

Kan Multidimensional Anxiety Scale for Children (MASC) og Beck Depression Inventory-II (BDI-II) predikere emosjonelle lidelser hos norsk ungdommer i poliklinisk behandling?

—
Stine Marie Skogly

Hovedoppgave ved profesjonsstudiet i psykologi – Mai 2015

Forord

Bakgrunn for valg av tema og oppgave var min interesse for og deltakelse i SMARTprosjektet som SMARTterapeut, gjennom veileder Veronica Lorentzen. Datamaterialet til studien fikk jeg fra fire ulike barne- og ungdomspsykiatriske poliklinikker (BUP) i Nord-Norge, som er en del av datainnsamlingen til veileders doktorgradsstudiet. Jeg er veldig takknemlig for at jeg fikk bruke en del av dette datamaterialet til min hovedoppgave.

Problemstilling, tips til utforming av oppgaven og litteratur på skriveprosessen fikk jeg fra Veronica. Søk av relevant litteratur til emnet, valg av statistiske analyser, samt tolkning av resultatet, og det skriftlige arbeidet har jeg stått for selv. Det hadde likevel ikke vært mulig å gjennomføre uten korrekturlesning, tilbakemeldinger og innspill fra Veronica og biveileder Per Håkan Brøndbo. Takk for uvurderlig hjelp og oppløftende kommentarer underveis. Vil også takke Kenneth for gjennomlesning av oppgaven og tips til struktur og formuleringer. Sist, men ikke minst vil jeg takke familie og venner som alltid har hatt troen på meg og oppmuntret meg når det har vært nødvendig.

Denne hovedoppgaven har blitt utformet som et ledd i profesjonsutdanningen i psykologi ved UiT Norges Arktiske universitet. Arbeidet med den har tidvis vært krevende, men også en veldig lærerik prosess som har gitt meg masse ny kunnskap og inspirasjon.



UiT / NORGES ARKTISKE
UNIVERSITET

**Kan Multidimensional Anxiety Scale for Children (MASC) og Beck
Depression Inventory-II (BDI-II) predikere emosjonelle lidelser hos norske
ungdommer i poliklinisk behandling?**

PSY-2901

Stine Marie Skogly, UiT

Veiledere:

Veronica Lorentzen, BUP Tromsø & UiT

Per Håkan Brøndbo, BUP Tromsø

Hovedoppgave for graden Cand. Psychol.

Institutt for psykologi

Det helsevitenskapelige fakultet

UiT, Norges Arktiske Universitet

Mai 2015

Sammendrag

Multidimensional Anxiety Scale for Children (MASC) og Beck Depression Inventory-II (BDI-II) er to veldokumenterte screeningskjema for henholdsvis angst og depresjon. Vi vet relativt lite om deres diagnostiske prediksjonsevner for norske ungdommer i poliklinisk behandling, og denne studien har som hovedformål å gi mer kunnskap på dette området. Data fra 57 ungdommer (14–18 år, 91.2% jenter) ved fire psykiatriske poliklinikker i Nord-Norge, ble analysert ved bruk av receiver operating characteristic (ROC). Konsensusdiagnose ble brukt som gullstandard og på bakgrunn av det ble pasientene delt inn i to grupper: angst (n = 12) og depresjon (n = 22). MASC og BDI-II skårer ble så sammenlignet med konsensusdiagnose og area under the curve (AUC) regnet ut. AUC for MASC totalskåre (angst vs. ikke-angst) var .58 med cut off skåre >50.5, det ga sensitivitet på 83% og spesifisitet på 36%. AUC for MASC ADI skåre (angst vs. ikke-angst) var .61 med cut off skåre >13.5, som ga sensitivitet på 92% og spesifisitet på 29%. AUC for BDI-II totalskåre (deprimert vs. ikke-deprimert) var .69 med cut off >26.5, og det ga sensitivitet på 86% og spesifisitet på 54%. Resultatene fra denne studien viser at verken MASC eller BDI-II er gode nok skjema for å predikere emosjonelle lidelser hos norske ungdommer i poliklinisk behandling.

Det finnes mange ulike utredningsskjemaer for selvutfylling ved kartlegging av emosjonelle lidelser, både for voksne og barn/unge. Å finne slike skjemaer som er lett å anvende, tidsbesparende og som gir et godt mål på psykisk helse er viktig for å raskt identifisere ungdommer med psykiske vansker, slik at man kan sette inn forebyggende tiltak eller igangsette tidlig behandling. To av selvrapporteringskjemaene som brukes mye i barne- og ungdomspsykiatrien er Multidimensional Anxiety Scale for Children (MASC; March, 1997) og Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996). Både MASC og BDI-II er godt validert i litteraturen som screeningverktøy, men hvor gode er de som diagnostiske verktøy? Det er en stor forskjell på screening og diagnostisering; et screeningverktøy skal kun identifisere at en lidelse *muligens* er tilstede for så å fremme videre undersøkelser av de som skårer positivt, mens et diagnostisk verktøy skal med stor sikkerhet fortelle at en lidelse *er* tilstede hos de med positivt resultat (Warner, 2004). Et diagnostisk verktøy vil aldri være hundre prosent sikkert, men den diagnostiske verdien til disse testene vil bli forbedret hvis vi bruker vår kunnskap om lidelsene til å begrense bruken av det til de pasientene vi mistenker har lidelsen. Testresultatet vil da være mer valid enn hvis vi brukte testen vilkårlig på alle pasienter (Akobeng, 2006a).

Konsekvenser av psykiske lidelser. Tidligere forskning har vist at milde psykiatriske symptomer, sub-terskel depresjon eller dystymi gir økt risiko for mer alvorlige psykiatriske lidelser og suicidal tanker (Blom, Larsson, Serlachius, & Ingvar, 2010). Det ser ut til at psykiske vansker i barne- og ungdomsårene har mer omfattende konsekvenser enn somatisk sykdom og gir i tillegg dårligere psykisk og fysisk helse senere i livet, (Houghton, Hunter, Trewin, Glasgow, & Carroll, 2014) øker risikoen for frafall i skolen, løsere tilknytning til arbeidslivet, økonomiske vansker, rusmisbruk og vansker i nære relasjoner (Skogen, Kjeldsen, Knudsen, Myklestad, & Nesvåg, 2014). Det er viktig å identifisere barn og unge i alderen 3-18 år med psykiske vansker tidlig fordi de er mer utsatt for psykososiale vansker enn andre aldersgrupper og det estimeres at mellom 15–20% rammes av psykososiale vansker (Mathiesen, Karevold, & Knudsen, 2009). Psykiske vansker har også samfunnsøkonomiske konsekvenser. Psykiske lidelser koster ca. 70 milliarder årlig i bare trygdeutgifter, behandlingsutgifter og sosiale utgifter. Av de 900 000 personene i yrkesaktiv alder som står utenfor arbeidslivet er det en fjerdedel av disse som hovedsakelig er det på grunn av psykiske plager (Norsk Psykologforening, 2014). Mer spesifikt er depresjon den ledende årsaken til arbeidsudyktighet og den viktigste årsaken til fullbyrdet selvmord. Det antas at i år 2020 vil depresjon være den nest største byrden av helseproblemer og folkehelseinstituttet beskriver det som den mest kostbare av alle psykiske lidelser (Wang & Nordahl, 2012).

Prevalens av angst og depresjon. I Norge i dag regnes det at det til en hver tid er 15–20% av barn og unge mellom 3 og 18 år som har nedsatt funksjon på grunn av psykiske plager og lidelser som atferdsproblemer, angst og depresjon. Omtrent 8% av barn og unge har så alvorlige problemer at de tilfredsstillende en psykiatrisk diagnose, (Skogen et al., 2014) mens det på et nasjonalt nivå i 2009 kun ble tilbydd behandling til 4.5% i spesialisthelsetjenesten (Mathiesen, 2009). Ungdommer som sliter psykisk har ofte sammensatte vansker og man har funnet at ca. 40% av ungdom som lider av en gruppe psykiske lidelser (for eksempel stemningslidelser) også oppfyller kriteriene for en annen gruppe lidelser (for eksempel atferdsforstyrrelser) (Skogen et al., 2014). Internasjonale studier rapporterer stor variasjon i årlig prevalens av angstlidelser, fra 2.8 til 18.4% hos 15- til 18-åringer (Blom et al., 2010) og mange unge går med ubehandlede angstproblemer som vedvarer uten behandling (Heiervang et al., 2007). En av de viktigste årsakene til dette er at de unge har problemer med å rapportere deres indre symptomer til sine foresatte og det er vanskelig for foresatte å sette seg inn i barnas indre verden (Baldwin & Dadds, 2007). Dette gjør det enda viktigere å bruke selvrapporteringsskjemaer som kan formidle de unges indre opplevelser av både angst- og depresjonssymptomer. Tre store undersøkelser av ungdommer i Oslo (Ung i Oslo, 1996, 2006 og 2012) har vist en tendens til økning i depresjonsplager (og nedgang i atferdsvansker) blant elever i 9. og 10. klasse samt 1. klasse videregående fra 2006 til 2012 (Skogen et al., 2014). Opptil 30% av alle ungdommer vil oppleve depressive symptomer som nedsatt stemningsleie eller fravær av glede (Dolle et al., 2012) og internasjonale studier viser at ved slutten av ungdomstiden vil mellom 9 og 20% ha oppfylt kriteriene for en depresjonsdiagnose (Blom et al., 2010). En stor bekymring er at depresjon ofte vedvarer inn i voksen alder (Dolle et al., 2012) og forringer livskvaliteten mer enn vanlige medisinske problemer som høyt blodtrykk og hjerteproblemer (Spitzer et al., 1995).

Kjønnsforskjeller i barne- og ungdomspsykiatrien. I følge aktivitetsdata-rapporten for psykisk helsevern (Krogh, 2015) ble det i Helse Nord RHF behandlet 6076 barn og unge i 2014. Kjønnsfordelingen var 52.5% gutter og 47.5% jenter. Nasjonalt var 58.9% 13 år eller eldre og i gruppen fra 16 år og oppover, er det markant flere jenter enn gutter. Gutter fra 7 til 12 år og jenter over 16 år utgjør de klart største pasientgruppene i BUP. For gutter fra 13 år og oppover var ”mistanke om depresjon” den mest dominerende henvisningsgrunnen, og i tillegg var det en del tilfeller av ”mistanke om angstlidelse”. For jenter fra 13 år og oppover var også ”mistanke om depresjon” den dominerende henvisningsgrunnen, tett fulgt av ”mistanke om angstlidelse”, og deretter ”mistanke om spiseforstyrrelse”. Av de 33 883 norske barn og unge som fikk en ICD-10 (World Health Organization, 2012) diagnose på Akse 1 (Klinisk

psykiatrisk syndrom), hadde 8% en diagnose innenfor F30-F39 Affektive lidelser stemningslidelser, og 16% en diagnose innenfor F40-F48 Nevrotiske, belastningsrelaterte og somatoforme lidelser (Krogh, 2015).

Multidimensional Anxiety Scale for Children (MASC)

MASC (March, 1997) ble utviklet for å kunne vurdere fire teoretisk meningsfulle domener av angstsymptomer hos barn og unge: fysiske, kognitive, affektive og atferdsmessige. Det kommer i et selvutfyllende skjema, en foreldreversjon og en lærerversjon.

Psykometriske egenskaper. MASC er validert gjennom tallrike studier.

