

Evidensbaserad screening och diagnostik inom BUP

Evidensbaserad behandling är numera rekommenderad och allmänt accepterad inom barn- och ungdomspsykiatri (BUP), men för screening, utredning och diagnostik verkar det inte finnas samma evidenskrav. Till exempel har telefonintervjumetoden BCFPI (Brief child and family phone interview) värderats högt och implementerats brett trots att en granskning av Socialstyrelsen visade att metoden riskerar att missa mer än hälften av barn och unga med pågående depression eller ångestsyndrom och att enbart cirka hälften av dem som faller ut faktiskt lider av dessa tillstånd [1]. Vidare används sällan semi- eller helstrukturerade diagnostiska intervjuer inom BUP, trots att de är betydligt mer träffsäkra än diagnostiska samtal som sker utan hjälpmedel [2]. Konsekvensen blir att psykiatriska tillstånd missas. Till exempel lider runt 40 procent av alla barn och unga som kommer till BUP av ett ångestsyndrom, men tillståndet identifieras hos färre än 1 av 10 [3-5]. Screening och diagnostik inom BUP försvaras av att ingen enskild metod är tillräcklig. Patientsäkert utredningsarbete måste därför vila på uppgifter från olika informanter (till exempel patient och anhörig) som samlas in med hjälp av olika metoder (till exempel intervju, skattningsskalor och observation). En sällan diskuterad svårighet är hur denna information ska vägas samman.

I denna artikel beskriver vi vad som bör karaktärisera evidensbaserad utredning och diagnostik inom BUP. Vi inleder med att summera vad som avses med evidensbaserad behandling. Därefter visar vi att evidensbaserad diagnostik kräver kunskap om 1) prevalensen av olika tillstånd i patientpopulationen samt 2) de psykometriska egenskaperna hos utredningsmetoderna som används. Vi avslutar med att visa hur diagnostisk information kan sammanvägas genom en enkel tillämpning av nomogram (med bakgrund i Bayes sats) och illustrerar hur detta evidensbaserade förhållningssätt till screening och diagnostik kan appliceras på två vanliga barnpsykiatriska tillstånd: ADHD och depression.

Evidensbaserad behandling inom BUP

Det nålsöga som varje ny eller gammal behandlingsmetod behöver passera består enligt GRADE, graderingen som Socialstyrelsen använder för att utvärdera evidensunderlaget för olika insatser inom hälso- och sjukvården [6], i att:

- behandlingen har utvärderats i en population som är relevant, väl karaktäriserad och på väsentliga områden jämförbar med målpopulationen
- behandlingen är så tydligt beskriven att den kan implementeras i klinisk vardag
- nytta och risker med behandlingen har jämförts med en grupp som inte fått behandlingen (till exempel

Tord Ivarsson, docent, institutionen för neurovetenskap och fysiologi, Göteborgs universitet
 ● tord.ivarsson@outlook.com

Matti Cervin, docent, specialitetspsykiolog, barn- och ungdomspsykiatri Skåne; medicinska fakulteten, Lunds universitet

empel fått placebo eller annan behandling)

- de utfallsmått som använts ger användbar och säker information
- forskningen är så väl utförd att vi kan lita på att skillnaden mellan behandlings- och jämförelsegrupp inte beror på slumpen, ett skevt urval av patienter eller tvivelaktiga studieprocedurer
- flera oberoende studier visar på ungefär samma nytta och risker.

En rad hårda krav ställs således på en behandling för att den ska betraktas som evidensbaserad. Tyvärr verkar kraven på utredning och diagnostik inom BUP vara lägre ställda, vilket är problematiskt eftersom evidensbaserad diagnostik är en förutsättning för och ett första steg i evidensbaserad behandling. I Socialstyrelsens riktlinjer finns viss vägledning kring diagnostik, om än inte tillräckligt breda, som i riktlinjerna för depression och ångestsyndrom [7], eller tillräckligt precisa, som i riktlinjerna för ADHD och autism [8].

