

Rosa Thrylsøe

EVIDENS OG DEPRESSION

om videnskabelige strukturer der
betinges behandlingspraksis

integreret speciale i socialvidenskab og filosofi & videnskabsteori
vejledere: Anni Greve og Klaus Frovin Jørgensen
Roskilde Universitet
marts 2015



Antal tegn: 239.888, Antal normalsider: 99,95

Studienummer: 40502, Modul: K2, Sprog: Dansk, Eksamenssprog: Dansk

Resumé

Dette integrerede speciale i Socialvidenskab og Filosofi & Videnskabsteori undersøger evidens baseret på randomiserede kontrollerede forsøg (RCT), som vidensgrundlag for den depressionsbehandling der udføres i den almene praksis. Metoden og rationalerne bag evidensbaseret medicin undersøges og det findes at en central forudsætning for at resultaterne af RCT-baseret forskning kan anvendes i behandlingsvalget for den enkelte patient, er at forsøgene baserer sig på prognostisk homogene patientgrupper. Evidens tager afsæt i de diagnostiske klasser, og den psykiatriske klassifikation undersøges med henblik på at afdække hvorvidt den danner grundlag for en prognostisk homogen depressionsdiagnose. Validiteten af de psykiatriske klasser undersøges og findes usikker, da det ikke kan påvises at diagnoserne svarer til faktisk forekommende afgrænsede patientgrupper, og da den definition der anvendes for psykisk forstyrrelse ikke sætter kriterier for hvad der kan forstås ved patologisk dysfunktion. Det konkluderes at der er en række forhold der problematiserer anvendelsen af RCT-baseret evidens som udgangspunkt for behandlingsvalg. Disse omfatter diagnosens manglende validitet og homogenitet, den subjektive karakter af de depressive kernesymptomer, evidensmetodologiens behov for standardiserede og objektive sygdomskriterier og den manglende sygdomsindsigt.

Abstract

This master thesis in Social Science and Philosophy & Science Studies, examines evidence based on randomized controlled trials (RCT) as scientific foundation for the treatment of depression that takes place in primary care. The method and the rationales are examined and it is found that a central precondition for the results of RCT-based evidence to be relevant in treatment decision for the individual patient is that the studies are based on prognostic homogeneous groups of patients. As the production of evidence relies on diagnostic categories, the psychiatric classifications of disorders are studied to evaluate whether they provide the basis for a prognostic homogeneous diagnosis. The validity of the psychiatric categories is inspected and found insecure as it cannot be rendered probable that the diagnoses correspond to actual delimited groups of patients and as the definition used to define mental disorder does not provide any criteria for defining disorder as pathological. It is concluded that a number of circumstances makes it problematic to base treatment decisions on RCT-based evidence. These comprise the diagnosis' lack of validity and homogeneity, the subjective nature of the depressive primary symptoms, the evidence methodology's demand for standardized and objective criteria of disease and the lack of insight into the depressive illness.

Evidens og depression

Om videnskabelige strukturer der betinger behandlingspraksis

Rosa Thrysøe

Indholdsfortegnelse

Kapitel 1:	Indledning og Problemstilling	5
	1.1 Projektets problemstilling	6
	1.2 Problemformulering	7
	1.3 Fremgangsmåde og arbejdsspørgsmål	7
	1.4 Begrebsafklaring og forkortelser	10
Kapitel 2:	Metodiske overvejelser	12
	2.1 Projektdesign	12
	2.2 Afgrænsning	14
	2.3 Empirisk materiale	15
	2.4 Videnskabsteoretiske antagelser	17
	2.5 Sygdomsbegrebet	18
	2.6 Alf Ross' sygdomsbegreb	19
Kapitel 3:	Evidensmetodologien	21
	3.1 Evidensbegrebets rødder	21
	3.2 The Cochrane Collaboration	23
	3.3 Evidenshierarkiet	24
	3.4 Randomiserede kontrollerede forsøg (RCT)	26
	3.5 Eksempel på et RCT	30
	3.6 Systematiske forskningsoversiger	31
	3.7 Eksempel på en systematiseret forskningsoversigt	32
Kapitel 4:	Analyse af EBM som vidensform	34
	4.1 Kritik af EBM's interne validitet	35
	4.2 Generaliserbarheden af EBM	38
	4.3 Begrænsninger ved den RCT-genererede viden	44
	4.4 Delkonklusion	49
Kapitel 5:	Sygdomsklassifikation	51
	5.1 Naturlig og konstrueret klassifikation	51

5.2 Reliabilitet	53
5.3 Validitet	54
5.4 Sygdomsklassifikation	57
5.5 Klassifikation i psykiatrien	59
5.6 The minimal medical model	61
5.7 Zones of rarity	62
5.8 Essentialisme	63
Kapitel 6: Analyse af klassifikationernes definition af mental forstyrrelse	66
6.1 Ross' svar på den antipsykiatriske kritik	66
6.2 ICD-10's definition af 'mental disorder'	67
6.3 DSM-5's definition af 'mental disorder'	69
6.4 Delkonklusion	71
Kapitel 7: Diagnosen Depression	73
7.1 De diagnostiske kriterier	73
7.2 Subkategorier	74
7.3 Diagnostiske redskaber	75
Kapitel 8: Analyse af diagnosen Depression	77
8.1 Spørgsmålet om dysfunktion	77
8.2 Kriteriernes karakter	78
8.3 Depressionsskalaerne MDI, HAMD og MADRS	80
8.4 Subkategorier	82
8.5 Delkonklusion	85
Kapitel 9: Diskussion	87
9.1 Homogenitet og prognostisk validitet	87
9.2 Objektivisering af det subjektive	89
9.3 Manglende kohærens mellem sygdomsniveauer	91
9.4 Subgrupper og generaliseringsspørgsmålet	93
Konklusion	95
Litteraturliste	98
Bilag:	
1. Major Depression Inventory	
2. Hamiltons Depression Rating Scale	
3. Montgomery & Åsberg Depression Rating Scale	

Kapitel 1: Indledning og problemstilling

Jeg vil indlede med at opridse de centrale bevæggrunde for specialets problemstilling og dermed underbygge undersøgelsens relevans.

Depression betegnes i dag som en folkesygdom. 100.000 danskere er diagnosticeret med depression og ifølge sundhedsstyrelsen har lige så mange en udiagnosticeret depression. Det anslås at 17-18% af den samlede befolkning på et tidspunkt i løbet af livet vil have en depression (Sundhedsstyrelsen 2007: 19, Sundhedsstyrelsen 2005: 2). Antallet af diagnosticerede depressioner er over de sidste 20 år steget eksplosivt og koster sammenlagt samfundet ca. 14 mia. kr. årligt i direkte omkostninger til behandling og understøttelse samt tabt arbejdsfortjeneste og nedsat produktivitet. (Danske Regioner et al. 2009: 3). Depression er desuden en alvorlig sygdom for den enkelte.

Verdenssundhedsorganisationen (WHO) placerer depression på en fjerdeplads over sygdomme, der giver anledning til størst tab af livskvalitet og leveår (Sundhedsstyrelsen 2007: 17).

Med stigningen i antallet af depressioner er medfulgt en vækst i brugen af antidepressiv medicin. Antidepressiv medicin anvendes bl.a. til behandling af angstlidelser og OCD (tvangslidelse) foruden depression (Statens Serum Institut 2012: 2). I 2012 var over 450.000 danskere i behandling med antidepressiva. Det svarer til mere end 8% af den danske befolkning (Danmarks Apotekerforening 2013: 1). Antidepressiva er blevet kritiseret for de alvorlige bivirkninger, de kan medføre samt de endnu mere alvorlige og langvarige bivirkninger, der kan opstå ved forsøg på at stoppe behandlingen. Blandt disse effekter er depression, panikanfald og hallucinationer (Gøtzsche 2014: 2). Ifølge prof. dr.med. Peter Gøtzsche er der endda tale om øget selvmordsrisiko for især yngre brugere af antidepressiva (DR P1 Debat, 17/1/2014; Irf 2010). Et stort studie af emotionelle bivirkninger forbundet med antidepressiva fandt blandt andet at 60% følte sig emotionelt følelsesløse, at 52% ikke ”føjte sig som sig selv” og at 39% bekymrede sig mindre om andre mennesker (Read et al 2014). Blandt de mest almindelige fysiske bivirkninger er desuden kvalme, svimmelhed, seksuel dysfunktion, øget træthæd og vægtforøgning.

Behandlingen af depression med antidepressiva er baseret på evidens. Langt den største del af behandlinger mod depression finder sted i almen praksis, dvs. hos den praktiserende læge (Sundhedsstyrelsen 2007: 37). Sundhedsstyrelsen har udarbejdet en klinisk vejledning for behandling af depression, hvori man på baggrund af en kortlægning af evidensen på området har udarbejdet en række anbefalinger for behandlingen. Heri anbefales antidepressiva som førstevalg mod depression af moderat og svær grad (Sundhedsstyrelsen 2007: 10; 44).

Begrebet evidens henviser til en særlig metodik for evaluering, analyse og sammenfatning af forskningsresultater. Evidens har til formål at samle, kvalitetsvurdere og syntetisere resultaterne af alle de enkelte studier i systematiske forskningsoversigter. Formålet er at isolere effekten af den variabel, der ønskes undersøgt, for eksempel en behandlingsindsats mod depression. I den forbindelse er der én forsøgsmetode, der anses for "the gold standard" i produktionen af evidens, nemlig det randomiserede kontrollerede forsøg (RCT). Evidensmetodikken er udviklet indenfor den somatiske medicin, men i vid udstrækning overført direkte til forskning i behandlingseffekter i psykiatrien (Gupta 2007: 104) The Cochrane Collaboration er en af de største producenter af evidens på sundhedsområdet. Her forsøger man i vid udstrækning at basere sig på denne type studier alene (Higgins & Green 2011: 5.5).

1.1 Projektets problemstilling

Der er en bekymrende stigning i forbruget af antidepressiv medicin til behandling af depression. Udskrivningen af denne medicin foregår dog på baggrund Sundhedsstyrelsens anbefalinger, der bygger på evidens, baseret på randomiserede kontrollerede forsøg, der anses for at være den metode, der producerer den mest pålidelige viden om behandlingseffekter.

I dette speciale vil jeg undersøge de videnskabelige og erkendelsesteoretiske strukturer, der ligger til grund for anvendelsen af en type psykofarmaka på en større befolkningsgruppe med risici for alvorlige og belastende bivirkninger. Specialet vil altså være en videnskabsteoretisk undersøgelse af dels evidens, som den metodologi der danner vidensgrundlaget for depressionsbehandlingen og dels diagnosen depression som metoden anvendes på. Jeg spørger derfor til hvilke forudsætninger og rationaler metodologien bygger på, hvad denne viden fortæller os og hvilke begrænsninger den har. Jeg undersøger desuden det teoretiske grundlag for den psykiatriske klassifikation og diagnosen depression, for derved at komme til en forståelse af depression som genstandsfelt for evidensmetodologiens anvendelse.

Evidensmetodologien er primært udviklet indenfor den somatiske medicin, og det er derfor vigtigt at undersøge, hvorvidt der findes forskelle mellem somatiske sygdomme og depression, der er relevante for, hvordan man skaber viden om og dokumenterer effekter af behandlingsindsatser for depression. Jeg mener, at det er nødvendigt at undersøge, hvordan den viden, vi baserer behandlingspraksis på, er tilvejebragt. Særligt i en situation hvor en enkelt forskningsmetode privilegeres, og vi samtidig oplever en vækst i antallet af personer som behandles for depression.

1.2 Problemformulering:

Hvilken betydning har det for den viden som depressionsbehandlingen i den kliniske praksis baserer sig på, at den primært er tilvejebragt på baggrund af et evidensbegreb som er udviklet indenfor den somatisk medicinske videnskab, og som baserer sig på randomiserede kontrollerede forsøg?

1.3 Fremgangsmåde og arbejdsspørgsmål

Jeg har nu introduceret, hvad specialet vil beskæftige sig med, samt hvorfor det er relevant at beskæftige sig videnskabsteoretisk med grundlaget for behandlingspraksis for depression. I det følgende afsnit vil jeg gøre rede for, hvordan jeg vil behandle problemstillingen ved at introducere tre arbejdsspørgsmål, der vil danne ramme om opgaven. De tre arbejdsspørgsmål retter sig mod videnskabelige strukturer, der ligger til grund for depressionsbehandlingen i den almene medicinske praksis.

1. Behandlingspraksis er for det første afhængig af forskning, der viser, at en bestemt gruppe, de depressionsdiagnosticerede, reagerer bedst eller oplever størst bedring ved behandling med antidepressiv medicin. Denne forskning kan udføres på en række forskellige måder. Fx kan forskningen variere i forhold til teoretisk og empirisk fundering, og den kan variere i forhold til, hvorvidt den kvalitative eller kvantitative indsigt prioriteres. Evidensbevægelsen repræsenterer et bestemt perspektiv på, hvordan man bedst tester og dokumenterer behandlingseffekten af en given indsats. I dette perspektiv privilegeres særligt én metode, nemlig det randomiserede kontrollerede forsøg (RCT) samt forskningsoversigter, der bygger på flere RCT. De forskellige typer studier er opstillet i et internationalt anerkendt hierarki, som udgør et tværvidenskabeligt instrument til at vurdere pålideligheden af forskningsresultater omhandlende indsatsers effekt. Dette hierarki benytter Sundhedsstyrelsen sig af, når de skal vurdere hvilke behandlingsindsatser, der er bedst evidens for. I deres 'Referenceprogram for behandling af unipolar depression hos voksne' skriver Sundhedsstyrelsen blandt andet:

Ved lettere grader af depression er behandling med antidepressiva ikke overbevisende bedre end placebo (evidensstyrke B). (...) En eventuel effekt af antidepressive lægemidler indtræder dog hurtigere end effekten af psykologisk behandling (evidensstyrke B). (...) Ved behandling af depression af moderat grad er der ingen overbevisende forskel på effekten af forskellige antidepressive lægemidler, som alle er bedre end placebo (evidensstyrke A). (Sundhedsstyrelsen 2007: 44).

Ved depression af svær grad hos indlagte patienter anbefales brug af TCA eller ECT (Ia). Til ikke-indlagte patienter anbefales SSRI, dual action-præparater eller TCA; konferer eventuelt med en speciallæge i psykiatri (Ia). (Sundhedsstyrelsen 2007: 50)

Antidepressiva af forskellige typer anbefales således til behandling af depression af moderat til svær sværhedsgrad. For depression af let grad anbefales ”watchful waiting”. Styrken af den evidens anbefalingerne baserer sig på, er angivet i parenteser efter hver anbefaling. ’Ia’ henviser til at den litteratur, der ligger til grund for anbefalingen har det højeste evidensniveau, hvilket gives til ”Cochrane-reviews, metaanalyser og andre systematiske oversigtsarbejder” (Sundhedsstyrelsen 2007: 10). ’A’ henviser ligeledes til at den dokumentationen har den højeste evidensstyrke, hvilket gives, hvis den baserer sig på systematiske forskningsoversigter eller RCT. Den hierarkiske inddeling af forskningstyper, betyder at vores viden om behandling af depression ideelt set baserer sig på én type studier og én metode. Evidensbaseret medicin er overtaget fra den somatiske medicinske videnskab. Metoden er designet til at måle de præcise effekter af behandlingsindsatser, men kan psykiske sygdomme virkelig studeres og måles på samme måde som somatiske? På baggrund af denne undren lyder mit første arbejdsspørgsmål:

Hvilken slags viden opnås gennem RCT-baseret evidens og hvilke begrænsninger har den?

2. Psykiatrien er et kompliceret felt, der som videnskab ligger i et krydsfelt mellem det biologiske, det psykologiske og det sociale. Dikotomien mellem den biologiske forklaringsmodel for psykisk sygdom og den psykosociale model har altid eksisteret (Ghaemi 2003: 4). I forhold til den somatiske medicin, er psykiatrien stadig i et præmaturot stadie. Endnu ved man ikke ret meget om, hvorfor og hvordan psykiske sygdomme opstår, hvad de specifikke lidelser er udtryk for, og hvordan de bedst muligt behandles. Selvom det biologisk-medicinske paradigme i dag er det dominerende, har ingen af paradigmerne endnu formået at overflødiggøre det andet perspektiv, og samtidig er det heller endnu ikke lykkedes for alvor at etablere en samlende teori, der integrerer biologiske, psykologiske og sociale perspektiver i en fælles forståelse af psykiske sygdomme (Ghaemi 2003: 9). Den neurobiologiske og neurogenetiske forskning gør for tiden store fremskridt, og dokumenterer de anatomiske strukturer og kemiske processer, der kan observeres i hjernen i forbindelse med psykiske lidelser. Dermed bidrager de med brikker til det puslespil, som forståelsen af psykisk sygdom er. Der er imidlertid lang vej igen og mange ubesvarede spørgsmål. Hvordan og hvorfor opstår disse neurologiske anomalier og er de resultatet af, eller årsagen til oplevelsen af psykisk sygdom? Endnu er der ikke noget, der tyder på, at der findes en simpel forklaring, eller at én tilgang (fx en molekylærgenetisk, neurologisk, traumepsykologisk eller psykosocial stressmodel)

alene kan gøre rede for fremkomsten af nogen psykiske lidelser, og udarbejdelsen af en tværvideenskabelig model for, hvordan de forskellige forklaringsrationaler integreres, er derfor essentiel. Det ikke at kende sygdommens oprindelse eller deres "natur", gør at vores forståelse af dem som fænomener også er kompromitteret. Hvad er depression præcis; er det én slags tilstand eller flere? kommer det "indefra" eller "udefra"? Hvad skal vi forstå ved psykisk sygdom og hvornår er der tale om problemer, der ikke er patologiske? I dag findes der stadig en konflikt mellem klinikere, der tilskriver biologiske forklaringsrationaler mest vægt og klinikere, der primært vægter psykosociale faktorer. Denne konflikt kommer til udtryk i uenighed ang. sygdomsforståelse, behandlingstilgang og i forståelsen af depression som fænomen. De rammer, der findes for konceptualisering af fænomenerne og erfaringerne, som udgør psykiske sygdomme har derfor en central betydning. Det drejer sig navnlig om de psykiatriske klassifikationer, der opstiller kriterierne for de enkelte diagnoser. Men på hvilket grundlag er disse diagnoser defineret, når der ikke hersker enighed, om hvad psykisk sygdom er? Jeg undersøger de psykiatriske klassifikationsmanualers definitioner af psykisk forstyrrelse for at komme til en forståelse af, hvilket sygdomsbegreb man arbejder med, og dermed hvad der gør en psykisk tilstand til en patologisk tilstand, der kan indskrives i et behandlingssystem. Mit andet arbejdsspørgsmål lyder derfor:

Hvordan konceptualiseres og defineres psykisk sygdom i de klassifikationssystemer?

3. Produktionen af evidens om depressionsbehandling forudsætter en begrebsmæssig ramme der definerer hvad depression er og hvordan den identificeres. Det er diagnosen der er afgørende for om det enkelte individ bestemmes som behandlingskrævende, diagnosen der skaber grundlaget for at der kan forskes i behandlingers virkning på specifikke sygdomme, og diagnosen der udgør grundlaget for Sundhedsstyrelsens anbefalinger om behandlingspraksis. Indenfor psykiatrien, såvel som for den øvrige medicin, foregår et omfattende internationalt samarbejde om at udarbejde diagnostiske klassifikationer, der angiver hvilke kriterier der må kunne identificeres for at hver enkelt diagnose kan stilles. I diagnosticeringen foregår således en gruppering, hvor man tillægger bestemte symptomer og tegn betydning i en kategorisering af forskellige psykiske diagnoser. Diagnosen skal sikre at de patienter der falder indenfor kategorien er sammenlignelige. Det gælder ikke mindst mht. hvordan de responder på behandling. Det kaldes at diagnosen har prognostisk validitet. Mit tredje og sidste arbejdsspørgsmål lyder derfor:

Hvordan er diagnosen depression defineret og i hvilken grad sikrer den prognostisk validitet?

Jeg afslutter specialet med en diskussion hvor jeg på baggrund af de tre analysedele diskuterer implikationerne af at basere praksis for depressionsbehandling på RCT-baseret evidens. Til sidst samler jeg op i specialets konklusion.

1.4 Begrebsafklaring og forkortelser

Jeg bruger gennem specialet betegnelsen *psykisk sygdom* om de diagnostiske klasser der fremgår af de internationale sygdomsklassifikationer. Det gør jeg, da dette er den mest anvendte generelle betegnelse i Danmark. Ved brug af ordet sygdom gør jeg mig ingen antagelser om diagnosens sygdomsstatus, blot ønsker jeg at markere, at der er tale om en lidelse der har status af diagnose, til forskel fra lidelse generelt set. I sygdomsklassifikationerne bruges derimod betegnelsen *mental disorder* (mental forstyrrelse), som betegnelse for de fænomener psykiatriske diagnoser dækker over. Jeg kommer nærmere ind på betydningen af disse betegnelser i kapitel 4.

Evidens er empirisk dokumentation. I dette speciale hvor en bestemt metodologi angående vidensproduktion om behandlingsindsatser undersøges, bruger jeg dog betegnelsen evidens på samme måde som den bruges i den videnskabelige diskussion generelt. Dvs. at betegnelsen evidens bruges om en bestemt type af empirisk data, der er frembragt gennem en bestemt metodologi og som er beskrevet i forskningslitteraturen som midlet til at opnå *Evidensbaseret Medicin (EBM)* og *Evidensbaseret Psykiatri (EBP)*.

Evidensmetodologi er ikke præcist det samme på tværs af alle organisationer og institutioner, der producerer eller anvender evidens. Når jeg omtaler evidensmetodologien forholder jeg mig til den metodologi, der beskrives i The Cochrane Collaborations håndbog til udarbejdelse af systematiske forskningsoversigter og i håndbogen *Evidence-Based Medicine. How to Teach and Practice EBM*.

