

En innføring i depresjonsvurdering med PHQ-9

Helsemyndighetene anbefaler strukturert kartlegging av depresjon. PHQ-9 er et alternativ som er enkelt å ta i bruk.

TEKST

Tomas Formo Langkaas

Even Rognan

Sverre U. Johnson

PUBLISERT 1. mars 2021

ABSTRACT:

An introduction to assessment of depression with PHQ-9

Introduction: Assessment of depression is a routine procedure in clinical practice in Norway. National guidelines (Helsedirektoratet, 2009) recommend the use of measurement instruments in assessment of depression. PHQ-9 is widely used in research and practice. The official PHQ-9 manual provides practical guidance on interpreting test results with the use of clinical cutoff scores and a diagnostic algorithm for DSM-IV.

Methods: Based on clinical practice and research, we summarise and provide guidance on the practical use of PHQ-9 beyond that offered in the official PHQ-9 manual, and applied to a Norwegian context.

Results: We provide a general introduction to diagnostic assessment of depression and the limited role of measurement instruments in such assessments. We describe how the original diagnostic algorithm can be adapted to ICD-10 criteria, how to apply clinical significance to using PHQ-9 as a feedback instrument to monitor treatment progress, and how to interpret results with missing answers.

Discussion: Finally, we discuss the shortcomings of relying on measurement instruments in assessment of depression, and conclude that PHQ-9 is better suited for ordinary practice than other instruments recommended in the national guidelines.

Keywords: PHQ-9, diagnostic assessment, progress and outcome monitoring

Depresjon er en av de vanligste psykiske lidelsene (World Health Organization, 2017). Vurdering av depresjon er følgelig en rutineoppgave i klinisk psykologisk arbeid i Norge, også fordi depresjon er en relevant risikofaktor å kartlegge ved vurdering av selvmordsrisiko (Helsedirektoratet, 2008). Nasjonale retningslinjer for diagnostisering og behandling av voksne med depresjon i primær- og spesialisthelsetjenesten (Helsedirektoratet, 2009) anbefaler bruk av strukturerte kartleggingsverktøy ved vurdering av depresjon. I klinisk praksis er det i hovedsak tre aktuelle bruksområder for slike verktøy: (1) som forhåndskartlegging (screening) for å identifisere pasienter med mulig udiagnostisert depresjon, (2) for strukturert kartlegging av symptomkriterier ved

diagnostisk vurdering av depresjon og (3) som kvantitativt mål på grad av depresjonssymptomer i vurdering av behandlingsutvikling. Et aktuelt alternativ i arbeidet med alle disse tre områdene er PHQ-9^[1] (Kroenke, Spitzer & Williams, 2001).

For praktisk bruk gir den offisielle instruksjonsmanualen til PHQ-9 (Pfizer, u.å.) begrenset informasjon. Manualen beskriver hvordan svarene på PHQ-9 kan summeres og hvordan resultatene kan sammenliknes med kliniske grenseverdier for å sannsynliggjøre mulig depresjon. I tillegg beskriver manualen en diagnostisk algoritme for å sammenholde svarene med diagnostiske kriterier for depresjon i Diagnostic and Statistical Manual of Mental Disorders (4th ed.; DSM-IV; American Psychiatric Association, 1994). Så langt vi kjenner til, finnes det ikke noen eksisterende norsk instruksjonsmanual eller manualer som tar for seg praktisk bruk av PHQ-9 utover det som allerede beskrives i den offisielle instruksjonsmanualen.

Med utgangspunkt i egen praktisk erfaring med utredning og behandling av depresjon i spesialisthelsetjenesten for voksne og vår egen bakgrunnskunnskap om forskning og praksis, sammenfatter vi her relevant praktisk informasjon som ikke står i den offisielle instruksjonsmanualen for PHQ-9, og som er tilpasset norske forhold. Vi beskriver hvordan strukturerte kartleggingsverktøy som PHQ-9 kan inngå som del av en helhetlig diagnostisk vurdering av depresjon; vi presenterer en diagnostisk algoritme tilpasset International Classification of Diseases, 10th Revision (ICD-10; World Health Organization, 1992), som er gjeldende diagnosesystem i Norge; vi viser hvordan PHQ-9 kan anvendes som tilbakemeldingsverktøy med bruk av klinisk signifikans (Jacobson & Truax, 1991), og vi viser hvordan resultater på PHQ-9 også kan tolkes ved ufullstendig besvarte skjemaer. Avslutningsvis drøfter vi sentrale begrensninger ved bruk av strukturerte kartleggingsverktøy i vurdering av depresjon.

Om PHQ-9

Patient Health Questionnaire (PHQ; Spitzer, Kroenke & Williams, 1999) ble opprinnelig utviklet for rask og effektiv kartlegging av fem vanlige psykiske helseproblemer i primærhelsetjenesten: depresjon, angst, somatisering, alkohol- og spiseproblemer. PHQ-9 er depresjonsmodulen i PHQ og er egnet som selvstendig spørreskjema for vurdering av depresjon. Skjemaet består av 9 spørsmål som er direkte knyttet til diagnosekriteriene i DSM-IV, og skjemaet er mye brukt internasjonalt, både i forskning og i klinisk praksis. Eksempelvis anbefaler offisielle britiske retningslinjer å bruke PHQ-9 ved depresjonsvurdering (National Institute for Health and Care Excellence, 2011). American Heart Association anbefaler både PHQ-9 og kortversjonen PHQ-2 for vurdering av depresjon ved samtidig hjertesykdom (Lichtman et al., 2008). PHQ-9 brukes som standard spørreskjema for måling av depresjonsgrad og endring i depresjon i det engelske tjenestetilbudet Increased Access to Psychological Therapies (IAPT; National Collaborating Centre for Mental Health, 2018). Etter modell fra IAPT er PHQ-9 også brukt i tjenestetilbudet Rask psykisk helsehjelp her i Norge (Smith, Alves & Knapstad, 2016). PHQ-9 er lisensfri, og offisielle oversettelser for ulike språk – inkludert norsk bokmål – ligger fritt tilgjengelig på nett (<https://www.phqscreeners.com>).