Evidensbaserad diagnostik inom BUP

Många behandlingar inom BUP är effektiva på sitt område, men samtidigt behäftade med risker eller biverkningar. Exempel på detta är läkemedelsbehandling av ADHD och depression, som kan ge stor nytta men också biverkningar. Det är därför viktigt att behandla rätt personer, alltså de som verkligen har ADHD respektive en depression av måttlig till svår natur, utan att sälla bort patienter under den diagnostiska processen på felaktiga grunder eller behandla dem som inte har diagnosen. Utan evidensbaserad diagnostik kommer inte rätt patient att matchas med rätt intervention, och då erhålls inte de effekter som beskrivs i behandlingslitteraturen.

Den diagnostiska processen inom psykiatri är komplicerad eftersom psykiatriska diagnoser ofta samexisterar [9], specifika symtom ingår i flera olika diagnoser [10] och psykiatriska svårigheter och tillstånd varierar mycket inom en och samma individ

HUVUDBUDSKAP

- Evidensbaserad diagnostik är en förutsättning för evidensbaserad behandling inom barn- och ungdomspsykiatri.
- Evidensbaserad diagnostik förutsätter kunskap om prevalens av olika tillstånd i patientpopulationen samt användning av utredningsmetoder med kända psykometriska egenskaper.
- Diagnostisk information kan på ett enkelt och systematiskt sätt vägas samman med hjälp av nomogram.

TABELL 1. Hur sensitivitet, specificitet och falskt positiva utfall beräknas för diagnostiska metoder

	Patienten har diagnos	Patienten har inte diagnos
Metoden indikerar att patienten har diagnos	A	C
Metoden indikerar att patienten inte har diagnos	B	D
Summa	A + B	C + D

Sannolikheten för att ett positivt utfall verkligen indikerar diagnos ökar om metoden har god sensitivitet, som räknas ut genom att dividera antalet individer med en diagnos som klassificeras korrekt av metoden med samtliga individer som har en diagnos: sensitivitet = $A / (A + B)$.

Dessutom måste metoden visa specificitet, det vill säga att de individer som inte har en diagnos blir korrekt klassificerade: specificitet = $D / (C + D)$.

Vidare finns falskt positiva utfall, alltså individer som av metoden felaktigt klassificeras med en diagnos. Falskt positiva ges av $C / (C + D) = 1 - \text{specificiteten}$.

över tid [9]. Inom BUP finns dessutom många informanter tillgängliga, och det är inte ovanligt att patient, föräldrar och lärare kan ge olika bilder. Flera informanter medför dock unika möjligheter, och forskning visar att föräldrar och lärare är goda informanter avseende externaliserade symtom (till exempel trots, hyperaktivitet, aggressivitet och impulsivitet), medan barn och unga ger säkrare information om internaliserade symtom (till exempel nedstämdhet, anhedoni, självmordstankar och ångest) [11]. Evidensbaserad diagnostik inom BUP bör därför i varje steg luta sig mot vetenskapligt utprovade metoder, som genomförs med den eller de informanter som kan ge mest användbar information.

Psykometri och diagnostik

En annan utmaning för psykiatrisk diagnostik är att ingen skattning eller intervju är perfekt. Detta medför att en skattning över eller under ett förutbestämt gränsvärde (till exempel jakande svar under en intervju eller höga poäng på en skattningsskala) inte är liktydigt med att individen har (eller inte har) en diagnos. För att värdera vad ett visst utfall betyder behövs kunskap om hur träffsäker utredningsmetoden

är, vilket visas genom metodens sensitivitet, specificitet samt dess falskt positiva utfall. I Tabell 1 definierar vi dessa begrepp.

Begreppen kan vara svåra att hålla isär, men de är centrala för evidensbaserad screening och diagnostik inom BUP eftersom de låter oss räkna ut något som kallas diagnostisk sannolikhetskvot (diagnostic likelihood ratio, DLR). Den ger oss sannolikheten för att en patient verkligen har en diagnos efter positivt utfall. DLR baseras på metodens psykometri och bestäms genom ekvationen $DLR = \text{sensitivitet} / (1 - \text{specificitet})$.

Sensitivitet och specificitet vid olika diagnoser finns angivna för vanliga utredningsinstrument inom BUP, till exempel BCFPI (föräldraskattning), CBCL (Child behaviour checklist; föräldraskattning) och YSR (Youth self report; skattning av barn/ungdom) [12, 13]. Dessa instrument mäter många olika symtom och tillstånd, men samma uppgifter finns för skalor utformade för specifika tillstånd, till exempel ADHD [14], ångestsyndrom [15] och depression [16]. Sannolikhetskvoten är speciellt användbar när flera utredningsmetoder behöver kombineras för att uppnå träffsäkerhet.