Undersøgelser der indsamler og bearbejder originaldata fra en gruppe af patienter kaldes *primærstudier*. Der findes en række forskellige studiedesign som man kan benytte sig af til at undersøge behandlingseffekter. *Randomiserede kontrollerede forsøg (RCT)* er en type af primærstudier. En *systematiseret forskningsoversigt* er et sekundært studie, dvs. det indebærer ikke indsamlingen af originaldata. I stedet samler studiet alle de primærstudier, der beskæftiger sig med besvarelsen af et bestemt forskningsspørgsmål, vurderer dem efter systematiske kriterier og syntetiserer deres resultater. Der findes flere metoder efter hvilke systematiske forskningsoversigter kan udarbejdes. The Cochrane Collaborations mest anvendte metode kaldes *metaanalyse* og de betegnes *Cochrane Reviews*.

Der findes to **diagnostiske sygdomsklassifikationer** over psykiske sygdomme, der udgør den centrale referenceramme for lægens beskrivelse, diagnosticering, behandling og prognose for sygdomstilfælde. I Danmark og resten af Europa anvendes World Health Organisations (**WHO**) *International Classification of Diseases and Related Health Problems*, der aktuelt findes i 10. reviderede udgave fra 1992 og går under betegnelsen **ICD-10**. I USA anvender man American Psychiatric Associations (**APA**) *Diagnostic and Statistical Manual of Mental Disorders*, der aktuelt findes i 5. reviderede udgave fra 2013 og går under betegnelsen **DSM-5**. Jeg anvender de nævnte forkortelser i resten af specialet.

Jeg tager desuden udgangspunkt i den behandlingspraksis der foregår i den **almene praksis**, dvs. hos den praktiserende læge hvor langt de fleste patienter med depression behandles. Termen **klinisk praksis** anvendes i øvrigt om den praksis der foregår i den praktiserende læges klinik.

Kapitel 2: Metodiske overvejelser

Jeg præsenterer i dette kapitel specialets projektdesign, afgrænsning fra relaterede problemstillinger og emner for yderligere undersøgelser, anvendelsen af empirisk materiale, samt specialets videnskabsteoretiske udgangspunkt. Til sidst introducerer jeg til en central dikotomi angående sygdomsbegreber og præsenterer den norske filosof Alf Ross' sygdomsbegreb, der vil danne baggrund for senere analyse.

2.1 Projektdesign

Med inspiration fra den svenske sociolog Thomas Brante vil jeg identificere de videnskabelige niveauer jeg arbejder på for at klargøre strukturen af undersøgelsen (Brante 2000). Brantes model er inspireret af den naturvidenskabelige niveaudeling i distinkte, men samtidig afhængige videnskaber. Biologien bygger på kemien, der bygger på fysikken, der bygger på matematikken. Hver videnskab er ligeledes bygget op af sub-niveauer baseret på forskellige strukturer og kausalmekanismer. Hvert niveau er en forudsætning for niveauet højere, og kan yde en delvis forklaring af dette niveau. Men samtidig har det højere niveau en vis autonomi og kan ikke reduceres til niveauet under. Endelig kan de højere niveauer udgøre nogle miljøer der kan have indflydelse på de lavere niveauer (Brante 2000: 11). I enhver undersøgelse vil der således være et niveau for interesse (I). Dette eksisterer som selvstændigt niveau, med egne strukturer og kausale mekanismer. Men det vil også være betinget af nogle underliggende niveauer (U) i de komponenter det bygger på, og desuden vil det udspille sig i en kontekst (O) som påvirker det 'oppefra'. I dette speciale kan jeg identificere kontekst, niveau for interesse og det underliggende niveau som følger:

(O) Konteksten for undersøgelsen er den stigende forekomst af danskere, der diagnosticeres og medicineres for depression. Udviklingen gælder for hele den vestlige verden og udgør et betydeligt samfundsmæssigt økonomisk problem. Der er flere (komplementerende) måder at anskue, hvad udviklingen skyldes. En måde er at se det stigende antal personer diagnosticeret med depression som udtryk for, at sundhedssektoren er blevet bedre til at opdage mennesker, der lider af depression. Der er således ikke tale om en stigning i antal deprimerede, men i antal diagnosticerede. Et andet bud er, at der er flere mennesker bliver deprimerede i forhold til tidligere, da tidens individualisme stresser individet med dets krav om evig selvrealisering, der fører til udmattelse og depression (Ehrenberg 2010). Et tredje bud er, at man diagnosticerer lidelse og problemer, der ikke hidtil har været anset for patologiske og dermed genstand for sundhedsvæsenet. Den engelske

sociolog Nikolas Rose (2006, 2010) taler således om en biologisk omformning af ubehaget, hvor normale livsproblemer gøres til medicinske anormaliteter. Et centralt spørgsmål på dette niveau er altså, om mennesker er syge, om mennesker bliver syge eller om mennesker italesættes/defineres som syge. Eller endnu bredere formuleret: Hvad forstår vi ved sygdom? Den kulturelle og sociale kontekst er altså af betydning for diagnosticeringen og behandlingen af depression.

(I) Niveaulet for interesse er den kliniske praksis, hvor diagnosticeringen og behandlingen foregår. Langt de fleste patienter med depression behandles i *almen praksis*, dvs. hos den praktiserende læge og jeg tager derfor udgangspunkt i den behandlingspraksis, der finder sted her. Lægen er naturligvis indført i den medicinske videnskab gennem sin uddannelse, der blandt andet indeholder bestemte måder at forstå sygdom, bestemte måder at opspore sygdom og bestemte metoder til at opnå viden om sygdom. Den praktiserende læges opgaver er meget forskelligartede, og der kræves derfor en bred medicinsk indsigt, da almen praksis er den almindelige borgers første og primære kontakt til sundhedsvæsenet. Lægens rolle er derudover ikke begrænset til det rent medicinske genstandsfelt, men kræver i højere grad end andre medicinske specialer, at fysiske, psykologiske, sociale, kulturelle og eksistentielle faktorer integreres i behandlingsindsatsen (Sundhedsstyrelsen 2013: 6). I behandlingsvalget for den enkelte patient skal der derfor indtænkes både rationel empirisk viden og humanistisk forståelse (Gøtzsche 2007: xii). Retningslinjerne for en sådan behandling for depression er beskrevet af sundhedsstyrelsen i deres Referencegram for behandling af unipolar depression fra 2007 og af Dansk Selskab for Almen Medicin i deres vejledning fra 2010.

(U) Det underliggende niveau er de videnskabelige strukturer, der betinger behandlingspraksis. De praktiserende læger uddannes i bestemte metoder til at anvende viden og opnå ny viden til brug i behandlingspraksis. Således indgår det i Sundhedsstyrelsens målbeskrivelse af speciallægeuddannelsen i almen medicin, at der skal udvikles kompetencer i anvendelsen af evidensbaserede behandlingsvejledninger og kendskab til evidensbaseret viden (Sundhedsstyrelsen 2013: 44; 60; 72). Desuden arbejder den praktiserende læge, og alle andre læger i øvrigt, med den diagnostiske sygdomsklassifikation ICD-10 som det centrale referencepunkt for at stille diagnose, vælge behandlingsform, stille en prognose, opsøge viden og kommunikere til patienter og andre parter om sygdom.

Denne undersøgelse kigger på det underliggende niveau. Mine tre arbejdsspørgsmål retter sig således hver mod undersøgelsen af en videnskabelig struktur. Det drejer sig dels om evidensmetoden som afgørende for, hvad der udgør pålidelig viden, dels om det sygdomsbegreb

man arbejder med i klassifikationen af psykiske sygdomme, og dels om de specifikke kriterier for diagnosen depression.

Specialet er bygget op om to gange tre kapitler. For hvert arbejdsspørgsmål er der et kapitel der undersøger den videnskabelige struktur, dens formål og teoretiske baggrund, efterfulgt af et kapitel der forholder sig analyserende til materialet med henblik på at besvare arbejdsspørgsmålet og afdække de specifikke betingelser for behandling af depression. Kapitel 3 og 4 drejer sig således om evidensmetodologien, kapitel 5 og 6 handler om klassifikation og sygdomsbegreb og kapitel 7 og 8 handler om diagnosen depression. I kapitel 9 samler jeg trådene fra de tre delanalyser i en diskussion af evidensmetodologiens anvendelse indenfor vidensproduktion om depressionsbehandling og i kapitel 10 konkluderer jeg på min undersøgelse.

2.2 Afgrænsning

Sundhedsstyrelsens anbefaling er ikke udelukkende baseret på RCT studier. Fx optræder der mange 'kliniske tips' i anbefalingerne, der i højere grad er erfaringsbaseret viden end evidensbaseret viden. Anbefalingerne er desuden ikke kun en vejledning i den bedste behandling, den angiver også rammerne for ressourceanvendelse i sundhedssystemet. Disse rammer er ikke blot resultatet af videnskabelig evidens, men også af politiske og økonomiske prioriteringer. De ressourcemæssige rammer for behandling kan være af stor betydning for, hvorvidt der tilbydes medicinsk behandling eller psykoterapeutisk behandling, da sidstnævnte er væsentligt dyrere for sundhedsvæsnet, og desuden for den enkelte patient, da der i langt de fleste tilfælde er en vis egenbetaling. Både samfundsøkonomiske prioriteringer og privatøkonomi kan altså have indflydelse på valg af behandling. Min undersøgelse favner derfor ikke det samlede spekter af faktorer, der betinger behandlingspraksis for depressionsbehandling, men begrænser sig til den videnskabelige legitimering for sundhedsvæsnets praksis.

Der er en stor og vidtrækkende diskussion i spørgsmålet om medicinalindustriens rolle ift. forskning og påvisning af behandlingseffekter. Peter Gøtzsche udgav i 2013 den stærkt kritiske bog *Dødelig medicin og organiseret kriminalitet*, om medicinalvirksomhedernes indflydelse på forbruget af antidepressiva, blandt andet i kraft af medicinalvirksomhedernes patenter på forskning i de medicinske præparater. Emnet er selvsagt af enorm betydning for en undersøgelse af det videnskabelige grundlag for forbruget af antidepressiva, men ikke et emne der vil blive behandlet i

dette speciale, hvor jeg ikke forholder mig til redeligheden af den eksisterende forskning, men til de metodiske retningslinjer for evidensbaseret medicin.

Produktionen af evidens er i det hele taget under indflydelse af interesser; politiske, professionsfaglige, kommercielle og civile fx i form af patientorganisationer. Dette speciale undersøger ikke de faktiske processer, hvorunder evidens bliver til, men koncentrerer sig om hvilket fundament denne viden hviler på, og hvilke implikationer det har for vores viden om depressionsbehandling. Jeg antager altså ikke, at der er et en-til-en forhold mellem evidensmetodologien og behandlingspraksis, men jeg mener, at den udgør en særdeles væsentlig faktor.

I og med at det store forbrug af antidepressiva er min motivation for denne undersøgelse, er det evidens om medicinske behandlingsmetoder for depression, som jeg tager udgangspunkt i. Der ligger imidlertid en meget interessant diskussion i spørgsmålet om, hvilken betydning evidensmetodologien har for dokumentationen af behandlingseffekter af psykoterapi samt andre psykologiske og sociale behandlingsindsatser. EBM er udviklet til at undersøge effekter af farmakologisk behandling, og flere centrale principper i metodologien kan være direkte modstridende med behandlingsrationalerne i psykologiske og sociale indsatser, hvilket kan have betydning for disse indsatsers mulighed for at konkurrere med de farmakologiske indsatser (Gupta 2007: 108). Dette kan naturligvis have stor betydning for, i hvilket omfang antidepressiva vælges over andre behandlingsformer for depression. Desværre vil en analyse af evidensproduktion om psykoterapeutisk behandling være for omfattende at foretage i dette speciale.

2.3 Empirisk materiale

Det anvendte empiriske materiale angår dels Evidensbaseret Medicin (EBM) og dels psykiatrisk klassifikation og diagnosen depression.

For den del der omhandler EBM, har jeg udvalgt materiale, der bedst muligt repræsenterer den metodiske position. De tekster jeg tager udgangspunkt i har været toneangivende for udbredelsen af EBM som forskningsmetode og redskab til behandlingsvalg i den kliniske praksis. Udtrykket Evidensbaseret Medicin blev første gang introduceret i 1992 af en gruppe forskere fra McMaster University i Canada, der kaldte sig Evidence-Based Working Group. I en indflydelsesrig artikel proklamerede gruppen indførelsen af et nyt paradigme i den medicinske praksis under overskriften *Evidence-Based Medicine. A New Approach to Teaching the Practice of Medicine* (1992). Jeg tager

desuden udgangspunkt i den evidens, der produceres af Cochrane Collaboration, der er en international organisation, som baserer sig på forskeres frivillige bidrag til samlingen af systematiske forskningsoversigter. Retningslinjerne, for hvordan de systematiske forskningsoversigter udarbejdes, er beskrevet i *Cochrane Handbook for Systematic Reviews of Interventions* (2011), der ligger tilgængelig på organisationens hjemmeside. Jeg benytter mig også af materiale på Nordic Cochrane Centers hjemmeside. Desuden kigger jeg nærmere på et enkelt Cochrane review og et enkelt RCT angående depressionsbehandling, der tjener til at illustrere evidensmetodologien som vidensform. Det valgte RCT indgår i det valgte Cochrane review og kan derfor forventes at leve op til Cochranes kvalitetskrav. Jeg tager desuden udgangspunkt i håndbogen *Evidence-Based Medicine. How to Practice and Teach EBM, 3.ed.* (2005) af Sharon E. Straus, W. Scott Richardson, Paul Glasziou og R. Brian Haynes, der er henvendt til praktiserende læger og er en indføring i, hvordan man indarbejder EBM's metoder og retningslinjer i den kliniske praksis. Haynes var del af Evidence-Based Working Group, ligesom flere andre medlemmer har været medforfattere på bogen i tidligere udgaver, blandt andre David Sackett, der har en meget central position i udviklingen af EBM. Angående beskrivelse af metoden bag RCT-studier, anvender jeg også bogen *Rational Diagnosis and Treatment. Evidence-Based Clinical Decision-Making* (2007) af formand for Nordisk Cochrane Center, prof. Peter Gøtzsches og prof. Henrik Wulff. Bogen er meget anvendt i dansk sammenhæng, men også internationalt og omtales således blandt andet i føromtalte håndbog. Endelig inddrager jeg også sekundære kilder omhandlende EBM, først og fremmest Olaf Rieper og Hanne Foss Hansens *Metodedebatten om evidens* (2007).

Angående psykiatrisk klassifikation og depression tager jeg udgangspunkt i de to omtalte klassifikationssystemer ICD-10 og DSM-5 af henholdsvis World Health Organisation og American Psychiatric Association. De to klassifikationer udgør internationale anvendte standarder for klassifikation af psykiske lidelser og er dermed den centrale referenceramme for al forskning og kommunikation i det hele taget ang. psykiske lidelser. Klassifikationerne revideres løbende og ICD-11 forventes at udkomme i 2017. Revideringsprocessen er åben for, at eksperter og interessenter kan stille forslag, kommentere og generelt deltage i revisionen. Jeg anvender begge klassifikationers indledende materiale angående definitioner af mental forstyrrelse og overvejelser om klassifikation generelt. For ICD-10 angår det indledningen til kapitel 5, der omhandler mentale og adfærdsmæssige forstyrrelser. Angående diagnosen depression anvender jeg kun ICD-10's diagnostiske kriterier, da det er disse man anvender i dansk sammenhæng. Desuden inddrager jeg tre redskaber til diagnostik og vurdering af sværhedsgrad for depression, kaldet psykometriske tests.

Det drejer sig om testen Major Depression Inventory test (MDI), der er meget udbredt i dansk sammenhæng og anbefales i forbindelse med diagnosticering i Sundhedsstyrelsens anbefalinger. Den anden hedder Hamiltons Depression Rating Scale (HAMD), der internationalt set er meget udbredt, også anbefales i Sundhedsstyrelsens anbefalinger til vurdering af sværhedsgrad og desuden anvendes i det eksempel på et Cochrane review, som jeg bruger. Den tredje hedder Montgomery & Åsberg Depression Rating Scale, der også er meget anvendt og bruges i det eksempel på et RCT som jeg inddrager.

Til sidst kan jeg nævne, at Sundhedsstyrelsens *Referenceprogram for unipolar depression hos voksne* (2007) og Dansk Selskab for Almen Medicins *Klinisk vejledning for almen praksis: Unipolar depression, diagnostik og behandling* (2010) løbende inddrages for at vise sammenhæng mellem de beskrevne videnskabelige strukturer og de retningslinjer som anvendes i klinisk praksis.

2.4 Videnskabsteoretiske antagelser

Dette speciale kan siges at være en undersøgelse af videnskabsteori i praksis. Gennem hele specialet vil jeg forsøge at sætte de videnskabelige rationaler og metoder, der afdækkes i forhold til den forståelsesramme, de er født af, og som derfor anviser omfanget af deres forklaringskraft. Jeg vil også forsøge at vise, hvordan evidensmetodologien har betydning for den indre konflikt, der ligger i psykiatrien, i kraft af dens splittelse mellem en grundlæggende psykologisk og en grundlæggende biologisk forståelse af psykisk sygdom.

En videnskabsteoretisk antagelse, der løber gennem specialet er således eksistensen af konkurrerende videnskabelige paradigmer. I en Kuhnsk terminologi kan man sige, at psykiatrien endnu ikke er modnet som videnskab. Der er ikke etableret en normalvidenskab; en overordnet teori, der formår at samle det videnskabelige felt. Derfor er en vis grad af inkommensurabilitet mellem forskellige måder at anskue psykiske problemer også forventelig (Kuhn 1978). Denne spænding mellem to tilgrundliggende videnskabelige paradigmer vil udgøre en central forklaringsramme gennem specialet.

En anden antagelse, der vil have betydning for min analyse af evidensmetodologien, er den empiriske metodes teoriafhængighed. Skiftende videnskabsteoretikere fra Popper (1934) til Kuhn (1962) og Feyerabend (1975) har op gennem det forrige århundrede dokumenteret, hvordan ingen videnskabelig metode sikrer fuld objektivitet i empiriske observationer (Harari 2001: 725, Bhaskar

2008: xxx). Forskeren bør derfor altid være bevidst om det teoretiske udgangspunkt der ligger i det videnskabelige felts synsvinkel.

Jeg tager endvidere udgangspunkt i Thomas Brantes variant af kritisk realisme (Bhaskar 1975), som han kalder kausal realisme. Den kritiske realisme kombinerer ontologisk realisme med epistemologisk relativisme i udformningen af en objektiv, men fejlbarlig teori om virkeligheden (Brante 2000:6). Det metodiske postulat, at al viden er fejlbarlig, lægger op til en metodisk og teoretisk pluralisme i en bestræbelse på ikke at ende i en naiv realisme. Denne erkendelsesteoretiske antagelse vil være en tematik som jeg vil underbygge i specialets analyse af evidensmetodologien i kapitel 4.

Brante lægger desuden vægt på udviklingen af et årsagsbegreb, der rummer mulighed for grader af kausalitet. I stedet for kun at forstå kausalitet som lovmæssigheder, må man også kunne tale om tendenser og tilbøjeligheder og forstå disse i relation til de omstændigheder, de optræder under. Den kausale realisme indebærer altså en erkendelse af, at langt de fleste årsagsforhold er multikausale, og at det væsentlige derfor bliver at afdække samspillet af samvirkende og modsatrettede tendenser. Dermed afvises også det Hume'ske årsagsbegreb, bestående i regularitetsteori, altså definitionen af kausalitet som associationer mellem observerede fænomener (Brante 2000: 7ff).

Jeg mener, at denne kausalitetsforståelse er frugtbar for en undersøgelse, der omhandler psykisk sygdom. Den somatiske sygdomsteori arbejder netop med multikausale net af årsagsrelationer, hvis betydning vil blive diskuteret i kapitel 5. Som tidligere nævnt indebærer psykiatriens tværvideenskabelige position integreringen af forskellige perspektiver på udviklingen af psykiske sygdomme, der indebærer, at vidt forskellige kausale relationer indgår i et eller andet form for samspil.

2.5 Sygdomsbegrebet

Man kan skelne mellem to overordnede traditioner indenfor debatten om sygdomsbegrebet: Den naturalistisk-objektivistiske og den relativistisk-subjektivistiske. I den naturalistisk-objektivistiske forståelse kan sygdomme defineres som en afvigelse fra artens design, ved at en del af organismen fungerer under artstypisk niveau. I den relativistisk-subjektive tradition er sygdom derimod defineret ved en tilstand, der er kulturelt og personligt uønskværdig. Hvad der gælder som sygdom, kan derfor bestemmes i læge-patient forholdet, og kan ikke sættes på fast formel (Thrysoe 2007: 129f). Indenfor de to traditioner vil der således også være forskel på, hvordan behandling skal

forstås. Indenfor den naturalistiske tradition forstås sygdomme ofte essentialistisk som sygdomsentiteter, biologiske anomalier eller fremmedlegemer, der kan observeres af lægen som legemligt afgrænsede entiteter, der udgør objekter for intervention. Indenfor den relativistisk-subjektivistiske tradition vægter derimod erfaringen af sygdom som et væsentligt perspektiv i behandlingssammenhæng. Sygdom bliver først sygdom i oplevelsen af nedsat funktion, af ikke at fungere i dagligdagen, af lidelse og af livskvalitet (Gannik & Launsø 2000: 7). Individet forholder sig til deres sygdom, den har bestemte betydninger for dem og de handler og indvirker på deres sygdomsforløb. Sygdom kan derfor ikke behandles uafhængigt af individets subjektivitet (Gannik 2008: 36). Diskussionen er ikke mindst interessant i psykiatrien, hvor den føjer perspektiver til den eksisterende konflikt mellem den biologiske og psykologiske forståelse af psykisk lidelse. Diskussionen vil således også komme til udtryk i min analyse af evidensmetodologien.