Svarene på PHQ-9 kan sammenfattes på to måter (Pfizer, u.å.): (a) Verdien av avgitte svar kan summeres, noe som gir et kvantitativt mål på depresjonsgrad på en skala fra 0 til 27, eller (b) svarene kan sammenholdes direkte med diagnosekriterier for depresjon ved bruk av en diagnostisk algoritme. Måleegenskaper ved PHQ-9 er undersøkt i flere studier og utvalg. Kroenke, Spitzer og Williams (2001) oppgir intern konsistens på 0,83 og en 48 timers test-retest-reliabilitet på 0,84. En nyere psykometrisk metaanalyse (Levis, Benedetti & Thombs, 2019) anslår sensitivitet på 0,88 og spesifisitet på 0,85 gitt en nedre klinisk grenseverdi (clinical cutoff score) på 10 og bruk av semistrukturert klinisk intervju som referansestandard.

Diagnostisk vurdering av depresjon

Depresjon er en tilstand som kjennetegnes av overdreven nedstemthet, nedsatt interesse og nedsatt glede i en grad og et omfang som begrenser daglig funksjon. Tilstanden overskrider hva som kan forstås som en naturlig reaksjon på den aktuelle livssituasjonen. Depresjon er slik sett noe annet enn å være misfornøyd, trist eller ulykkelig som en naturlig og rimelig konsekvens av aktuelle omstendigheter.

Omfattende forskningsinnsats har hittil ikke bidratt til å identifisere noen pålitelige biologiske markører for depresjon (Strawbridge, Young & Cleare, 2017), og det er omdiskutert om det er realistisk å skulle forklare depresjon som en underliggende biologisk tilstand (Bekhuis, Hartman, Boschloo & Lucassen, 2019). Klinisk vurdering av om diagnosekriteriene er oppfylt er det som er referansestandard for å diagnostisere depresjon, slik kriteriene er beskrevet i ICD-10 og DSM-5. I praksis dreier dette seg om å vurdere tre spørsmål: (1) Er symptomer som kjennetegner depresjon til stede, (2) bidrar symptomene til å forklare aktuell funksjonsbegrensning, og (3) er depresjon den beste forklaringen på symptom bildet? Strukturerte kartleggingsverktøy – som PHQ-9 – kan primært bistå for å avklare det første spørsmålet.

Diagnosesystem	ICD-10	DSM-IV/DSM-5
Symptomvarighet	Minst 2 uker	Minst 2 uker
Kjernesymptomer	Tap av interesse eller glede (1)	Tap av interesse eller glede (1)
	Nedstemthet (2)	Nedstemthet (2)
	Redusert energi eller økt tretthet (4)	Redusert energi eller økt tretthet (4)
Tilleggssymptomer	Søvnvansker (3)	Søvnvansker (3)
	Endret matlyst (5)	Endret matlyst (5)
	Tap av selvtillit og selvfølelse	Følelse av verdiløshet eller overdreven eller upassende skyldfølelse (6)
	Overdreven selvbebreidelse eller overdreven eller upassende skyldfølelse (6)	Konsentrasjonsvansker (7)
	Konsentrasjonsvansker (7)	Endret psykomotorisk tempo (8)
	Endret psykomotorisk tempo (8)	Selvmondsatferd eller tilbakevendende tanker om død eller selvmord (9)
	Selvmondsatferd eller tilbakevendende tanker om død eller selvmord (9)	
Symptomkrav	Minst 2 kjernesymptomer og minst 4 symptomer totalt	Minst ett kjernesymptom og minst 5 symptomer totalt

TABELL 1 Symptomer på depressiv episode med tilhørende spørsmål i PHQ-9 angitt i parentes

Tabell 1 beskriver symptomer som kjennetegner depressiv episode i ICD-10, DSM-IV og DSM-5. Tabellen viser at symptomer på depresjon har vært uendret i de to siste versjonene av DSM, og at det er stor grad av sammenfall mellom DSM og ICD-10. Hovedkjennetegn ved depresjon – nedstemthet og tap av interesse eller glede – gjenspeiles i kjernesymptomer i begge diagnosesystemene. ICD-10 regner i tillegg redusert energi eller økt tretthet som et kjernesymptom, mens dette er kun et tilleggssymptom i DSM. De fleste tilleggssymptomene er også felles, med unntak av tap av selvtillit og selvfølelse, som ikke inngår i DSM. Det er også en liten forskjell i hvilken kombinasjon av symptomer som er minstekrav for en depressiv episode. I ICD-10 holder det med 4 symptomer, der minst 2 må være kjernesymptomer. I DSM er det krav om 5 symptomer, og minst ett av dem må være et kjernesymptom.

Det andre spørsmålet er om pasientens aktuelle funksjonsbegrensninger er knyttet til depresjonssymptomer. Dersom det ikke foreligger noen funksjonsbegrensning, eller den ikke er knyttet til depressive symptomer, faller naturlig nok grunnlaget for å vurdere depresjon bort.

Det siste spørsmålet – om depresjon er den beste forklaringen på symptom bildet – er det som stiller størst krav til faglig vurdering, og her kommer begrensningene ved strukturerte kartleggingsverktøy tydelig frem. Nedstemthet og mangel på tiltak og glede er en normalreaksjon på vanlige livsbelastninger og livskriser uten å være tegn på en sykdomstilstand som krever behandling av helsepersonell. På samme måte kan disse symptomene være en normalreaksjon på funksjonstap knyttet til andre samtidige lidelser, for eksempel funksjonsbegrensende angsttilstander. Enkelte somatiske tilstander kan gi depresjonsliknende symptomer (slik som stoffskifteproblemer). Det samme kan hjerneorganiske lidelser (som demens) eller aktiv bruk av rusmidler. Fysisk utmattelse og søvndeprivasjon kan også gi depresjonsliknende symptomer. Når det er flere mulige forklaringer på aktuelle depresjonssymptomer, er det derfor nødvendig med en vurdering av om disse tilstrekkelig kan forklare tilstandsbildet, eller om tilstanden bedre forstås som uttrykk for en samtidig depresjon. Det er følgelig alltid behov for faglig vurdering, og strukturerte kartleggingsverktøy som PHQ-9 kan først og fremst bidra med å samle relevant informasjon som danner grunnlag for slik vurdering.

