliditet gjorde Aasen (2001) en konfirmatorisk faktoranalyse som fant flere modeller som ga en tilfredsstillende fit. Den som til slutt ble foretrukket var en modell med tre ulike underfaktorer; negative holdninger, prestasjonsvansker og somatiske plager. Denne faktorstrukturen har også blitt rapportert i en amerikansk faktoranalyse (Osman et al., 1997). Spørsmålet om tap av seksuell interesse var vanskelig å plassere innen denne faktorstrukturen. At tap av seksuelle interesse er vanskelig å plassere i en faktorstruktur er også kjent fra den engelske utgaven av BDI II (Beck, Steer & Brown, 1996).

Reliabilitet

Indre konsistens var svært god ($\alpha = 0,91$) for normalgruppen og god ($\alpha = 0,86$) for studentgruppen. Tre ukers test-retest reliabilitet blant studenter var god ($r = 0,77$) (Aasen, 2001). Den svenske studien rapporterte også om svært god indre konsistens ($\alpha = 0,90$) (Holländare et al., 2010).

Diskusjon og konklusjon

BDI II er et mål på depressive symptomer som er mye brukt i internasjonal forskning og har solid dokumentasjon på gode psykometriske egenskaper. Av psykometriske egenskaper for den norske utgaven er det reliabilitet i form av indre konsistens og test-retest som er best undersøkt, og her fremstår BDI II som et reliabelt instrument. Selv om det fra et psykometrisk synspunkt er ønskelig med høy indre konsistens, drøfter Aasen (2001) hvorvidt høy indre konsistent er ønskelig for et mål på depresjon. Siden depresjon er et mangefasettert fenomen vil skalaer med høy indre konsistens kunne kritiseres for å være for smale mål på depresjon, og det vil være en kritikk som vil kunne rettes også mot den norske versjonen av BDI II.

I tillegg til reliabilitet er også konvergent og divergent validitet sett i relasjon til andre selvrapporerings skjema undersøkt og funnet stort sett adekvate i normalpopulasjoner. BDI II skårer imidlertid høyt på sammenheng med andre selvrapporeringsmål på angst. Den høye korrelasjonskoeffisienten mellom BDI II og BAI-skårer må ikke nødvendigvis tolkes som svak diskriminant validitet, men kan reflektere at angst

Tabell 2. Forekomster av depresjon målt med BDI II i generelle norske populasjoner

Referanse	Utvalg/gruppe	N	Mild depresjon		Moderat depresjon		Alvorlig depresjon	
			n	%	n	%	n	%
Aasen 2001	Normalpopulasjon	875	93	10,6	52	6,0	21	2,4
Aasen 2001	Studenter	303	16	5,8	12	4,2	3	1,0

og depresjon er nært beslektede fenomener som forventes å korrelere høyt (Clark & Watson, 2006; Tyrer, 2001). De høye korrelasjonene kan imidlertid også være et uttrykk for svak diskriminant validitet.

Et kjent problem med selvrapporteringskjemaer for depresjon er at de ofte er lite egnet til å skille personer som har midlertidige psykiske stress-symptomer, fra personer som har en underliggende depressiv lidelse (Coyne, 1994). Midlertidige stressymptomer går ofte over av seg selv uten behandling eller krever enklere psykososiale intervensjoner. Depresjon er derimot behandlingstrengende og noen ganger en alvorlig tilstand. Hvor egnet BDI II er til bruk i screening for depresjon og hvor godt BDI II skiller mellom ulike alvorlighetsgrader av depresjon er ikke undersøkt for den norske utgaven. For å undersøke dette måtte skårer fra BDI II blitt sammenlignet med resultater fra kliniske intervjuer, noe som ikke har blitt gjort. De viktigste spørsmålene omkring klinisk nytteverdi er dermed ikke undersøkt ved den norske utgaven.

BDI II fremstår altså i norsk utgave som et kartleggingskjema som har grunnleggende gode måleegenskaper men de viktigste spørsmålene for klinisk bruk av BDI II på norsk er uavklarte. Måleegenskapene ved BDI II i den norske ungdomspopulasjonen er ikke dokumentert.

Referanser

Aasen, H. (2001). *An empirical investigation of depression symptoms: norms, psychometric characteristics and factor structure of the Beck Depression Inventory II*. Bergen: Det psykologiske fakultet, Universitetet i Bergen.

Beck, A. T. & Steer, R. A. (1993). *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.

Clark, L. A. & Watson, D. (2006). Distress and fear disorders: an alternative empirically based taxonomy of the 'mood' and 'anxiety' disorders. *British Journal of Psychiatry, 189*, 481-483.

Coyne, J. C. (1994). Self-reported distress: Analog or ersatz depression? *Psychological Bulletin, 116*(1), 29-45.

Derogatis, L. R. (1983). *SCL-90-R. Administration, scoring and procedures manual*. Baltimore, MA: Clinical Psychometric Research.

European Federation of Psychologist Association (EFPA). (2008). EFPA Review model for the description and evaluation of psychological tests: Test review form and notes for reviewers (3.42 ed.): EFPA.

Holländare, F., Andersson, G., & Engström, I. (2010). A comparison of psychometric properties between internet and paper versions of two depression instruments (BDI-II and MADRS-S) administered to clinic patients. *Journal of Medical Internet Research, 12*, e49.

Hynninen, M. J., Pallesen, S., & Nordhus, I. H. (2007). Factors affecting health status in COPD patients with co-morbid anxiety or depression. *International Journal of COPD, 2*, 323-328.

Osman, A., Downs, W. R., Barrios, F. X., Kopper, B. A., Gutierrez, P. M., & Chiro, C. E. (1998). Factor structure and psychometric characteristics of the Beck Depression Inventory-II. *Journal of Psychopathology and Behavioral Assessment, 19*, 359-376.

Tyrer, P. (2001). The case for cothymia: mixed anxiety and depression as a single diagnosis. *British Journal of Psychiatry, 179*, 191-193.

Zung, W. W. K. (1965). A Self-rating Depression Scale. *Archives of General Psychiatry, 12*, 63-70.