Bayes sats, nomogram och diagnostik

Men vad betyder ett utfall på en viss skala eller en intervju i praktiken? Om vi tar ADHD-skalan på BCFPI som exempel kan vi starta med det vi redan vet när vi kommer till väntrummet, nämligen att lite mer än var femte patient (22 procent) som kommer till BUP har ADHD, vilket också blir oddsen för att den patient klinikern ska möta faktiskt har ADHD. Denna siffra är tagen från Högberg och kollegors utvärdering av den diagnostiska intervjun Mini-kid (Mini international neuropsychiatric interview för barn och ungdomar), som validerades med hjälp av diagnoser baserade på metoden LEAD (Longitudinal, expert, all data). Vid LEAD-diagnostik görs all information om patienten, även den som uppkommer under behandling, tillgänglig för experter som sedan bedömer förekomst av olika diagnoser. Högberg och kollegor [17] analyserade patienter från flera olika BUP-mottagningar, vil-

FAKTA 1. Bayes sats tillämpad på ADHD vid ett nybesök på en allmän BUP-mottagning

$$P(\text{ADHD}|\text{BCFPI-ADHD positiv}) = \frac{P(\text{BCFPI-ADHD positiv}|\text{ADHD}) \times P(\text{ADHD}|\text{BUP})}{(P(\text{BCFPI-ADHD positiv}|\text{ADHD}) \times P(\text{ADHD}|\text{BUP})) + (P(\text{BCFPI-ADHD positiv}|\text{ej ADHD}) \times P(\text{ej ADHD}|\text{BUP}))}$$

- $P(\text{BCFPI-ADHD positiv}|\text{ADHD}) = \text{Andel sant positiva BCFPI-ADHD} = \text{sensitivitet} = 0,74$
- $P(\text{ADHD}|\text{BUP}) = \text{Andel av ADHD i BUP bland nybesöken} = 0,22$
- $P(\text{BCFPI-ADHD positiv}|\text{ej ADHD}) = \text{Andel falskt positiva BCFPI-ADHD} = 1 - \text{spec} = 0,32$
- $P(\text{ej ADHD}|\text{BUP}) = \text{Andel ej ADHD bland BUP-nybesök} = 0,78$

- Sannolikheten (P) för att en patient med en positiv skattning på BCFPI verkligen har ADHD:

$$P(\text{ADHD}|\text{BCFPI-ADHD positiv}) = \frac{0,74 \times 0,22}{0,74 \times 0,22 + 0,32 \times 0,78} = 0,41$$

- Av 100 patienter i väntrummet kan 22 förväntas ha ADHD. BCFPI fångar upp 74 procent av dem, det vill säga 16 individer. Om inga falskt positiva hade funnits hade oddsden för ADHD bland dem som utfallit varit $0,22 \times 0,74 / 0,22 \times 0,74 = 1,0$, men eftersom det finns falskt positiva (proportionen är 0,32 hos dem som inte har ADHD) kommer BCFPI att identifiera 25 falskt positiva individer. Detta medför att bara 40 procent av dem som klassificeras som ADHD-positiva faktiskt har ADHD, det vill säga $0,22 \times 0,74 / (0,22 \times 0,74 + 0,32 \times 0,78)$.

FAKTA 2. DLR (diagnostic likelihood ratio) för BCFPI (Brief child and family phone interview), CBCL (Child behaviour checklist) och YSR (Youth self report)