Hvor ofte har du vært plaget av ett eller flere av de følgende problemene i løpet av de siste 2 ukene?	Ikke i det hele tatt	Noen dager	Mer enn halvparten av dagene	Nesten hver dag
1. Lite interesse for eller glede over å gjøre ting	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input checked="" type="checkbox"/> 2	<input type="checkbox"/> 3
2. Følt deg nedfor, deprimeret eller fylt av håpløshet	<input type="checkbox"/> 0	<input checked="" type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
3. Vansker med å sovne eller med å sove natten gjennom uten å våkne – eller med at du sover for mye	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input checked="" type="checkbox"/> 2	<input type="checkbox"/> 3
4. Følt deg trett eller slapp	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input checked="" type="checkbox"/> 3
5. Dårlig appetitt eller å spise for mye	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input checked="" type="checkbox"/> 2	<input type="checkbox"/> 3
6. Vært misfornøyd med deg selv eller følt deg mislykket – eller følt at du har sviktet deg selv eller familien din	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input checked="" type="checkbox"/> 2	<input type="checkbox"/> 3
7. Vansker med å konsentrere deg om ting, slik som å lese avisen eller se på TV	<input type="checkbox"/> 0	<input checked="" type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
8. Beveget deg eller snakket så langsomt at andre kan ha merket det? Eller motsatt – følt deg så urolig eller rastløs at du har vært mye mer i bevegelse enn vanlig	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input checked="" type="checkbox"/> 3
9. Tanker om at du like gjerne kunne vært død eller på annen måte ville skade deg selv	<input type="checkbox"/> 0	<input checked="" type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3
Hvis du har opplevd ett eller flere av de problemene som nevnes, i hvor stor grad har problemene gjort det vanskelig for deg å utføre arbeidet ditt, ordne med ting hjemme eller å komme overens med andre?	<input type="checkbox"/> Ikke vanskelig i det hele tatt	<input type="checkbox"/> Litt vanskelig	<input checked="" type="checkbox"/> Veldig vanskelig	<input type="checkbox"/> Ekstremt vanskelig

PHQ-9 er utviklet av Dr. Robert L. Spitzer, Dr. Janet B. Williams, Dr. Kurt Kroenke og medarbeidere med et utdanningsstipend fra Pfizer Inc. Det er ikke nødvendig med tillatelse til reproduksjon, oversettelse, fremvisning eller distribusjon

FIGUR 1 Eksempel på utfylt PHQ-9

Merknad: Rettighetshaverne til PHQ-9 har gjort PHQ-9 fritt tilgjengelig for reproduksjon, oversettelse, fremvisning og distribusjon (www.phqscreeners.com). PHQ-9 kan derfor gjengis og anvendes fritt.

PHQ-9 i praktisk bruk

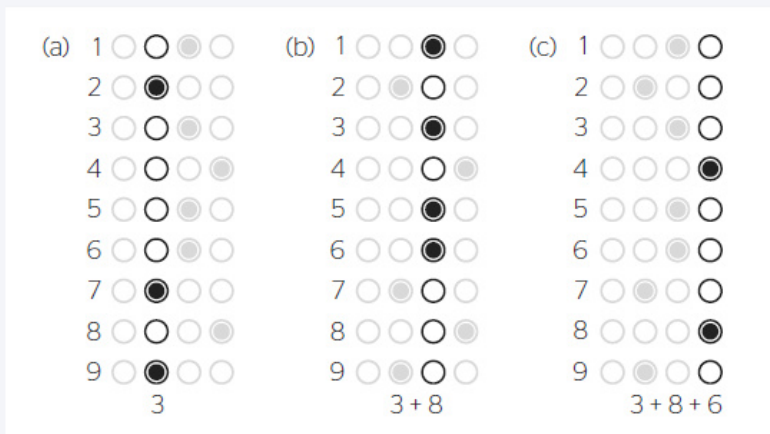
PHQ-9 besvares som et vanlig spørreskjema. Figur 1 viser et eksempel på et besvart skjema. Det tar normalt noen få minutter å fylle ut skjemaet, og dette kan inngå som en del av en konsultasjon, eller skjemaet kan fylles ut i forkant. For fremmedspråklige pasienter kan det være aktuelt å besvare spørsmålene på eget morsmål der det finnes tilgjengelige oversettelser. PHQ-9 kan også gjennomføres som et kort klinisk intervju

med pasienter, der det vurderes som mer aktuelt. Dette kan være i saker der man ønsker å være ekstra oppmerksom på at pasienten får med seg og forstår spørsmålene, for eksempel hos pasienter med lesevansker, konsentrasjonsvansker eller ved mulig redusert kognitiv funksjon.

Spørsmålene dreier seg om omfang av symptomer på depresjon i løpet av de siste to ukene. Hvert spørsmål har 4 faste svaralternativer: «Ikke i det hele tatt», «Noen dager», «Mer enn halvparten av dagene» og «Nesten hver dag». Svaralternativene gir henholdsvis 0, 1, 2 og 3 poeng, og svarene i figur 1 kan slik sett gjengis i kortform som 2 1 2 3 2 2 1 3 1. Tilleggsspørsmålet om funksjonsvansker bidrar med relevant informasjon for å vurdere om det foreligger noen funksjonsbegrensende tilstand eller ikke.