For at skabe en ramme omkring, hvilken jeg kan diskutere de sygdomsbegreber, der kommer til udtryk i evidensproduktionen og i den psykiatriske sygdomsklassifikation, har jeg brug for at se nærmere på hvordan sygdom mere specifikt kan defineres. Mellem de to traditioner indenfor sygdomsbegrebet findes en række forsøg på brobygninger. Jeg har valgt at tage udgangspunkt i den norske filosof Alf Ross der netop placerer sig mellem de to traditioner og forsøger at løse op for modsætningen. Alf Ross' sygdomsbegreb er desuden udviklet til både at gælde for somatisk og psykisk sygdom.

2.6 Alf Ross' sygdomsbegreb

Hos Alf Ross defineres sygdom funktionelt. Ligesom det naturalistisk-objektivistiske sygdomsbegreb taler han om væsentlige afvigelser fra den menneskelige arts normalplan, men i den forstand, at normalplanet afspejler et normativt ideal for menneskets funktioner, til forskel fra et statistisk eller gennemsnitligt begreb for det normale. Dette ideal omhandler den menneskelige organismes udfoldelse af sine livsfunktioner. Normalplanet udtrykker således de optimale betingelser for organismens livsduelighed. Ofte viser afvigelser fra normalplanet sig i behov – en mangel i organismen, som den reagerer på ved at søge afhjælpning eller blot ved at føle ubehag, mens det i enkelte andre tilfælde er andre mennesker, der konstaterer manglen ved at sammenligne med en idealforestilling om menneskets funktionalitet (Ross 1979: 122ff). Dermed gøres sygdom til et anliggende, der står i forhold til subjektet og ikke til arten, uden at det af den grund bliver relativistisk. Normalplanet angiver en ramme for *”hvad individet skulle have ”ret” til at beklage sig over og ønske afhjulpning”* (Ross 1979: 126). Dette ønske om afhjælpning, siger Ross, varierer med

de forestillinger, man gør sig om sygdommens natur og de muligheder man besidder i behandlingshenseende. Desuden har sygdom en kvantitativ dimension. *"Sygdomsulemperne må overskride en vis toleranstærskel for individet vil blive anset for syg. Og der kan næppe være nogen tvivl om, at denne tærskel ligger højst forskelligt i forskellige samfund."* (Ross 1979: 126). Ross påpeger altså, at der må være tale om en væsentlig afvigelse fra normalplan, men samtidig kan der ikke sættes et mål for størrelsen af denne, og den vil variere afhængigt af kultur.

Det giver ifølge Ross ikke mening at spørge, hvad sygdommen er i sig selv, til forskel fra dens manifesterede symptomer. Sygdom er netop defineret i forhold til den funktionelle betydning den har for individet (Ross 1979: 128). Der er ikke fra naturens hånd nogen adskillelse mellem det normale og det patologiske. Der er ikke noget særegent ved de "syge" biologiske processer; de følger naturens love ligesom de "normale". Der findes så at sige ikke sygdomme, men kun syge mennesker. Derfor kan et sygdomsbegreb kun anlægges ud fra et *"i idealisering grundet normativt synspunkt"* (Ross 1979: 117). Det er interessant at bemærke, at de materielle strukturer således ikke spiller nogen direkte rolle i definitionen af sygdom, de materielle strukturer i kroppen er kun relevante som bærere af funktioner. Dette giver også mulighed for et sygdomsbegreb, hvor der strukturelt set ikke er nogen forskel på psykisk og somatisk sygdom (Pedersen 2006: 10).

For de mentale lidelser opridser han således en række funktioner, der er definerende for et mentalt normalplan. Ross peger på følgende funktioner, men listen er ikke nødvendigvis fuldkommen: *"sansning, perception, læreproces, tænkning og sprog, motivation, følelse og emotion samt adfærd"* (Ross 1980: 5). Ved en mental funktionsforstyrrelse, er der ifølge Ross tale om en kompromittering af patientens mulighed for at kommunikere med andre mennesker. Grundlaget for en god interpersonel kommunikation er en vis enighed i opfattelsen af os selv og den omgivende verden. Depressionen er derfor *"den umotiverede (eller helt urimeligt motiverede) følelse af ulyst, ulykke, sorg, nedslåethed stigende til fortvivlelse og dyb melankoli, og forbundet med svækkelse eller lammelse af beslutnings- og handledygtighed."* (Ross 1980: 10) Der er altså tale om en forstyrrelse af emotionen, fordi følelserne er umotiverede eller urimelige, og derfor står personen ude af stand til at kommunikere meningsfuldt med sine omgivelser. Hvis den almindelige opfattelse af virkeligheden er forstyrret, er der ikke mulighed for forståelse, og dermed ikke noget grundlag for kommunikationen. Ross indfører således en norm i forhold til hvilken, det bliver muligt at skelne mellem patologisk eller "syg" lidelse og anden ikke-patologisk lidelse.

Kapitel 3: Evidensmetodologien

Dette kapitel har til formål at redegøre for, hvad det er for en metodisk tilgang, der menes, når man taler om evidens. Jeg tager afsæt i nogle af de mest centrale tekster omhandlende grundlaget for evidensbaseret medicin (EBM), materiale fra The Cochrane Collaboration's guidelines til udarbejdelse af reviews og øvrigt materiale tilgængeligt på Nordisk Cochrane Centers hjemmeside. Jeg kigger desuden nærmere på et RCT og en systematiseret forskningsoversigt, der begge undersøger effekter af antidepressiva. Håndbogen i evidensbaseret klinisk praksis *Evidence-Based Medicine. How to Practice and Teach EBM* inddrages også.

3.1 Evidensbegrebets rødder

Begrebet om evidensbaseret medicin (EBM) tager sit afsæt i den britiske læge og forsker Archibald Cochrane, der i 1972 med bogen *Effectiveness and Efficiency: Random Reflections on Health Services* efterlyste en systematisk opsummering af alle velgennemførte kliniske forsøg (Hjørland 2011: 1301). Cochrane mente at den lægelige praksis og viden ofte hvilede på teorier eller traditioner, der ikke var empirisk efterprøvede, hvilket gav anledning til forskelligartet og ineffektive eller direkte skadelig behandlingspraksis. De kliniske forsøg skulle fungere som en fair test af behandlingseffekten. Det skulle sikre patienter, læger og andre beslutningstagere et grundlag for at kunne tage kvalificerede beslutninger angående valg af behandling og forebyggelse og gøre det muligt at foretage ressourcemæssige prioriteringer i sundhedsvæsenet (Cochrane Collaboration, 3; Hansen & Rieper 2010: 13).

Metodisk tager EBM sit afsæt i den kliniske epidemiologi. Den kliniske epidemiologi studerer forekomsten af sygdom og sygdomsårsager i populationer og analyserer de statistiske sammenhænge mellem disse (Sørensen 2004: 39). Evidensbaseret medicin beskrives som en integrering af den kliniske forskning med lægens ekspertise og vurdering af den enkelte patient. Den oftest anvendte definition af EBM er formuleret af epidemiologen David Sackett og lyder: ”*Evidence-based medicine is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of the individual patient*” (Sackett et al 1996: 71).

Evidensbaseret medicin introduceredes i forlængelse af Cochranes tanker af en gruppe forskere fra McMaster University i Ontario, der kaldte sig ”Evidence Based Medicine Working Group”. I deres artikel *Evidence-Based Medicine. A new approach to teaching the practice of medicine* fra 1992,

bliver EBM omtalt som et nyt paradigme indenfor medicinsk praksis. Følgende citat beskriver, hvordan EBM adskiller sig fra tidligere medicinsk praksis:

Evidence-based medicine de-emphasizes intuition, unsystematic clinical experience, and pathophysiological rationale as sufficient grounds for clinical decision making and stresses the examination of evidence from clinical research. (Evidence-based medicine working group 1992: 2420).

Forfatterne anskuer EBM som et paradigmeskifte indenfor den medicinske praksis' metodologi, idet den formulerer en række problemer ved den nuværende praksis og angiver en vej for, hvordan der fremover skal findes løsninger for disse problemer. Oparbejdningen af en omfattende database over kliniske forsøg skal danne grundlag for klinisk praksis, frem for "intuition, usystematisk klinisk erfaring og patofysiologiske rationaler". Det er blandt andet en ukritisk brug af overleveret viden, gruppen sigter mod. Med "usystematisk klinisk erfaring" henviser man for eksempel til det forhold, at medicinske guidelines og tekstbøger kan basere sig på ekspertvurderinger, der ikke tager den seneste forskning i betragtning, og dermed potentielt sætter den medicinske praksis tilbage. Når der desuden gøres op med "patofysiologiske rationaler", er der tale om det som Jeremy Howick kalder "mechanistic reasoning", altså konklusioner om behandlingsmetoder, der afledes af den aktuelle viden (eller formodede viden) om sygdommens oprindelse. Det er derfor en afvisning af teoretisk viden om sygdommens mekanisme og årsag, som det primære fundament for behandlingsvalg. Den nye evidensbaserede medicin skal derimod bygge på direkte observationer fra den kliniske forskning, foretaget gennem empiriske forsøg (Howick 2011: 16).

I den aktuelle 3. udgave af den klassiske introduktion til EBM 'Evidence-Based Medicine. How to Practice and Teach EBM' fra 2005 (Straus et al.), er forfatterne aldeles eksplicitte angående denne pointe. I kapitlet 'How to find the best evidence' lyder første punkt: "*Burn your traditional textbooks*". Forklaring følger:

"Although we may find some useful information in texts about "background questions", such as the pathophysiology of clinical problems, it is best not to use them for seeking the answers to "foreground questions", such as the causal (risk) factors, diagnosis, prognosis, prevention or treatment of a disorder if there is an up-to-date, evidence-based alternative. About now, you might be asking if we really mean that you should burn your traditional textbooks. We do." (Straus et al 2005: 32)

Patofysiologisk viden er altså ifølge Straus et al. udmærket som "baggrundsviden", men den bør hverken anvendes til at besvare spørgsmål om diagnose, prognose, forebyggelse, behandling eller endog kausalfaktorer - da man med kausalfaktorer lægger vægt på epidemiologiske metoder til at

forstå sygdommes årsagsforhold, altså observationer af sammenfaldende faktorer, jf. David Sacketts *Clinical Epidemiology. A Basic Science for Clinical* fra 1991. Det er således uklart, hvilken rolle kendskabet til de patofysiologiske sygdomsmekanismer skal spille i den kliniske praksis ifølge evidensmetodologiens fortalere. Skønt den originale artikel af Evidence-Based Working Group gør opmærksom på nødvendigheden af et kendskab til de patofysiologiske sygdomsmekanismer (1992: 2421), er det heller ikke her klart hvilken vægt denne viden tillægges.

For at kunne basere sine beslutninger på evidens må klinikerens besidde en forståelse af ”certain rules of evidence” der sætter ham i stand til at tolke og anvende forskningsresultaterne i praksis (Evidence-Based Working Group 1992: 2421). I den forbindelse må der også oparbejdes nogle specifikke færdigheder, der sætter klinikerens i stand til at foretage en kritisk tilegnelse af forskningslitteraturen. Det handler først og fremmest om at kunne foretage en litteratursøgning, anvende ’the rules of evidence’ til at vælge det bedste af de relevante studier og vurdere dets validitet, og endelig udvinde den information, der er relevant for klinisk praksis og anvende den på patientens problem (Evidence-Based Working Group 1992: 2421).

Det er altså den enkelte læge der i sit behandlingsvalg skal konsultere forskningsdatabasen og vurdere de tilgængelige studiers kvalitet og anvendelighed. Allerede i 1992 kritiseredes denne målsætning dog for at være urealistisk. Med den mængde af publikationer den almindelige læge skulle læse for at holde sig opdateret, ville han ikke have tid til andet. I 2010 ansloges det at der alene for RCT-studier udgives 75 artikler om dagen og 11 systematiske reviews (Bastian et al. 2010: 1). Bluhm og Borgerson fremhæver at på trods af at et centralt aspekt ved evidensbevægelsens idegrundlag var ønsket om at klinikerne blev mindre afhængige af autoriteter, har bevægelsen selv resulteret i oprettelsen af en række institutioner der udgør autoriteter indenfor evidensbaseret medicin. Herunder er en central institution The Cochrane Collaboration. Den kritiske evaluering af forskningen som krævedes af den enkelte læge, er således i praksis udliciteret til en række evidensproducerende organisationer. (Bluhm & Borgerson 2011: 215f).

3.2 The Cochrane Collaboration

The Cochrane Collaboration blev grundlagt i 1993 for at fremme produktionen af systematiske forskningsoversigter. Forskningsoversigterne der kvalitetsvurderer og syntetiserer resultaterne af de enkelte forsøg, tillægges den højeste grad af evidens – det vil sige at man anser den viden der er resultatet af en systematisk forskningsoversigt for at være den viden man kan være mest sikker på.

Flere af medlemmerne fra Evidence-Based Medicine Working Group blev således også ved stiftelsen en del af Cochrane-organisationen (Hansen & Rieper 2010: 10). Cochrane reviews (som forskningsoversigterne også hedder) har altså en central status, når man ser på, hvilken slags viden der anses for god evidens. Eksempelvis anføres det i *Referenceprogram for unipolar depression hos voksne*, at ”cochrane-reviews, metaanalyser eller andre systematiske forskningsoversigter er tillagt det højeste evidensniveau” (Sundhedsstyrelsen 2007: 10). Der er således et formelt system, der angiver cochrane-reviews og andre systematiserede oversigter af samme slags, som den mest sikre viden at basere sin praksis på.

3.3 Evidenshierarkiet

En central ”rule of evidence” er rangeringen af forskningsdesign i et hierarki. Det gør klinikerer i stand til at vurdere forskningsresultaternes værdi, således at hvis to studier viser modstridende resultater, er der klare rammer for hvilke resultater, der kan tillægges størst vægt på baggrund af typen af forskningsdesign, der er anvendt (Rieper & Hansen 2007: 18). Rangeringen er designet til at afspejle forsøgstypernes metodologiske styrke. Det antages at jo højere placering en forsøgstype har i hierarkiet, des mindre sandsynlighed er der for at den er påvirket af bias, des højere sandsynlighed er der for at den afspejler en kausal sammenhæng korrekt, og mere sandsynligt er det at resultaterne kan generaliseres til at gælde en bredere patientpopulation end forsøgspopulationen (Bluhm & Borgerson 2011: 210).

Evidenshierarkiet findes i forskellige varianter, men er altid bygget op således at randomiserede kontrollerede forsøg ligger øverst, efterfulgt af ikke-randomiserede kontrollerede forsøg, forløbsundersøgelser og nederst ”anekdotisk” viden, som omfatter ekspertvurderinger og eksempler på god klinisk praksis (Rieper & Hansen 2007: 19).

Sundhedsstyrelsens anbefalinger er således lavet på baggrund af en vurdering af de forskellige studier ud fra evidenshierarkiets rangering. Sundhedsstyrelsen angiver nedenstående hierarki for forskningsmetoder:

Publikationstype	Evidens	Styrke
Metaanalyse, systematisk oversigt Randomiseret kontrolleret studie (RCT)	Ia Ib	A
Ikke-randomiseret kontrolleret studie Kohorteundersøgelse Diagnostisk test (direkte diagnostisk metode)	IIa IIb	B
Case-kontrol-undersøgelse Diagnostisk test (indirekte nosografisk metode) Beslutningsanalyse Deskriptiv undersøgelse	III	C
Mindre serier. Oversigtsartikel. Ekspertvurdering. Ledende artikel	IV	D

Kilde: Referenceprogram for unipolar depression hos voksne, SfR, Sundhedsstyrelsen.

Gradueringen angiver hvilken viden der skal tillægges hvilken grad af evidens og dermed i hvor høj grad denne viden bør danne udgangspunkt for den anbefalede behandlingspraksis. Hierarkiet er, for at være præcis, ikke egentligt et hierarki for evidens. Hierarkiet forholder sig altså ikke til værdien af den frembragte viden. Det siger således ikke noget om hvilke kvalitetskriterier der fx adskiller et godt casestudie fra et dårligt casestudie, og dermed hvilke kvalitetskrav der udløser en højere grad af evidens. (Bluhm & Borgerson 2011: 221). Hierarkiet forholder sig til de forskningsmetoder, der er anvendt til at frembringe viden, og det vil sige at et RCT-studie altid nyder en højere grad af evidens og udgør mere sikker viden end eksempelvis et casestudie, uanset hvilken viden de ellers måtte frembringe.

De systematiske forskningsoversigter og metaanalyser bygger på de enkelte studier, der er foretaget ang. fx en bestemt behandlingsindsats' virkning. Men givet den kvalitetsvurdering der går forud for udarbejdelsen af oversigterne, er det langt fra alle typer studier der inddrages. I Cochrane reviews baserer man sig som sagt så vidt muligt kun på RCT-studier (Higgins & Green 2011: chapter 5.5).

I *Evidence-Based Medicine. How to Practice and Teach EBM* er man på linje med denne praksis. Straus et al. anbefaler, at klinikerne, der skal finde evidens om effekten af en behandlingsindsats, altid starter med at lede efter forskningsoversigter. Forskningsoversigter der udelukkende baserer sig på RCT-studier skal ifølge Straus et al. foretrækkes, og det anbefales ikke at benytte sig af forskningsoversigter der inddrager ikke-randomiserede studier, medmindre disse analyseres separat. Risikoen er at enkelte af studierne er misledende og derfor resulterer i dårligere evidens (Straus et al. 2005: 149). I gennemgangen af primærstudier foreslår Straus et al. ligeledes at man springer alle

artikler over der ikke baserer sig på RCT:” *We can begin to rapidly critically appraise articles by scanning the abstract to determine if the study is randomized; if it isn't, we can bin it*” (Straus et al. 2005: 118)

De randomiserede kontrollerede forsøg samt metaanalyser af samme har altså en utrolig stærk position i evidensbaseret medicin og forskning i behandlingseffekter. Jeg vil derfor i dette kapitel fokusere på netop disse studier, frem for at beskrive hver enkelt af metoderne der optræder i evidenshierarkiet. Jeg ser derfor i de følgende afsnit på metoderne bag RCT og Cochrane review og gennemgår et eksempel på hver.

3.4 Randomiserede kontrollerede forsøg

Randomiserede kontrollerede forsøg, eller RCT (randomized controlled trials), anses i evidensmetodologien for at være den bedste metode til producere sikker viden om en given indsats effekt og omtales som 'the gold standard' indenfor produktionen af evidens. Et RCT foregår ved en tilfældig udvælgelse af mindst én patient-/klientgruppe der skal modtage den givne indsats og mindst én kontrolgruppe. Kontrolgruppen udgør sammenligningsgrundlaget og modtager den normalt tilbudte indsats, en placebo-indsats eller ingen behandling.

For at eliminere alle andre variable end indsatsen, skal interventionsgruppen og kontrolgruppen være så ens som muligt. Ved at foretage en randomiseret udvælgelse af grupperne undgår man selektionsbias. For forsøg der ikke er randomiserede, kan en risiko være, at lægerne placerer patienterne i kontrol og interventionsgruppe på baggrund af hvem han tror vil have mest gavn af indsatsen. Dermed vil resultatet af forsøget have tendens til at blive for positivt (Cochrane Collaboration, 1). Desuden sikrer den randomiserede udvælgelse af grupperne for alle forudsete og uforudsete faktorer der kan have indflydelse på resultatet. Ideen er, at idet medlemmerne af indsatsgruppe og kontrolgruppe er tilfældigt udvalgt, vil de individuelle karakteristika hos deltagerne der kan have betydning for forsøget, statistisk set være gældende både for interventionsgruppen og kontrolgruppen (Rieper & Hansen 2007: 21). Ved at kunne sammenligne indsatsgruppen med en kontrolgruppe, sikrer man sig at den forskel man ser i effekten af indsatserne, rent faktisk kan tilskrives den indsats man ønsker at vurdere, derved isoleres årsags-virkningsforholdet mellem indsats og effekt mest muligt.

Desuden skelnes der mellem ublændede, enkeltblændede og dobbeltblændede RCT. De bedste forsøg er dobbeltblændede, dvs. at hverken forsøgslederne eller grupperne ved, hvem der er

interventionsgruppe og hvem der er i kontrolgruppe. Ideen her er ligeledes at sikre sig mod menneskelige bias der kunne få betydning for objektiviteten af de målte effekter. Bias behøver ikke have noget at gøre med at forsøgslederen bevidst forsøger at manipulere forsøget, det kan lige såvel være en ubevidst bias, der har indflydelse på eksempelvis hvilke symptomer han lægger mærke til og hvilke af patientens udtalelser han vægter (Sehon & Stanley 2003: 8).

Effekt mål

Der er mange måder at opgøre den målte effekt. Et mål kaldes *relativ risiko*, der angiver et estimat for hvor stor risikoen for et bestemt udfald, fx remission (aftagende til ingen sygdomstilstand) er for indsatsgruppen ift. kontrolgruppen (Gøtzsche 2007: 135). Ordet risiko forekommer forkert i dette tilfælde hvor der er tale om en bedring, men begreberne stammer fra sandsynlighedsregning og knytter sig ikke eksklusivt til negative udfald. Hvis den relative risiko er 1 er risikoen den samme, om man er i kontrol- eller indsatsgruppe. Er den relative risiko 2 er risikoen for remission dobbelt så stor for indsatsgruppen, altså 100% mere effektiv. Den relative risiko kan dermed vise imponerende procentsatser, selv hvor behandlingseffekten ikke er ret stor. Er der fx kun 2 ud af 100 placebopatienter der oplever effekt, vil en effekt på 6 ud af 100 i behandlingsgruppen betyde at indsatsen er 300% mere effektiv, på trods af at 94% af forsøgspopulationen ikke havde nogen gavn af behandlingen. Omvendt kan en mindre relativ risiko sagtens være udtryk for klinisk signifikante resultater, fx hvis kontrolgruppen modtager den normalt anvendte behandling. En relativ risiko på eksempelvis 1,1 kan således betyde at den testede behandling er effektiv i 10% flere tilfælde end den normalt tilbudte behandling, hvilket må siges at være en stor gevinst.