Faktorstrukturen til skjemaet har blitt kryssvalidert på tvers av alder og kjønn i internasjonale studier, både i kliniske- (March, 1997; Rynn et al., 2006) og normalpopulasjoner, (March, 1997; Baldwin & Dadds, 2007) og i normalpopulasjon i nordiske land som Island (Olason, Sighvartsson, & Smári, 2004) og Sverige (Ivarsson, 2006). March et al. (1999) fant at MASC kunne skille mellom barn (7–9 år) med ADHD diagnose og angstlidelser og dette ble senere replisert i Australia med ungdom (13–17.7 år) av Houghton et al. (2014). Intern reliabilitet viser seg å være fra moderat til strek på tvers av alle delskalaer og respondenter (Baldwin & Dadds, 2007; Dierker et al., 2001; March Parker, Sullivan, Stallings, & Conners, 1997). Test-retest reliabiliteten (Pearson's r) varierer fra tilfredsstillende til utmerket (March et al., 1997; March, Sullivan, & Parker, 1999). MASC har konvergent validitet ved at det korrelerer signifikant med andre mål på angst, som for eksempel Revised Children's Manifest Anxiety Scale (RCMAS) (March et al., 1997; Muris, Merckelbach, Ollendick, King, & Bogie, 2002; Olason et al., 2004; Rynn et al., 2006), Spence Children's Anxiety Scale (SCAS) (Baldwin & Dadds, 2007; Muris et al., 2002), State-Trait Anxiety Inventory for Children (STAIC) (Muris et al., 2002; Rynn et al., 2006) og Screen for Child Anxiety Related Emotional Disorders (SCARED) (Muris et al., 2002). Divergent validitet bekreftes ved at MASC og Children's Depression Inventory (CDI) ikke korrelerer signifikant (March et al., 1997; Olason et al., 2004; Rynn et al., 2006).

Diagnostisk evne. For å undersøke MASCs diagnostiske nøyaktighet, må skjemaet sammenlignes med en gullstandard (Greenhalgh, 1997). Gullstandarden kan være ulike diagnostiske intervju, kliniker-, forsker- eller konsensusdiagnoser. Det finnes ikke i psykiatrien diagnostiske tester som kan lene seg på blodprøver, kjemisk patologi eller radiologi for å finne svar, slik den medisinske tradisjonen kan i mye større grad (Warner, 2004). De ulike målene som brukes som gullstandard kan bare være så gode som det klassifiseringssystemet (ICD-10, DSM-IV) som er konstruert for å operasjonalisere diagnosene. Per i dag er det de kliniske strukturerte intervjuene utviklet av epidemiologer som

klarer dette best og derfor anbefales som gullstandard for psykiatiske diagnoser (Costello, Egger, & Angold, 2005). I studiene som nevnes nedenfor ble ulike former av det strukturerte intervjuet "The Anxiety Disorders Interview Schedule" for DSM-IV (ADIS-C, ADIS-R, ADIS-C/R) brukt som gullstandard. Mål på diagnostisk evne ble testet ved hjelp av receiver operating characteristic (ROC) analyse og blir rapportert som area under the curve (AUC). Generelt blir AUC bedømt som følger: .50–.70 = lav; .70–.90 = moderat, og over .90 indikerer høy nøyaktighet (Pintea & Moldovan, 2009). I en skolebasert studie med amerikanske ungdommer mellom 13–15 år, fant Dierker et al. (2001) at MASC skåren kunne predikere angstlidelser, spesielt hos kvinner og spesifikt for generalisert angstlidelse (GAD) (AUC = .82). Rynn et al. (2006) undersøkte et klinisk utvalg av amerikanske barn og unge med enten diagnostisert angstlidelse eller depresjon, og MASC var i stand til å identifisere individene med angstlidelse (AUC = .742). En nederlandsk studie av totalskåre og alle delskalaene på MASC (van Gastel & Ferdinand, 2008) undersøkte barn og ungdommer mellom 8–18 år i en klinisk setting. Resultatene viste best resultat for delskalaen "Fysiske symptom" som predikerte panikkangst (AUC = .84), "Sosial angst" predikerte sosial angst (AUC = .76), og "Separasjonsangst" predikerte separasjonsangst (AUC = .74). Wei et al. (2014) gjorde en studie på 488 amerikanske ungdommer mellom 13 og 17 år med en diagnostisert angstlidelse og resultatene viste at underfaktorene "Sosial angst" (AUC = .86) og "Separasjonsangst" (AUC = .70) var signifikant prediktive for tilstedeværelsen og alvorlighetsgraden av disse to lidelsene.

Norske studier. Villabø, Gere, & Torgersens (2012) studie av MASCs prediktive evne til diagnose hos norske barn og unge (7–13 år) i poliklinisk behandling, konkluderte med at MASC skåren kan skille mellom de med og de uten angstlidelse (AUC = .70), i tillegg til å korrekt identifisere unge med separasjonsangst (AUC = .75) og sosial fobi (AUC = .73), men mindre presist for GAD. Videre fant de at MASC hadde god intern reliabilitet på tvers av delskalaer og informanter (barn/mor/far). Den norske versjonen av MASC, oversatt av Arnulf Kolstad, er enda ikke ferdig validert i norske utvalg, men flere studier er underveis. Neumer, Richter, & Villabø (2014) tar for seg faktorstruktur og psykometriske egenskaper i et klinisk utvalg med barn fra 7–13 år (N = 279). Resultatene støtter en 3-faktor modell, uavhengig av kjønn, og gir dermed ikke støtte til den opprinnelige 4-faktor modellen til March et al. (1997). De tre faktorene var "Fysiske symptomer", "Sosial angst", og "Unngåelse" (Harm avoidance). En andre studie tar for seg de diagnostiske evnene til MASCs delskalaer, også her på et klinisk utvalg med barn fra 7–13 år i Norge (N = 325), (Neumer, Richter, Villabø,

Arnberg, & Gere, 2014) og resultatene støtter funnene til Villabø, et al. (2012). Ingen studier er så vidt forfatteren bekjent gjort med MASC på ungdommer i klinisk populasjon.

Beck Depression Inventory-II (BDI-II)

BDI-I ble oppdatert i 1996 til BDI-II (Beck et al., 1996) og er ikke basert på en spesifikk teori om depresjon, slik tidligere versjon var (Wang & Gorenstein, 2013). Spørreskjemaet inneholder 21 symptomer (tabell 1) som dekker de ni kriteriene på klinisk depresjon slik det beskrives i American Psychiatric Association's (1994) Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV).

Tabell 1

BDI-II symptomliste

Tristhet	Selvforakt	Ubesluttsomhet
Pessimisme	Selvkritisisme	Verdiløshet
Tidligere nederlag	Suicidale tanker eller ønsker	Tap av energi
Tap av glede	Gråt	Søvnvansker
Skyldfølelse	Agitasjon	Irritabilitet
Følelse av straff	Tap av interesse	Apetittendring
Konsentrasjon	Slitenhet	Tap av seksuell interesse

Selv om BDI-II er ment å vurdere alvorlighetsgraden hos individer allerede diagnostisert med depresjon, viser eksisterende resultater at det også kan anbefales som et screeningsverktøy (Wang & Gorenstein, 2013).

Psykometriske egenskaper. BDI-II har vært i bruk i over 15 år og i en stor review-artikkel av Wang & Gorenstein (2013) samlet de 118 studier gjort på de psykometriske egenskapene til BDI-II fra 1. januar 1996 til 10. oktober 2012. Faktoranalyser viser en robust modell av generell depresjon bestående av to faktorer: kognitiv-affektiv og somatisk-vegetativ. Intern reliabilitet rapporteres til å ligge rundt gjennomsnittlig alfa koeffisient på 0.9, som er en forbedring fra original versjonens gjennomsnitt på 0.85. Resultater for test-retest reliabiliteten viser koeffisienter fra bra til utmerket ($r = 0.73-0.96$), med et gjennomsnittlig retest intervall på 2 uker. Konvergent validitet er høy mellom BDI-I og BDI-II ($r = 0.82-0.94$) og med andre selvutfyllingsskjema som måler depresjon, blant annet Center for Epidemiologic Studies of Depression (CES-D), Hamilton Depression Rating Scale (HAM-D), Zung Self-Rating

Depression Scale (SDS), Montgomery-Åsberg Depression Rating Scale (MADRS), og Geriatric Depression Scale (GDS) ($r = 0.66-0.94$). BDI-II er også signifikant korrelert og har konvergent validitet med skjemaer som måler angst. Beck Anxiety Inventory (BAI), Hamilton Anxiety Rating Scale (HAM-A), og State-Trait Anxiety Inventory (STAI) ($r = 0.37-0.83$), noe forfatterne tolker har sammenheng med at angst og depresjon har mye av det samme symptombildet og at lidelsene har høy komorbiditet med hverandre. Når det kommer til divergent validitet, viser studier med skjemaer som måler substansmisbruk og kronisk smerte, lav korrelasjon ($r < 0.4$). BDI-IIs begreps validitet er snevrere enn hos BDI-I, men fortsatt tilfredsstillende (Wang & Gorenstein, 2013). BDI-II ble opprinnelig konstruert for voksne, men det har også blitt validert i ungdomspopulasjonen, blant andre av Steer, Kumar, Ranieri, & Beck (1998) på polikliniske pasienter. I denne studien ble intern validitet bekreftet, med alfa koeffisient på 0.92 og faktormodellen kunne sammenlignes med tidligere funn for voksne. Symptomoppbyggingen med en kognitiv-affektiv og somatisk-vegetativ inndeling var også lik den funnet for voksne polikliniske pasienter. Etnisitet ble heller ikke funnet å signifikant korrelere med BDI-II skårer og jenter rapporterte høyere skårer enn gutter, gjennomnittlig ca. 5 poeng mer. BDI-IIs evne til å skille mellom diagnostiske kategorier eller grad av depresjon ble ikke testet i denne studien. Osman, Kopper, Barrios, Gutierrez, & Bagge (2004) sin studie av psykiatrisk innlagte ungdommer fant god reliabilitet for totalskåre og delskårer på BDI-II med koeffisienter mellom .72-.91. Konvergent, divergent og diskriminativ validitet ble også bekreftet. Også de fant jenter å ha ca. 5 poeng høyere skåre enn gutter.

Diagnostisk evne. Fordi BDI-II er konstruert for å fange opp DSM-IVs ni kriterier for depresjon, er spesifisiteten blitt bedre enn hos forgjengeren, med konsekvensen at sensitiviteten er gått noe ned. Den rapporteres fortsatt >70% i de fleste studier (Wang & Gorenstein, 2013). AUC rapporteres rundt .75 og høyere. Variasjonen i nøyaktighet avhenger av utvalget (klinisk eller normalpopulasjon), prosent av depressive pasienter og hvilken gullstandard man bruker for klinisk depresjon (Wang & Gorenstein, 2013). Den optimale cut off skåren for å få best mulig sensitivitet og spesifisitet kalkuleres ved hjelp av Youden's Index i ROC analyse (Youden, 1950). Ulike cut off skårer rapporteres i de forskjellige utvalgene. Laveste cut off skåre for å identifisere depresjon fant man i ikke-kliniske utvalg (fra 10 til 16), mellomliggende cut off skårer hos somatisk syke (fra 7 til 20) og høyeste cut off skårer fant man i de psykiatriske utvalgene (fra 19 til 31) (Wang & Gorenstein, 2013). Dette forteller noe om at man må være forsiktig når man tolker totalskåren, alt etter hvilket utvalg man har. I den finske normalpopulasjonen (N = 311) fant Nuevo, Lehtinen, Reyna-

Liberato, & Ayuso-Mateos (2009) en cut off skåre på 17/18 med AUC = .80 (BDI-I), mens Aalto, Elovainio, Kivimäki, Uutela, & Pirkola (2012) noen år senere og med et betydelig større normalutvalg (N = 5561) konkluderte med en optimal cut off skåre på 11/12 med AUC = .92. Viinamäki et al. (2004) undersøkte voksne polikliniske pasienter i Finland og fant cut off skåre på 14/15 med AUC = .81 (BDI-I). Når det gjelder ungdommer, fant Dolle et al. (2012) i en studie av tyske ungdommer som fikk psykiatrisk helsehjelp, en optimal cut off skåre på >23 med AUC = .93 (BDI-II), mens en amerikansk studie som sammenlignet ungdommer i normalpopulasjonen med psykiatriske innlagte ungdommer, konkluderte med at en cut off skåre på 10 (AUC = .77) kunne skille disse to gruppene (BDI-II) (Osman, Barrios, Guttierrez, Williams, & Bailey, 2008). Kumar, Steer, Teitelman, & Villacis (2002) undersøkte også innlagte ungdommer i sin studie og fant AUC = .92 for totalskåre på BDI-II, med cut off på 21 og over. I de forskjellige studiene brukes ulike mål for gullstandarden og i ovenfor nevnte studier finner vi; the Schedules for Clinical Assessment in Neuropsychiatry (SCAN; Nuevo et al., 2009), Structured Clinical Interview for DSM-III-R (SCID-I; Viinamäki et al., 2004), Composite International Diagnostic Interview (CIDI; Aalto et al., 2012), Kinder-DIPS (Dolle et al., 2012), Mood Module from the Primary Care Evaluation of Mental Disorders (PRIME-MD; Kumar et al., 2002) og konsensusdiagnose (Osman et al., 2008). Sistnevnte metode anbefales for klinisk forskning da medlemmene i teamet som setter diagnose har best kunnskap til pasientens psykiatriske historie (anamnestiske opplysninger) og bakgrunn for henvisning. I tillegg bruker de multiple metoder for å innhente og evaluere opplysninger, blant annet semistrukturerte intervju, direkte observasjon, familiesamtaler og pasientenes egen selvrappport (Osman et al., 2008).