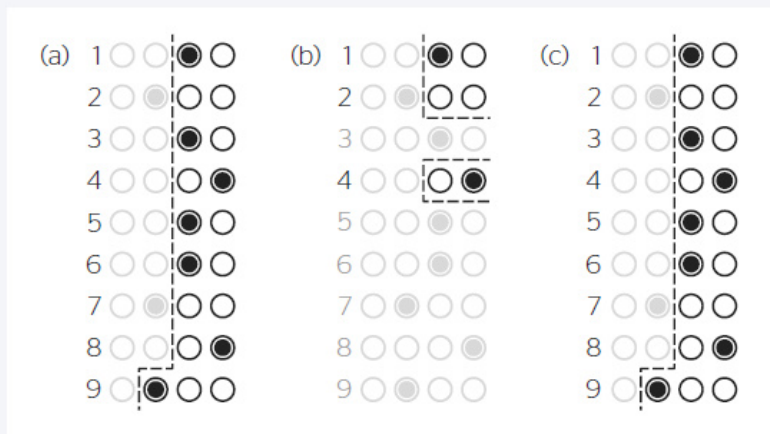
PHQ-9 for forhåndskartlegging

For bruk ved forhåndskartlegging kan sammenlagt resultat av avgitte svar på PHQ-9 sammenliknes med kliniske grenseverdier for å gi en omtrentlig pekepinn på om, og i hvilken grad, en pasient er deprimert. Omtrentlige nedre grenseverdier er 5, 10 og 20 for mulig mild, moderat og alvorlig depresjon (Kroenke et al., 2001), mens resultater under 5 tyder på ingen depresjon. For kun dikotom vurdering av mulig depresjon er 10 en egnet nedre grenseverdi (Kroenke et al., 2001; Levis, Benedetti & Thombs, 2019; Manea, Gilbody & McMillan, 2012). Figur 2 illustrerer hvordan svar kan sammenfattes og vurderes på denne måten.



FIGUR 2 Forhåndskartlegging av mulig depresjon med PHQ-9

Merknad: Summering av svarene fra figur 1 for å illustrere forhåndskartlegging (screening) med PHQ-9. Sammenlagt resultat av (a) alle ettpoengs svar, (b) alle topoengs svar og (c) alle trepoengs svar blir her $3 + 8 + 6 = 17$. Omtrentlige nedre grenseverdier er henholdsvis 5, 10 og 20 for mulig mild, moderat og alvorlig depresjon. Et resultat på 17 sammenfaller med mulig moderat depresjon.



FIGUR 3 Vurdering av diagnosekriterier for depresjon i ICD-10 med PHQ-9

Merknad: Vurdering av eksempelsvarene i figur 1 opp mot diagnostiske kriterier i ICD-10. (a) For hvert topoengs eller trepoengs svar regnes ett symptom å være til stede, med unntak av spørsmål 9, hvor også ettpoengs svar teller. (b) Spørsmål 1, 2 og 4 angir kjernesymptomer, og minst 2 av disse må være til stede, slik som her. (c) Om nok kjernesymptomer er til stede, telles totalt antall symptomer opp. Her er totalt 7 til stede. Mild og moderat depresjon krever minst 2 kjernesymptomer og henholdsvis 4 eller 6 symptomer totalt. Alvorlig depresjon krever alle kjernesymptomer og minst 8 totalt. Her er kriteriene for moderat depresjon oppfylt.

Bruk av PHQ-9 for diagnostisk kartlegging

Instruksjonsmanualen (Pfizer, u.å.) beskriver også en diagnostisk algoritme for å vurdere svarene på PHQ-9 direkte opp mot symptomkriteriene for depresjon i DSM. For hvert spørsmål besvart med «Mer enn halvparten av dagene» (topoengssvar) eller høyere regnes ett symptom å være til stede, med unntak av spørsmål 9 hvor «Noen dager» (ettpoengssvar) regnes som tilstrekkelig. Som vist i tabell 1 stiller ICD-10 litt andre krav til symptomkombinasjon enn DSM, noe som gir grunnlag for en tilsvarende diagnostisk algoritme tilpasset ICD-10:

1. For hvert spørsmål besvart med «Mer enn halvparten av dagene» eller mer (topoengssvar), regnes ett symptom å være til stede. For spørsmål 9 er det nok med «Noen dager» (ettpoengssvar).
2. Spørsmål 1, 2 og 4 angir kjernesymptomer ved depresjon. Minst 2 av disse må være til stede ved mild og moderat depresjon, mens alle 3 må være til stede ved alvorlig depresjon.
3. I tillegg må minst totalt 4, 6 eller 8 symptomer være til stede – inkludert kjernesymptomer – ved henholdsvis mild, moderat og alvorlig depresjon.

Figur 3 illustrerer vurdering av diagnostiske kriterier fra svarene i figur 1.

Selvmondsrisiko

Spørsmål 9 dreier seg om tanker om død eller selvskaade og kan brukes som informasjon for å vurdere nærmere kartlegging av selvmordsrisiko. Dersom spørsmålet er ubesvart, eller svaret er noe annet enn «Ikke i det hele tatt», kan det være grunn til å undersøke dette spørsmålet mer. Det kan være en god rutine å alltid undersøke svaret på dette spørsmålet, og kanskje spesielt legge merke til om det er ubesvart.

PHQ-9 som tilbakemeldingsverktøy

For pasienter som følges opp for depresjon, kan summen av avgitte svar brukes som et kvantitativt mål på depresjonsgrad. Ved å svare på PHQ-9 gjentatte ganger underveis i behandling gir dette grunnlag for å følge med på behandlingsutvikling. Klinisk signifikans (Jacobson & Truax, 1991) er en standardisert metode brukt i psykoterapiforskning for å anslå den praktiske kliniske betydningen av endringer hos enkeltpasienter. Med utgangspunkt i to tall som er beregnet fra normdata, (a) minste målbare endring og (b) grenseverdi som angir område for normalt funksjonsnivå, kan målt endring siden behandlingsstart kategoriseres som enten verre, uendret, bedre eller frisk (Langkaas, Johnson, Stenshorne, Collin-Tiller & Vrabel, 2017). For PHQ-9 kan minste målbare endring anslås til ± 6 og område for normal funksjon til ≤ 6 (se Langkaas et al., 2017, for detaljer og beregningsgrunnlag). Ved å sammenholde en pasients resultat med pasientens resultat ved behandlingsstart kan dermed pasientens aktuelle tilstand vurderes som

- verre ved økning på 6 poeng eller mer

- uendret ved mindre enn 6 poengs endring
- bedre ved nedgang på 6 poeng eller mer og siste resultat over 6
- frisk ved nedgang på 6 poeng eller mer og siste resultat lik 6 eller lavere

Tabell 2 viser resultatområder for de ulike endringskategoriene med utgangspunkt i resultat ved behandlingsstart, og tabellen kan brukes som referanse for å vurdere endring hos en pasient. Eksempelvis må en pasient med et utgangspunkt på 17 ved behandlingsstart oppnå et resultat i området 0–6 for å regnes som frisk. Pasientens tilstand vil regnes som bedre ved resultater i området 7–11, uendret ved resultater i området 12–22, og som verre ved resultater over 22.