Men ét er, hvor god en behandling er i forhold til en kontrolindsats, et andet er hvor mange der rent faktisk opnår bedring, dvs. hvad den *absolutte risiko* er. Den absolutte risiko findes ved at tage den forholdsmæssige andel af patienter der har opnået bedring indsatsgruppen, og fratække den forholdsmæssige andel i kontrolgruppen (Cochrane Collaboration, 2). Tag fx to grupper med 80 patienter i hver. 60 patienter i indsatsgruppen oplever bedring, mod kun 40 i kontrolgruppen. Dermed er den absolutte risiko for bedring 25% ved brug af behandlingen $((60/80)-(40/80)=0,25)$. Den absolutte risiko er i særdeleshed relevant i forsøg hvor der anvendes placebo i kontrolgruppen, da den så kan forventes at give et billede af den reelle effekt, når man modregner den psykologiske effekt af behandlingsindsatsen.

Endelig vil jeg fremhæve målet Numbers Needed to Treat (NNT), der giver et intuitivt godt billede af effekten. NNT dækker over antallet for, hvor mange patienter man gennemsnitligt skal behandle for, at én patient har gavn af behandlingen. Den ideelle indsats har en NNT på 1, da det ville betyde, at alle patienter der modtager behandlingen opnår bedring og at ingen patienter behandles unødvendigt. NNT findes ved at beregne den reciprokke værdi af den absolutte risiko (Cochrane Collaboration, 2). I eksemplet fra før vil NNT være 4 ($1/0,25$). For hver fire patienter man behandler med indsatsen, opnår én bedring der kan tilskrives indsatsen.

Relativ risiko, absolut risiko og NNT anvendes til såkaldte *dikotomiske* resultatmålinger. Dvs. at der angives et succeskriterium, fx remission eller respons (der dækker over enhver signifikant bedring), og måles hvor mange patienter i indsatsgruppen sammenlignet med kontrolgruppen, der opnåede dette mål. Dvs. at det angår andelen af patienter. Effekten kan dog også opgøres i *kontinuerlige* resultatmål. Dvs. at i stedet for at have et succeskriterium, opgøres den gennemsnitlige effekt for den samlede gruppe af patienter.

Risikoberegning

For at være sikker på at den observerede effekt på indsatsgruppen skyldes indsatsen, arbejder man med en nul-hypotese. Nul-hypotesen siger at indsatsgruppen vil opleve samme effekt som placebo-gruppen. For at vise at en behandlingsindsats har en effekt, skal nul-hypotesen altså kunne afvises. Nul-hypotesen antager at der er to og kun to forhold, der kan øve indflydelse på resultatet af forsøget. Dels selve indsatsen, og dels den variation der er mellem individerne i de to grupper (Williams 2010: 107). For at nul-hypotesen skal kunne afvises, skal sandsynligheden for at den observerede effekt skyldes tilfældig variation være mindre end 5%. Jo flere personer der er med i forsøget, des mindre er muligheden for at indsatsgruppen og kontrolgruppen, trods randomisering tilfældigt adskiller sig signifikant fra hinanden pga. individuelle karakteristika. Ligeledes vil der, jo større effekten er, være mindre sandsynlighed for at forskellen kan tilskrives denne tilfældige variation mellem patienterne. Det vil sige at for forsøg med mange deltagere og stor effekt, vil det være meget usandsynligt at effekten ikke kan tilskrives indsatsen. Den beregnede sandsynlighed kaldes p-værdien og grænsen på 5% bruges i langt de fleste RCT til at angive om resultatet af testen kan siges at være *statistisk signifikant*. P-værdien angiver altså sandsynligheden for at nul-hypotesen ikke kan afvises, dvs. jo mindre p-værdien er, des mere sikkert er resultatet. De to variable; antal deltagere og effektens størrelse, kan bruges til at veje op for hinanden i beregningen af p-værdien. Et forsøg med en forholdsvis lille effekt, kan ved at inkludere tilstrækkelig mange

deltagere, opnå en lav p-værdi under 5%. Dermed er resultatet statistisk signifikant, men effekten af indsatsen er vel at mærke ikke blevet større af den grund. Man bør derfor skelne mellem statistisk signifikans og klinisk signifikans, hvor det sidste drejer sig om hvilken betydning resultatet bør have for patienterne, dvs. hvorvidt behandlingen kan forventes at virke eller ej (Gøtzsche 2007: 131ff).

Hvor p-værdien er et mål for resultatets styrke, siger det ikke noget om resultatets præcision. For at bedømme hvor *præcis* den målte effekts størrelse er, dvs. hvor stor den terapeutiske gevinst ved at vælge behandlingsindsatsen frem for placebo, må man beregne et *risikointerval*. Ligesom for p-værdien benytter man sig som oftest af et mål på 95% sandsynlighed. Har man fx målt den absolutte risiko til 10%, kan risikointervallet på baggrund af antal patienter i forsøget, fx ligge på 3%-17%. Det betyder at der er 95% sandsynlighed for, at den reelle effekt af indsatsen er 3-17% bedre end placeboindsatsen. De 95% sikkerhed kaldes *confidence interval* (CI) og det findes ved at beregne sandsynligheden af at opnå resultatet på 10%, uanset hvad den reelle procentdel måtte være. Pointen er, at hvis den reelle procentdel af patienter der kunne have gavn af en behandling er 1%, så ville der kun være en meget lille sandsynlighed for, at vi i forsøget har udvalgt en gruppe deltagere, hvoraf hele 10% har gavn af behandlingen. Der er endnu mindre sandsynlighed for, at den reelle andel af patienter der ville have gavn af behandlingen er 90%. Der vil imidlertid altid være en vis sandsynlighed, og det er derfor man i intervallet kun har 95% sandsynlighed. Derfor er der således altid en tyvende dels chance for at den reelle effekt af behandlingen ligger udenfor risikointervallet. Hvordan risikointervallet præcist beregnes, er desværre for matematisk kompliceret at gøre rede for her, men det vigtigste i denne sammenhæng er også at forstå hvilken form for statistisk sikkerhed risikointervallet dækker over. Jo flere patienter der indgår i forsøget, jo større sandsynlighed er der for, at den reelle effekt ligger tæt på det målte resultat, og des mere snævert vil risikointervallet således være. (Gøtzsche 2007: 101ff, 133f).

Grunden til at jeg gennemgår disse statistiske mål er ikke at jeg nu vil gennemgå en masse beregninger. Derimod ønsker jeg at vise hvilken form for mål man opgør effekterne i indenfor evidensmetodikken, og hvilken viden de bidrager med. I det følgende eksempel skal vi således se hvordan effekten af en behandlingsindsats mod depression opgøres.

3.5 Eksempel på et RCT

For at eksemplificere RCT-metoden vil jeg nu kigge nærmere på et specifikt forsøg. Jeg har valgt studiet “Escitalopram 10 mg/day is effective and well tolerated in a placebo-controlled study in depression in primary care”, der tester effekten af præparatet Escitalopram, et antidepressiva af typen SSRI i forhold til effekten af placebo, når det anvendes i den primære sundhedssektor. Studiet er foretaget af en britisk samt to danske forskere, der begge er ansat ved medicinalvirksomheden Lundbeck A/S. Studiet er ligeledes sponsoreret af Lundbeck.

Studiet er foretaget i 40 konsultationer i Canada, Estland, Frankrig, Holland og Storbritannien og løb over 8 uger. Det er udført som dobbeltblændet, med indsatsgruppe og kontrolgruppe kørende parallelt og med fast dosering af præparatet. Der deltog i alt 380 patienter i forsøget, 191 i indsatsgruppen og 189 i kontrolgruppen. Alle patienterne i forsøget var diagnosticeret med moderat til svær depression, og sværhedsgraden var målt på skalaen Montgomery Åsberg Depression Rating Scale (MADRS).¹ Alle deltagere var ud over depressionslidelsen i øvrigt sunde og raske, og vurderet egnede til at indgå i forsøget på baggrund af helbredstjek, sygehistorie, test med elektrokardiogram og laboratorietests. Patienter var i øvrigt ekskluderet fra forsøget hvis de led af andre psykiske sygdomme eller allerede modtog behandling for depression (Wade et al. 2002: 96).

Ved at sammenligne score før og efter forsøget var det muligt at måle effekten af indsatsen. Behandling med escitalopram viste sig gennemsnitligt at reducere scoren på MADRS med 2,7 point mere end placeboindsatsen. Resultatet var statistisk signifikant med en p-værdi på 0,2%. På det dikotomiske resultatmål oplevede 55% af patienterne i behandlingsgruppen en bedring svarende til mindst en halvering af scoren på MADRS, mod 42% i placebogruppen (Wade et al 2002: 99). Dvs. at den absolutte risiko for at opnå bedring med behandlingsindsatsen er 13%. Den relative risiko er derimod 1,31, dvs. behandlingen er 31% mere effektiv end placebo. Vi kan på baggrund af den absolutte risiko beregne NNT, der ligger på 7,7 ($1/0,13=7,69$). For hver 7,7 patient der behandles, oplever 1 patient altså effekt. Risikointervallet er beregnet til 95% CI 1,06-1,62 (Arroll et al 2014: 56). Dvs. at der er 95% sandsynlighed for at den reelle effekt af indsatsen er et sted mellem 6% og 62% bedre end placeboindsatsen.

¹ Montgomery Åsberg Depression Rating Scale angiver et numerisk mål fra 0-60 for depressionens sværhedsgrad, ved at vurdere sværhedsgraden af hver enkelt af 10 symptomer på en skala fra 0-6, se evt. bilag 3. Jeg kommer nærmere ind på brugen af depressionsskalaer i kapitel 7.

Forfatterne konkluderer at kombinationen af en høj effekt og høj tolerabilitet (relativt få bivirkninger) gør escitalopram til et glimrende førstevalg for behandling af depression i den primære sundhedssektor (Wade et al 2002: 101).

3.6 Systematiske forskningsoversigter

De systematiserede forskningsoversigter udgør blomsten af den videnssystematisering Archie Cochrane igangsatte. Forud for produktionen af en systematisk forskningsoversigt går en udvælgelse af hvilke primærstudier der skal medtages. Udvælgelsen af studier der skal indgå i en forskningsoversigt sker på baggrund af en kvalitetsvurdering. Som det også fremgår af evidenshierarkiet, anses de forskellige typer primærstudier imidlertid ikke som lige gode til at vurdere indsatserne. Før udarbejdelsen af et review anlægges man derfor et snit i hierarkiet, der angiver grænsen for hvilke designtyper der kan medtages i forskningsoversigten. I kvalitetsvurderingen identificerer man således først typen af design der er anvendt. Hvis det pågældende design ligger over snittet, kan man derefter gå videre til en kvalitetsvurdering af selve primærstudiet (Hansen & Rieper 2007: 18).

Beslutningen om hvilke typer studier der skal indgå i udarbejdelsen af et bestemt cochrane-review, samt de øvrige inklusionskriterier og kvalitetskrav angives i udarbejdelsen af en protokol, der går forud for reviewet. Protokollen beskriver baggrunden for udarbejdelsen af reviewet, præcis hvilke typer indsatser de inkluderede studier skal beskæftige sig med, hvilke typer af resultatmåling der anvendes, hvordan søgningen skal udføres, hvilken metode syntetiseringen af primærstudierne skal foretages efter, osv. Hensigten er at have så systematisk en proces som muligt i udarbejdelsen af reviewet for at minimere enhver form for bias. Vurderingen af hvorvidt det enkelte studie skal inkluderes samt selve ekstraktionen af data, skal foretages individuelt af mindst to forskere (Higgins & Green 2011: 2.3.4.1). Det væsentligste objekt for cochrane-reviews er således at producere oversigtsartikler der samler den samlede tilgængelige viden om et specifikt forskningsspørgsmål, der lever op til et sæt af smalle kriterier for inklusion og syntetisere denne viden efter en systematik, der minimerer muligheden for at bias kan opstå. Når alle studier der skal inkluderes i reviewet er udvalgt, skal studierne analyseres ved at syntetisere resultaterne i en kvantitativ statistisk metaanalyse (Higgins & Green 2011: 9.1.2).

Jeg vil nu se nærmere på et review der omhandler behandling af depression, for at få en bedre forståelse af, hvad det er for en viden der opnås i Cochrane Collaboration's systematiske reviews.

3.7 Eksempel på en systematiseret forskningsoversigt

Jeg vil illustrere metodologien i evidensproduktionen ved at se på et specifikt review. Jeg har valgt det nyeste tilgængelige Cochrane-review omhandlende behandling af depression, der ikke er bundet til en specifik gruppe (fx depression ifm. fødsel eller demens). Reviewet er publiceret første gang i 2009 og siden opdateret og publiceret i 2014. Reviewet hedder 'Antidepressants versus placebo for depression in primary care' og undersøger to forskellige slags antidepressivas effekt når de anvendes specifikt i den primære sundhedssektor til patienter der har fået stillet diagnosen depression. De to slags antidepressiva hedder 'tricyclic antidepressants' (TCA) og selective serotonin reuptake inhibitors (SSRI). Reviewet inkluderer i øvrigt studiet af Wade, Lemming og Hedegaard som jeg netop har gennemgået.

En række kriterier skulle være opfyldt for at en givent studie kunne inkluderes i reviewet.

1. Studiet skulle være et randomiseret kontrolleret forsøg (RCT) der testede enten TCA eller SSRI mod en placeboindsats hos voksne patienter
2. Patienterne skulle være mellem 18 og 65 år og være rekrutteret fra den primære sundhedssektor.
3. Diagnosen unipolar depression var stillet på baggrund af strukturerede interviews og i overensstemmelse med internationale kriterier såsom ICD eller DSM.
4. Studierne skulle have en varighed af mindst 4 uger.
5. Studiernes primære udkom skulle være reduktionen af depressive symptomer målt vha. anerkendte rating-scales for depression som Hamiltons Depressionsskala eller Montgomery Åsberg depressionsskala.
6. Studierne måtte ikke inkludere patienter med comorbiditet (flere diagnoser).

Efter at kriterierne er fastsat og en omfattende søgning i alle relevante databaser er gennemført, kan kvalitetsvurderingen af studierne begynde. Først læser to forskere, uafhængigt af hinanden, alle de fundne studiers abstracts for at vurdere om de er relevante ift. kriterierne for inklusion. 37 artikler blev fundet relevante efter den første screening. Dernæst vurderes kvaliteten af hvert enkelt studie ved brug af en "Quality Rating Scale". Skalaen sætter blandt andet mål for hvorvidt antallet af forsøgspersoner er passende, om den kilde man har hvervet forsøgspersonerne fra er repræsentativ, om fordelingen har været blændet og om behandlingsindsatsen er tilstrækkelig klart beskrevet (Arroll et al 2014: 5). På baggrund af denne kvalitetsvurdering fandt forskerne i alt 14 studier der

mødte inklusionskriterierne. De 14 studier dækkede en række depressive lidelser, mens enkelte specifikt omhandlede kategorien Major Depression.² 10 studier omhandlede effekten af TCA, 2 studier effekten af SSRI (herunder Wade et al.) og 2 studier omhandlede begge typer (Arroll et al 2014: 7).

Studierne omfatter tilsammen 1364 patienter i indsatsgrupper og 919 patienter i kontrolgrupper. Studierne er næsten alle af kort varighed, dvs. typisk 6-8 uger og de er foretaget i tidsperioden 1971 til 2002. Det nævnes endvidere at de fleste studier var støttet med midler fra medicinalvirksomheder.

Det konkluderes at både TCA og SSRI er statistisk signifikant mere effektivt end placebo. I de dikotomiske resultatmål varierer succeskriteriet mellem studierne fra respons til remission. Det betyder at man på baggrund heraf ikke ved præcis *hvor meget* indsatsen virkede på de patienter der oplevede bedring. Den relative risiko for TCA lå på 1,24 og for SSRI på 1,28. Dvs. at for hver 4 patienter der oplevede effekt af placebobehandlingen var der ca. 5 der oplevede effekt af indsatsen. Dog skal man huske at have risikointervallet med, der ligger på 95% CI 1,11-1,38 og 1,15-1,43 for hhv. SSRI og TCA. Groft sagt er der dermed 95% sandsynlighed for at effekten af indsatserne er mellem 11% og 43% mere effektive end placebo. Forfatterne har ikke angivet procentsatserne for effekt af hhv. indsatser og placebo, men på baggrund af datagrundlaget kan man regne sig frem til at den gennemsnitlige effekt af TCA er 62% mod 49% i placebogrupperne. For SSRI er den gennemsnitlige effekt 58% mod 45% i placebogrupperne³ (Arroll et al. 2014: 34; 56). Den absolutte risiko for de to typer antidepressiva er samlet set 12,7%.⁴ NNT kan da beregnes ved at tage den reciprokke af den absolutte risiko ($1/0,127=7,8$). Det vil sige at behandling med antidepressiva gennemsnitligt medførte en reduktion af depressive symptomer for ca. 1 ud af 8 patienter (Arroll et al 2014).

² Major depression er en amerikansk betegnelse og svarer til de europæiske betegnelser moderat depression og svær depression.

³ SSRI: Indsats: $410/707=0,58$. Kontrol: $253/562=0,45$.

TCA: Indsats: $352/568=0,6197$ Kontrol: $242/490=0,4939$.

⁴ SSRI: Absolut Risiko= $0,58-0,47=0,1299$.

TCA: Absolut Risiko= $0,62-0,49=0,1258$.

Kapitel 4: Analyse af EBM som vidensform

Dette kapitel har til formål at undersøge vidensformen, forudsætningerne og begrænsningerne ved EBM og lede frem mod en undersøgelse af hvordan diagnosen depression lever op til disse forudsætninger.

EBM har betydet at behandlingsmetoder er blevet genstand for øget kontrol, således at postulerede effekter af nye behandlingsmetoder møder langt skrapere krav til dokumentation. Den øgede skepsis hos klinikerne betyder at det er blevet sværere at komme igennem med obskure behandlingsformer og medicinske produkter uden den fornødne empiriske dokumentation og patienter kan derfor være mindre udsatte for skadelige behandlingspraksisser end tidligere. RCT'et er effektivt til at kunne udpege generelle egenskaber ved medikamenter og sige noget om sandsynligheden for forskellige behandlingsudfald, samt sideeffekter og bivirkninger og dermed sortere uvirksom og skadelig medicin fra (Gomory 2013: 21). Behandlinger der bygger på grundige empiriske forsøg er også en sikkerhed for gennemsigtigheden af beslutninger og dermed lægens legitimering af sin praksis ift. juridiske krav, forsikringsvæsen, politikere og samfundet generelt (Maier 2006: 326).

Kritikken af EBM retter sig ikke mod den påstand at den medicinske praksis bør basere sig på empirisk dokumentation. Det er der naturligvis ingen der er uenige i. Kritikken drejer sig derimod om reglerne for hvad denne dokumentation skal bestå i, altså hvad der udgør evidens.

Jeg vil i dette afsnit se på kritikken af EBM. Afsnittet vil være struktureret omkring 3 følgende punkter:

- Kritik af EBMs videnskabsteoretiske grundlag
- Kritik af EBMs anvendelighed i forhold til generalisering af resultaterne
- Kritik af begrænsningerne ved den RCT-genererede viden

Der er tale om tre store og omdiskuterede kritikpunkter af EBM, og jeg behandler dem her hverken i deres fulde bredde eller dybde. Jeg trækker på en række kritikere af EBM, der anskuer metodologien fra både medicinske, videnskabsfilosofiske og socialvidenskabelige perspektiver. Desuden inddrager jeg igen håndbogen *Evidence-Based Medicine. How to Practice and Teach EBM* samt materiale fra The Cochrane Collaboration. Eksemplerne på hhv. RCT og Cochrane review tjener til at illustrere pointerne.

4.1 Kritik af EBM's interne validitet

Udgangspunktet for at privilegere RCT har været under kritik fra et videnskabsfilosofisk perspektiv, da det hævdes at RCT-genereret viden ikke reelt har den sikkerhed, eller validitet, som tilhængere tilskriver den. Der skelnes mellem metodens interne og eksterne validitet. Den interne validitet omhandler forsøgets evne til at kortlægge den faktiske effekt af indsatsen hos forsøgspopulationen. Den eksterne validitet omhandler forsøgets evne til at kortlægge hvad den faktiske effekt af indsatsen ville være for den generelle patientpopulation, som man ønsker at repræsentere i forsøget. I dette afsnit vil jeg først se på den videnskabsteoretiske kritik af RCT-forsøgs interne validitet. Denne kritik tager udgangspunkt i den centrale komponent der adskiller RCT fra trinnet lavere i evidenshierarkiet, nemlig randomisering.

Effekten af randomisering

Randomisering er en central komponent i EBM, der skal yde sikkerhed for at kontrol- og indsatsgruppe er ens i forhold til alle faktorer der kan være af betydning for den målte effekt, og dermed agere garant for at den observerede effekt alene skyldes indsatsen. Andre former for studiedesign, der er placeret lavere i evidenshierarkiet, benytter andre metoder til at få adgang til et sammenligningsgrundlag og dermed foretage kontrollerede forsøg. Fx kan man anvende naturligt forekommende grupper eller man kan matche deltagerne i forsøget således at grupperne designes så ens som muligt på baggrund af de faktorer man mener kan have indflydelse på resultatet (Rieper & Hansen 2007: 24). Disse forsøg minder i høj grad om RCT i udformningen, og benyttes typisk hvor randomisering ikke er mulig, eksempelvis pga. etiske hensyn. Naturlige grupper kan også betyde at forsøgsdeltagerne ikke behøver hverves til forsøget som i et RCT. Med randomisering mener man dog at opnå en højere sikkerhed, da den tilfældige udvælgelse balancerer fordelingen af både faktorer man kender og faktorer man ikke havde forudset kunne have en betydning (Straus 2005: 118). Som videnskabsfilosofferne Nancy Cartwright og John Worrall gør opmærksom på, er randomisering dog ikke en garanti for at kontrol- og indsatsgruppe er ens på alle variable af relevans for effekten, i den specifikke fordeling der foretages (Worrall 2010: 359; Cartwright 2007: 19). Dette anerkendes også i 'Evidence-based medicine. How to practice and teach EBM', hvor det anføres at på trods af at fordelingen ved randomisering er den balancerede fordeling, kan ubalanceret alligevel opstå ved den tilfældige fordeling (Straus 2005: 120). Derfor tjekker man typisk for 'baseline imbalances' efter fordelingen af de to grupper er foretaget, dvs. at det kontrolleres hvorvidt de to grupper rent faktisk er ens på de parametre man ved, eller har formodning om kan

have kausal betydning for effekten. Hvis der findes sådanne statistisk signifikante 'baseline imbalances' kan man gentage randomiseringen og håbe på et bedre resultat (Worrall 2010: 359). I det beskrevne RCT er der således også foretages et tjek for baseline imbalances, her kontrollerer man for alder, køn, vægt og BMI, som antages at kunne have betydning for effekten (Wade 2002: 98). I og med at der på trods af randomisering kan være tale om forskelle indenfor faktorer man kender kausaleffekten af, kan der imidlertid også være forskelle indenfor faktorer man ikke har forudset kausalvirkningen af, eller som ikke er mulige at opgøre kvantitativt. Disse forskelle er det vel at mærke ikke muligt at tjekke for. Ifølge Worrall er randomisering genstand for en misforstået opfattelse af sandsynligheder. Sandsynligheden for at *gennemsnittet* af n antal tilfældige fordelinger nærmer sig en balanceret fordeling, stiger i takt med antallet n. Dvs. at hvis man kørte forsøget gentagende gange med samme gruppe patienter og randomiserede dem på ny for hver gang, ville der være statistisk belæg for at konkludere at gennemsnittet for effekten i indsatsgrupperne kontrollerede for ubalancer i fordelingen. Det er imidlertid ikke sådan randomisering i RCT foregår, og det ville heller ikke kunne lade sig gøre, da man naturligvis ikke kan behandle de samme patienter igen og igen. (Worrall 2010: 360).