Norske studier. Den norske versjonen av BDI-II (Nordhus, Pallesen, & Aasen, 2001) har i en tidligere studie blitt validert i den norske normal- og studentpopulasjonen (Aasen, 2001), men ikke i klinisk populasjon og ikke med ungdommer. Det finnes heller ikke, så langt forfatteren kan finne, norske studier på skjemaets diagnostiske evne per dags dato.

Hensikt og mål for studien

Hensikten med nåværende studie var å undersøke hvor godt de norske versjonene av MASC og BDI-II kan predikere emosjonelle lidelser i et norsk poliklinisk ungdomsutvalg (14 til 18 år), som deltar i et behandlingsprogram basert på kognitiv atferdsterapi. Det er også ønskelig å undersøke hvilken cut off skåre som er optimal for å oppnå best mulig sensitivitet og spesifisitet for skjemaene i dette utvalget, slik at de kan være til nytte i screeningøyemed ved å oppdage diagnose eller behov for videre behandling (Warner, 2004). Dette for å sikre at

barn og unge med angst- og depresjonsplager blir oppdaget tidligst mulig og kan starte med målrettet behandling.

Metode

Deltakere

Deltakerne (heretter kalt pasientene eller ungdommene) er hentet fra SMARTprosjektet ved fire poliklinikker innen psykisk helsevern for barn og unge (BUP) i Nord Norge; BUP Stokmarknes, BUP Tromsø og omegn, BUP Harstad og BUP Bodø, i tidsperioden fra våren-14 til våren-15. Foreldre/foresatte til alle ungdommer mellom 14 til 18 år som ble henvist til en av fire BUP'ene fikk tilsendt en privat kode til det databaserte semi-strukturerte diagnostiske intervjuet Development and Well-being Assessment (DAWBA; Goodman, Ford, Richards, Gatward, & Meltzer, 2000) som inkluderer Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997). I de tilfellene hvor ungdommen skåret seks eller mer på emosjonelle vansker på SDQ fikk ungdommen og de foresatte tilbud om å være med i et seks ukers langt behandlingsprogram basert på kognitiv atferdsterapi for angst og depresjon, kalt SMART (Neumer & Junge-Hoffmeister, 2010). Ekskluderingskriterier fra studien var mindre enn seks på emosjonelle vansker skåre, diagnose på gjennomgripende utviklingsforstyrrelse spekteret, psykotiske symptomer, akutt suicidalitet, bruk av angstdempende medikamenter eller foreldre/foresatte som ikke behersket norsk.

Fra det opprinnelige datasettet ($N = 75$), ble tretten ekskludert på grunn av manglende registrert diagnose, fire på grunn av ukomplett data på MASC/BDI-II (T1), og en fordi det kun var registrert Z-diagnose. Endelig antall med i analyse var $N = 57$. Aldersspennet var fra 14 til 17 år og gjennomsnittsalderen var 15.82 ($SD 1.13$). Av disse var 19.3% 14–14.11 år, 14% var 15–15.11 år, 29.8% var 16–16.11 år, 35.1% var 17–17.11 år, og 1.8% var ukjent. Kjønnfordelingen var 3.5% gutter, 91.2% jenter og 5.3% ukjent. Alle pasientene har fylt ut MASC og BDI-II ved oppstart i SMART (T1) og fått diagnose satt i konsensus.

Måleinstrument

MASC. Multidimensional Anxiety Scale for Children er et selvutfyllingsskjema som inneholder 39 ledd og ble utviklet for å måle angst hos barn og unge mellom 8–19 år (March et al. 1997). Hvert ledd kan skåres på en 4-poengs likert-skala der 0 = stemmer aldri for meg, 1 = stemmer sjelden for meg, 2 = stemmer noen ganger for meg, og 3 = stemmer ofte for meg. MASC har fire hovedfaktorer som er ment å måle forskjellig type angst: Fysiske symptomer, 12 ledd med underfaktorer anspent/rastløs og somatisk/autonom; Unngåelse, 9 ledd med underfaktorer perfeksjonisme og angstmestring; Sosial angst, 9 ledd med underfaktorer

ydmykelse/avvisning og prestasjonsangst, og Separasjons-/Panikkangst, 9 ledd. Summen av alle leddene gir en totalskåre som rapporteres og sammenlignes med standardiserte skårer, der $T = 50$ er gjennomsnitt. I tillegg utviklet March et al. (1997) en skala som består av de ti leddene som skiller best mellom barn med og uten en angstlidelse. Denne heter Anxiety Disorder Index (ADI) og har sin egen totalskåre.

BDI-II. Beck Depression Inventory-II (Beck et al., 1996; Norsk versjon: Nordhus et al., 2001) er et selvutfyllingsskjema og består av 21 ledd av symptomer med en 4-poengs likert-skala fra 0–3. Man skårer etter hvordan man har hatt det de to siste ukene inkludert samme dag. Jo høyere skåre, jo mer er man plaget av symptomet. Høyeste skåre for hvert symptom legges sammen og gir en totalskåre mellom 0–63 poeng. Den reviderte manualen foreslår følgende cut off skårer for å vurdere alvorlighetsgrad for depresjon: 0–13 er ingen til minimal, 14–19 er mild, 20–28 er moderat og 29–63 er alvorlig (Beck et al., 1996).

Prosedyre

SMARTprosjektet er godkjent av Regional komite for medisinsk og helsefaglig forskningsetikk (REK). Det er en kontrollert, randomisert studie hvor to tredjedeler får umiddelbar behandling, mens resterende tredjedel venter i seks uker og utgjør kontrollgruppen (T0), før de inkluderes i behandlingsgruppen. Ventetiden er like lang som behandlingstiden. Ungdommer fra T0 som ikke lengre oppfylte inklusjonskriteriene fikk ordinært tilbud fra BUP og de som fortsatt oppfylte kriteriene inngikk i studien. Før oppstart av SMART måtte alle (også venteliste-pasientene) gjennomgå en utredning (T1) med utfylling av DAWBA med SDQ på nytt og flere selvrapporteringsmål, deriblant MASC og BDI-II, og flere kartleggings skjema som ikke nevnes her. Foreldre/foresatte ble innkalt til en behandler ved BUP for et semi-strukturert intervju hvor anamnesticke opplysninger ble innhentet. For ungdommer over 16 år, som regnes som myndige i helsevesenet i følge norsk lov, ble det innhentet særskilt samtykke for dette intervjuet. Disse kunne også delta i studien uten foresattes samtykke. Alle pasientene ble informert om studiens konfidensialitet, frivillighet og mulighet for å tilbaketrekke samtykke når som helst underveis, uten konsekvenser eller tap av tilbud på BUP. Etter endt behandling gjennomgikk ungdommene en etterundersøkelse (T2). En oppfølgingsundersøkelse seks måneder etter avsluttet behandling (T3) avsluttet datainnsamlingen.

Pasientenes diagnoser ble satt i konsensusmøte bestående av behandlere med ulik profesjonsbakgrunn, blant annet psykologer, pedagoger, sosionomer og leger. Det var alltid en spesialist til stede som har hovedansvaret for diagnosesettingen, enten en psykologspesialist eller en psykiater. Under konsensus ble det lagt frem innsamlet

informasjon om pasienten, hentet fra blant annet DAWBA, anamneseintervju, kartleggingsverktøy som MASC og BDI-II og samtale med foreldre/foresatte og pasienten selv. Informasjonen drøftes tverrfaglig, men spesialist har hovedansvar for diagnosesettingen. Diagnoser ble stilt etter kriterier fra den internasjonale statistiske klassifikasjonen av sykdommer og beslektede helseproblemer versjon 10 (ICD-10; World Health Organization, 2012).

Inndeling av pasientene for analyse. I studien ble de ulike diagnosene delt inn i forskjellige kategorier for analyse; Angst (ulike typer angstlidelser), Depresjon (depressive lidelser av ulik alvorlighetsgrad), Tilpasningsforstyrrelser (ulike F43.xx diagnoser), Atferdsforstyrrelser (ulike F90.xx diagnoser), Andre (obsessiv-kompulsiv lidelse, R-diagnose, ikke-diagnose), og Komorbiditet (blandet angst- og depresjonslidelse eller annen kombinasjon av lidelser). Komorbiditet ble i denne oppgaven registrert når kriteriet for mer enn en diagnose var oppfylt, uten å ta hensyn til ICD-10s retningslinjer om forrang.

Statistiske analyser

Alle statistiske analyser ble utført med IBM SPSS versjon 22 for Macintosh. I denne studien vil MASC og BDI-IIs evne til diagnostisk prediksjon sammenlignes (T1 mål) med konsensusdiagnose (gullstandarden) ved hjelp av receiver operating characteristic (ROC) analyse. En uavhengig t-test analyse ble utført for å sammenligne gjennomsnittsskårene til MASC og BDI-II hos de diagnostisert med angst eller depresjon og de uten disse diagnosene. Fordi det var så få gutter med i studien, var det ikke hensiktsmessig å se på forskjeller mellom kjønnene. Pearson korrelasjonstest ble kjørt mellom MASC skårer og BDI-II skåre for å undersøke samsvar mellom skjemaene.

Analyser av sensitivitet, spesifisitet og diagnostisk nøyaktighet. For å måle MASC og BDI-IIs evner til screening og prediksjon av diagnose, ble sensitivitet, spesifisitet, positiv prediktiv verdi (PPV), negativ prediktiv verdi (NPV), positive likelihood ratio (LR^+), negative likelihood ratio (LR^-) (tabell 2), og area under the curve (AUC) regnet ut.

Tabell 2

Sensitivitet, spesifisitet og prediktive verdier i en 2 x 2 tabell

Gullstandard

Screeningtest	Pasienter med diagnose	Pasienter uten diagnose	Total
Test er positiv	a	b	Positive tester (a + b)
Test er negativ	c	d	Negative tester (c + d)
	Total antall pasienter med diagnose (a + c)	Total antall pasienter uten diagnose (b + d)	Total antall pasienter (a + b + c + d)

Sensitivitet: Andel personer med diagnose som vil få et positivt resultat [$a/(a + c)$].

Spesifisitet: Andel personer uten diagnose som vil få et negativt resultat [$d/(b + d)$].

Positiv prediktiv verdi: Andel personer med et positiv testresultat som faktisk har diagnose [$a/(a + b)$].

Negativ prediktiv verdi: Andel personer med negativt testresultat som ikke har diagnose [$d/(c + b)$].