PHQ-9 ved behandlings-start	PHQ-9 resultatområder			
	Frisk	Bedre	Uendret	Verre
0			0-5	6-27
1			0-6	7-27
2			0-7	8-27
3			0-8	9-27
4			0-9	10-27
5			0-10	11-27
6	0		1-11	12-27
7	0-1		2-12	13-27
8	0-2		3-13	14-27
9	0-3		4-14	15-27
10	0-4		5-15	16-27
11	0-5		6-16	17-27
12	0-6		7-17	18-27
13	0-6	7	8-18	19-27
14	0-6	7-8	9-19	20-27
15	0-6	7-9	10-20	21-27
16	0-6	7-10	11-21	22-27
17	0-6	7-11	12-22	23-27
18	0-6	7-12	13-23	24-27
19	0-6	7-13	14-24	25-27
20	0-6	7-14	15-25	26-27
21	0-6	7-15	16-26	27
22	0-6	7-16	17-27	
23	0-6	7-17	18-27	
24	0-6	7-18	19-27	
25	0-6	7-19	20-27	
26	0-6	7-20	21-27	
27	0-6	7-21	22-27	

TABELL 2 Resultatområder for å vurdere klinisk signifikant endring av depresjon med PHQ-9

Merknad: Tallene er basert på en minste målbar endring på 6 poeng og et normalt funksjonsområde med en øvre grenseverdi på 6.

Vurdering av ufullstendig besvarte skjemaer

Fra tid til annen oppstår det behov for å sammenfatte spørreskjemaer der pasienter av ulike grunner unnlater å svare på enkeltpørsmål. Dette representerer en praktisk utfordring siden de ulike måtene å sammenfatte spørreskjemaer på normalt forutsetter fullstendig besvarte skjemaer. Ufullstendig besvarte skjemaer gir likevel grunnlag for å

avgrense og vurdere mulige resultater. Ved kvantitativ sammenfatning angir det faktiske totalresultatet laveste mulige resultat, mens høyeste mulige resultat er gitt ved faktisk resultat + antall ubesvarte $\times 3$. Eksempelvis vil svarene 2 _ 0 2 0 _ _ 3 2 ha et faktisk resultat på 9 og et mulig resultat i området 9–18. Dette indikerer mild til moderat depresjon som mer sannsynlig resultat enn alvorlig eller ingen depresjon.

På samme måte er det også mulig å avgrense hva som er aktuelle diagnostiske muligheter ved ufullstendig besvarte skjemaer. I eksempelet med svarene 2 _ 0 2 0 _ _ 3 2 er allerede to kjernesymptomer til stede og fire symptomer totalt. Dette betyr at disse svarene i det minste passer med diagnostiske kriterier for mild depresjon. Med tre ubesvarte spørsmål ville det være mulig med høyst syv symptomer totalt, noe som ville passe med kriterier for moderat depresjon. Slik sett utelukker dette alvorlig og ingen depresjon som muligheter.

PHQ-9 sammenliknet med aktuelle alternativer

Fried (2017) viser hvordan spørsmålene i ulike strukturerte kartleggingsverktøy for depresjon i begrenset grad samsvarer på tvers av slike verktøy, noe som potensielt bidrar til ulike resultater i kartlegging av depresjon. PHQ-9 har her en generell fordel av å ha spørsmål som er direkte knyttet til diagnosekriterier for depresjon, og kan dermed sammenholdes direkte med disse.

Nasjonale retningslinjer for utredning og behandling av depresjon (Helsedirektoratet, 2009) anbefaler ulike strukturerte kartleggingsverktøy som Montgomery-Åsberg Depression Rating Scale (MADRS; Montgomery & Åsberg, 1979), Beck Depression Inventory-II (BDI-II; Beck, Steer & Brown, 1996), Geriatric Depression Scale (GDS; Yesavage et al., 1982), WHO-5 (Topp, Østergaard, Søndergaard & Bech, 2015), Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983), Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I; First, Spitzer, Gibbon & Williams, 1996) og Mini International Neuropsychiatric Interview (M.I.N.I.; Sheehan et al., 1998).

BDI-II, HADS og SCID-I er rettighetsbeskyttet, og av disse har vi kun klart å finne offisiell norsk versjon for BDI-II. I norske versjoner av GDS (Yesavage, 1983) og WHO-5 (World Health Organization, u.å.) er det anbefalt å følge opp med MADRS eller PHQ-9 ved behov for å vurdere depresjon, noe som i praksis betyr at GDS og WHO-5 er overflødige for dette formålet. BDI-II, MADRS og M.I.N.I. ser slik sett ut til å være de aktuelle alternativene til PHQ-9 i praksis.

Sammenliknet med BDI-II har PHQ-9 fordelene av å være kostnadsfri i bruk, ha langt færre spørsmål (BDI-II har 21 spørsmål), alle spørsmål er knyttet til diagnostiske kriterier, og tolkning er ikke kun basert på grenseverdier. Sammenliknet med MADRS har PHQ-9 fordelene av å være laget for selvutfylling, med klart definerte spørsmål og svaralternativer, noe som gjør at psykometriske egenskaper ikke avhenger av en godt trent intervjuer. Sammenliknet med M.I.N.I. er PHQ-9 enkel å sammenholde med diagnostiske kriterier for ICD-10, mens vurdering av depresjon i M.I.N.I. kun er basert på DSM-IV og ikke lar seg tilpasse til ICD-10. Sammenliknet med alle disse alternativene har PHQ-9 fordelene av å ha offisielle oversettelser for ulike språk fritt tilgjengelig på

nett, noe som ofte gjør det mulig å la fremmedspråklige pasienter besvare spørsmålene på eget morsmål.