Eksempel: Effekten af tilbagevirkende bøn

Usikkerheden ved randomisering er illustreret i et stort dobbeltblændet RCT ang. blodbaneinfektion af den israelske professor i medicin Leonard Leibovici. Forsøget baserede sig på de patienter der i perioden 1990-96 havde været indlagt for blodbaneinfektion på et universitetshospital, hvilket beløb sig til 3393 patienter. Forsøget fordelte efterfølgende patienterne ved randomisering, trak lod mellem forsøgs- og indsatsgruppe og testede for 'baseline imbalances', uden at finde nogen signifikante forskelle i blandt andet køn, alder, blodværdier og infektionskilde. I 2001 gav man listen over personer i indsatsgruppen til en person der bad en kort bøn på vegne af den samlede gruppe. Forsøget viste at denne såkaldte "remote, retroactive intercessory prayer" nedsatte indlæggelsestiden på hospitalet, samt varigheden af feber signifikant. Der fandtes ligeledes en forskel til fordel for indsatsgruppen mht. dødelighed, dog var denne ikke signifikant. Det konkluderedes derfor, fuldt ud i tråd med reglerne for EBM at denne procedure burde overvejes som del af behandlingspraksis: "*Remote intercessory prayer said for a group of patients is associated with a shorter hospital stay and shorter duration of fever in patients with a bloodstream infection, even when the intervention is performed 4–10 years after the infection*" (Leibovici 2001: 1451).

Leibovicis forsøg viser at der må have været mindst én, men sandsynligvis flere, gennemsnitlige forskelle på de to grupper i kausalfaktorer der gav sig til udslag i signifikante forskelle i sygdomsforløb. Disse forskelle kan fx have angået patienternes helbred før sygdomsforløbet, om sygdommen blev konstateret hurtigt eller langsomt, personalets erfaring eller patientens egen håndtering af situationen. Den statistiske forskel mellem de to grupper er opstået tilfældigt, på trods af randomisering og på trods af forsøgets størrelse. Randomisering var ikke en garanti for en ligelig fordeling på tværs af alle relevante faktorer, og den har derfor ikke nødvendigvis en højere sikkerhed end andre metoder til etablering af kontrol og indsatsgruppe. Endvidere bliver det klart at det ikke er hensigtsmæssigt at basere sig alene på denne type studier som eneste vidensgrundlag, da det ikke er muligt at opdage denne form for usikkerheder ved RCT'et. Cartwright mener således at den bedste mulighed for at kunne opnå en højere sikkerhed for den interne validitet er at lade eksperter på det specifikke forskningsområde der undersøges, følge forsøgets proces, frem for at det udføres udelukkende af forskere der er specialister i metodologien. Derved øges muligheden for at kunne spotte relevante forskelle på indsats- og kontrolgruppe (Cartwright 2007: 19). Samtidig åbner man processen op for at der kan hentes flere indsigter i hvordan, hvornår og hvorfor behandlingen har effekt eller ikke har effekt, hvilket jeg kommer nærmere ind på i et følgende afsnit. Ved at kombinere metodologien med ekspertviden baserer man sig ikke udelukkende på en metodisk mekanisme for at undgå fejl, men også på den erfaringsmæssige, teoretiske og empiriske viden man har om det specifikke felt. EBM lader dog ikke megen plads til ekspertvurderingers rolle i forskningssammenhæng, der er placeret nederst i evidenshierarkiet.

Kritik af EBM som empiristisk

Leibovicis forsøg gør også opmærksom på at gennemførelsen af et RCT altid forudsætter en hypotese, valget af forskningsspørgsmål, der afspejler en forventning til hvilke slags indsatser der virker, hvordan de skal testes, hvilke informationer der er relevante (fx hvilke baseline imbalances der testes for) og hvad der på den anden side udgør 'støj' (Williams 2010: 108). Når man i den medicinske forskning ikke er villig til at tage indsatsens kliniske effekt seriøst, skyldes det at man netop bringer den generelle teoretiske ramme og forståelse man har om kausaleffekter i klinisk praksis ind i vurderingen af resultatet. Den rent empiriske observation kan altså ikke stå alene, i vurderingen af studiets resultater. Man er ikke tilbøjelig til at have tiltro til resultatet eftersom det ikke er bakket op teoretisk og ikke er biologisk plausibelt (Sehon & Stanley 2003: 7). Filosof Scott Sehon og patolog Donald Stanley mener derfor at den teoretiske baggrundsviden opnået gennem

grundforskning, bruges som kompas i tolkningen af kliniske forsøg. ”*Statistical information from an RCT is virtually uninterpretable and meaningless if stripped away from the backdrop of our basic understanding of physiology and biochemistry.*” (Sehon & Stanley 2003: 7). Man må altså gøre sig klart at også EBM metodologien er afhængig af teoretisk viden.

Flere kritikere mener at EBM gør sig skyldig i empirisme, og da metodologien indebærer en antagelse om at videnskabelige observationer kan foretages objektivt og uafhængigt af forskerens teoretiske forståelsesramme og bias. I det givne eksempel, bliver de baseline-imbbalances man vælger at tjekke for, til en teoretisk antagelse, og i øvrigt en udpræget bio-medicinsk af slagsen, da kun køn, alder og vægt forventes at kunne give anledning til ubalancer i sammenligningsgrundlaget.

4.2 Generaliserbarheden af RCT

Det andet kritikpunkt handler om anvendeligheden af RCT-baseret evidens til brug i den kliniske praksis. Her kommer vi ind på en kritik af metodens eksterne validitet.

RCT og behandlingseffekter

Det konkluderes i det gennemgåede RCT at den testede indsats er et glimrende førstevalg til behandling af depression i den primære sundhedssektor. Også i Cochrane reviewet konkluderes det at både SSRI og TCA er effektive for behandling af depression i den primære sundhedssektor. I såvel det enkelte RCT som i reviewet viser resultaterne at lægemidlerne er 24-30% bedre end en placeboindsats. Konklusionen har sin berettigelse i det forhold at det er vist at der ikke er tale om et tilfældigt sammenfald. Der er under 5% sandsynlighed for at den observerede forskel i antal patienter i bedring skyldes en tilfældighed eller andre faktorer end de testede indsatser, og der er 95% sandsynlighed for at denne bedre effekt ved TCA og SSRI er et sted mellem 11% og 43% relativt bedre end placeboindsatsen.

I RCT'et af Wade et al. oplevede 55% af deltagerne i indsatsgruppen effekt. Men 42% kunne lige så godt have modtaget en placeboindsats, og nogle af dem kunne formentlig have oplevet samme bedring uden en egentlig behandlingsindsats, da de 42% naturligvis også dækker over personer hvor sygdommen var gået over af sig selv (spontan remission). Man udskriver ikke normalt placeboindsatser i den kliniske praksis, og diskussionen af placebo og dets effekter er omfangsrig og involverer også etiske aspekter. Ikke desto mindre peger resultatet på at 42% af patienterne i forsøgets indsatsgruppe overmedicineres, og måske ville have gavn af enten ingen behandling eller

så mildt et præparat som muligt, især hvis det også betød lavere risiko for bivirkninger og afhængighed. For 45% kunne der ikke påvises en nævneværdig positiv effekt af behandlingen.

Den gruppe for hvem indsatsen udgjorde en passende behandlingsform er altså 13%. Dette resultat er konsistent med metaanalysen. Ifølge lederen af Nordisk Cochrane Center, dr.med. Peter Gøtzsche er lignende resultater opnået i en amerikansk metaanalyse af US Food and Drug Administration, der involverede 100.000 forsøgsparticipanter, og hvor det konkluderedes at 10% flere patienter oplevede effekt af antidepressiva end placebo (Gøtzsche 2014: 9). Det forholdsvist lave resultat der ses i de gennemgåede eksempler er altså ikke en undtagelse. Problemet er naturligvis at man ikke på forhånd ved, hvilke 13% behandlingen var effektiv for.

Evidensmetodologien kritiseres derfor for gennem hierarkiet at privilegere forskningsmetoder der fokuserer på den generelle statistiske viden om behandlingers effekt. Et RCT producerer gruppe-aggregerede gennemsnit som mål for behandlingsmetodens effekt. Men hvordan denne viden skal anvendes til at informere behandlingsvalget i den enkelte patients tilfælde er ikke klart (Gomory 2013: 20). Især i forhold til RCT og metaanalyser der påviser effekter af den størrelsesorden vi har set i de to eksempler, bliver spørgsmålet, hvilke beslutninger klinikerne skal drage på baggrund af studierne. At hver 8. patient diagnosticeret med en depressiv lidelse har gavn af et præparat, er ikke nogen god indikator for hvorvidt den enkelte patient i lægens praksis vil have gavn af det.

Fra RCT til patient – EBM's retningslinjer

I *Evidence-Based Medicine. How to Practice and Teach EBM*, anføres det på første side at EBM indebærer at den bedst tilgængelige evidens integreres i den kliniske praksis sammen med lægens vurdering af den individuelle patient, og dennes personlige værdier (Straus et al. 2005: 1). Straus et al. beskriver appliceringsprocessen i 4 trin. Første trin går ud på at lægen skal bruge sin kliniske ekspertise til at vurdere hvorvidt patienten er så anderledes fra forsøgspopulationen at studiets resultater ikke kan anvendes (2005: 132). Her anbefales det at vurdere kvalitative forskelle på patient og forsøgspopulation, navnlig sociodemografiske forhold og patobiologi. Der er imidlertid få og sjældne tilfælde hvor sådanne forskelle gør sig gældende. Som mulige sociodemografiske og patobiologiske forskelle angives comorbiditet, forskelle i farmakogenetik og manglende respons i immunsystemet. Relevante forskelle beløber sig altså til biologiske komplicerende faktorer der indvirker på det kausalforhold der er gjort rede for i forsøget. Mht. andre kvalitativt forskellige subgrupper, anføres det at tilsvarende forskelle i behandlingseffekt er ”ekstremt sjældne”, og det

anbefales derfor at man er meget forsigtig med at konkludere noget på baggrund af kvalitative forskelle blandt patientsubgrupper og accepterer RCT'ets generelle resultat som udgangspunkt for vurderingen af behandlingens anvendelse i det specifikke tilfælde, med mindre man kan svare bekræftende på alle fem følgende spørgsmål:

Does it really make biological and clinical sense? 2. Is the qualitative difference both clinically (beneficial for some but useful or harmful for others) and statistically significant? 3. Was it hypothesized before the study began (rather than the product of dredging the data)? 4. Was it one of just a few subgroup analyses carried out in the study? 5. Has the result been confirmed in other independent studies? (Straus et al 2005: 133).

Det er altså i høj grad den teoretiske og praktiske forforståelse, samt den allerede etablerede (og dermed hypotetiserede og empirisk testede) viden der skal bringes i spil i vurderingen af resultatet når det kommer til påvisningen af kvalitativt forskellige patientsubkategorier. Ifølge Horwitz et al. skyldes modstanden mod subgruppeanalyser til at identificere forskelle i behandlingseffekt, risikoen for at disse forskelle kun er et udslag af tilfældighed, da en subgruppe i et RCT kan bestå af relativt lille antal patienter (Horwitz et al. 1996: 395). Det er således mere anbefalelsesværdigt at kigge på subgruppeanalyser i reviews, hvor det således kan vise sig om samme resultat er opnået på tværs af de forskellige studier. Dog skal 5 ovenstående kriterier stadig mødes (Straus et al 2005: 153). I The Cochrane Handbook til udarbejdelse af reviews, anføres det imidlertid at det sjældent er muligt at foretage analyser af patientsubgrupper, da de enkelte RCT som oftest ikke vil indeholde nok og ensartet information (Higgins & Green 2011: 9.6.2). Det vil derfor sjældent være muligt at drage nogen erfaringer ang. kvalitative forskelle mellem patientsubgrupper fra RCT og RCT-baserede reviews.

Et andet trin i appliceringen af forsøgsresultatet på patienten, er en vurdering af risikoen for de potentielle fordele og ulemper patienten kunne opnå som følge af behandlingen. Lægen skal vurdere om sandsynligheden for at behandlingen vil virke, er stor nok i forhold til sandsynligheden for at den ikke vil virke, eller at det vil skade patienten. Til det formål anvendes 'Numbers Needed to Treat' (og 'Numbers Needed to Harm' for bivirkninger). Den relative risiko for både positiv og negativ effekt antages at være konstant, men lægen kan foretage en individuel vurdering, ved at vurdere patientens 'baseline risk' i forhold til den gennemsnitlige forsøgsdeltager i kontrolgruppens risici (Straus et al. 2005: 137). Det nævnes i Straus et al. at forskellig 'baseline risk' fx kan skyldes comorbiditet (Ibid.: 129), men ikke hvad det ellers kan skyldes, eller hvad det præcis betyder for patientens risici for positiv og negativ effekt, at vedkommende også har en anden diagnose.

Patienter med comorbiditet ekskluderes som oftest fra forsøgene, ligesom det også er tilfældet i de to eksempler. Grundlaget for lægens vurdering af hvorvidt patienten har højere eller lavere risiko for det ene eller det andet udfald, må derfor hvile på hans ekspertise og erfaring. I det tilfælde at patienten ikke lider af andre sygdomme, og i øvrigt ligner den (ikke særligt specifikke) forsøgspopulation, foreskriver håndbogen ingen værktøjer til at fastslå hvorvidt patienten er den ene ud af otte der vil have fordel af behandlingen.⁵

I langt de fleste tilfælde bør det derfor accepteres at et resultat som 'Escitalopram er virkningsfuldt for hver 8. Patient med depression', ikke kan specificeres nærmere, og at kvalitativt distinkte patientsubgrupper altså ikke har nogen betydning for behandlingseffekten. I RCT'et af Wade et al. og reviewet af Arroll et al. er der ikke foretaget nogen subgruppeanalyser af patientgrupper eller eventuelle forskelle i den måde forsøgene udførtes i de enkelte praksisser. Vi er altså på denne baggrund ikke blevet klogere på hvordan den enkelte læge undgår at medicinere 7/8 af de deprimerede patienter med uvirksom, skadelig eller unødvendig medicin.

Ekstern validitet

En forudsætning for generaliserbarheden af resultaterne opnået i RCT eller Cochrane review, er at de patienter der skal modtage behandlingsindsatsen er sammenlignelige med patienterne der indgik i forsøgene. Forsøgspopulationen skal altså have været repræsentativ for patientpopulationen. Dette forhold kaldes ekstern validitet.

Flere kritikere har anfægtet den eksterne validitet af RCT. En årsag er, at det kan vise sig svært at rekruttere patienter til RCT. For patienten der plages af sygdom er udsigten til enten at modtage en placebo-behandling eller, i bedste fald, en behandling hvor behandlingseffekten er ukendt, ikke specielt tiltalende. Når patienter så alligevel deltager i forsøgene, kan der nemt opstå en overrepræsentation af patienter der har tidligere fejlslagne behandlingsforsøg bag sig, og som derfor er mere desperate. Ifølge Williams er deltagerne i et RCT ofte netop ikke et randomiseret udvalgt udsnit af den population man ønsker at generalisere over, og de kan derfor ikke forventes at være

⁵ Endnu et trin hos Straus et al. er desuden at patientens værdier og præferencer tages til overvejelse. Dette punkt tildeles en vis vægt eftersom manglende samarbejdsvillighed kan have negative konsekvenser for dens effekt. Til vurderingen findes der en udregningsmetode, der på baggrund af størrelsen af patientens frygt for et negativt behandlingsresultat, vægten af præferencerne, samt risiciene for forskellige udfald og bivirkninger udregner patientens individuelle værditillæggelse af indsatsen (Straus 2005: 140).

repræsentative, tværtimod kan særlige karakteristika være gældende for forsøgsdeltagerne (Williams 2010: 109).

Lignende observationer går igen hos Bluhm og Borgerson, der anfører at RCT forsøg ofte er karakteriseret ved, i kraft af in- og eksklusionskriterier, at have en sundere og yngre befolkning end den befolkning behandlingerne skal anvendes på. Det kan fx opnås ved en aldersrestriktion og eksklusion af patienter med andre sygdomme, som det ses i eksemplerne, eller ved at ekskludere patienter der også tager anden medicin. For en sundere og yngre forsøgspopulation er det mindre sandsynligt at der forekommer bivirkninger, og der er højere sandsynlighed for et generelt positivt resultat (Bluhm & Borgerson 2007: 218). Ifølge Rieper og Hansen har det således vist sig at forsøgspopulationen sjældent er repræsentativ for den samlede patientpopulation, og at in- og eksklusionskriterier medfører at kun 10-20% af patienterne i den normale kliniske praksis er egnede til at deltage i forsøget (Rieper & Hansen 2007: 74). Det at RCT altså er *opstillet* som forsøg, med krav om særligt kontrollerede rammer og patienter der skal acceptere at indgå i en behandling der er anderledes end den behandling de ellers havde fået tilbudt, med mulighed for placebo, udgør endnu en udfordring for den eksterne validitet af RCT. Ved ikke at stille krav om randomisering, kan man foretage forsøg der sammenligner forskellige behandlingsindsatser, ved at tage udgangspunkt i naturligt forekommende grupper. Eller man kan foretage kohortestudier, der undersøger store patientgrupper der deler et fælles karakteristika, fx medicinering med antidepressiva. Der udføres således ikke interventioner som del af undersøgelsen, men man baserer sig i stedet på data fra de patienter der i forvejen modtager indsatsen (Rieper & Hansen 2007: 27). Kohortestudier kan løbe over meget længere perioder end RCT der som oftest er udført over en kort periode på 4-8 uger. I virkeligheden er patienter ofte i behandling med antidepressiva i meget længere tid, og eventuelle langtidseffekter af behandlingerne kan derfor ikke undersøges i RCT.

Den eksterne validitet af RCT bemærkes også i Sundhedsstyrelsens anbefalinger der blandt andet anfører om den eksterne validitet, at *”Det er velkendt, at randomiserede undersøgelser under velkontrollerede forhold af en udvalgt gruppe af patienter typisk viser en større behandlingseffekt end under naturalistiske forhold.”* (Sundhedsstyrelsen 2007: 41).

Antagelsen om resultaters universalitet

Håndbogens påstand om at kvalitativt distinkte subgrupper der reagerer forskelligt på behandlingsindsatser, er ”ekstremt sjældne”, indebærer en antagelse om at RCT’ets konklusioner

må antages at være tæt på universelle. Hvis sygdomstilfældet kan defineres som værende af en bestemt type, vil det (medmindre der findes andre biologiske anomalier der kan indvirke på kausalforholdet som ved comorbiditet) reagere på behandling på samme måde som andre tilfælde af samme slags reagerer på samme behandling. Denne antagelse afspejler altså et overvejende biomedicinsk sygdomsbegreb, ideen om sygdomme som enheder med naturlige forløb, der for så vidt er uafhængige af patienten som individ.

Netop denne påstand om resultaternes universalitet, påpeger sociolog Hanne Foss Hansen, er afhængig af forskningsfeltet. Den formodede universalitet kan blive en udfordring indenfor indsatser hvor konteksten for indsatsen kan have indflydelse på effekten (Hansen 2014: 15).

Antagelsen om resultatets universalitet er afhængig af at der er ekstern validitet, og at forsøgspopulationen altså er repræsentativ for patientpopulationen. Men antagelsen om universalitet er også afhængig af at den patientpopulation man undersøger udgør en homogen gruppe angående faktorer der er centrale for sygdommens karakter og forløb, og følgelig må forventes at kunne behandles med samme indsats. Forsøgets inklusionskriterier må afgrænse sin forsøgspopulation til en gruppe patienter som man forventer i det store hele reagerer ens. Dette synes at være en selvfølgelighed. Fx ville det ikke give mening at teste en behandlingsindsats for depression blandt alle patienterne tilknyttet et psykiatrisk afsnit. Nogle af disse patienter ville lide af depression, men andre af skizofreni, og igen andre af tvangstanker, og forsøget ville derfor ikke vise den reelle effekt af behandlingsindsatsen på depression. Et sådant forsøg ville antage at alle patienter tilknyttet psykiatrisk afsnit er sammenlignelige angående hvilken behandlingsindsats der virker. På samme måde antager RCT og Cochrane review at den patientgruppe der lever op til inklusionskriterierne er sammenlignelige angående hvilken behandlingsindsats der virker.