Utregning av sensitivitet, spesifisitet og prediktive verdier beregnes ut fra informasjon i tabell 2, der kolonnene viser resultatet fra gullstandarden (i denne studien konsensusdiagnose), og radene viser resultatet fra screening testen (i denne studien MASC og BDI-II). Bokstaven 'a' representerer sann positiv, 'b' representerer falsk positiv, 'c' representerer falsk negativ, og 'd' representerer sann negativ. En tests sensitivitet er testens evne til å identifisere de *med* diagnose, fra de som *har* diagnose. Altså andel av de sanne positive riktig identifisert. I tabell 2 vil sensitivitet være $a/(a + c)$. Sensitivitet kan ikke si noe om hvorvidt de som *ikke* har diagnose også tester positivt, eller eventuelt hvor mange. Spesifisitet er testens evne til å identifisere de *uten* diagnose, fra de som *ikke* har diagnose. Altså andel av de sanne negative riktig identifisert. I tabell 2 vil spesifisitet være $d/(b + d)$ (Akobeng, 2006a).

Sensitivitet og spesifisitet kan ikke brukes for å estimere sannsynligheten for diagnose gitt et bestemt svar på screeningtesten, og har derfor redusert klinisk nytteverdi. Til å hjelpe klinikeren i denne bedømmingen kan man regne ut prediktive verdier (Akobeng, 2006a). PPV er andelen personer med positive testresultat, som er korrekt diagnostisert (de *har* diagnose). Fra tabell 2 vil PPV være $a/(a + b)$. NPV er andelen personer med negativt testresultat, som er korrekt diagnostisert (de *har ikke* diagnose). Fra tabell 2 vil NPV være $d/(c + b)$. (Altman &

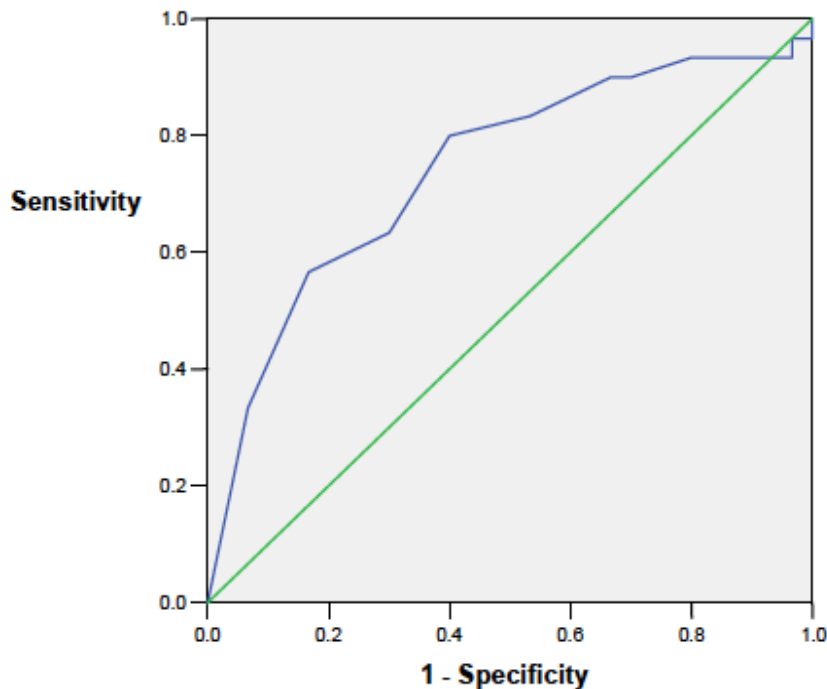
Bland, 1994a). Det er viktig å være klar over at både PPV og NPV varierer i henhold til prevalens av diagnosen i populasjonen de anvendes i. Testens PPV øker ved økt prevalens av lidelser, og NPV minsker ved økt prevalens av lidelser. Dette betyr at man ikke kan generalisere resultat av PPV og NPV fra en populasjon til en annen, hvis prevalensen av lidelsen man ser etter er forskjellig hos de to (Akobeng, 2006a).

Verdier for sensitivitet, spesifisitet, PPV og NPV vil i denne oppgaven transformeres fra tallverdier til prosent.

For å undersøke hvor godt MASC og BDI-II's totalskårer kan skille mellom pasientene med og uten diagnosen vi er interessert i, testenes diagnostiske evne, anvendes receiver operating characteristic (ROC) analyse. Diagnostisk evne refererer til kvaliteten av informasjon innhentet med valgt instrument (Pintea & Moldovan, 2009). ROC analyse ble først utviklet i 1950-årene og ble opprinnelig brukt til evaluering av oppdagede radarsignaler for å skille virkelige objekter fra støy. I nyere tid er det blitt vanlige å bruke denne analyse metoden innen den medisinske tradisjonen (Altman & Bland, 1994b). Litteraturen på ROC analyse indikerer tre hovedbruksområder: 1. Undersøke en tests evne til å skille mellom grupper, 2. Velge den optimale cut off skåren til testen, og 3. Sammenligne prestasjonen til to eller flere tester (Pintea & Moldovan, 2009). Det tredje punktet er ikke aktuelt i denne studien, da de to selvrapporteringskjemaene brukt her er designet til å måle symptomer på to forskjellige lidelser. I en ROC graf (figur 1) blir sensitiviteten (den sanne positive raten) plottet mot den falske positive raten (1 minus spesifisitet: $1 - \text{spesifisitet}$) for de ulike cut off skårene til testen (Warner, 2004). En fordel ved å bruke ROC graf er dets evne til å grafisk beskrive skjemaers evner uten å miste viktig informasjon eller å kreve statistisk ekspertise av leseren (Rynn et al., 2006).

Figur 1

Et eksempel på en ROC graf fra SPSS



Det laveste venstre punktet i grafen (0.0) er punktet som ikke har noen falske positive, men heller ikke noen sanne positive. Det høyeste høyre punktet av grafen (1.1) identifiserer alle sanne positive, men inneholder også alle falske positive. Det øvre venstre punktet (0.1) er det ideelle punktet hvor alle sanne positive er identifisert og ingen falske positive eller falske negative er inkludert. Det laveste høyre punktet (1.0) er det dårligste utfallet hvor alle identifiserte positive er falske positive. Den diagonale linjen i grafen representerer utførelsen fra en tilfeldig test som riktig identifiserer halvparten av de positive og halvparten av de negative. Alle cut off skårer *over* den diagonale linjen vil prestere bedre enn tilfeldig gjetting ($AUC > .50$) og motsatt vil cut off skårer *under* diagonallinjen prestere dårligere enn tilfeldig gjetting ($AUC < .50$) (Pintea & Moldovan, 2009). En ideell test vil ha en cut off skåre med sensitivitet på 1 (100% identifisering av sanne tilfeller), og spesifisitet på 1 (100% ekskludering av ikke-tilfeller), som er det øverste venstre punktet i grafen, men slike tester er sjeldne (Warner, 2004). En god test vil ha en kurve som er godt over den diagonale linjen og så nært det øverste venstre punktet i grafen som mulig. Dette rommet mellom diagonallinjen og det øverste venstre punktet, blir kalt area under the curve (AUC), og måler testens

diagnostiske evne. Jo større området er, jo bedre er testens evne. Som nevnt tidligere tolkes AUC mellom .50 og .70 som lav, AUC mellom .70 og .90 som moderat, og AUC over .90 som god (Pintea & Moldovan, 2009). Youden's index, cut off skåren med den høyeste kombinasjonen av sensitivitet og spesifisitet (Youden, 1950), ble valgt som optimal cut off skåre.

Et annen statistisk mål for å undersøke diagnostisk evne til et selvrappoteringskjema, er sannsynlighetsrater (likelihood ratios). Disse ratene gir for hvert testresultat, en indikasjon på hvor mange flere eller færre ganger det er sannsynlig at en pasient med diagnose er til å ha det gitte testresultatet, enn pasienter uten diagnose (Deeks & Altman, 2004). De kan også brukes til å beregne sannsynligheten for diagnose til individuelle pasienter (Akobeng, 2006b). Sannsynlighetsrater kan oppgis som positive eller negative avhengig av om testresultatet gir positivt eller negativt svar, respektive LR^+ og LR^- . En sannsynlighetsrate *større* enn 1 for et gitt testresultat, indikerer at diagnose er berettiget, og sannsynlighetsrate *under* 1, indikerer at diagnose er fraværende. Rate over 10 og under 0.1 regnes som gode indikasjoner på henholdsvis tilstedeværelse eller fravær av diagnose (Deeks & Altman, 2004). For å beregne LR^+ og LR^- for et testresultat med bare to utfall (diagnose vs. ikke diagnose her) kan man bruke sensitivitet og spesifisitet: $LR^+ = \text{sensitivitet} / (1 - \text{spesifisitet})$, og $LR^- = (1 - \text{spesifisitet}) / \text{sensitivitet}$. Verdiene vil bli tolket etter retningslinjer fra Jaeschke og kollegaer, der tester med LR over 10 eller under 0.1 tyder på store og ofte betydelige endringer fra pre-test til post-test sannsynlighet, LR mellom 5–10 eller 0.1–0.2 tyder på moderat endring, mellom 2–5 eller 0.5–0.2 tyder på liten, men mulig betydningsfull endring og LR mellom 1–2 eller 0.5–1 tyder på liten, sjelden betydelig endring (Jaeschke, Guyatt, & Sackett, 1994).

Konfidensintervall (KI), som uttrykker intervallet der den sanne verdien med stor sannsynlighet befinner seg (Greenhalgh, 1997), vil også bli presentert for alle verdiene i oppgaven. Utregningene ble gjort ved hjelp av internettkalkulator:

<https://www.med.emory.edu/EMAC/curriculum/diagnosis/calculat.htm>

Resultat

Gjennomsnitt og standardavvik for MASC og BDI-II i de ulike diagnostiske gruppene

Gjennomsnittsskåre og standardavvik for MASC totalskåre, MASC ADI skåre og BDI-II totalskåre for hele utvalget ($N = 57$) og de to gruppene med diagnostisert angstlidelse ($n = 12$) og depressiv lidelse ($n = 22$) kan ses i tabell 3.

Tabell 3

Gjennomsnitt og standardavvik for MASC og BDI-II skårer for hele utvalget og pasienter diagnostisert med angst- og depresjonslidelse

	<u>Total (N = 57)</u>	<u>Angst</u>	<u>Depresjon</u>
	M (SD)	M (SD)	M (SD)
MASC totalskåre	58.81 (SD 16.31)	62.08 (SD 14.37)	60.55 (SD 16.61)
MASC ADI skåre	16.44 (SD 4.39)	17.75 (SD 3.14)	16.91 (SD 4.78)
BDI-II totalskåre	30.05 (SD 11.73)	17.50 (SD 10.80)	34.68 (SD 7.13)

M: Gjennomsnittsskåre.

SD: Standardavvik.

Angst: Pasienter diagnostisert med en angstlidelse.

Depresjon: Pasienter diagnostisert med en depressiv lidelse.

MASC: Multidimensional Anxiety Scale for Children.

BDI-II: Beck Depression Inventory-II.

Konsensusdiagnose

Ved tidspunkt for konsensusmøte hadde alle foruten om en pasient (1.8%) grunnlag for ICD-10 diagnose. Diagnosene ble satt på bakgrunn av blant annet innsamlet informasjon fra DAWBA (inkludert SDQ), anamnestic intervju, kartleggingsinstrumenter og samtale. Det var 21.1% av utvalget som tilfredsstilte kriteriene for en angstlidelse, og 38.6% som oppfylte kriteriene for en depressiv lidelse. For mer detaljert oversikt se tabell 4.

Tabell 4

Frekvens av diagnoser hos utvalget

	<i>n</i>	(%)
Angst	12	21.1
<i>Sosial angst</i>	(9/12)	-
<i>Generalisert angst (GAD)</i>	(3/12)	-
Depresjon	22	38.6
<i>Mild depresjon</i>	(8/22)	-
<i>Moderat depresjon</i>	(13/22)	-
<i>Alvorlig depresjon</i>	(1/22)	-
Tilpasningsforstyrrelser	4	7
Atferdsforstyrrelser	2	3.5
Andre	3	5.3
Komorbide Lidelser	14	24.5
Total	57	100

n: Antall personer i utvalget.