Hva med ICD-11?

ICD-11 er nå utgitt i en første forberedende versjon (<https://icd.who.int/>), men det kan fortsatt ta mange år før ICD-11 er innført i Norge og overtar som gjeldende diagnosesystem. Så langt inneholder ICD-11 enkelte mindre endringer i diagnosekriteriene for depresjon. Redusert energi eller økt tretthet vil ikke lenger regnes som et kjernesymptom, noe som vil bidra til bedre harmonisering med DSM-5. Håpløshet innføres som nytt symptom på depresjon. Reed et al. (2019) beskriver dessuten at minste antall symptomer som kreves, vil økes til 5 – i samsvar med DSM-5 – men dette gjenspeiles så langt ikke i den siste versjonen av ICD-11 på nett (<https://icd.who.int/>). Slik diagnosekriteriene ser ut så langt, vil PHQ-9 fortsette å være egnet for bruk i klinisk praksis også etter at ICD-11 offisielt tas i bruk.

Sentrale begrensninger ved bruk av PHQ-9 i praksis

Et godt validert selvrapportskjema med relativt høy sensitivitet og spesifisitet – slik som PHQ-9 – har fortsatt omfattende begrensninger dersom det brukes alene i forsøk på å identifisere klinisk depresjon. Positiv prediktiv verdi er et mål på hvor sannsynlig det er at et positivt testresultat faktisk stemmer. Levis og kollegaer (2019) anslår at for PHQ-9 gir den anbefalt nedre kliniske grenseverdien på 10 en positiv prediktiv verdi på 0,49 ved bruk i primærhelsetjenesten. I tilfeller der slik forhåndskartlegging med PHQ-9 indikerer mulig klinisk depresjon, vil det derfor være 51 % sannsynlig at dette faktisk ikke stemmer. Gitt at vi kan anta at det er høyere forekomst av depresjon i klinisk psykologisk arbeid i spesialisthelsetjenesten vil positiv prediktiv verdi her sannsynligvis være noe høyere.

Anslagene til Levis og kollegaer (2019) tilsvarer en negativ prediktiv verdi på 0,98. Det vil si at i tilfeller der slik forhåndskartlegging ikke indikerer mulig klinisk depresjon, vil det være 98 % sannsynlig at dette faktisk stemmer. Oppsummert viser da anslagene til Levis og kollegaer (2019) at PHQ-9 brukt alene i hovedsak egner seg for å utelukke klinisk depresjon. Mitchell, Yadegarfar, Gill og Stubbs (2016) kom til samme konklusjon og viste samtidig til at PHQ-2 – som kun består av de to første spørsmålene på PHQ-9 – med en nedre klinisk grenseverdi på 2 er tilstrekkelig godt egnet for å utelukke klinisk depresjon i praksis. Dette samsvarer godt med de diagnostiske algoritmene for PHQ-9 – det kreves minst ett topoengs svar på de to første spørsmålene for å oppfylle diagnostiske krav til kjernesymptomer ved depresjon både i DSM og ICD-10.

Det kan være verdt å merke seg at en vurdering basert på kliniske grenseverdier ikke nødvendigvis sammenfaller med vurdering av diagnostiske kriterier. I spørsmål om diagnostikk gir derfor PHQ-9 bedre informasjon når svarene sammenholdes direkte med diagnosekriteriene, enn ved å tolke resultatet ut fra de omtrentlige nedre grenseverdiene 5, 10 og 20. I teorien kan forskjellen være stor, noe vi kan illustrere med svarene 1 1 3 3 3 3 3 3 0. Med et sammenlagt resultat på 20 kunne disse svarene indikere

mulig alvorlig depresjon, men ved å sjekke svarene opp mot diagnosekriteriene ser vi at tilstrekkelig antall kjernesymptomer på depresjon ikke er til stede. Det eneste kjernesymptomet til stede er redusert energi eller økt tretthet (spørsmål 4), noe som ikke holder for å stille en depresjonsdiagnose. Ved slike svar vil det være naturlig å søke etter andre forklaringer enn depresjon for det høye resultatet. Dette illustrerer en sentral begrensning ved alle strukturerte kartleggingsverktøy som kun bruker grenseverdier for å indikere sannsynlig diagnose – det er ikke gitt at et resultat over en klinisk grenseverdi faktisk representerer et tilfelle der diagnosekriteriene er oppfylt.

I klinisk praksis betyr dette at det er begrenset nytte av å anvende strukturerte kartleggingsverktøy med kliniske grenseverdier for å identifisere depresjon. Siden svarene på PHQ-9 kan sammenholdes direkte med diagnosekriterier for depresjon, vil vi derfor hevde at det er større praktisk nytte av å anvende en diagnostisk algoritme i kombinasjon med en helhetlig klinisk vurdering fremfor å anvende kliniske grenseverdier.

En generell begrensning ved strukturerte kartleggingsverktøy er at de ikke er tilstrekkelige for å gjøre en diagnostisk avklaring. Det er i tillegg behov for en faglig vurdering av all tilgjengelig informasjon, inkludert andre mulige forklaringer, for å avgjøre om det foreligger en behandlingstrengende depresjonstilstand. Dette samsvarer også med konklusjonene ovenfor om at PHQ-9 brukt alene i hovedsak egner seg for å utelukke depresjon, men ikke identifisere depresjon. Vi har beskrevet at tilstandsbildet må vurderes opp mot øvrig livssituasjon, og vi har beskrevet hvordan bruk av grenseverdier for å sannsynliggjøre diagnoser i enkelte tilfeller kan være misvisende. Fra eget klinisk arbeid har vi også erfaring med pasienter som besvarer PHQ-9 på en måte som samsvarer med diagnostiske kriterier for alvorlig depresjon, men der dette ikke understøttes av det kliniske inntrykket og øvrig informasjon. I slike tilfeller er det naturlig å undersøke andre forklaringer på pasientens svar.