DLR kvantifierar hur mycket sannolikheten för en diagnos förändras vid positivt utfall. Sannolikheten räknas ut med hjälp av metodens psykometriska egenskaper i enlighet med ekvationen:

$$DLR = \frac{\text{sensitivitet}}{(1 - \text{specificitet})}$$

- För BCFPI blir $DLR^{ADHD} = 0,74 / (1 - 0,68) = 2,3$
- För CBCL blir $DLR^{ADHD} = 0,81 / (1 - 0,70) = 2,7$
- För YSR blir $DLR^{ADHD} = 0,65 / (1 - 0,70) = 2,2$
- För BCFPI blir $DLR^{depression} = 0,56 / (1 - 0,81) = 2,9$
- För CBCL blir $DLR^{depression} = 0,75 / (1 - 0,70) = 2,5$
- För YSR blir $DLR^{depression} = 0,67 / (1 - 0,76) = 2,8$

ket medför att vi kan ha en relativt hög tilltro till att de prevalenssiffror som redovisas är rimligt nära sanningen för patienter inom BUP:s öppenvård i Sverige. Eventuellt har denna andel ökat under åren som gått sedan studien genomfördes, vilket är viktigt att avgöra på lokal nivå.

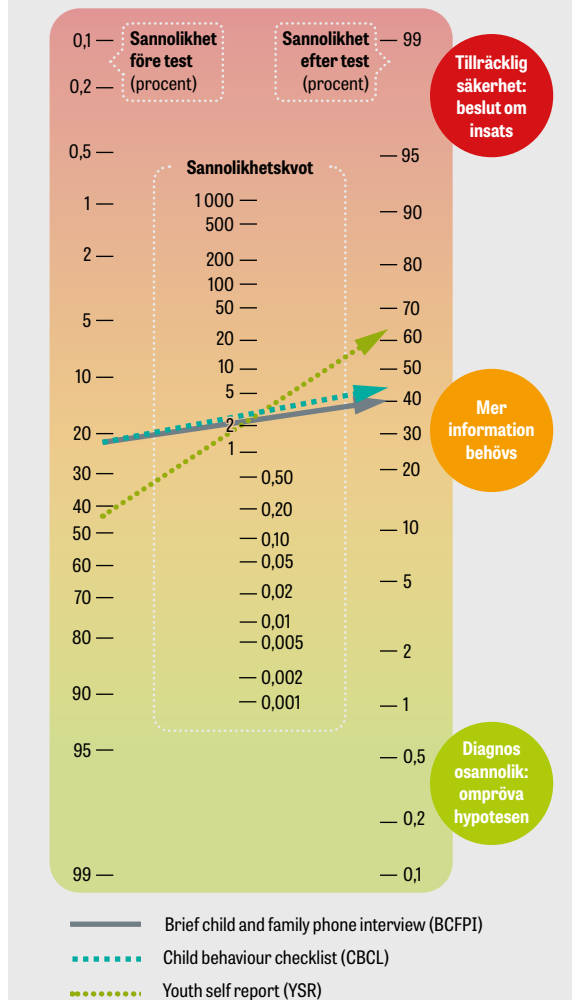
När ADHD-prevalensen i patientpopulationen är känd kan vi använda DLR för att värdera vad ett positivt utfall på BCFPI-skalan för ADHD betyder. Detta går att räkna ut med en form av statistik som utvecklades under 1700-talet av pastorn Thomas Bayes och som fått en alltmer vidgad användning under de senaste 30 åren. I korthet handlar statistiken om att kvantifiera hur vår tilltro till förekomsten av ett särskilt fenomen (till exempel att någon uppfyller kriterierna för en diagnos) förändras när vi får ny information. I Fakta 1 använder vi Bayes sats för att räkna ut hur ett positivt utfall på ADHD-skalan på BCFPI påverkar sannolikheten för att patienten faktiskt har ADHD. Vi ser att sannolikheten för att patienten uppfyller diagnoskriterierna är cirka 40 procent efter ett positivt utfall på BCFPI.

Det kan ses som förvånande att inte ens en majoritet av patienterna med positivt utfall har ADHD, men det beror på att sensitiviteten är relativt låg (0,74), vilket medför att 26 procent av de med ADHD inte ger positivt utfall, samt metodens relativt stora andel falskt positiva (0,32), vilket medför att de flesta med positivt utfall (cirka 60 procent) faktiskt inte uppfyller kriterierna för ADHD. Ytterligare utredning måste således ske.