(%): Prosent av utvalget.

MASC og BDI-II skårer for de ulike angst- og depresjonslidelsene

For å undersøke MASC og BDI-II skårer for de ulike angst- og depresjonslidelsene ble gjennomsnittsskårer og standardavvik regnet ut for hver av disse. Resultatene kan ses i tabell 5 og 6.

Tabell 5

MASC skårer for de ulike angstlidelsene

	<u>MASC totalskåre</u>	<u>MASC ADI skåre</u>
	M (SD)	M (SD)
Sosial angst	62.33 (15.69)	17.78 (3.31)
Generalisert angst	61.33 (12.22)	17.67 (3.22)

MASC: Multidimensional Anxiety Scale for Children.

M: Gjennomsnittsskåre.

SD: Standardavvik.

Tabell 6

BDI-II skårer for de ulike depressive lidelsene

	<u>BDI-II totalskåre</u>
	M (SD)
Mild depresjon	33.00 (4.41)
Moderat depresjon	35.92 (8.58)
Alvorlig depresjon	32.00 (-)

BDI-II: Beck Depression Inventory-II.

M: Gjennomsnittsskåre.

SD: Standardavvik.

Korrelasjonsanalyse mellom MASC og BDI-II skårer

Ved bruk av Pearson korrelasjon ble skårene til MASC og BDI-II undersøkt. Her forventes det at MASC totalskåre og MASC ADI skåre korrelerer positivt på et signifikant nivå og at de ikke korrelerer med skårer på BDI-II. Som det fremkommer av tabell 7 var

eneste signifikante korrelasjonen mellom MASC totalskåre og MASC ADI skåre ($r = .90$). BDI-II hadde ikke signifikante korrelasjoner med MASC.

Tabell 7

Pearson korrelasjon mellom MASC totalskåre, MASC ADI skåre og BDI-II totalskåre

	MASC totalskåre	MASC ADI skåre	BDI-II totalskåre
MASC totalskåre	–	.90*	.19
MASC ADI skåre	.90*	–	.21
BDI-II totalskåre	.19	.21	–

* $p < 0.01$ (2-halet).

MASC: Multidimensional Anxiety Scale for Children.

BDI-II: Beck Depression Inventory-II.

Uavhengig t-test på gjennomsnittsskårer fra MASC og BDI-II

Som man kan se av tabell 8 var det en signifikant forskjell ($t(55) = 4.98, p = .00$) mellom gjennomsnittsskåren til BDI-II totalskåre for de diagnostisert med en angstlidelse og de uten. Denne forskjellen kunne ikke påvises i verken total- eller ADI skåren på MASC. Det var og en signifikant forskjell ($t(55) = -2.47, p = .02$) mellom gjennomsnittet på BDI-II totalskåren for de diagnostiserte med depressiv lidelse og de uten. Heller ikke her fantes det signifikant forskjell på noen av MASC skårene.

Tabell 8

Uavhengig t-test mellom MASC totalskåre, MASC ADI skåre og BDI-II totalskåre for pasienter med og uten diagnose

	M/U angstlidelse	M/U depressiv lidelse
MASC totalskåre	$t(55) = -.78,$ $p = .44$	$t(55) = -.64$ $p = .53$
MASC ADI skåre	$t(55) = -.17,$ $p = .25$	$t(55) = -.64$ $p = .53$
BDI-II totalskåre	$t(55) = 4.98$ $p = .00^{**}$	$t(55) = -2.47,$ $p = .02^*$

* $p < .05$ (2-halet).

** $p < .001$ (2-halet).

M/U angstlidelser: Pasienter diagnostisert med en angstlidelse sammenlignet med de uten.

M/U depressive lidelser: Pasienter diagnostisert med en depressive lidelse sammenlignet med de uten.

MASC: Multidimensional Anxiety Scale for Children.

BDI-II: Beck Depression Inventory-II.

Diagnostisk evne

For å undersøke MASC og BDI-II's evne til å skille mellom de med og uten angst- og depresjonsdiagnose, ble ROC analyse utført og area under the curve (AUC) regnet ut for totalskårene på begge skjema, i tillegg til ADI skåren fra MASC. Sensitivitet og spesifisitet for de ulike cut off skårene, samt PPV, NPV, LR^+ og LR^- med 95% konfidensintervall kan ses oppsummert i tabell 9. ROC graf for MASC og BDI-II kan ses i figur 2, 3 og 4.

Tabell 9

Cut off skåre, AUC, sensitivitet, spesifisitet, PPV, NPV, LR⁺ og LR⁻ for MASC skårer og BDI-II skåre

	MASC totalskåre		MASC ADI skåre		BDI-II totalskåre	
	Skårer	95% CI	Skårer	95% CI	Skårer	95% CI
Cut off skåre	>50.5	-	>13.5	-	>26.5	-
AUC	0.58	[0.41 – 0.76]	0.61	[0.45 – 0.77]	0.69	[0.55 – 0.83]
Sensitivitet	83%	[78% – 88%]	92%	[88% – 96%]	86%	[83% – 89%]
Spesifisitet	36%	[33% – 39%]	29%	[26% – 32%]	54%	[50% – 58%]
PPV	26%	[22% – 29%]	26%	[23% – 29%]	54%	[50% – 58%]
NPV	89%	[85% – 92%]	93%	[90% – 96%]	86%	[83% – 89%]
LR ⁺	1.30	[1.20 – 1.40]	1.30	[1.20 – 1.40]	1.87	[1.70 – 2.0]
LR ⁻	0.47	[0.35 – 0.65]	0.28	[0.17 – 0.44]	0.26	[0.20 – 0.33]

AUC: area under the curve.

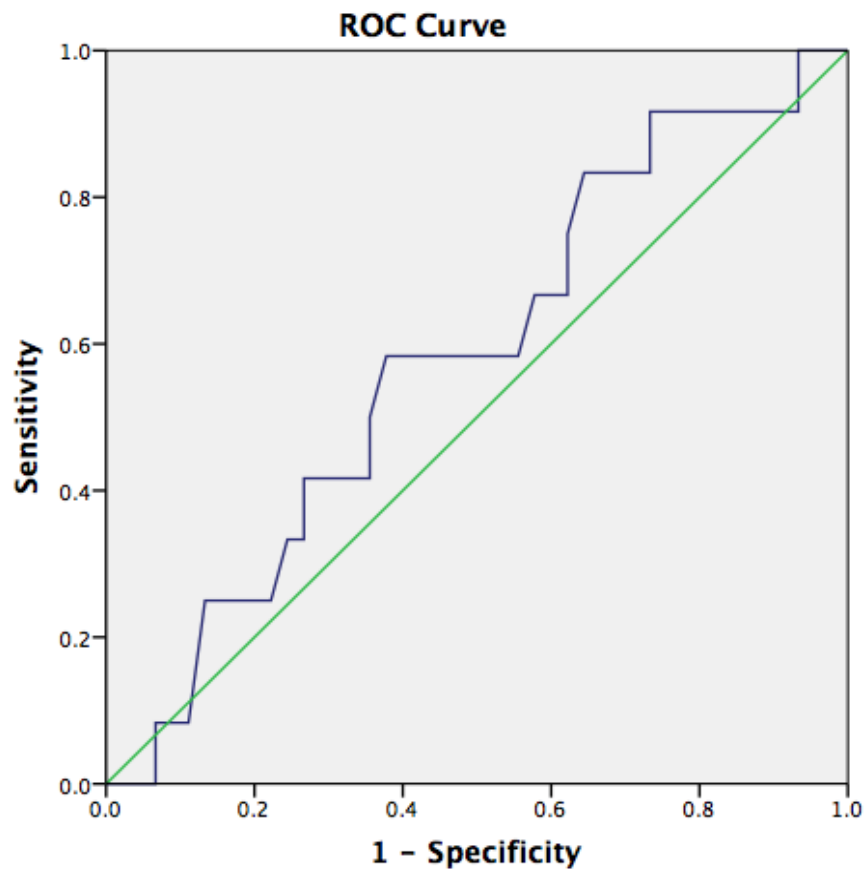
PPV: positiv prediktiv verdi.

NPV: negativ prediktiv verdi.

LR⁺: positiv sannsynlighetsrate.

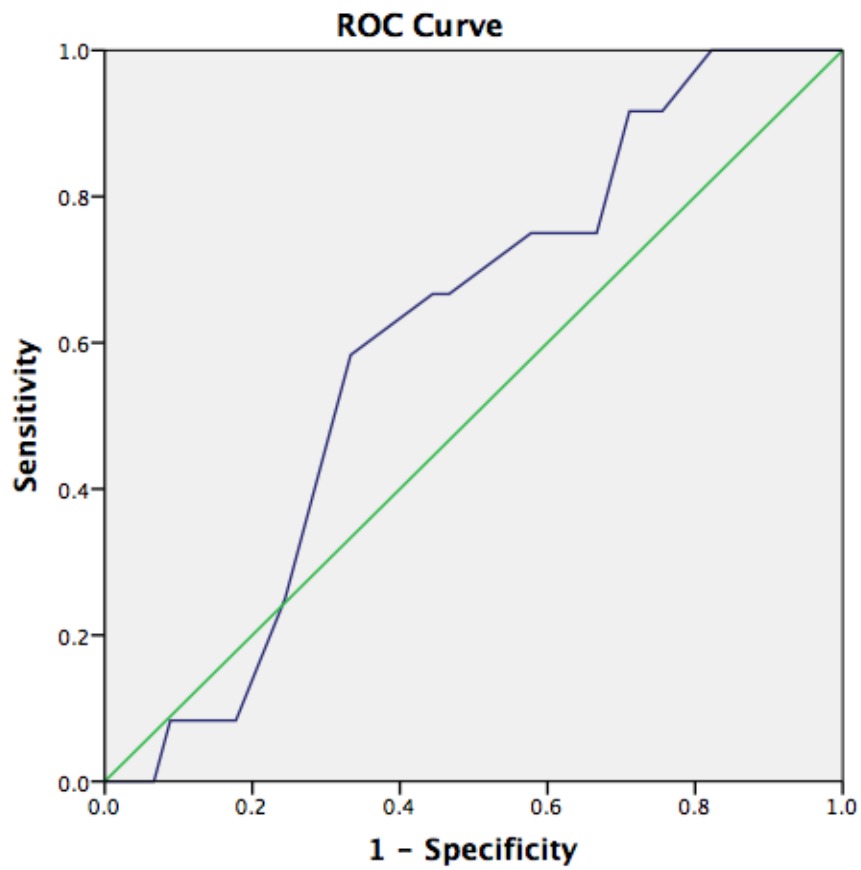
LR⁻: negativ sannsynlighetsrate.