Ifølge Bluhm kræver behandlingsvalg derfor en fundamentalt anden form for viden end den RCT-forsøg kan generere. RCT-forsøg giver kendskab til de fundamentale processer omkring behandlingers effekt, dvs. behandlingsindsatsers kapaciteter, eller hvor ofte de virker. Valg af behandling for den enkelte patient bør dog tage udgangspunkt i en anden viden, nemlig viden om hvordan behandlingseffekten er forskellig under forskellige omstændigheder (Bluhm 2007: 149). Viden af denne karakter kræver en langt mere forståelsesorienteret tilgang til behandlingseffekter. Man går dermed potentielt glip af vigtige indsigter i hvordan behandlingsindsatser bedre kan målrettes ved at have en generel prioritering af RCT. Især hvis denne prioritering kommer til udtryk i fordelingen af forskningsmidler og i publiceringen af artikler. Healy peger eksempelvis på at flere

af de største medicinske tidsskrifter slet ikke publicerer kvalitative casestudieartikler, selv hvis de er udført af ledende forskere, udsat for usikkerhedstest og viser resultater der er i overensstemmelse med kendte patologiske mekanismer (Healy 2007: 134).

4.3 Begrænsninger ved den RCT-genererede viden

I dette afsnit fortsætter jeg med at se på den kritik der omhandler den viden som EBM ikke bibringer, og som man dermed potentielt udelukker fra vidensgrundlaget for behandlingspraksis.

Placebo og non-specifikke behandlingselementer

42% af deltagerne i RCT-eksemplet responderede på placeboindsatsen. Som psykiater David Healy gør opmærksom på giver det placeboindsatser et langt lavere NNT, i dette tilfælde på 2,38.

Benyttede man en placeboindsats som førstevalg i behandlingen af depression ville man ifølge RCT'ets resultater kun opnå 76,4% af den bedring der kunne være opnået ved at anvende SSRI. Til gengæld har placeboindsatser langt færre negative effekter, og er derfor et sikrere behandlingsvalg (Healy 2007: 132). Thomas et al. har foretaget en analyse af metaanalyser af RCT-studier omhandlende hhv. antidepressiva og kognitiv adfærdsterapi for depression. Her peger de på at de non-specifikke elementer af behandlingsindsatserne, var langt vigtigere end de specifikke. Non-specifikke elementer henviser til elementer af behandlingsforløb der ikke menes at have indflydelse på underliggende årsager til sygdommen, som fx forholdet mellem læge og patient, men som altså alligevel viste sig mere effektive end de specifikke elementer, nemlig de aktive stoffer i antidepressiva og de særskilte terapeutiske elementer ved kognitiv adfærdsterapi, der retter sig specifikt mod behandling af depression (Thomas et al. 2012: 152). I studier der anvender såkaldte aktive placeboindsatser, der simulerer nogle af sideeffekterne ved antidepressiva (fx tørhed i munden) er placeboeffekten endnu højere, og den relative effekt af antidepressiva endnu lavere. Studier med aktiv placebo regnes for at være en mere fair test eftersom patienterne blændes mere effektivt (Gøtzsche 2014: 9). Man har dog stadig til gode at undersøge de præcise effekter af effektiv blænding (Moncrieff et al. 2012: 8). Uanset den præcise effekt af antidepressiva er det tydeligt at behandling af depression ikke for alle og ikke udelukkende drejer sig om indholdet af den pille der udleveres. Alligevel har der ifølge Cohen og Jacobs ikke været udvist megen interesse i at afdække årsagerne og virkningerne af de store placeboeffekter (2007: 47). En af årsagerne er måske, at det er meget vanskeligt kvantitativt at opgøre effekten af de enkelte faktorer der indgår under samlebetegnelsen placebo. Hvordan måler man fx effekten af at blive set, hørt og taget

alvorligt af en læge? Beslutningen om at opsøge behandling kan i sig selv have en effekt (Healy 2007: 131). Lægens råd om livsstilsfaktorer og spørgsmål til arbejds- og familieliv kan have betydning for patientens refleksion over og bearbejdning af sygdommen. Disse faktorer er svære at kvantificere og undersøge i et RCT, da de først om fremmest drejer sig om patientens oplevelse og håndtering af sygdomsforløbet.

Konteksten for behandling

Sociolog Dorte Effersøe Gannik mener i forlængelse heraf at der mangler en erkendelse af, at RCT-studier kun belyser bestemte faktorer ved behandlingspraksis, og at andre metoder har en lige så vigtig rolle at spille i vidensgrundlaget (Gannik 2008: 49). Nærmere bestemt indebærer EBM de facto en biologisk reduktionisme i dens fokusering på specifikke behandlingsindsatser. Gannik peger på at der findes et misforhold mellem, på den ene side en tiltagende kompleks sygdomsvirkelighed, hvor det i de kliniske beslutninger om behandlingsvalg er nødvendigt at medtænke konteksten for indsatsen, og på den anden side de forskningsmetoder der anses for legitime indenfor EBM, altså navnlig RCT hvor man så vidt muligt forsøger at udelukke kontekstuelle faktorer i forskningen (Gannik 2008: 33).

EBM's metodologiske tilgang, mener Gannik, er konsistent med en biomedicinsk opfattelse af sygdomme som entiteter, da det antages at ethvert sygdomsbillede svarer til en specifik diagnose med dertilhørende sygdomsmekanisme og prognose (Gannik 2008: 36). I praksis udmønter det sig i en ide om standardisering og objektivisering. Med standardisering menes at sygdomme kan beskrives, defineres og behandles på samme måde, uafhængigt af personen der er syg. Dvs. at der findes én viden om sygdommen og dens behandling. Med objektivisering menes at mennesket kan betragtes som genstand for sygdommen, og at indgreb bedst foretages ad teknisk, mekanisk eller kemisk vej, uden inddragelse af den syge som subjekt med vilje, følelser, hensigter, intentionalitet og handlinger (Gannik 2008: 36).

Ved kun at interessere sig for indsatsens isolerede virkning mister man indsigt i den interaktion der foregår mellem den specifikke behandling og de omstændigheder den udføres under, som har betydning for behandlingens effekt (Gannik 2008: 43). Konteksten søges neutraliseret og afskrevet som placebo, og RCT bliver dermed en forhindring i at opnå kontekstuel eller situationel viden (Launsø & Gannik 2000: 246f). Hanne Foss Hansen er inde på samme problematik. Hun anfører at evidensmetodologien kan betragtes som en 'black box', hvor input og outcome registreres, men

hvor der ikke opnås nogen indsigt i hvordan, hvorfor og under hvilke omstændigheder indsatsen virker (Hansen 2014: 15). Dermed går man glip af viden om hvordan behandlingsindsatser kan skræddersys bedre den enkelte, ligesom man også går glip af viden om eventuelle forudsætninger for at opnå en større effekt, i kraft af behandlingselementer der med fordel kan kombineres. Da RCT er designet til at undersøge én faktor, er komplekse behandlingsindsatser virkning desuden svære at teste, da ideen her er at de forskellige elementer støtter op om og forstærker effekten af hinanden.

Evidens med smallere sigte

Som jeg har redegjort for i det foregående afsnit anses subgruppeanalyse for metodisk problematisk i RCT og i praksis umuligt i Cochrane-reviews. En anden mulighed for at teste mere specifikke patientpopulationer indenfor EBM er imidlertid at anlægge et smallere fokus i RCT såvel som review som hele, i undersøgelsen af indsatsers effekt. Som det ses af nedenstående tabel, er der dog forholdsvis få Cochrane-reviews der undersøger effekterne af behandlingsindsatser på mere specifikke patientpopulationer der lider af depression, hvilket muligvis kan skyldes mangel på RCT-studier der undersøger præcis samme smalle patientgruppe. Denne betragtning findes også hos Wulff og Gøtzsche der anfører at ideelt set ville man producere RCT-baseret evidens for patientsubgrupper med forskellige symptomer. De skriver dog også at dette ”*naturligvis er urealistisk*” (1997: 187). Det er urealistisk at skulle udføre randomiserede kontrollerede forsøg for hver subgruppe af symptomsammensætninger, fordi disse forsøg er omkostningstunge og kræver meget store antal deltagere for at kunne blive tilstrækkeligt sikre. Hvis deltagerantallet er for lille, vil usikkerheden blive for stor, dvs. man skal kunne påvise meget store effekter hvis man skal usandsynliggøre at en given effekt ikke kan skyldes et tilfælde. Hvis nul-hypotesen kan afvises og effekten tilskrives indsatsen vil et meget bredt usikkerhedsinterval gøre det svært at sige noget om effektens størrelse. Som det ses i review-eksemplet, fandt man at kun 14 RCT produceret tilbage til 1971 levede op til kvalitetskriterierne i en undersøgelse af både SSRI og TCA, de to største typer antidepressiva, anvendt i den primære sundhedssektor, der tager sig af langt størstedelen af behandlingstilfældene indenfor depression. Det er derfor ikke mærkeligt at det kan vise sig umuligt at producere metaanalyser over langt mindre subgrupper, baseret på fx sociodemografiske data. Der findes i The Cochrane Library 70 reviews der omhandler depression. Ud af dem omhandler 56 reviews behandling af depression. Som det ses af nedenstående skema omhandler enkelte specifikke subgrupper af patienter med depression, men de fleste undersøger en eller flere

behandlingsindsatsers effekt på depression generelt. Behandlingerne testes enten mod en placeboindsats eller i sammenligning med andre behandlingsformer.

Patientpopulation	Antal reviews
Ældre med depression	4
Patienter med fødselsdepression	2
Børn og Unge med depression	3
Patienter med depression og anden comorbiditet (fysisk + psykisk)	7
Patienter med psykotisk depression	1
Patienter med sæsonbetinget depression	1
Patienter med depression generelt	38

Kilde: The Cochrane Library

Nogle faktorer undersøges altså i særskilte RCT og metaanalyser. Som det ses af tabellen er aldersgrupper en variabel der laves særskilte undersøgelser indenfor. Depression der optræder i forbindelse med særlige omstændigheder undersøges også særskilt som det ses med comorbiditet, fødselsdepression og sæsonbetinget depression. Endelig udgør psykotisk depression en kvalitativ særskilt subkategori indenfor kategorien depression, der altså (i et enkelt review) betragtes alene, hvilket giver mulighed for at opdage forskelle i behandlingseffekter mellem denne subtype og depression generelt. Det hypotetiseres således at disse forskellige patientgrupper kan være forskellige fra patienter med depression i almindelighed, og at det kan have betydning for effekten af behandlingsindsatsers effekt og hensigtsmæssighed. De fire første af patientgrupperne er således også karakteriseret ved at være dem, der i eksemplet er udelukket fra de mere generelt orienterede forsøg og reviews i kraft af inklusions- og eksklusionskriterier.

Ikke-biologiske kausalfaktorer

De faktorer man undersøger er primært af biologisk karakter. Interessen for forskellige aldersgrupper afspejler at depressionen kan være afhængig af hjernens udvikling, ligesom kroppen er udsat for massive hormonforandringer i forbindelse med fødsel og graviditet. Dermed ikke sagt at psykologiske og sociale faktorer ikke kan udgøre baggrund for valg af subgruppe, men subgrupperne er netop *også* biologisk forankret. Når Gannik og Foss Hansen imidlertid taler om patientens subjektivitet og konteksten som afgørende for behandlingseffekter, begrænser det sig dog ikke til biologiske omstændigheder.

Hvis man forsøger at tænke ud over de primært biologisk bestemte subgrupper og tager det forhold i betragtning at vi her beskæftiger os med en psykisk sygdom, der også er bestemt af sociale faktorer, kunne sociodemografiske variable af betydning for behandlingseffekten fx være etnicitet, nationalitet, socialklasse, beskæftigelse, civilstand og eventuelle tidligere depressioner. På baggrund af disse variable kunne der måske findes forskelle i patienternes værditilskrivelse, tro på - og frygt for - forskellige behandlingsformer, motivation, bivirkningernes vægt og betydning, behandlingens 'fit' ind i patientens dagligdag osv. Der er tale om kvalitativt forskellige, men let kategoriserbare variable, som kunne have betydning statistisk set for behandlingseffekt. Det ville eksempelvis ikke principielt være umuligt at foretage et RCT der undersøgte hvorvidt enlige eller fraskilte reagerede bedst på antidepressiva eller adfærdsterapi. Ved at basere sig på smallere patientpopulationer, ville man muligvis kunne opnå mere specifikke behandlingsanbefalinger baseret på et lavere NNT, således at færre patienter behandles unødvendigt eller uvirksomt. Samtidig er det at producere RCT-baseret evidens for alle patientsubgrupper en uendelig proces, da grupperne altid vil kunne defineres finere og i uendelige kombinationer med andre variable. RCT er omkostningstunge og kræver mange patienter, der alle skal indvilge i at deltage i forsøget. Selv om der principielt er mulighed for at man gennem RCT-forsøg kan opnå subgrupper der er mere prognostisk homogene, er det altså svært at efterkomme praktisk, og altså af Gøtzsche og Wulff dømt urealistisk.

Et andet problem er, at selvom det muligvis ved inddragelse af mere snævert definerede grupper, ville være muligt at indkredse relevante patientpopulationer og dermed i højere grad målrette behandlingsindsatsen, giver EBM ikke mulighed for at opnå viden om hvorfor behandlingen virker eller ikke virker. Derfor vil der fortsat være et problem omkring generalisering til den enkelte patient. Lad os forestille os at en bestemt etnicitet rent faktisk har betydning for effekten af en bestemt behandlingsindsats mod depression. På baggrund af RCT og metaanalyse vil det ikke være muligt at sige noget om hvorvidt forskellen i behandlingseffekt enten skyldes kulturelt indlejrede værdier omkring sundhed og behandling, eller en genetisk disposition - eller det forhold at depressionen er relateret til at have forladt hjemland og familie. Denne viden er relevant fx ift. hvorvidt forskellen primært gælder en gruppe flygtninge og indvandrere eller om den også gælder deres efterkommere, om resultatet er unikt for den specifikke etniske gruppe eller man ved fordel kan undersøge om samme effekt gælder for andre grupper. Uden en dybere forståelse af de kausale sammenhænge, vil behandlingsvalget i det enkelte tilfælde stadig være under-informeret. Hvis man derimod ved *hvorfor* etniciteten *generelt* har betydning for behandlingseffekten, vil det være muligt at arbejde med denne viden i forhold til behandlingsvalg for den enkelte patient. Lægen får dermed

mulighed for at undersøge om de karakteristika og årsagsfaktorer der er typiske for patienter der på den ene eller anden måde ligner den patient han skal behandle, også rent faktisk er tilstede hos det pågældende individ, eller om en anden mindre typisk årsag til depressionen er mere sandsynlig i det specifikke tilfælde. For at undersøge hvorfor behandlinger virker eller ikke virker er det dog ikke tilstrækkeligt at basere sig på den kvantitative metode. Det bliver nødvendigt at anlægge en forståelsesorienteret tilgang til, hvilke omstændigheder der er af betydning, og derfor involvere forsøgspersonerne ikke blot som standardobjekter der kan måles på, men som subjekter der oplever sygdom og behandling.

Når evidenshierarkiet placerer kvalitative studier nederst eller helt uden for hierarkiet, får det altså konsekvenser for hvilken form for viden der opnås. De fænomener som RCT belyser, altså den statistiske effekt af farmakologiske indsatser, betragtet isoleret og i en forsøgssetting, bliver dermed også vigtigere end andre vidensformer. Dermed bliver hierarkiet for forskningsdesign også et hierarki over forskningsspørgsmål og over hvilken viden vi helst skal besidde, eftersom kun nogle spørgsmål kan undersøges med de metoder der anses for god evidens. Vigtig viden kan således være umulig at opnå evidens for, og dermed basere praksis og anbefalinger på.

4.4 Delkonklusion

Det guidende spørgsmål i kapitel 3 og 4 har været:

Hvilken slags viden opnås gennem RCT-baseret evidens og hvilke begrænsninger har den?

I første afsnit argumenterede jeg for at randomiserede kontrollerede forsøg ikke sikrer den interne validitet som det antages hos eksempelvis Straus et al. og i Cochrane handbook. Randomisering udgør ikke en garanti for fuld sammenlignelighed mellem indsats- og kontrolgruppe. Muligheden for skævvridning betyder at man bliver nødt til at søge indsigt i forsøgenes processer, frem for at forlade sig på den mekaniske kontrol der ligger i metodologien. Når man tjekker for 'baseline imbalances' er det således et udtryk for at metodens statistiske fundament ikke yder tilstrækkelig sikkerhed for den interne validitet. De 'baseline imbalances' som man kontrollerer for, er desuden et udtryk for en bestemt forståelse af hvad der udgør relevante kausalfaktorer, og man bør derfor være bevidst om denne teoriafhængighed, der desuden går igen i forsøgets hypotese, i valget af effektmål og i fortolkningen af resultater. I og med at der netop ligger en teoretisk forståelsesramme til grund for forsøget (der ikke kan undgås, men bør ekspliciteres), er det nødvendigt at informere

forsøgene ved brug af eksperter fra det specifikke forskningsfelt, frem for at forsøgene kun udføres af metodiske eksperter.

I andet afsnit argumenterede jeg for at de retningslinjer der angives hos Straus et al. for anvendelsen af evidens i behandlingsvalget er utilstrækkelige. Mens resultaterne af RCT-baseret evidens er af statistisk generel karakter, er det ikke klart hvordan denne viden informerer valget for den enkelte patient. Én forudsætning for evidensens relevans er dens eksterne validitet. Flere kritikere peger imidlertid på at den forsøgssetting RCT kræver, samt inklusions- og eksklusionskriterier, kompromitterer den eksterne validitet. EBM-håndbogen nedtoner betydningen af kvalitativt behandlingsmæssigt distinkte patientsubgrupper. Dermed antages det at den RCT-baserede evidens er universel og patientpopulationen er sammenlignelige angående behandlingsrespons. Dermed er den anden forudsætning for RCT-baseret evidens patientpopulationens homogenitet.

I tredje afsnit undersøgte jeg hvilke begrænsninger for viden den RCT-baserede evidensmetodologi medfører. En central kritik af metoden er at den kun kan undersøge de specifikke elementer af behandlingspraksis, mens de non-specifikke elementer, der tilsammen betegnes placebo, har større betydning for behandling af depression. De non-specifikke elementer kan ikke testes i RCT da de ikke udgør en standardiseret indsats, og er afhængige af patientens oplevelse og håndtering af sygdommen. RCT kritiseres desuden for at have en standardiserende og objektiviserende sygdomsforståelse, der udelukker indsigter i hvordan kontekster har betydning for behandlingseffekt. I kraft af RCT'ets fokus på at isolere enkelte behandlingsfaktorer, underkendes desuden betydningen af samvirkende faktorer. Mulighederne af foretage undersøgelser der kunne afdække mere specifikke grupper behandlingsrespons er praktisk vanskelige i EBM og desuden vil der altid være et generaliseringsproblem eftersom der ikke opnås indsigt i kausalmechanismen bag behandlingseffekten. Endelig er der mulige årsagsfaktorer der ikke kan undersøges i RCT, da de omhandler patienters subjektive oplevelse af sygdom og behandling, der ikke kan undersøges som standardiserede faktorer.

Den centrale forudsætning for at den RCT-baserede evidens er relevant som grundlag for behandlingspraksis, er at patientpopulationen der undersøges er homogen. De tre grupper af kritik som jeg har gennemgået, vil alle være af mindre relevans, hvis patientgrundlaget kan betragtes som homogent. Som Bluhm og Borgerson skriver er den højest rangerende forskning designet til at afdække simple kausalforhold mellem behandlinger og effekter (Bluhm & Borgerson 2011: 218). I de to eksempler er der imidlertid ikke påvise noget simpelt kausalforhold mellem behandling af

depressionsdiagnosticerede patienter med SSRI eller TCA og en positiv effekt, da langt fra alle patienter oplevede respons eller remission. Men det betyder ikke nødvendigvis at behandlingen ikke var den rette. Der er flere sygdomme hvor det er tilfældet at selv den bedste behandling som lægevidenskaben kan mønstre, for nogle patienter ikke er tilstrækkelig. Strålebehandling kan eksempelvis ikke altid behandle kræftsvulster. Ikke desto mindre ville det være urimeligt at sige at strålebehandlingen derfor var en fejlbehandling, i nogle tilfælde er kræftsvulsten simpelthen bare for omfangsrig, placeret vanskeligt etc. til at man med de redskaber man har tilgængeligt, er i stand til at behandle den. Patientgrupper kan derfor godt reagere forskelligt på behandling, men alligevel være homogen på den måde, at de responderer bedst eller har størst chance for at respondere, på samme behandling. I det følgende kapitel undersøger jeg derfor grundlaget for klassifikation i psykiatrien, med henblik på at afdække hvorvidt diagnosen depression med rimelighed kan forventes at skabe en homogen patientpopulation.

Kapitel 5: Sygdomsklassifikation

Klassifikation er en hjørnesten i al videnskabelig praksis. Klassifikationen er et redskab til at skabe orden i det kendte, og en ramme til at forstå nye fænomener og observationer. Klassifikation er vigtigt for vores evne til at kunne forudsige og forklare. Hvis et nyt fænomen kan antages at være af samme slags som tidligere registrerede fænomener, fortæller det noget om hvordan det kan forventes at opføre sig både prospektivt (forudsigelse) og retrospektivt (forklaring) (Thornton 2007: 167). Klassifikation er således et redskab til at skabe system i enkelttilfældene og bestemme hvilke der kan sammenlignes og over hvilke der kan generaliseres.

I dette kapitel vil jeg starte med at se på nogle centrale begreber i forhold til sygdomsklassifikation. Det drejer sig om distinktioner mellem naturlige og konstruerede klasser og mellem kategoriers reliabilitet og validitet. Jeg diskuterer efterfølgende nogle centrale problemstillinger ang. sygdomsklassifikation ved at tage udgangspunkt i den somatiske medicin, for herefter at kunne karakterisere den psykiatriske sygdomsklassifikation og forklare hvilket grundlag og hvilke rationaler der ligger bag klassifikationen. Endelig analyserer jeg de to klassifikationssystemers definitioner af psykisk forstyrrelse.