Det ter sig dock inte lätt att sitta med en kalkylator framför patienten. Vi vill därför lyfta fram hur användning av DLR tillämpade via nomogram [18] kan underlätta för den tidspressade kliniker. Nomogram utvecklades av kliniska epidemiologer i Kanada för att förstärka evidensbaserat tänkande inom medicinen [19] och är en tillämpning av statistik baserad på Bayes sats. Vi går till publicerade data för att räkna ut DLR för BCFPI, CBCL och YSR avseende ADHD och depression [12, 13]. Andra skalor hade lika gärna kunnat användas, men psykometriska data för CBCL och YSR fanns lätt tillgängliga och medgav dessutom möjligheten att illustrera sammanvägning av skattningar. Uträkningarna presenteras i Fakta 2.

I Figur 1 visar vi med ett nomogram hur ett positivt

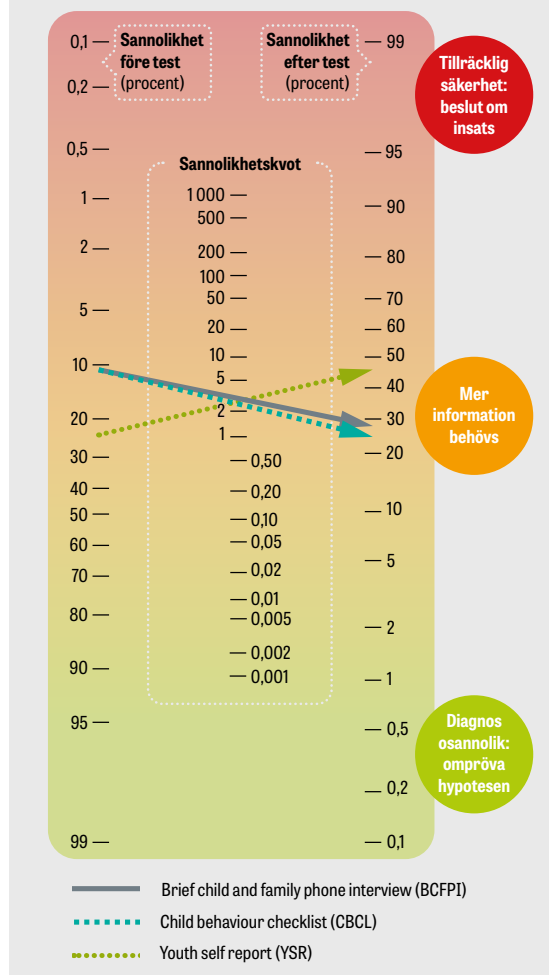
FIGUR 1. Nomogram för ADHD baserade på BCFPI (Brief child and family phone interview) samt kombinationen CBCL (Child behaviour checklist) och YSR (Youth self report)



utfall för ADHD på BCFPI påverkar sannolikheten för faktisk diagnos. Till vänster för vi in prevalensen/sannolikheten i patientpopulationen (22 procent, enligt Högberg et al). Sedan dras ett streck genom DLR-värdet (mittenaxeln). Där pilen markerar på axeln till höger ligger den uppdaterade sannolikheten. Vi ser att pilen indikerar en sannolikhet på cirka 40 procent, precis som när vi räknade för hand. Vid kombinationen CBCL och YSR följer vi samma principer. Vi börjar på 22 procent och drar strecket genom DLR för CBCL. Den uppdaterade sannolikheten blir cirka 43 procent. När vi lägger till YSR uppdaterar vi sannolikheten till vänster med vår nya sannolikhet (43 procent) och drar ett streck genom DLR för YSR. Den uppdaterade sannolikheten blir cirka 62 procent.

I Figur 2 upprepar vi samma förfarande för egentlig depression (prevalensen är cirka 11 procent, enligt studien av Högberg et al). Positivt utfall på enbart BCFPI ger en uppdaterad sannolikhet på cirka 27 procent, medan positivt utfall på såväl CBCL som YSR ger en uppdaterad sannolikhet på cirka 47 procent. Se

FIGUR 2. Nomogram för egentlig depression baserade på BCFPI (Brief child and family phone interview) samt kombinationen CBCL (Child behaviour checklist) och YSR (Youth self report)



Youngstrom [20] för mer information om hur nomogram kan användas.