Figur 2

ROC graf for MASC totalskåre

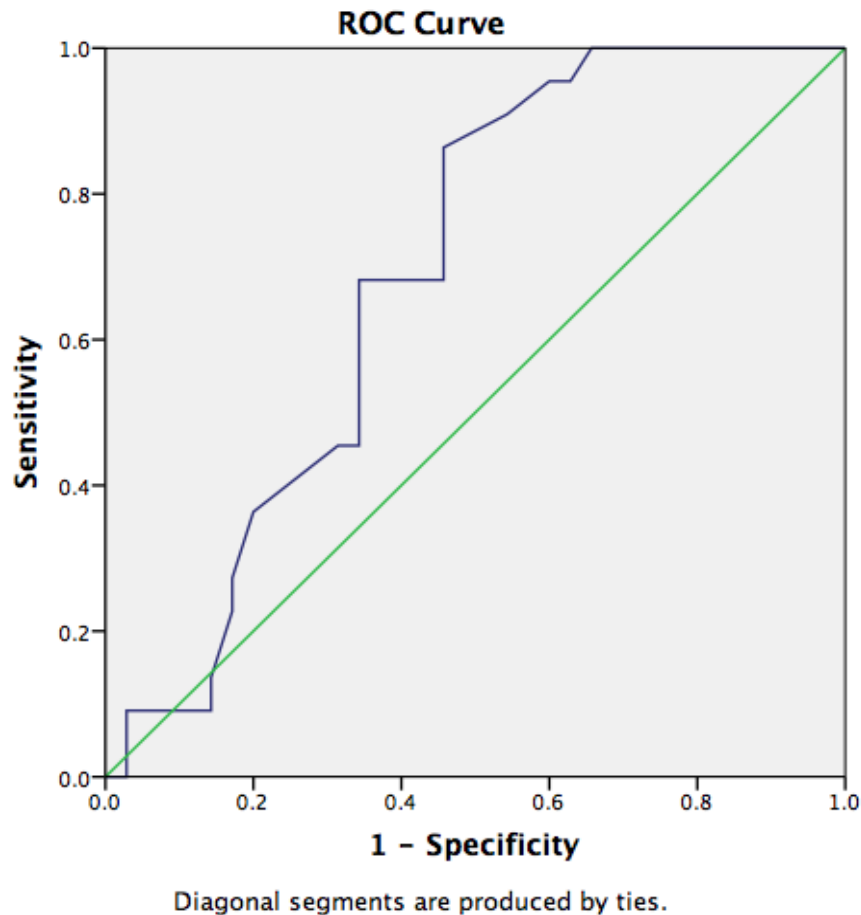
Diagonal segments are produced by ties.

Figur 3

ROC graf for MASC ADI skåre

Diagonal segments are produced by ties.

Figur 4

ROC graf for BDI totalskåre

Diskusjon

Hensikten med denne studien var å undersøke MASC og BDI-II's diagnostiske evne til å predikere henholdsvis angst- og depresjonslidelser i et poliklinisk utvalg av ungdommer (14–18 år) som deltok i et behandlingsprogram basert på kognitiv atferdsterapi. Dette ble gjennomført i en randomisert studie ved å sammenligne MASCs total- og ADI skåre, og BDI-II's totalskåre opp mot diagnoser satt i konsensus (gullstandard) ved fire psykiatriske barne-og ungdomspoliklinikker i Nord-Norge. Per dags dato er det få studier på MASC og BDI-II som er gjort i et norsk poliklinisk ungdomsutvalg og denne studien er et forsøk på å bidra til mer kunnskap om skjemaenes egenskaper her. Denne studien har høy økologisk validitet fordi den utføres i en poliklinisk setting med de selvrapporterings skjema og metode for diagnosesetting som brukes i det daglige. I tillegg finnes det ikke en kontrollgruppe uten diagnoser i denne

studien. Alle deltakerne er pasienter ved BUP, noe som gir et mer reelt bilde av den populasjonen vi ønsker å undersøke MASC og BDI-IIs diagnostiske evner i. Pasientene med diagnostisert angstlidelse hadde ikke høyere gjennomsnittsskåre på noen av MASC skårene sammenlignet med de uten angst. Kun 2 av 10 ungdommer hadde i følge konsensusdiagnose angstlidelse, likevel hadde nesten 7 av 10 ungdommer i studien høyere skåre enn cut off for totalskåren og nesten 8 av 10 høyere på ADI skåren. Resultatene for gjennomsnittsskårer i denne studien er generelt noe høyere enn funn i andre kliniske studier med samme aldersutvalg. Utvalget vårt som helhet hadde en gjennomsnittsskåre for MASC totalskåre på 58.81 (*SD* 16.31) og gjennomsnittsskåre for ADI skåre på 16.44 (*SD* 4.39). Angstpasientene hadde noe høyere skårer med henholdsvis 62.08 (*SD* 14.37) og 17.75 (*SD* 3.14). Wei et al. (2014) fant i sin amerikanske studie av 13- til 17-åringer som deltok i Child/Adolescent Anxiety Multimodal Study (CAMS), en gjennomsnittsskåre på 56.44 (*SD* 19.10), men her var det ikke ”rene” angstpasienter i analysene, slik som i vårt utvalg, men også komorbiditet av angstlidelser med andre lidelser. Den nederlandske studien til van Gastel & Ferdinand (2008) av 11- til 18-åringer henvist til en universitetspoliklinikk for angst- og depresjonslidelser, fant gjennomsnittsskåre på 52.9 (*SD* ukjent), men det nevnes ikke her eventuell komorbiditet av diagnoser. Både Wei et al. og van Gastel & Ferdinand har i tillegg henholdsvis 49.6% og 51.9% gutter i sitt utvalg, og vårt utvalg besto av kun 3.5% gutter. Jenter er i noe grad mer tilbøyelige til å rapportere en angstlidelse (Costello, Egger, & Angold, 2005). Studien til Rynn et al. (2006) av amerikanske barn og unge henvist til ulike poliklinikker for angst- og depresjonsplager er lik vår metodiske fremgangsmåte og resultat. Også her undersøkte de MASC totalskåre og MASC ADI skåre hver for seg og så på resultatene separat for angst- og depresjonspasientene. For hele utvalget fant de gjennomsnittsskåre for totalskåre på 57.8 (*SD* 17.3), som kun er ett poeng lavere enn i denne studien, og gjennomsnittsskåre for ADI skåre på 15.9 (*SD* 4.9), som kun er et halvt poeng lavere enn i vårt utvalg. Hos angstpasientene var også gjennomsnittsskårene tinærmet lik vår studie med henholdsvis 61.2 (*SD* ukjent) og 16.5 (*SD* ukjent). Barbosa, Tannock, & Manassis (2002) hadde også et utvalg ($n = 15$) med kun angstpasienter (dog var alderen 8–12 år) og fant gjennomsnittsskåre på 62.3 (*SD* 14.4), noe som kan tyde på at rene angstpasienter vil skåre høyere enn andre polikliniske pasienter. Dette fant Grills-Taquechel, Ollendick, & Fisak (2008) å være tilfellet i sin studie av 7–17 åringer som var henvist til poliklinikk av ulik årsak, med 45.7 (*SD* 22.87) på MASC totalskåre og 12.29 (*SD* 6.24) på MASC ADI skåre.

Når det gjelder korrelasjon mellom MASC og BDI-II skårer fant vi ikke noen signifikante funn, slik som Rynn et al. (2006) fant i sin studie av ungdommer over tolv år.

Forskjellene i resultat kan bero på at Rynn et al. brukte BDI-I i sitt studie og at utvalget besto av barn og unge som hadde takket ja til å delta i et psykofarmakologisk studie og dermed skiller seg fra vårt utvalg i så måte. Det kan også bety at disse to skjemaene måler forskjellige ting, nemlig symptomer på henholdsvis angst og depresjon. Likevel viste uavhengig t-test at MASC ikke kunne skille mellom de som hadde angstdiagnose og de uten. BDI-II derimot klarte å skille både mellom de med og uten depressive lidelser og mellom de med og uten angstlidelse. Debatten har versert rundt hvorvidt angst og depresjon kun er variasjoner av samme fenomen som ikke kan skilles fra hverandre, eller om de er beslektede fenomen, men med ulik sammensetning av følelser (King, Ollendick, & Gullone, 1991). Senere forskning har samlet bevis for det sistnevnte, og mener angst er sammensatt av mye negative følelsesuttrykk, høy fysiologisk aktivering og normalt nivå av positive følelser, mens depresjon karkarakteriseres av mye negative følelsesuttrykk, normal fysiologisk aktivering, og lavt nivå av positive følelser (Chorpita, Daleiden, Moffitt, Yim, & Umemoto, 2000; Daleiden, Chorpita, & Lu, 2000). I vårt utvalg hadde ti ungdommer en komorbid angst- og depresjonslidelse og det er sannsynlig at flere ungdommer led av symptomer på begge lidelsene uten å oppfylle kriteriene for diagnose på begge. En studie av Haine, Brookman-Frazer, Tsai, Roesch, & Garland (2007) avdekket at mange klinikere satte diagnose basert på deres brede forståelse rundt kategorien til lidelsene ungdommene beskrev. Hvorvidt de oppfylte kriteriene for denne lidelsen var mindre viktig. Mer spesifikt viste det seg for emosjonelle lidelser at verken ungdommens funksjonsnivå eller foreldrenes beskrivelse av problemene hadde noen signifikant sammenheng med diagnosen som ble satt. For angstlidelser var diagnosen positivt relatert til ungdommens rapporterte funksjonsnivå og symptomer, noe som indikerer at for å få angstdiagnose er det essensielt at man rapporterer sine subjektive plager. I en studie av Angold, Costello, Farmer, Burns, & Erkanli (1999) argumenterte forfatterne for at mange av de unge i psykiatrien som ikke oppfyller kriteriene til dokumenterte DSM-III diagnoser, men som har symptomer assosiert med psykososial svekkelse, også tilfredsstiller en psykiatrisk diagnose. Komorbiditeten mellom angst og depresjon kan muligens forklares av at det er lignende etiologiske mekanismer involvert i begge fenomen, (Biederman, Rosenbaum, Chaloff, & Kagan, 1995, ref. i Muris et al., 2002) eller at en av lidelsene øker risikoen for den andre (Cole, Peeke, Martin, Truglio, & Seroczynski, 1998). De deprimerte ungdommene i vårt utvalg hadde nesten like høye gjennomsnittsskårer på MASC som de med angstlidelse, mens de med angst hadde betydelig lavere skårer på BDI-II. En mulig forklaring er at det viser seg at mellom 25–50% av unge med depresjon har komorbid angstlidelse, mens bare 10–15% av unge med angstlidelse har

komorbid depresjon (Axelson & Birmaher, 2001). En annen mulig forklaring på dette fenomenet skisseres i Muris et al. (2002), der skårer fra CDI som måler depresjonssymptomer var høyest korrelert med delskaler som måler mer alvorlig angst, som sosial angst og GAD. Det var nettopp disse to angstlidelsene som var tilstede i vårt utvalg. Angold, Castello, & Erkanli (1999) presiserer også i deres store review av komorbiditet at depresjon og GAD er sannsynlig å overlappe i stor grad fordi de deler mange av de samme symptomene.

Ved valgt optimal cut off for MASC totalskåre på >50.5 ($AUC = .58$) endte vi med en rimelig høy sensitivitet på 83%, men med lav spesifisitet på 36%. Det betyr at en skåre over 50.5 vil korrekt identifisere 83% av de som har en angstlidelse, men bare 36% av de som ikke har en angstlidelse vil skåre under cut off. ADI skåren med optimal cut off på >13.5 ($AUC = .61$) hadde god sensitivitet på 92%, men enda litt lavere spesifisitet enn totalskåren med sine 29%. Rynn et al. (2006) fant i sin studie en cut off >46.67 for GAD, noe som ga sensitivitet på 80% og spesifisitet på 34%. PPV for vår totalskåre og ADI skåre var 26%, som er veldig lavt. Det betyr at bare ca. en fjerdedel av alle de som skårer over cut off faktisk har en angstlidelse. Sagt på en annen måte; hvis man har en skåre over cut off har man likevel bare 26% sjans for å ha en angstlidelse, noe som betyr at det vil være mange falske positive. På den andre siden var NPV rimelig høy for både totalskåren (89%) og ADI skåren (93%), som betyr at 89 av 100 med skåre under cut off på totalskåren og 93 av 100 med skåre under cut off på ADI skåren – ikke har en angstlidelse.