5.1 Naturlig og konstrueret klassifikation

Fænomener kan klassificeres på vidt forskellige måder, og et centralt spørgsmål vil ofte være i hvor høj grad klassifikationen er naturlig, og i hvor høj grad den er pålagt eller kunstig. Når man taler om naturlige klasser, menes at enkelttilfældene der kan tilskrives klassen deler nogle naturlige underliggende ensartede karaktertræk, der er definerende for klassen. Konstruerede klasser derimod er udtryk for, at vi har en interesse i at klassificere enkeltelementerne på en bestemt måde. Et eksempel på en klassifikation der er udtryk for naturlige klasser, er det periodiske system. Man kan være temmelig sikker på at det periodiske system rent faktisk afspejler naturens opbygning af grundstoffer. Et utal af eksperimenter og forsøg har afdækket og bevist at grundstoffernes egenskaber følger af deres strukturelle opbygning af protoner, neutroner og elektroner. Vi ved så meget om kemiske strukturer at vi endda er i stand til manipulere med dem og at forudse virkningerne. Spørgsmålet om hvorvidt et stof er et grundstof og hvor det i så tilfælde skal placeres, kan afklares alene ved den naturvidenskabelige forskning.

Naturlige klasser siges at eksistere uafhængigt af vores viden og klassifikation af dem, og det er således ikke et spørgsmål om at skabe klassen, men om at opdage den. Alle medlemmerne af en

naturlig klasse er ikke nødvendigvis komplet ens, men de deler nogle definerende essentielle egenskaber, der er determinerende for andre klasse-specifikke egenskaber. Dermed udgør naturlige klasser basis for stærke induktive slutninger. Hvis en enhed kan identificeres som hørende til en bestemt klasse, udløser denne erkendelse yderligere viden om enheden, da medlemmerne af en klasse kan forventes at opføre sig ens (Varga 2013: 84). Hvis det fx kan registreres at et stof består af atomer med 79 protoner, opfylder stoffet den definerende egenskab for klassen guld. På baggrund af vores kendskab til andre guld-enheder, kan der nu sluttes induktivt til viden om enhedens øvrige egenskaber, fx smeltepunkt og tyngde.

Anderledes forholder det sig imidlertid med de klasser vi refererer til når vi eksempelvis snakker om urter eller ukrudt. Der findes ikke et sæt af biologiske kriterier der kan adskille urter fra ikke-urter. Urter har det tilfælles at være mindre planter der dør efter blomstring, men dette kriterium adskiller ikke urter fra eksempelvis ukrudt. Forskellen er funktionel, det skel der er draget mellem urter og ukrudt hviler på menneskets forhold til planten, af vores værditilskrivning. Urter og ukrudt er derfor udtryk for konstruerede klasser (Thornton 2007: 168ff). Klassifikation kan således foretages med forskellige formål for øje. Hertil knytter der sig især to centrale begreber, nemlig reliabilitet og validitet.

5.2 Reliabilitet

Reliabilitet er et udtryk for datas pålidelighed, og er således et mål for en metodes præcision. At en metode har en høj præcision, indebærer at de resultater den frembringer, er reproducerbare. I den kliniske praksis handler reliabilitet derfor om at flere observatører på et givent tidspunkt skal kunne nå de samme konklusioner om en given patient. Når det angår psykiatrien er metoden der anvendes psykiaterens iagttagelsesevne, og denne afhænger i høj grad af den enkeltes viden, erfaring og indlevelsessevne. Der kan være stor forskel på hvilken vægt psykiateren tillægger bestemte symptomer eller udtalelser og dermed hvilke observationer der tillægges størst vægt og får betydning for diagnosen (Parnas, Kragh-Sørensen & Mors 2010: 44f). Der er derfor relativt høj risiko i psykiatrien for at resultater, fx en diagnose, ikke er reproducerbar, og dermed mangler diagnosen reliabilitet. For at sikre en høj reliabilitet skal reglerne for hvordan en diagnose skal stilles derfor ikke være for komplekse (Ibid.: 45).

Klassifikationen skal forsyne lægen med pejlemærker for hvilken diagnose der skal anvendes. Det gør den ved at angive en række diagnostiske kriterier. Kombinationen af kriterier skal udgøre

tilstrækkelig information til at patientens tilstand kan identificeres med en bestemt diagnose, og til at andre diagnoser samtidig kan udelukkes. De diagnostiske kriterier skal derimod ikke være en udtømmende beskrivelse af en sygdom, da det i praksis ville blive alt for besværligt at skulle forholde sig til i diagnosticeringen af den enkelte patient. Det fulde sygdomsbillede, inklusiv den aktuelle empiriske viden om behandling og forekomst, samt patogenetiske og ætiologiske hypoteser angående sygdommens årsag, oprindelse og mekanisme, hører for så vidt ikke hjemme i klassifikationssystemet (Ibid.: 47). Den diagnostiske manual tilbyder derfor ikke en sygdomsforståelse og den er ikke fuldkommen i sin beskrivelse af sygdommen.

Hensynet til at sikre reliabilitet i det psykiatriske klassifikationssystem har ifølge Parnas, Kragh-Sørensen og Mors, de seneste årtier resulteret i en tendens til forsimpling og reduktion af den psykopatologiske viden. Dermed, mener de, er hensynet til klassifikationens validitet til gengæld blevet skubbet i baggrunden (Ibid.: 45).

5.3 Validitet

Validitet betegner nogets gyldighed, fx et begreb, et fund eller en diagnose. Validiteten er en betegnelse for i hvilket omfang der er overensstemmelse mellem begrebet og det som begrebet refererer til, altså en virkelighed der ligger udenfor begrebet. For at en diagnose er valid må den altså være i overensstemmelse med det fænomen som den beskriver. Fx kan flere psykiatere godt komme til den konklusion at en patient er suicidal. De har måske brugt samme metode til at komme frem til deres konklusion, hvorpå man kan slutte at metoden har høj reliabilitet. Eller de har brugt forskellige metoder til at komme frem til samme diagnose, hvortil man kan sige at diagnosen 'suicidal' har høj reliabilitet. Dette fortæller dog ikke noget om hvorvidt diagnosen er sandfærdig. Diagnosen er kun sandfærdig eller valid hvis patienten rent faktisk er selvmordstruet (Parnas, Kragh-Sørensen & Mors 2010: 43).

Spørgsmålet om validitet er naturligvis kompliceret, for ofte kan det være svært at sætte kriterier for, hvornår man kan vide noget som helst med sikkerhed. I psykiatrien bliver spørgsmålet om hvorvidt klassifikationen og de psykiatriske diagnoser er valide, ekstra svært at svare på, da man ikke kan observere et kausalt forløb. Dvs. at man ikke kan udpege specifikke ætiologiske faktorer, der udvikler sig på bestemte måder og giver sig til udtryk i bestemte symptomer og sygdomsbilleder. Måske derfor er det ikke klarlagt præcis hvilken betydning validitet har i forhold

til psykiske diagnoser. Validitet drejer sig som Kendell og Jablensky siger, om *"the nature of reality"*, men hvordan denne afdækkes er mere kompliceret (Kendell & Jablensky 2003: 5).

Når man konceptualiserer et symptombillede, ved at klassificere det som en diagnostisk kategori, er det således spørgsmålet, hvordan dette begreb henter sin gyldighed, hvad det er for et fælles forhold der henvises til når man klassificerer fænomener som del af samme gruppe. Den virkelige modpart som diagnosen refererer til, kan defineres på forskellige niveauer (Parnas, Kragh-Sørensen & Mors 2010: 43). De fleste somatiske sygdomme er defineret på patogenetisk niveau, dvs. at diagnosen henviser til den biologiske mekanisme der menes at ligge til grund for sygdomstilstanden, og ideelt set ville det samme kunne gøres i psykiatrien. Der er forsøgt etableret en model der fokuserer på de biologiske aspekter af psykisk sygdom, således at diagnoserne valideres i forhold til fx molekylær genetik, neurokemi, neuroanatomi og kognitiv neurovidenskab (Kendell & Jablensky 2003: 5). Men indtil videre er der ikke tilstrækkelig sikkerhed for relationen mellem det man kan observere i hjernen og den sygdomstilstand der opleves og udvises hos patienten.

Derfor lægger man det validerende kriterium i symptomerne. De mest anvendte former for validitet i psykiatrien afspejles i fem kriterier der angives hos Parnas, Kragh-Sørensen og Mors (2010: 43f). Jeg refererer her tre.

1. Begrebsvaliditet (concept validity), henviser til de oprindelige overvejelser som ligger til grund for etableringen af en bestemt diagnose. Hvis de diagnostiske kriterier ikke svarer til indholdet af begrebet, mangler diagnosen begrebsvaliditet. Omfatter diagnosen depression fx ikke 'nedsat stemningsleje', er den ikke i overensstemmelse med det oprindelige begreb for depression.

2. Teoretisk validitet (construct validity), der lægger det eksterne kriterium i en teoretisk forståelsesramme. Teoretisk validitet spørger til, om diagnosen rent faktisk er i stand til at teste det den siger den tester, dvs. i hvilket omfang afspejler diagnosen depressionstilfældene. Vi har altså en teoretisk konstruktion (depression), og en diagnose der har til formål at omfatte alle tilfælde af denne tilstand. Teoretisk validitet angår altså diagnosens evne til at omfatte netop depressionstilfældene. Denne form for validitet forudsætter et stærkt teoretisk fundament og testes ved at symptomer og sygdomsforløb må være i overensstemmelse med det teorien om diagnosen.

3. Prognostisk validitet, der henviser til det fremtidige forløb, evt. som resultat af behandling. Hvis sygdommen forløber i overensstemmelse med det oprindelige sygdomsbegreb kan diagnosen siges at have prognostisk validitet.

Det eksterne kriterium til validering af diagnosen lægges enten i begrebet, i den teoretiske konstruktion eller i det teoretiserede forløb. Validitetsformerne validerer diagnosen i forhold til dens evne til at indfange eksempelvis 'det vi forstår ved depression', dvs. de depressive symptomer, men derved valideres diagnosen ikke efter et naturligt forekommende fænomen. Der gives ingen garanti for at den grænse der drages omkring diagnosens genstandsfelt, er udtryk for en bestemt underliggende mekanisme. Der er altså ingen sikkerhed for at den diagnostiske kategori svarer til en egentlig uafhængig eller naturlig sygdomsentitet, altså med andre ord, om den opdeling der er foretaget afspejler underliggende patologiske tilstande der adskiller sig fra andre former for patologiske tilstande og fra ikke-patologiske tilstande. (Thornton 2007: 174;180). Dvs. om den med et ofte anvendt udtryk "carves nature at its joints".

Prognostisk validitet er dog en særdeles nyttig egenskab for en diagnose og i forhold til denne undersøgelse, da den søger at sikre en population der kan forventes at reagere sammenligneligt på samme behandling. Prognostisk validitet af diagnosen svarer altså til homogenitet mht. behandlingsrespons som jeg i kapitel 4 konkluderede var en forudsætning for at man kan anvende resultaterne af RCT. Men hvordan opnås så denne prognostiske validitet?

Ifølge Kendell må man i mangel af kendskab til den mekanisme der forårsager sygdommen, forsøge at validere diagnoserne ved at føre dokumentation for, at diagnoserne er afgrænset fra hinanden på relevant vis, således at hver kategori afspejler forskelle der er væsentlige for forståelsen af sygdommens årsag, forløb eller behandling. I en meget omtalt artikel argumenterer Kendell & Jablensky derfor for at de psykiatriske diagnosers validitet kun kan bestemmes ved at dokumentere såkaldte "zones of rarity" mellem de forskellige diagnostiske tilstande (Kendell & Jablensky 2003: 9). Dokumentationen af "zones of rarity", skal tjene til at vise at der netop kan skelnes mellem forskellige psykiske lidelser, og at kriterierne for hver diagnose trækker grænserne for diagnosernes omfang således at de omfatter en population hvis lidelser har fælles træk og som adskiller sig fra andre former for lidelser. Hvis kategorierne rent faktisk svarer til forskellige tilstande der kan anses for at udgøre homogene grupper angående årsag eller karakteristika der er af betydning for deres forløb eller behandling, må flertallet af patienter have enten den ene eller den anden lidelse. 'Zones of rarity' er altså påvisningen af at majoriteten af psykiatriske patienter falder indenfor netop de diagnostiske grupper der beskrives i klassifikationen, og ikke falder mellem to relaterede kategorier i kraft af deres specifikke symptomsammensætning, eller mellem en diagnostisk kategori og ingen kategori. Hvis dette er tilfældet giver det også mening at behandlingen retter sig specifikt mod en

bestemt diagnose, for diagnosen vil angive at et bestemt underliggende karakteristika gør sig gældende. Hvis kategorierne altså afspejler distinkte og naturligt forekommende forskellige lidelser, må det være muligt at validere kategorierne ved påvisning af at de fleste patienter enten lider af det ene eller det andet, med andre ord, at der er få ”grå” tilfælde i forhold til de typiske ”hvide” og ”sorte” tilfælde (Kendell 2013: 8).

Jeg vender tilbage til diskussionen af diagnosen depressions validitet, efter jeg har set på de omstændigheder den somatiske medicinske klassifikation er bygget op om.

5.4 Sygdomsklassifikation

Sygdomstilstande kan klassificeres på forskellige niveauer alt efter hvilken viden og interesse man har. Der findes tre overordnede niveauer. Symptomniveauet siger noget om hvad det er for et problem der skal løses, men ikke noget om hvordan det løses, eller hvorfor det er opstået. Det patogenetiske niveau angår sygdomsmekanismen og siger altså noget om hvad der genererer problemet. Endelig er der det ætiologiske niveau, der fortæller noget om hvorfor problemet er opstået, hvilket fx kan være væsentligt i et forebyggelsesperspektiv (Wulff, Pedersen & Rosenberg 1993: 90).

Klassifikationen af somatiske sygdomme er overvejende patogenetisk, men det er ikke muligt at lave en klassifikation der giver en udtømmende beskrivelse af sygdomsmekanismen. Som oftest kender man blot et udsnit af de faktorer der indgår i det kausale net der udgør sygdomsmekanismen (Wulff, Pedersen & Rosenberg 1993.: 87). Over de sidste 200 år har sygdomsklassifikationen således ændret sig i takt med lægevidenskabens nyopdagelser og skiftende fokus. I begyndelsen af 1800-tallet begyndte franske patologer at identificere sygdomme med patoanatomiske læsioner der resulterede i en anatomisk sygdomsteori. Nogle årtier senere udviklede lægevidenskaben sig mod en større indsigt i human fysiologi og forståelsen af, hvilke forhold der for bestemte sygdomstilstande var de essentielle, revurderedes, hvilket giver sig til udtryk i sygdomsklassifikationen. Samme mekanisme gentog sig ved udviklingen af den moderne mikrobiologi. Den klassifikation vi har i dag, der bredt betegnes den biologiske sygdomsmodel, bygger derfor på flere medicinske specialer (Ibid.: 92). Der er behov for i klassifikationen af forskellige sygdomme at lægge vægten på forskellige biologiske niveauer. Det betyder dog ikke at tidligere klassifikationer nødvendigvis har været decideret forkerte. Det er snarere sådan at man i takt med at man opnår mere og mere viden, bliver bevidst om at enkelte faktorer er væsentligere,

fremtræder hyppigere, eller er mere relevante ift. vores viden om sygdommen og dens behandling, end andre faktorer er, i det kausalnet der fører til udviklingen af sygdom. Sygdomsklassifikationen dækker over det forhold, at alle patienter er forskellige, udviser forskellige kliniske sygdomsbilleder og har forskellige sygdomsmekanismer. Klassifikationens formål er på trods heraf at finde mønstre og ligheder i alle enkelttilfældene.

Et led i den øvelse der handler om at skabe orden i det uordentlige, er en analyse af de forskellige faktorerens betydning i kausalnettet. Ideelt set bygger en sygdomsklasse på en række betingelser der er både nødvendige og tilstrækkelige. Dette er tilfældet for grundstoffer, hvis tilhørsforhold til en bestemt klasse udelukkende afhænger af klart definerede kriterier for stoffets struktur. Med en sygdomsklassifikation hvor hver klasse er defineret ved et antal nødvendige og tilstrækkelige betingelser, ville det altid være muligt at drage en klar grænse for diagnosens genstandsfelt. I analysen af årsager til udviklingen af en bestemt sygdom er det dog sjældent muligt at identificere årsagsfaktorerne som hverken nødvendige eller tilstrækkelige. Fx er mavesår hverken en nødvendig eller tilstrækkelig årsag til øvre abdominal smerte. Den er ikke nødvendig da smerten kan opstå af andre årsager end mavesår, og den er ikke tilstrækkelig da ikke alle patienter med mavesår får smerter. Når en patient henvender sig med øvre abdominal smerte, kan det derfor ikke sluttes at vedkommende har mavesår før der foreligger en specifik undersøgelse af hvorvidt dette er tilfældet. I det tilfælde at patienten rent faktisk har mavesår, vil man til gengæld normalt sige at mavesåret er årsagen til smerten, selv om mavesåret ikke alene kan have forårsaget smerten. Mavesåret er da en non-redundant (ikke-overflødig) årsagsfaktor, dvs. at smerten ikke var forekommet hvis ikke der havde været et mavesår (Ibid.: 84). Det vil sige at diagnosticeringen af mavesår medfører et valg om at fokusere på én non-redundant faktor, frem for at tage hele det diagnostiske årsagskompleks i betragtning. Endelig kan mavesåret være opstået af forskellige årsager, dvs. der findes heller ingen nødvendige eller tilstrækkelige betingelser på det ætiologiske plan. Valget af årsagsfaktor afspejler en vurdering af mavesår som den væsentligste kausalfaktor i sygdomstilstanden, og samtidig en faktor der kan behandles, dvs. valget af årsagsfaktor der får betydning for diagnosen, er altså udtryk for en interesse (Pedersen 1990: 8;16). Dermed afspejler kategoriseringen en kausal mekanisme. Som Stig Andur Pedersen gør opmærksom på, behøver denne mekanisme ikke at være naturlig, i betydningen 'naturlig klasse'. Den kan ligeledes være social, psykologisk eller noget helt tredje (Pedersen 1990: 8), det væsentlige er at mekanismen indgår som forklarende faktor for sygdomstilstanden.

Den delvise specificering betyder, at det som udgør en naturlig klasse på ét niveau, ikke nødvendigvis fremstår som en naturlig klasse på et andet niveau, og at det, som er definerende for den naturlige klasse, ikke nødvendigvis betegner alle karakteristika ved det patologiske niveau.

Jeg håber med disse eksempler at have vist at diagnostik og behandling er komplekse processer, og at det ikke er muligt at opstille regler eller principper der er universelt gyldige. Klassifikationen af sygdomme er menneskeskabt og skal ikke forstås som begreber der modsvares af enheder der naturligt eksisterer i sig selv. Når man således taler om ”opdagelsen” af en ny sygdom, er der ikke noget man som sådan har opdaget. Man har derimod valgt at sætte en ny etiket på en gruppe af patienter, da man har vurderet at der et relevant skel at drage. Klassifikationen af sygdomme kan derfor ikke forstås på samme måde som klassifikationen af grundstoffer eller eksempelvis klassifikationen af plantearter. En relevant forskel er, at sygdomme har et forløb, og at dette forløb er påvirkeligt (Wulff, Pedersen & Rosenberg 1993: 96ff).

Selvom at klassifikationen af somatiske sygdomme døjer med disse vanskeligheder, vil de fleste somatiske diagnoser alligevel leve op til Kendell og Jablenskys begreb om validitet ved kravet om ’zones of rarity’. I en undersøgelse af hvorvidt patienterne fordelte sig indenfor diagnoserne eller på tværs af diagnoserne, ville man naturligvis finde, at enten lider patienten af mavesår eller også gør han ikke. Kriteriet for at diagnosen mavesår kan stilles er netop tilstedeværelsen af samme. Patienten kunne også klassificeres anderledes, hvis man vægtede andre sygdomskriterier, men det ændrer ikke ved at diagnosen mavesår udgør en valid klasse.

5.5 Klassifikation i psykiatrien

Vi skal nu se på hvordan man etablerer diagnostiske kategorier i psykiatrien.

I Danmark og resten af Europa benytter man sig af diagnosemanualen ICD (International Classification of Diseases) der udarbejdes af WHO (World Health Organisation), og som findes i 10. udgave fra 1992. Der foregår et løbende revideringsarbejde og ICD-11 forventes at kunne udgives i 2017. ICD-10 klassificerer alle sygdomme, ulykkestilfælde og dødsårsager, og det er således kun kap. 5 ud af 21 sektioner der handler om ”psykiske lidelser og adfærdsmæssige forstyrrelser”. I USA bruger man i stedet diagnosemanualen DSM (Diagnostic and Statistical Manual of Mental Disorders) der udarbejdes af den amerikanske psykiatriorganisation (APA) og findes i 5. udgave fra 2013, DSM-5. Der er enkelte forskelle mellem de to manualer. Eksempelvis tillader den amerikanske diagnosemanuals diagnostiske regler at samme patient kan have flere

psykiske diagnoser samtidig, mens det kun er tilfældet i begrænset omfang i ICD-10 pga. veldefinerede hierarkiske regler for diagnostik. De to manualer adskiller sig dog ikke væsentligt fra hinanden i grundlag og opbygning. Gennem de seneste årtier har WHO og APA arbejdet på at nærme sig hinanden, og de to manualer er i dag temmelig ens (Cooper 2013: x). Jeg vil i dette kapitel derfor læne mig op af publikationer der refererer til både det ene og det andet system, så længe der ikke er tale om tematikker og problemstillinger der er specifikke for den amerikanske model.

Ambitionen for den nye klassifikation var ifølge Parnas et al. oprindeligt at lave en rent beskrivende, ateoretisk klassifikation over psykiske lidelser. Ideen med at udarbejde en sådan klassifikation var at vedkende sig, at den ætiologiske og patogenetiske viden indenfor psykiske lidelser var alt for ufuldstændig til at det var muligt at konstruere en sygdomsklassifikation der er inddelt på baggrund af de bagvedliggende årsager til sygdomsbilledet. I stedet skulle klassifikationen baseres på det der kan observeres direkte, nemlig symptomerne. Hver klasse af tilstande defineres gennem et antal specificerede symptomer eller tegn, der ofte vægtes lige meget og er udviklet med fokus på diagnosens reliabilitet. (Parnas, Kragh-Sørensen & Mors 2010: 32).