I båda dessa exempel visar nomogrammet för BCFPI dilemmat med att förlita sig på enbart föräldern, medan kombinationen CBCL och YSR illustrerar nytan av information från såväl förälder som patient. Exempelen visar också att inte ens en kombination av patient- och föräldraskattade metoder med relativt goda psykometriska egenskaper uppnår sannolikheter som är tillräckligt höga för att ställa en diagnos. Ytterligare utredningsinsatser behövs, men genom att utgå från metodernas kända psykometriska egenskaper tillsammans med nomogram får vi en tydligare vägledning om hur olika instrument kan kombineras för att uppnå en högre tillförlitlighet. Dessutom kan de sammanvägda oddsen för till exempel depression visa vad som inte får missas, till exempel en suicidriskbedömning. Här bör det noteras att en av de viktigaste metoderna vid psykiatrisk diagnostik är en semi- eller helstrukturerad diagnostisk intervju, till exempel K-SADS (Kiddie schedule for affective disorders and schizophrenia), ADIS (Anxiety

disorders interview schedule) eller MINI. Intervjuer som sker utan stöd av sådana hjälpmedel är inte tillförlitliga – något som noterats i såväl internationella [2] som svenska studier [21]. Icke-strukturerade diagnostiska intervjuer utgör därför endast ett begränsat bidrag i den diagnostiska processen, vilket enkelt kan visas med hjälp av nomogram. I exemplet för ADHD och depression skulle en icke-strukturerad diagnostisk intervju inte medföra någon märkbar skillnad för sannolikheterna för de två diagnoserna. För att uppnå patientsäker screening och diagnostik är det därför viktigt att använda metoder med kända och för ändamålet goda psykometriska egenskaper.

Diskussion

Screening, utredning och diagnostik är en viktig och krävande del av sjukvården och innefattar hela processen, från att efterfråga patientens upplevda symptom, observera tecken på sjukdom och bedöma påverkan på funktion inom livets alla områden till en slutlig diagnos, som är så säker att evidensbaserad behandling kan erbjudas. Kvaliteten på den information som samlas in är i varje steg avgörande. Kvalitet beror på flera saker, men inom psykiatri är psykometrin hos de metoder som används särskilt viktig, liksom att i varje steg använda informanter och metoder som kan ge relevant och unik information. Som vi visade för depression ökar sannolikheten avsevärt om föräldraskattning läggs till barnets självskattning, liksom när barnets självskattning adderas till föräldraskattning [13]. Vad gäller ADHD brukar självskattning av barn och unga sällan ingå, men av de data vi använde [13] framgår att sannolikheten för en ADHD-diagnos ökar avsevärt när man till positivt utfall på CBCL adderar ett positivt utfall på YSR.

Sammantaget visar dessa exempel att den diagnostiska processen inom BUP är komplicerad, men att den måste vägledas av ett systematiskt förhållningssätt, som tar sin utgångspunkt i utredningsmetodernas kända psykometriska egenskaper. För de flesta barnpsykiatriska tillstånd pekar forskningen mot att information från såväl patient som viktiga vuxna i dennas närhet, kombinerad med träffsäkra diagnostiska intervjuer, är det bästa sättet att uppnå en patientsäker screening och diagnostik. I Tabell 2 ger vi en kort beskrivning av vilka metoder som kan kombineras vid olika frågeställningar och moment under den diagnostiska processen.

Hur kan bättre diagnostiska rutiner och metoder införas inom BUP? Sveriges Kommuner och regioner (SKR) kunde som övergripande intresseorganisation för huvudmännen tillämpa kriterierna för evidensbaserad diagnostik som vi angett ovan. Om det saknas kunskapsunderlag för något steg bör SKR se till att kunskapsbristen rättas till. Om någon metod brister bör den ersättas med en metod eller uppsättning metoder som har bättre psykometriska egenskaper, alternativt kan metoden kompletteras så att den får bättre egenskaper. Inom ett så eftersatt område som screening och diagnostik inom BUP kunde SKR finansiera forskningsstudier i samverkan med huvudmännen, som också har ett ansvar. Vidare har forskare inom BUP ett ansvar att studera vanliga hjälpmedels psykometri vid olika gränsvärden och frågeställningar. Även i Socialstyrelsens riktlinjearbete skulle ett

TABELL 2. Förslag på hur information från olika informanter kan kombineras för att uppnå ökad säkerhet i en utredningsprocess