LR^+ for både totalskåre og ADI skåre var ubetydelig (1.3). Pasienter som skårer over cut off har bare en minimal større sjans for å ha en angstlidelse, enn de som skårer under. LR^- for totalskåren var også ubetydelig (0.47), noe som betyr at det bare er ca. to ganger mer sannsynlig at pasienter uten en angstlidelse skårer under cut off, enn de som faktisk har angst. For ADI skåren var LR^- bedre, men liten (0.28), noe som indikerer en tre til fire ganger større sannsynlighet for skåre under cut off for pasienter uten angst, sammenlignet med de med diagnose. ADI skåren på MASC er derfor rimelig godt egnet til å identifisere de som ikke har diagnose. Samlet sett tyder resultatene på at ADI skåren skiller bedre enn totalskåren på pasienter med og uten angstlidelse, noe Villabø et al. (2012) ikke fant for norske barn mellom 7–13 år i poliklinisk behandling. Både cut off for total- og ADI skåre ga dårlig sensitivitet- og spesifisitetsbalanse og konklusjonen er at MASC ikke er et godt nok instrument for å predikere angstlidelser i vår polikliniske ungdomspopulasjon.

Det var 4 av 10 ungdommer i utvalget som ble diagnostisert med en depressiv lidelse og gjennomsnittsskåren på BDI-II var signifikant høyere for dem enn for de uten denne diagnosen. Gjennomsnittsskåren for hele utvalget var 30.05 (SD 11.73) og for pasientene

diagnostisert med en depressiv lidelse var skåren 34.68 (*SD* 7.13). Gjennomsnittsskåren i vår studie er sammenlignbar med den tyske studien av unge mellom 13-16 år i poliklinisk behandling til Dolle et al. (2012), hvor pasientene diagnostisert med depresjon eller dystymi hadde gjennomsnittsskåre på 31.58 (*SD* 9.64). Lignende resultat fant Kumar et al. (2002) for ungdommer innlagt på en psykiatrisk enhet i USA (32.80, *SD* 12.19), mens Steer et al. (1998) fant betydelig lavere skåre (23.04, *SD* 11.40) i deres studie av ungdommer i poliklinisk behandling. Hvis vi ser på gjennomsnittsskåren for de ulike alvorlighetsgradene av depresjon, mild (33.00), moderat (35.92), og alvorlig (32.00) har alle skåre over det som i følge manualen til Beck et al. (1996) skulle tilsi alvorlig depresjon (29–63). Det er tydelig en diskrepans mellom skårene og alvorlighetsgraden på diagnose satt i konsensus, og lignende funn er rapportert av Dolle et al. (2012). De satte spørsmålstegn ved bruk av samme cut off skårer for alvorlighetsgrad i ungdomspopulasjon som er beregnet for voksne i BDI-II manualen. Deres mulige årsaksforklaring på denne diskrepansen er at ungdommer kanskje ikke skiller like godt mellom alvorlighetsgraden på svaralternativene, eller forstår spørsmålene annerledes enn voksne. En annen mulig forklaring i dette utvalget kan være forbundet med diagnostiseringsprosessen. I konsensus hvor diagnosene blir satt brukes informasjon fra det semi-strukturerte diagnostiske intervjuet DAWBA. Litteraturen viser at samsvartaten mellom strukturerte intervju og klinikerbedømmelser generelt er moderat til dårlig og holder seg også slik når klinikerne blir presentert for resultatene fra disse intervjuene (slik som i konsensusdiagnose) (Lewczyk, Garland, Hurlburt, Gearity, & Hough, 2003). Faktorer som kan påvirke denne diskrepansen fra klinikerens ståsted er blant annet fortrolighet med diagnoser, hvilke diagnoser som er sosialt aksepterte, prevalens i det geografiske området, alder og kjønn (Brøndbo, Mathiassen, Martinussen, Handegård, & Kvernmo, 2013). De to sistnevnte fant Brøndbo og kollegaer betydningsfulle for skjevheten i vurderingen mellom DAWBA- og klinikerdiagnoser for emosjonelle lidelser. De konkluderte likevel med at enigheten mellom diagnoser satt av DAWBA og klinikere var på et akseptabelt nivå.

Ved valgt optimal cut off for BDI-II totalskåre på >26.5 (*AUC* = .69) fikk vi en rimelig høy sensitivitet (86%), som betyr at totalskåren vil korrekt identifisere 86 av 100 med en depressiv lidelse. Spesifisiteten var den høyeste i vår studie med 54%, men fortsatt vil ca. halvparten av pasientene som ikke har en depressiv lidelse skåre over cut off. Det vil med andre ord være mange falske positive som også lav PPV (54%) tyder på. NPV for totalskåren på BDI-II var 86%, hvilket betyr at 86% av de som skårer under cut off er sanne negative. De resterende 14% med skåre under cut off har en depressiv lidelse som dermed ikke vil bli

identifisert. Sammenligner vi våre resultat med studien til Dolle et al. (2012), som er den nærmeste med tanke på populasjon og utvalg, fant de med nesten identisk cut off skåre, >26 (for jenter), en AUC på .91 med sensitivitet på 87% og spesifisitet på 89%. Altså et sammenlagt mye bedre resultat enn vårt. Subica et al. (2014) fant samme AUC (.695) som denne studien, men med en lavere cut off skåre på >19, for voksne. Fordi pasienter som ikke hadde diagnostisert depresjon også hadde høy totalskår, var en hypotese at BDI-II i tillegg til kjernesymptomer på depresjon også kartlegger en mer generalisert følelse av lidelse som kan opptre på tvers av flere diagnostiske kategorier. Dette gjør det ekstra utfordrende å skille mellom lidelser som har lignende karaktertrekk, som de internaliserende lidelsene depresjon og angst, på grunn av diagnostisk overlapp. Våre pasienter som kun var diagnostisert med angst hadde gjennomsnittsskåre på 17.50 (*SD* 10.80), hvilket i manualen til Beck et al. (1996) tilsier mild depresjon. Hypotesen til Subica og kollegaer om BDI-II kan muligens dermed adopteres til våre funn.

LR^+ for BDI-II totalskåre var ubetydelig (1.87), hvilket betyr at det bare er to ganger så sannsynlig at pasienter med en depressiv lidelse skårer over cut off, enn de uten. LR^- for totalskåren var liten (0.26), men pasienter som ikke har en depressiv lidelse har fire ganger mer sannsynlig en skåre under cut off enn de som er deprimerte. Dette tyder på, i likhet med MASC, at BDI-II er bedre på å utelukke diagnose, enn å identifisere depresjonslidelse hos ungdommer i vårt polikliniske utvalg. I tillegg ga vår optimale cut off skåre en dårlig sensitivitet-og spesifisitetsbalanse.

Når vi ikke har en AUC verdi som når akseptabelt nivå, kan vi heller ikke bestemme cut off skårer for å identifisere angst- eller depresjonsdiagnose. Ideelt sett vil vi at diagnostiske verktøy skal identifisere alle med diagnose, og hvis vi endret våre cut off skårer ned til >31.5 for MASC totalskåre, >11.5 for MASC ADI skåre og >21.5 for BDI-II, vil vi få en sensitivitet på 100%, og dermed inkludere alle med angst- og depresjonsdiagnose. Problemet er at det vil gå utover spesifisiteten vår som vil synke betraktelig og vi vil få en stor andel falske positive. Dette er ikke akseptabelt for et diagnostisk skjema, men nyttig i screeningøyemed for å velge hvilke pasienter man skal gjøre videre undersøkelser på. Både MASC og BDI-II virker etter resultatene i denne undersøkelsen å dømme og gjøre det best i sistnevnte kategori.

Denne studien har flere svakheter. For det første det relativt få antall individer i hver diagnostiske gruppe, angstlidelser ($n = 12$) og depressive lidelser ($n = 22$). Fordi det var så få individer med ren angstdiagnose var det ikke mulig å se på de ulike delskalaene og få en mer nyansert beskrivelse av den diagnostiske evnen til MASC, slik det er gjort i tidligere norske

studier (Neumer et al., 2014; Villabø et al., 2012). Det begrensede utvalget av depressive pasienter gjorde det ikke mulig å velge cut off skårer for de ulike alvorlighetsgradene av depresjon. For det andre består utvalget nesten utelukkende av jenter så kjønnsforskjeller i skåringen var dermed ikke mulig å analysere. Kjønnsforskjeller er funnet i tidligere studier med barn og unge, både på MASC (Dierker et al., 2001; Villabø et al., 2012) og BDI-II (Dolle et al., 2012; Steer et al., 1998). Videre ble det i denne studien kun brukt informasjon fra selvutfyllingsversjonen til MASC (ungdommen selv) i ROC analysen. Selv om ungdom generelt regnes som den viktigste informanten når det kommer til emosjonelle problemer, spesielt angst (Muris et al., 2002), har tidligere studier rapportert en økt diagnostisk presisjon ved bruk av tilleggsinformasjon fra foreldre/omsorgspersoner (Villabø et al., 2012; Wei et al., 2014).

Videre undersøkelser av MASC og BDI-IIs diagnostiske evne bør replikeres i et større utvalg med alle typer pasienter fra BUP og med en jevnere kjønnsfordeling for mer generaliserbare resultater. Spesielt med tanke på at vi vet at gutter over 13 år hyppigst er henvist til BUP for mistanke om depresjon og angstlidelser.

Vår konklusjon fra denne studien er at verken MASC eller BDI-II presterer godt nok som diagnostiske verktøy til å predikere emosjonelle lidelser i vår polikliniske ungdomspopulasjon. Ved bruk av ovenfornevnte cut off skårer derimot kan skjemaene brukes i en screeningprosess for å identifisere ungdommer med angst- og depresjonssymptomer til videre diagnostisk utredning, i tråd med March (1997) sin anbefaling, og deretter inn i et passende behandlingstilbud.

Referanser

- Aalto, A.-M., Elovainio, M., Kivimäki, M., Uutela, A., & Pirkola, S. (2012). The Beck depression inventory and general health questionnaire as a measure of depression in the general population: A valid study using the composite international diagnostic interview as the gold standard. *Psychiatric Research*, *197*, ss. 163-171.
- Aasen, H. (2001). An empirical investigation of depression symptoms: Norms, psychometric characteristics and factor structure of the Beck Depression Inventory-II (Hovedoppgave i psykologi). Hentet 8 januar, 2015 fra:
<https://bora.uib.no/bitstream/handle/1956/1773/Masteroppgave-aasen.pdf?sequence=1>
- Akobeng, A. K. (2006a). Understanding diagnostic tests 1: Sensitivity, specificity and predictive values. *Acta Pædiatrica*, *96*, ss. 338-341.
- Akobeng, A. K. (2006b). Understanding diagnostic tests 2: Likelihood ratios, pre- and post-test probabilities and their use in clinical practice. *Acta Pædiatrica*, *96*, ss. 487- 491.
- Altman, D. G., & Bland, M. (1994a). Diagnostic tests 2: Predictive values. *British Medical Journal*, *309*, s. 102.
- Altman, D. G., & Bland, M. (1994b). Diagnostic tests 3: Receiver operating characteristic plots. *British Medical Journal*, *309*, s. 188.
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental health disorders* (4th ed.). Washington, DC: Forfatter.
- Angold, A., Costello, E., & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry*, *40* (1), ss. 57-87.
- Angold, A., Costello, E., Farmer, E., Burns, B., & Erkanli, A. (1999). Impaired but undiagnosed. *Journal of the American Academy of Child and Adolescent Psychiatry*, *38* (2), ss. 129-137.
- Axelson, D. A., & Birmaher, B. (2001). Relation between anxiety and depressive disorders in childhood and adolescence. *Depression and Anxiety*, *14*, ss. 67-78.
- Baldwin, J., & Dadds, M. (2007). Reliability and validity of parent and child versions of the multidimensional anxiety scale for children in community samples. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46* (2).
- Barbosa, J., Tannock, R., & Manassis, K. (2002). Measuring anxiety: Parent-child reporting differences in clinical samples. *Depression and Anxiety*, *15*, ss. 61-65.
- Beck, A. T., Steer, R., & Brown, G. (1996). *Beck depression inventory - second edition. Manual*. San Antonio, TX: The Psychological Corporation.