I likhet med mange andre oversatte kartleggingsverktøy foreligger det så langt ikke noen norsk valideringsstudie av PHQ-9. Av kartleggingsverktøy for depresjon kjenner vi kun til at det er gjort norsk validering av BDI-II (Aasen, 2001). En rekke faktorer peker mot at PHQ-9 likevel egner seg for bruk under norske forhold. Innholdsmessig fremgår det at spørsmålene på PHQ-9 samsvarer med kriteriene i diagnosesystemene, og det er kjent at resultater på PHQ-9 i stor grad sammenfaller med resultater på BDI-II (Titov et al., 2011). PHQ-9 er allerede i bruk i Rask psykisk helsehjelp, og vi er kjent med at Modum Bad og Folkehelseinstituttet har foretatt en uavhengig oversettelse og tilbakeoversettelse av PHQ-9 (tilgjengelig på <https://www.kognitiv.no/>). Denne oversettelsen ser ut til å samsvare godt med den offisielle norske versjonen.

Når det gjelder bruk av kartleggingsverktøy for å vurdere endring, er det verdt å påpeke at observasjon av endring i seg selv ikke kan gi noe informasjon om virkning av behandling. Dette gjelder også ved observasjon av eventuell klinisk signifikant endring. Klinisk signifikans er kun en metodisk tilnærming for å vurdere endring. Metodikken er ikke egnet til å si noe om virkning (Langkaas, Wampold & Hoffart, 2018).

Tall å huske	Forklaring
9	Spørsmål 9 gir informasjon om behov for å undersøke selvmordsfare.
5, 10, 20	Nedre grenseverdier for rask omtrentlig vurdering av henholdsvis mild, moderat eller alvorlig depresjon.
1, 2, 4	Spørsmål 1, 2 og 4 angir kjernesymptomer for depresjon i ICD-10.
2, 4, 6, 8	I ICD-10 må minst 2 kjernesymptomer være til stede og samlet sett minst 4, 6 eller 8 symptomer for henholdsvis mild, moderat eller alvorlig depresjon (alvorlig depresjon krever alle kjerne-symptomene).
± 6, ≤ 6	For vurdering av klinisk signifikant endring regnes minste målbare forskjell å være på 6, og øvre grense for normal funksjon er på 6. Ved en målbar nedgang der siste resultat er innenfor området for normal funksjon, regnes endringen som klinisk signifikant.

TABELL 3 Viktige tall for å vurdere svar på PHQ-9

Konklusjon

PHQ-9 er et spørreskjema som er allsidig og anvendelig i vanlig klinisk praksis. Skjemaet er enkelt å besvare, er gratis å bruke og tar kort tid å vurdere. Tabell 3 gir en kort sammenfatning av viktige tall en må huske for å vurdere svarene. PHQ-9 kan fint brukes i papirform, og anvendes både som spørreskjema og intervju. Bruksområdet er både i konsultasjon og mellom konsultasjoner, for forhåndskartlegging av depresjon, diagnostisk vurdering, samt vurdering av endring. For fremmedspråklige pasienter kan skjemaet ofte gis på morsmålet.

For depresjonsvurdering, som ofte er en rutineoppgave i klinisk praksis, bidrar PHQ-9 etter vårt syn til raskere, mer fokusert og mer fleksibel kartlegging enn annen strukturert kartlegging av depresjon, som med MADRS, M.I.N.I. eller BDI-II. Når det gjelder rutinemessige kliniske oppgaver – som depresjonsvurdering – vil selv små forbedringer som dette kunne gjøre en forskjell i en travel klinisk hverdag.

Takk

Forfatterne ønsker å takke Pål Nielsen Glad for nyttige innspill på et tidligere utkast av denne artikkelen.

Merknader

Forfatterne har avgitt signert interessekonflikterklæring og oppgir ingen interessekonflikter. PHQ-9 ble opprinnelig utviklet med økonomisk støtte fra legemiddelfirmaet Pfizer. Forfatterne har ingen tilknytning til Pfizer.

TEKST

Tomas Formo Langkaas, Poliklinikken sentrum, Drammen DPS, Vestre Viken HF

KONTAKT: tomas.langkaas@gmail.com

Even Rognan, Poliklinikken sentrum, Drammen DPS, Vestre Viken HF

Sverre U. Johnson, Psykologisk institutt, Universitetet i Oslo og Forskningsinstituttet, Modum Bad.

+ Vis referanser

- American Psychiatric Association. (1994). (4.)
utg.). Washington, DC: American Psychiatric Association.
- American Psychiatric Association. (2013). (5.)
utg.). Washington, DC: American Psychiatric
Association. <https://doi.org/10.1176/appi.books.9780890425596>
- Beck, A. T., Steer, R. A. & Brown, G. K. (1996). . San Antonio,
TX: Psychological Corporation. <https://doi.org/10.1037/t00742-000>
- Bekhuis, E., Hartman, T. C. O., Boschloo, L. & Lucassen, P. L. (2019). A new approach to
psychopathology: the example of depression. , (680), 146-147.
<https://doi.org/10.3399/bjgp19X701717>
- First, M. B., Spitzer, R. L, Gibbon M. & Williams, J. B.W. (1996)
. Washington, D.C.: American Psychiatric
Press, Inc.
- Fried, E. I. (2017). The 52 symptoms of major depression: Lack of content overlap among seven
common depression scales. , , 191-197.
<https://doi.org/10.1016/j.jad.2016.10.019>
- Helsedirektoratet. (2008).
. Oslo: Helsedirektoratet.
- Helsedirektoratet. (2009).
. Oslo: Helsedirektoratet.
- Jacobson, N.S. & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful
change in psychotherapy research. (1), 12-19.
<https://doi.org/10.1037/0022-006X.59.1.12>
- Kroenke, K., Spitzer, R. L. & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity
measure. , (9), 606-613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Kroenke, K., Strine, T. W., Spitzer, R. L., Williams, J. B., Berry, J. T. & Mokdad, A. H. (2009). The
PHQ-8 as a measure of current depression in the general population.
, (1-3), 163-173. <https://doi.org/10.1016/j.jad.2008.06.026>
- Langkaas, T. F., Johnson, S. U., Stenshorne, K., Collin-Tiller, C. & Vrabel, K. R. (2017). Klinisk
signifikans i praksis. (12), 1160-1167.
- Langkaas, T. F., Wampold, B. E. & Hoffart, A. (2018). Five types of clinical difference to monitor in
practice. (3), 241-254. <https://doi.org/10.1037/pst0000194>
- Levis, B., Benedetti, A. & Thombs, B. D. (2019). Accuracy of Patient Health Questionnaire-9 (PHQ-9)
for screening to detect major depression: individual participant data meta-analysis. , ,
11476. <https://doi.org/10.1136/bmj.l1476>