	Självskattning	Föräldraskattning	Lärarskattning	Kliniker
Screening/triagering	Måste ingå: bred symtomskattning.	Måste ingå: bred symtomskattning.		Symtom som utgör sökanledning samt deras påverkan på livet.
Diagnostik	Måste ingå: specifik symtomskattning för de symtomområden som utgör sökanledning. Bör ingå: skattning av funktionspåverkan och livskvalitet.	Måste ingå: specifik symtomskattning för de symtomområden som utgör sökanledning. Bör ingå: skattning av funktionspåverkan/livskvalitet.	Bör ingå vid ADHD: symtom-/funktionsskattning.	Måste ingå: intervju/samtal som är systematisk och brett med semi- eller helstrukturerad intervju, där både patient och föräldrar intervjuas separat och tillsammans. Systematisk skattning av funktionsnivå-/påverkan.
Sammanvägning	Självskattning har särskild vikt vid till exempel ångestsyndrom, depression och tvångssyndrom.	Föräldraskattning har stor vikt vid alla tillstånd, men tillskrivs olika vikt beroende på vilka symtom som utreds.	Lärarskattning har särskild vikt vid till exempel ADHD, autism, utagerande beteenden, social ångest och inlärningssvårigheter.	Stor vikt i alla diagnostiska processer.

systematiskt förhållningssätt till screening och diagnostik vara välkommet, vilket framgår i de nya riktlinjerna för ADHD och autism [8], där det helt saknas resonemang om och vägledning kring hur diagnostiken kan evidensbaseras. Forskningsprojekt med syfte att utvärdera hela den diagnostiska processen, från screening till slutlig diagnos, och hur olika instrument och informationskällor kan sammanvägas (inklusive mer nyanserad information om hur väl varje informant upplever sig ha insyn i de symtom som undersöks) vore välkomna.

Avslutningsvis är det viktigt att poängtera att utredningsmetoder inom BUP, som till exempel BCFPI, CBCL och YSR, inte har en enda psykometri utan flera: en för varje tillstånd. För att komplicera saken ytterligare avgörs ett instruments användbarhet inte bara av instrumentet utan även av tillståndens prevalens i den population som undersöks. Vanliga tillstånd som ADHD och depression är mindre krävande, medan mer ovanliga tillstånd som tvångssyndrom och paniksyndrom är mer utmanande. En duktig diagnostiker inom BUP behöver således inte bara ha omfattande kunskap om psykiatriska diagnoser och god förmåga att vinna förtroende hos barn och deras fa-

miljer, utan också hur vanliga olika tillstånd är i patientpopulationen, vilka utredningsmetoder som ska användas och när samt vilka psykometriska egenskaper dessa metoder har vid just den frågeställning som är aktuell. När allt detta är på plats kan nomogram användas för att sammanfoga och värdera den information som framkommer.

Avslutande kommentar

Vi vill med denna artikel betona att psykiatrisk diagnostik alltid utgör en förskjutning i vår tro på vad patienten har för diagnos, från en mycket osäker tro när vi kommer till väntrummet till en allt säkrare sådan efter att flera diagnostiska moment slutförts. Detta skifte, från en initial till en uppdaterad tro, kan tydliggöras med Bayes sats. Vi hoppas att detta tänkesätt, som kan framstå som komplicerat men som varken är särskilt tidskrävande eller svårt att utföra, kan bidra till att stärka screening, utredning och diagnostik inom BUP. ○

- Potentiella bindningar eller jävsförhållanden: Inga uppgivna.
- Gudmundur Skarphedinsson, professor vid psykologiska fakulteten, Háskóli Íslands, har faktagranskat artikeln.

Citera som: *Läkartidningen*. 2023;120:22144



Läs mer!