- Blom, E., Larsson, J.-O., Serlachius, E., & Ingvar, M. (2010). The differentiation between depressive and anxious adolescent females and controls by behavioral self-rating scales. *Journal of Affective Disorders, 122*, pp. 232-240.
- Brøndbo, P. H., Mathiassen, B., Martinussen, M., Handegård, B., & Kvernmo, S. (2013). Agreement on diagnoses of mental health problems between an online clinical assignment and a routine clinical assignment. *Journal of Telemedicine and Telecare, 19*, ss. 113-119.
- Chorpita, B. F., Daleiden, E., Moffitt, C., Yim, L., & Umemoto, L. (2000). Assessment of tripartite factors of emotion in children and adolescents I: Structural validity and normative data on an affect and arousal scale. *Journal of Psychopathology and Behavioral Assessment, 22*, ss. 141-160.
- Cole, D. A., Peeke, L., Martin, J., Truglio, R., & Seroczynski, A. (1998). A longitudinal look at the relation between depression and anxiety in children. *Journal of Consulting and Clinical Psychology, 66*, ss. 451-460.
- Costello, E. J., Egger, H. L., & Angold, A. (2005). 10-year research update review: The epidemiology of child and adolescent psychiatric disorders: I. Methods and public health burden. *Journal of the American Academy of Child and Adolescent Psychiatry, 44* (10), ss. 972-986.
- Costello, E. J., Egger, H. L., & Angold, A. (2005). The developmental epidemiology of anxiety disorders: Phenomenology, prevalence, and comorbidity. *Child and Adolescent Psychiatric clinics of North America, 14*, ss. 631-648.
- Daleiden, E., Chorpita, B., & Lu, W. (2000). Assessment of tripartite factors of emotion in children and adolescents II: Concurrent validity of the affect and arousal scales for children. *Journal of Psychopathology and Behavioral Assessment, 22*, ss. 161-182.
- Deeks, J. J., & Altman, D. (2004). Diagnostic tests 4: Likelihood ratios. *British Medical Journal, 329*, ss. 168-169.
- Dierker, L. C., Albano, A., Clarke, G., Heimberg, R., Kendall, P., Merikangas, K., et al. (2001). Screening for anxiety and depression in early adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry, 40* (8).
- Dolle, K., Schulte-Körne, G., O'Leary, A., von Hofacker, N., Izat, Y., & Allgaier, A.-K. (2012). The Beck depression inventory-II in adolescent mental health patients: Cut-off scores for detecting depression and rating severity. *Psychiatric Research, 200*, ss. 843-848.

- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*, 581-586
- Goodman, R., Ford, T., Richards, H., Gatward, R., & Meltzer, H. (2000). The development and well-being assessment: Description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry*, *41*, 645-655.
- Greenhalgh, T. (1997). How to read a paper. Papers that report diagnostic or screening tests. *British Medical Journal*, *315*.
- Grills-Taquechel, A. E., Ollendick, T., & Fisak, B. (2008). Reexamination of the MASC factor structure and discriminant ability in a mixed clinical outpatient sample. *Depression and Anxiety*, *25*, ss. 942-950.
- Haine, R. A., Brookman-Fraze, L., Tsai, K. H., Roesch, S. C., & Garland, A. F. (2007). Clinician perspectives of diagnosis and perceived client change in "real world" psychotherapy for youth emotional and behavioral disorders. *Journal of Child and Family Studies*, *16*, ss. 712-728.
- Heiervang, E., Stormark, K., Lundervold, A., Heinmann, M., Goodman, R., Posserud, M-B., et al. (2007). Psychiatric disorders in Norwegian 8- to 10-year-olds: An epidemiological survey of prevalence, risk factors, and service use. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46* (4), ss. 438-447.
- Houghton, S., Hunter, S., Trewin, T., Glasgow, K., & Carroll, A. (2014). The multidimensional anxiety scale for children: A further validation with Australian adolescents with and without ADHD. *Journal of Attention Disorders*, *18* (5), ss. 402-411.
- Ivarsson, T. (2006). Normative data for the multidimensional anxiety scale for children (MASC) in Swedish adolescents. *Nordic Journal of Psychiatry*, *60* (2).
- Jaeschke, R., Guyatt, G., & Sackett, D. (1994). User's guide to the medical literature III. How to use an article about a diagnostic test B. What are the results and will they help me in caring for my patients? *The Journal of the American Medical Association*, *271* (9), ss. 703-707.
- King, N. J., Ollendick, T., & Gullone, E. (1991). Negative affectivity in children and adolescent: Relations between anxiety and depression. *Clinical Psychology Review*, *11*, ss. 441-459.
- Krogh, F. (2015). *Aktivitetsdata for psykisk helsevern for barn og unge 2014. Norsk pasientregister*. Oslo: Helsedirektoratet.

- Kumar, G., Steer, R., Teitelman, K., & Villacis, L. (2002). Effectiveness of Beck depression inventory-II subscales in screening for major depressive disorders in adolescent psychiatric inpatients. *Assessment, 9* (2), ss. 164-170.
- Lewczyk, C. M., Garland, A., Hurlburt M. S., Gearity, J., & Hough, R. (2003). Comparing DISV-IV and clinician diagnoses among youths receiving public mental health services. *Journal of the American Academy of Child and Adolescent Psychiatry, 42* (3), ss. 349-356.
- March, J. S. (1997). *Manual for the multidimensional anxiety scale for children (MASC)*. Toronto: Multi-Health Systems.
- March, J. S., Conners, C., Arnold, G., Epstein, J., Parker, J., Hinshaw, S., et al. (1999). The multidimensional anxiety scale for children (MASC): Confirmatory factor analysis in a pediatric ADHD sample. *Journal of Attention Disorders, 3* (2).
- March, J. S., Parker, J., Sullivan, K., Stallings, P., & Conners, K. (1997). The multidimensional anxiety scale for children (MASC): Factor structure, reliability and validity. *Journal of the American Academy of Child and Adolescent Psychiatry, 36* (4).
- March, J. S., Sullivan, K., & Parker, J. (1999). Test-retest reliability of the multidimensional anxiety scale for children. *Journal of Anxiety Disorders, 13* (4), ss. 349-358.
- Mathiesen, K. S. (2009). *Psykiske lidelser i Norge: Et folkehelseperspektiv. Del 2: Barn og unge (Report No. 2009:8)*. Nasjonalt folkehelseinstitutt.
- Mathiesen, K. S., Karevold, E., & Knudsen, A. (2009). *Psychiatric impairment among children and adolescents in Norway*. Oslo: Nasjonalt folkehelseinstitutt.
- Muris, P., Merckelbach, H., Ollendick, T., King, N., & Bogie, N. (2002). Three traditional and three new childhood anxiety questionnaires: Their reliability and validity in a normal adolescent sample. *Behavior Research and Therapy, 40*, ss. 753-772.
- Neumer, S.-P., Richter, J., & Villabø, M. (2014). Factor structure and psychometric properties of the multidimensional anxiety scale for children (MASC) in a clinical sample. *Posterpresentasjon ved EABCT konferanse*. Abstractband.
- Neumer, S.-P., Richter, J., Villabø, M., Arnberg, K., & Gere, M. (2014). Cross validation of the MASC subscales in identifying separation anxiety disorder (SAP), social phobia (SOP), and generalized anxiety disorder (GAD). *Posterpresentasjon ved EABCT konferanse*. Abstractband.
- Nordhus, I. H., Pallesen, S., & Aasen, H. (2001). *Norsk versjon av Beck depression inventory-II*. Universitetet i Bergen, Institutt for Klinisk Psykologi, Psykologisk fakultet. Universitetet i Bergen.

- Nuevo, R., Lehtinen, V., Reyna-Liberato, P., & Ayuso-Mateos, J. (2009). Usefulness of the Beck depression inventory as a screening method for depression among the general population of Finland. *Scandinavian Journal of Public Health*, 37, ss. 28-34.
- Olason, D. T., Sighvartsson, M., & Smári, J. (2004). Psychometric properties of the multidimensional anxiety scale for children (MASC) among Icelandic schoolchildren. *Scandinavian Journal of Psychology*, 45, ss. 429-436.
- Osman, A., Barrios, F., Guttierrez, P., Williams, J., & Bailey, J. (2008). Psychometric properties of the Beck depression inventory-II in non-clinical adolescent samples. *Journal of Clinical Psychology*, 64 (1), ss. 83-102.
- Osman, A., Kopper, B., Barrios, F., Gutierrez, P., & Bagge, C. (2004). Reliability and validity of the Beck depression inventory-II with adolescent psychiatric inpatients. *Psychological Assessment*, 16 (2), ss. 120-132.
- Pintea, S., & Moldovan, R. (2009). The receiver-operating characteristic (ROC) analysis: Fundamentals and applications in clinical psychology. *Journal of Cognitive and Behavioral Psychotherapies*, 9 (1), ss. 49-66.
- Norsk Psykologforening (u.d.). *Fakta om psykiske lidelser*. Hentet 7 november, 2014 fra: <http://www.psykol.no/Foreningen/For-pressen/Fakta-om-psykiske-lidelser>
- Rynn, M. A., Barber, J., Khalid-Khan, S., Siqueland, L., Dembiski, M., McCarthy, K., et al. (2006). The psychometric properties of the MASC in a pediatric psychiatric sample. *Anxiety Disorders*, 20, ss. 139-157.
- Skogen, J. C., Kjeldsen, A., Knudsen, A., Myklestad, I., & Nesvåg, R. *Folkehelseinstituttet.no*. Hentet 24 oktober, 2014 fra Kapittel i Folkehelse rapporten 2014: <http://www.fhi.no/artikler/?id=110703>
- Spitzer, R. L., Kroenke, K., Linzer, M., Hahn, S., Williams, J., deGruy III, F., et al. (1995). Health-related quality of life in primary care patients. Results from the PRIME-MD 1000 study. *Journal of the American Medical Association*, 274, ss. 1511-1517.
- Steer, R. A., Kumar, G., Ranieri, W., & Beck, A. (1998). Use of the Beck depression inventory with adolescent psychiatric outpatients. *Journal of Psychopathology and Behavioral Assessment*, 20 (2), ss. 127-137.
- Subica, A. M., Fowler, J., Elhai, J., Frueh, B., Sharp, C., Kelly, E., et al. (2014). Factor structure and diagnostic validity of the Beck depression inventory-II with adult clinical inpatients: Comparison to a gold-standard diagnostic interview. *Psychological Assessment*, 26 (4), ss. 1106-1115.

- van Gastel, W., & Ferdinand, R. (2008). Screening capacity of the multidimensional anxiety scale for children (MASC) for DSM-IV anxiety disorders. *Depression and Anxiety, 25* (12), ss. 1046-1052.
- Viinamäki, H., Tanskanen, A., Honkalampi, K., Koivumaa-Honkanen, H., Haatainen, K., Kaustio, O., et al. (2004). Is the Beck depression inventory suitable for screening major depression in different phases of the disease? *Nordic Journal of Psychiatry, 58* (1), ss. 49-53.
- Villabø, M., Gere, M., & Torgersen, S. (2012). Diagnostic efficiency of the child and parent versions of the multidimensional anxiety scale for children. *Journal of Clinical Child & Adolescent Psychology, 41* (1), ss. 75-85.
- Wang, C. E., & Nordahl, H. (2012). Nye utfordringer for psykologer i arbeidet med depresjon? *Tidsskrift for Norsk Psykologforening, 49*, ss. 2-3.
- Wang, Y.-P., & Gorenstein, C. (2013). Psychometric properties of the Beck depression inventory-II: A comprehensive review. *Revista Brasileira de Psiquiatria, 35*, ss. 416-431.
- Warner, J. (2004). Clinician's guide to evaluating diagnostic and screening tests in psychiatry. *Advances in Psychiatric Treatment, 10*, ss. 446-454.
- Wei, C., Hoff, A., Villabø, M., Peterman, J., Kendall, P., Piacentini, J., et al. (2014). Assessing anxiety in youth with the multidimensional anxiety scale for children. *Journal of Clinical Child & Adolescent Psychology, 43* (4), ss. 566-578.
- Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer, 3*, ss. 32-35.