- Lichtman, J. H., Bigger Jr, J. T., Blumenthal, J. A., Frasure-Smith, N., Kaufmann, P. G., Lespérance, F., ... & Froelicher, E. S. (2008). Depression and coronary heart disease: recommendations for screening, referral, and treatment: a science advisory from the American Heart Association Prevention Committee of the Council on Cardiovascular Nursing, Council on Clinical Cardiology, Council on Epidemiology and Prevention, and Interdisciplinary Council on Quality of Care and Outcomes Research: endorsed by the American Psychiatric Association. *Circulation*, *117*(17), 1768-1775. <https://doi.org/10.1161/CIRCULATIONAHA.108.190769>
- Manea, L., Gilbody, S. & McMillan, D. (2012). Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): A meta-analysis. *British Medical Journal*, *364*(8210), E191-E196. <https://doi.org/10.1503/cmaj.110829>
- Mitchell, A. J., Yadegarfar, M., Gill, J. & Stubbs, B. (2016). Case finding and screening clinical utility of the Patient Health Questionnaire (PHQ-9 and PHQ-2) for depression in primary care: a diagnostic meta-analysis of 40 studies. *Journal of Clinical Pharmacy and Therapeutics*, *41*(2), 127-138. <https://doi.org/10.1192/bjpo.bp.115.001685>
- Montgomery, S. A. & Åsberg, M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry*, *134*(4), 382-389. <https://doi.org/10.1192/bjp.134.4.382>
- National Collaborating Centre for Mental Health. (2018). *Depression: Clinical Guidelines for the Primary Care Setting*. England: National Health Service.
- National Institute for Health and Care Excellence. (2011). *Depression: Clinical Guidelines (CG123)*. Hentet fra <https://www.nice.org.uk/guidance/cg123>
- Pfizer. (u.å.). *PHQ-9 and PHQ-2 Instructions*. Hentet fra <https://phqscreeners.pfizer.edrupalgardens.com/sites/g/files/g10016261/f/201412/instructions.pdf>
- Reed, G. M., First, M. B., Kogan, C. S., Hyman, S. E., Gureje, O., Gaebel, W., ... & Claudino, A. (2019). Innovations and changes in the ICD-11 classification of mental, behavioural and neurodevelopmental disorders. *World Psychiatry*, *18*(1), 3-19. <https://doi.org/10.1002/wps.20611>
- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., ... Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, *59*(Suppl 20), 22-33.
- Smith, O. R. F., Alves, D. E. & Knapstad M. (2016). *Depression: A Manual for the General Practitioner*. Oslo: Folkehelseinstituttet.
- Spitzer, R. L., Kroenke, K. & Williams, J. B. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ Primary Care Study. *Archives of General Psychiatry*, *56*(10), 1737-1744. <https://doi.org/10.1001/jama.282.18.1737>
- Strawbridge, R., Young, A. H. & Cleare, A. J. (2017). Biomarkers for depression: recent insights, current challenges and future prospects. *Journal of Clinical Pharmacy and Therapeutics*, *42*(3), 1245-1262. <https://doi.org/10.2147/NDT.S114542>
- Titov, N., Dear, B. F., McMillan, D., Anderson, T., Zou, J. & Sunderland, M. (2011). Psychometric comparison of the PHQ-9 and BDI-II for measuring response during treatment of depression. *Journal of Clinical Pharmacy and Therapeutics*, *36*(2), 126-136. <https://doi.org/10.1080/16506073.2010.550059>
- Topp, C. W., Østergaard, S. D., Søndergaard, S. & Bech, P. (2015). The WHO-5 Well-Being Index: a systematic review of the literature. *Journal of Clinical Pharmacy and Therapeutics*, *40*(3), 167-176. <https://doi.org/10.1159/000376585>
- World Health Organization. (u.å.). *WHO-5 Well-Being Index*. Hentet fra https://www.psykiatri-regionh.dk/who-5/Documents/WHO5_Norwegian.pdf
- World Health Organization. (1992). *International Classification of Diseases (ICD-10)*. Geneva: World Health Organization.

World Health Organization. (2017).
. Geneva: World Health Organization.

Yesavage, J. A. (1983). . Hentet fra <https://ah-web-prod.s3.amazonaws.com/documents/GDS.pdf>

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M. & Leirer, V. O. (1982).
Development and validation of a geriatric depression screening scale: a preliminary report.
, (1), 37-49. [https://doi.org/10.1016/0022-3956\(82\)90033-4](https://doi.org/10.1016/0022-3956(82)90033-4)

Zigmond, A. S. & Snaith, R. P. (1983). The hospital anxiety and depression scale.
, (6), 361-370. <https://doi.org/10.1111/j.1600-0447.1983.tb09716.x>

Aasen, H. (2001). . Bergen: Det psykologiske
fakultet, Universitetet i Bergen. Hentet fra <https://bora.uib.no/handle/1956/1773>