Fullständig referenslista på Läkartidningen.se

REFERENSER

- Nationella riktlinjer för vård vid depression och ångestsyndrom. Stöd för styrning och ledning. Stockholm: Socialstyrelsen; 2021. Artikelnr 2021-4-7339.
- Rettew DC, Lynch AD, Achenbach TM, et al. Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. *Int J Methods Psychiatr Res*. 2009;18(3):169-84.
- Esbjörn BH, Hoeyer M, Dyrborg J, et al. Prevalence and co-morbidity among anxiety disorders in a national cohort of psychiatrically referred children and adolescents. *J Anxiety Disord*. 2010;24(8):866-72.
- Olofsdotter S, Vadlin S, Sonby K, et al. Anxiety disorders among adolescents referred to general psychiatry for multiple causes: clinical presentation, prevalence, and comorbidity. *Scand J Child Adolesc Psychiatr Psychol*. 2016;4(2):55-64.
- Hansen BH, Oerbeck B, Skirbekk B, et al. Non-obsessive-compulsive anxiety disorders in child and adolescent mental health services - Are they underdiagnosed, and how accurate is referral information? *Nord J Psychiatry*. 2016;70(2):133-9.
- BMJ Best Practice. What is GRADE? <https://bestpractice.bmj.com/info/toolkit/learn-ebm/what-is-grade/>.
- Nationella riktlinjer för vård vid depression och ångestsyndrom. Stöd för styrning och ledning. Stockholm: Socialstyrelsen; 2017. Artikelnr 2017-12-4.
- Nationella riktlinjer för vård och stöd vid adhd och autism. Prioriteringsstöd till beslutsfattare och chefer 2022. Stockholm: Socialstyrelsen; 2022. Artikelnr 2022-10-8100.
- Caspi A, Houts RM, Ambler A, et al. Longitudinal assessment of mental health disorders and comorbidities across 4 decades among participants in the Dunedin Birth cohort study. *JAMA Netw Open*. 2020;3(4):e203221.
- Forbes MK, Neo B, Nezami OM, et al. Elemental psychopathology: distilling constituent symptoms and patterns of repetition in the diagnostic criteria of the DSM-5. *Psyarxiv*. Epub 21 mar 2023. doi: 10.31234/osf.io/u56p2.
- Makol BA, Youngstrom EA, Racz SJ, et al. Integrating multiple informants' reports: how conceptual and measurement models may address longstanding problems in clinical decision-making. *Clin Psychol Sci*. 2020;8(6):953-70.
- Andersson M, Bäckström M, Ivarsson T, et al. Validity of the Brief child and family phone interview (BCFPI) by comparison with Longitudinal expert all data (LEAD) diagnoses in outpatients. *Scand J Child Adolesc Psychiatr Psychol*. 2018;6(2):83-9.
- Skarphedinsson G, Jarbin H, Andersson M, et al. Diagnostic efficiency and validity of the DSM-oriented Child behavior checklist and Youth self-report scales in a clinical sample of Swedish youth. *PLoS One*. 2021;16(7):e0254953.
- Costa DS, de Paula JJ, Malloy-Diniz LF, et al. Parent SNAP-IV rating of attention-deficit/hyperactivity disorder: accuracy in a clinical sample of ADHD, validity, and reliability in a Brazilian sample. *J Pediatr*. (Rio J). 2019;95(6):736-43.
- Ivarsson T, Skarphedinsson G, Andersson M, et al. The validity of the Screen for child anxiety related emotional disorders revised (SCARED-R) scale and sub-scales in Swedish youth. *Child Psychiatry Hum Dev*. 2018;49(2):234-43.
- Jarbin H, Ivarsson T, Andersson M, et al. Screening efficiency of the Mood and feelings questionnaire (MFQ) and Short Mood and feelings questionnaire (SMFQ) in Swedish help seeking outpatients. *PLoS One*. 2020;15(3):e0230623.
- Högberg C, Billstedt E, Björck C, et al. Diagnostic validity of the MINI-KID disorder classifications in specialized child and adolescent psychiatric outpatient clinics in Sweden. *BMC Psychiatry*. 2019;19(1):142.
- Fagan TJ. Letter: Nomogram for Bayes theorem. *N Engl J Med*. 1975;293(5):257.
- Straus SE, Glasziou P, Richardson WS, et al. Evidence-based medicine. How to practice and teach it. New York: Churchill Livingstone/Elsevier; 2011.
- Youngstrom EA. A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: we are ready to ROC. *J Pediatr Psychol*. 2014;39(2):204-21.
- Delling N, Ivarsson T, Andersson M, et al. Rätt diagnos på BUP - hur kommer vi dit? [poster]. Svenska föreningen för barn- och ungdomspsykiatri, årsmöte 2017.