Denne form for klassifikation indførtes med DSM-III i 1980 og betegnes som nominalistisk. Nominalistisk vil sige en forståelse af diagnoserne som vedtagne benævnelser; menneskeskabte konventioner med udspring i de observationer der er gjort og de praktiske erfaringer der er opnået. Det nominalistiske udgangspunkt er bevidst om at diagnoserne ikke er uafhængige, men derimod indlejret i vores sproglige vaner, vores historiske erfaringer og samfundsmæssige forhold, og dvs. konstrueret og afhængige af hvordan vi opfatter verden (Parnas, Kragh-Sørensen & Mors 2010: 37).

Klassifikationen kan siges at forholde sig neutralt til konflikten mellem den biologiske og den psykosociale forklaringsmodel. Ved at klassifikationssystemet netop forholder sig ateoretisk til de psykiske lidelsers oprindelse, og baserer sig på symptomer som udgangspunkt for klassifikationen, undgår man at klassifikationen afgrænser den psykiatriske videnskab til kun at omfatte forklaringsmodeller der ligger indenfor naturvidenskaben (Ghaemi 2003: 7). Samtidig er det empiriske udgangspunktet for klassifikationen – registreringen af observerbare symptomer, dog baseret på en naturalistisk forståelse af sygdomsenheder, som årsager til bestemte sygdomsbilleder.

5.6 The minimal medical model

Det nominalistiske udgangspunkt beskrives hos Goodwin og Guze, i deres indflydelsesrige lærebog 'Psychiatric Diagnosis' der første gang udkom i 1974. Samuel Guze og Donald Goodwin var blandt de forskere der udarbejdede DSM-III (North & Yutsy 2010: xi). Goodwin og Guze indskærper at klassifikation har to formål: kommunikation og forudsigelse (Goodwin & Guze 2010: xv).

Kommunikation handler om klassifikationens reliabilitet – at sikre at man taler om samme fænomen når man bruger en diagnostisk kategori. Forudsigelse handler derimod om den prognostiske validitet, dvs. klassifikationens formål er at definere diagnoserne således at de svarer til bestemte forløb og behandlingsrespons. Eftersom det gøres på rent empirisk baggrund antages det altså at en bedre differentiering af symptombilleder kan medføre bedre prognostisk validitet. De indleder således bogen med de følgende linjer:

Diseases are conventions and may not "fit" anything in nature at all. Through the centuries, diseases have come and gone, some more useful than others, and there is no guarantee that our present "disease" – medical or psychiatric – will represent the same clusters of symptoms and signs a hundred years from now that they do today. On the contrary, as more is learned, more fruitful clusters surely will emerge. (Goodwin & Guze 2010: xvii).

Hvis man sammenligner med de somatiske sygdomsklasser, har også de ændret sig i takt med den medicinske videnskabs udvikling. Men ændringer i de medicinske klasser er sket med den teoretiske udvikling og opblomstringen af nye specialer indenfor videnskaben. Hvis den somatiske medicins diagnose identificeres på "clusters of symptoms and signs", er disse baseret på en teori om det kausalkompleks der går forud for symptomerne. Desuden er det ikke tilfældet for den somatiske medicin at prognostisk validitet nødvendigvis vil afspejles i differentieringen af symptomer, da naturlige klasser på patogenetisk niveau ikke garanterer naturlige klasser på symptomniveau.

Goodwin og Guze er eksponenter for en empiristisk position der også omtales "the minimal medical model". Psykiatriens genstandsfelt afgrænser sig for empiristerne til den empiriske viden der kan opnås om disse fænomener, gennem observationer, men det anses dog samtidig for indlysende at psykiske lidelser er et medicinsk fænomen, idet de repræsenterer manifestationer af funktionelle forstyrrelser i en del af kroppen, altså hjernen (Murphy 2013). I det omfang der findes en fyldestgørende forklaring på den psykiatriske diagnose, er sygdommen ikke længere at betragte som psykiatrisk, men er i stedet blevet biologisk. Ætiologi og patogenese er dermed ikke et ærinde for psykiatrien, og diagnoserne er konstrueret på baggrund af et praktisk behov (Murphy 2013).

Zones of rarity

I indledningen til ICD-10 afgrænses klassifikationens ambitioner i overensstemmelse med Goodwin og Guzes model:

The descriptions and guidelines carry no theoretical implications, (...) They are simply a set of symptoms and comments that have been agreed (...) to be a reasonable basis for defining the limits of categories in the classification of mental disorders. (WHO 1992: 2)

Som jeg har gjort rede for i afsnit 5.3 er det netop validiteten af diagnosernes grænser som Kendell og Jablenskys begreb om 'zones of rarity' omhandler, altså at påvise diagnoserne som svarende til nogle faktiske klasser som de optræder i virkeligheden. I og med at det ætiologiske og patogenetiske årsagskompleks ikke er kortlagt for langt de fleste psykiske sygdomme, ville påvisningen af en naturligt afgrænset klasse af sygdomstilfælde indikere en specifik ætologi eller patogenese der er forskellig fra de tilstødende klasser, og dermed ville diagnosernes grænser afspejle information om sygdommens underliggende realitet, og dermed diagnosens validitet.

Hvis der ikke kan findes 'zones of rarity' betyder det at patientgrundlaget for de forskellige diagnoser flyder sammen. Der kan fx være en gruppe af patienter der udviser alle de typiske symptomer på depression og ikke oplever andre symptomer, men der er ligeså mange der oplever overvejende depressionssymptomer plus en 2-3 angstsymptomer, og der er igen en tilsvarende gruppe der oplever lige mange symptomer fra hver diagnose. Hvis dette er tilfældet, er det et problem for diagnosernes validitet, da grænserne dermed ligeså godt kunne være trukket andre steder, og dermed er der ikke grundlag for at tro at de sygdomstilfælde der kan diagnosticeres under en enkelt diagnose har noget specifikt tilfælles, og dermed er mere sammenlignelige med hinanden end med tilstødende diagnoser.

Forsøgene på at finde 'zones of rarity' mellem de relaterede diagnoser og mellem patologisk og ikke-patologisk tilstand, og dermed demonstrere naturlige grænser, har ikke været ret mange endnu. De forsøg der er gjort, har imidlertid ikke haft meget held med sig. Kendell har fx undersøgt hvorvidt patienter med skizofrene psykoser og affektive psykosers symptomsammensætninger fordelte sig binært eller om de fordelte sig kontinuerligt, og kunne ikke påvise nogen diskontinuitet og dermed heller ingen naturlig fordeling mellem to diagnostiske klasser (Kendell 2009: 9). Der er ligeledes foretaget mindst tre studier af grænsen mellem bipolar lidelse (maniodepressivitet) og depression, hvor man ikke har kunnet identificere en binær fordeling ang. antallet af bipolare symptomer (Parker 2013: 137). Desuden tyder studier på at mindre ændringer i definitionen af

syndromer, blandt andet depression, kan resultere i store ændringer i prævalensen, hvilket ligeledes indikerer at definitionen ikke svarer til en naturlig afgrænset lidelse (Kendell 2009: 9).

I indledningen til den nyere udgave af DSM, DSM-5 er man opmærksom på det problem, at der ikke kan forudsættes klare grænser for diagnosernes genstandsfelt:

Although some mental disorders may have well-defined boundaries around symptom clusters, scientific evidence now places many, if not most, disorders on a spectrum with closely related disorders that have shared symptoms, shared genetic and environmental risk factors, and possibly shared neural substrates (...) In short, we have come to recognize that the boundaries between disorders are more porous than originally perceived. (APA 2013: Introduction).

Forsøgene kan ikke vise at der ingen grænse findes, da det er muligt at 'the point of rarity' befinder sig indenfor kriterier eller symptomer man ikke har testet for. Samtidig må man konstatere at så længe man ikke har fundet nogen naturlig fordeling mellem de diagnostiske grupper, er der ikke belæg for at antage at diagnosen rent faktisk er valid, i det at den flyder sammen med andre diagnoser. En reel mulighed er derfor at der ikke findes kategorier af psykiske sygdomme, eller i det mindste ikke i det omfang klassifikationen giver udtryk for. En dimensional sygdomsklassifikationsmodel, svarer måske bedre til det kontinuum af varianter af lidelser og symptomer som de empiriske data synes at udtrykke.

5.8 Essentialisme

På trods af at de diagnostiske klasser altså ikke afspejler specifikke underliggende sygdomsmekanismer, og desuden ikke kan adskilles som distinkte syndromer på baggrund af symptomsammensætningen, bliver de psykiatriske diagnoser ifølge flere kritikere (Kendell & Jablensky 2003, Parnas et al 2010), alligevel opfattet som sygdomsenheder så snart de får plads i DSM eller ICD. Ifølge Parnas et al. tillægges de diagnostiske kategorier forestillinger om deres essentielle eller naturlige karakter. Dermed forstås de psykiske lidelser som velafgrænsede sygdomme, der kan forstås på linje med somatiske lidelser, som havende en objektiv væren i verden, uafhængigt af den menneskelige forståelse af sygdommen, dvs. som naturlige kategorier (Parnas, Kragh-Sørensen & Mors 2010: 37). Det at en diagnose optræder med en præcis definition i den officielle klassifikation over sygdomme har det med at medføre en tingsliggørelse af den diagnostiske konstruktion (Kendell & Jablensky 2003: 5). Diagnosen tillægges forklarende egenskaber, men indeholder reelt ikke andet end konstateringen af en række symptomer der optræder samtidigt. Kendell retter med kravet om 'zones of rarity' en kritik mod

klassifikationssystemet, eller måske især anvendelsen af klassifikationssystemet. Diagnosen afspejler ikke sygdomsintensiteter, men ikke desto mindre betragtes de som sygdomsintensiteter, og man glemmer at psykiatriske klasser ikke har den samme validitet som somatiske klasser.

Den antipsykiatriske kritik

Hvis den diagnostiske kategori ikke kan valideres ud fra andet end vores konstruerede begreber, baseret på empirisk tilgængelige symptomer, er der ingen sikkerhed for at kategorien ikke dækker over vidt forskellige underliggende forskelle, herunder ikke-patologiske problemer. Som rene konventioner har empirismen altså den svaghed, at den ikke adresserer problematikken omkring hvad der definerer en tilstand som patologisk i forhold til normal. Et vigtigt eksempel på denne problematik, er klassificeringen af homoseksualitet i DSM som en psykisk diagnose op til 1973.

Her berører vi et andet aspekt af kritik mod psykiatrien, nemlig retten til at definere hvad der kan forstås som sygdom. I 1960'erne satte den antipsykiatriske bevægelse spørgsmålstegn ved validitetens af hele psykiatriens fundament, ved at anfægte eksistensen af de psykiske sygdomme som sygdomme i medicinsk forstand. Den amerikanske psykiater Thomas Szasz kaldte i 1972 psykisk sygdom for en myte. Szasz benægtede ikke eksistensen af de virkelige fænomener der går under betegnelsen psykisk sygdom, men reagerede mod anskuelsen af psykiske lidelser som hørende under det medicinske domæne.

The concept of illness, whether bodily or mental, implies deviation from some clearly defined norm. In the case of physical illness, the norm is the structural and functional integrity of the human body. (...) What is the norm, deviation from which is regarded as mental illness? This question cannot be easily answered. But whatever the norm may be, we can be certain of only one thing: namely that it must be stated in terms of psychological, ethical and legal concepts (Szasz 1972: 15 gengivet i Thornton 2007: 16)

Psykiske sygdomme er defineret ved afvigelse fra psykologiske, etiske og juridiske normer. At definere bestemte afvigelser fra normerne som sygdom, er derfor at foretage en værdidom. At forestille sig at psykisk sygdom kan forstås i et medicinsk perspektiv og behandles med medicinske metoder, mente Szasz derfor var absurd, eftersom medicinske virkemidler aldrig vil kunne løse problemer, der per definition er ikke-medicinske (Thornton 2007: 16). Den anti-psykiatriske bevægelse er radikal i sin afvisning af et egentligt sygdomsbegreb for psykiske tilstande, hvilket også betyder at der ikke er noget 'objektivt' eller i naturvidenskabelig forstand 'virkeligt' ved psykiske lidelser. Det værdimæssige aspekt ved psykisk sygdom, er dog yderligere anledning til at stille spørgsmålstegn ved validiteten af psykiatrisk klassifikation. For hvis psykiske sygdomme kun

kan defineres i forhold til en række symptomer der opfattes som adfærd eller tanker der er afvigende fra det normale, så er der nødvendigvis et værdimæssigt aspekt i definitionen af hvad der udgør psykisk sygdom.

Denne kritik er i dag udbredt indenfor sociologien, blandt andre hos den engelske sociolog Nikolas Rose, der mener at der er sket en psykiatrisk omformning af ubehag generelt. Ligesom Szasz mener han at omformningen er afhængig af en norm i forhold til hvilken, oplevelser kan vurderes som unormale, samt desuden af et sprog, der gør det muligt at kommunikere ubehaget. Hvis normerne foreskriver det aktive, ansvarlige, vælgende selv, og der findes et sprog omkring depression, angst osv., kan tilstande, der tidligere har været defineret som almindelige livsproblemer, omkodes til depression (Rose 2006: 480). Hvis omkodningen sker i et medicinsk perspektiv, vil vi også i stigende grad forstå os selv som biologiske væsener, hvilket betyder en ændring for vores grænser for normalitet og patologi (Rose 2010: 51f).

Der er dermed tale om en konstruktivistiske kritik af den psykiatriske klassifikation, der går på dens evne til at skelne mellem normal og patologisk lidelse. Hvis det ikke er muligt at vise med 'points of rarity', at de diagnostiske klasser indfanger naturligt forekommende grupper af patienter, er grænsen mellem, hvad der omfattes af klassifikationen og hvad der ikke omfattes, sat arbitrært, og dermed åben for kritik. Hvis man således skal imødekomme den antipsykiatriske og konstruktivistiske kritik må man forsøge begrebsmæssigt at afgrænse patologisk lidelse fra ikke-patologisk lidelse. Derfor kigger jeg i det følgende afsnit på hvordan ICD-10 og DSM-5 definerer psykisk sygdom.

Kapitel 6: Analyse af klassifikationernes definitioner af mental forstyrrelse

Jeg har introduceret Alf Ross' sygdomsbegreb da jeg mener at det kan bruges til at mediere mellem det naturalistisk-objektivistiske sygdomsbegreb og det relativistisk-subjektivistiske sygdomsbegreb.

6.1 Ross' svar på den antipsykiatriske kritik

I forhold til Szasz' kritik af psykiske lidelser som rene værdidomme ville Ross være uenig. Afvigelsen fra en psykologisk, etisk eller moralsk norm er ikke det der definerer sygdommen. Det der kan betegnes sygdom, er derimod den afvigelse fra normalen der består i en funktionsnedsættelse der opleves hæmmende for individets mulighed for at udfolde sine livsmuligheder. En klassifikation over psykiske lidelser kan derfor godt bestå i en række beskrevne syndromer, der for de fleste patienter der oplever at lide under dem, er forbundet med en funktionel kompromittering som ønskes afhjulpet. Sygdom er derfor ikke et spørgsmål om hvad andre mener er afvigende og afvigelse er ikke sygdom i sig selv. Afvigelse kan dog medføre funktionsnedsættelse og i kraft heraf - ikke i kraft af den måde samfundet forholder sig til den - kan afvigelsen betegnes som sygdom.

Den nominalistiske klassifikations udgangspunkt i symptomerne, er ikke uforeneligt med Ross sygdomsbegreb, der også netop forstår sygdom i kraft af dens funktionelle betydning. Men angående psykiske sygdomme er normen for rigtig funktionsforløb dog afhængig af patientens mulighed for at kommunikere med andre mennesker. Ross forstår psykisk sygdom som funktionel forstyrrelse og afvigelse fra et mentalt normalplan, og denne forstyrrelse er af en sådan alvor at den resulterer i en kompromittering af individets mulighed for at kommunikere med andre mennesker (Ross 1980: 5). Grundlaget for en god interpersonel kommunikation er en vis enighed i opfattelsen af os selv og den omgivende verden, og dermed bliver konteksten for funktionsforstyrrelsen væsentlig, især i forhold til depression, da den motiverede og dermed forståelige depressive tilstand ikke er udtryk for en funktionsforstyrrelse.

En definition af psykisk forstyrrelse der kan afgrænse patologiske funktionsforstyrrelser fra ikke-patologiske funktionsforstyrrelser, må derfor angive rammerne for at de psykiske symptomer kan forstås som reelle funktionsforstyrrelser og dermed patologiske.

I det følgende vil jeg derfor se på hvordan DSM-5 og ICD-10 definerer psykisk sygdom, eller rettere psykiske forstyrrelser. Jeg medtager DSM-5, da den er 20 år yngre end ICD-10, og fordi det i DSM-5 anføres at man i revisionen af manualen har samarbejdet med den gruppe der reviderer ICD-11 der forventes at udkomme i 2017, med henblik på at harmonisere de to diagnosesystemer mest muligt (APA 2013: Introduction, Harmonization with ICD-11). Definitionen af psykiske forstyrrelser i DSM-5, er derfor muligvis også et fingerpeg mod hvilken retning ICD-11 vil tage, og den er derfor interessant at tage med i betragtning.

Både ICD-10 og DSM-5 bruger termen 'disorder', der på dansk oversættes til "forstyrrelse", om de psykiske diagnoser. Det skyldes at de psykiske diagnoser ikke kan siges at have en sikker sygdomsstatus pga. den ringe indsigt i sygdomsmekanismen (Parnas, Kragh-Sørensen & Mors 2010: 36) Endvidere betegnes forstyrrelserne som syndromer. Et syndrom er medicinsk defineret som en samling af symptomer og tegn der ofte optræder sammen.

6.2 ICD-10's definition af 'mental disorder'

I ICD-10 findes det bedste bud på en definition af psykisk forstyrrelse under overskriften 'Problems of terminology', 'Disorders'. Den lyder:

The term "disorder" is used throughout the classification, so as to avoid even greater problems inherent in the use of terms such as "disease" and "illness". "Disorder" is not an exact term but it is used here to imply the existence of a clinically recognizable set of symptoms or behaviour associated in most cases with distress and with interference with personal functions. Social deviance or conflict alone, without personal dysfunction, should not be included in mental disorder as defined here. (WHO 1992: 5).

I første omgang adresserer ICD-10 altså ikke specifikt behovet for en definition af psykisk sygdom, altså hvilke kriterier der må være opfyldt for at der kan tales om en patologisk tilstand. I stedet forholder man sig til vanskelighederne ved at vælge en passende term for de diagnostiske klasser. De diagnostiske klasser kan ikke kaldes disease (sygdom) og heller ikke illness (lidelse). Disorder (forstyrrelse, uorden) er ikke et eksakt ordvalg, men må bruges i mangel af bedre. Forstyrrelse dækker ikke over konfliktende eller afvigende adfærd. Definitionen vægrer sig mod den antipsykiatriske kritik, ved at udelukke den blotte afvigelse fra normalen som forstyrrelse. Den eneste positive definition af det fænomen som klassifikationen skal beskrive, er "*the existence of a clinically recognizable set of symptoms or behaviour associated in most cases with distress and with interference with personal functions*".

I ICD-10 benytter man altså en subjektiv markør i form af 'distress' (bekymring/lidelse) sammen med en objektiv markør i form af 'dysfunktion'. Problematisk adfærd er ikke et tilstrækkeligt kriterium for at tale om disorder. Personlig dysfunktion derimod er en nødvendig betingelse for at problematisk adfærd kan diagnosticeres. Personlig dysfunktion vil som oftest være akkompagneret af 'distress', men det er tilsyneladende ikke et nødvendigt kriterium.

Personlig dysfunktion virker altså til at være kodeordet her. Men hvordan skal personlig funktion og personlig dysfunktion forstås? Der siges ikke, som i DSM-5 definitionen, noget om, at der skal være tale om en *underliggende* dysfunktion. Dermed kommer der ikke i ICD-10's definition samme forventning om en biologisk funktionsfejl eller en psykologisk brist, som årsagsforklaring til udtryk. Personlig dysfunktion kunne indikere den *personlige oplevelse* af nedsat funktionsevne, det kunne fx være nedsat koncentrationsevne, nedsat evne til at indgå i sociale sammenhænge eller nedsat evne til at kontrollere sit temperament. At dysfunktionen er personlig kunne betyde at der er tale om en afvigelse fra et individuelt defineret normalplan – altså hvad personen normalt er i stand til.

Denne definition er i overensstemmelse med Alf Ross' sygdomsbegreb, så længe der er tale om en væsentlig afvigelse fra et funktionelt normalplan (hvor end denne grænse så måtte føres). Det funktionelle normalplan for psykiske funktioner er imidlertid hos Ross defineret i forhold til evnen til at kommunikere med andre mennesker hvilket indebærer en grundlæggende fælles forståelsesramme. Angående emotionelle lidelser som depression må den nedsatte funktionsevne være umotiveret eller urimeligt motiveret for at der reelt er tale om en forstyrrelse af funktionen og ikke blot funktion der er i overensstemmelse med normalplanet for emotionel reaktion.

Dermed er den blotte oplevelse af nedsat funktionsevne altså ikke per se patologisk, på baggrund af Ross' sygdomsbegreb, mens oplevelsen af den funktionsnedsættelse der ikke er kommunikerbar og dermed åben for interpersonel forståelse, må anses for patologisk. Derfor er de omstændigheder, funktionsnedsættelsen forekommer under, afgørende for hvorvidt funktionsnedsættelse kan betragtes som dysfunktion. Definitionen i ICD-10 definerer ikke videre deres begreb om dysfunktion, dvs. den angiver ikke "*hvilken norm der er bestemmende for, hvad der er tilfredsstillende og hvad der er utilfredsstillende funktionsforløb*" (Ross 1980: 17). Den adresserer dermed ikke spørgsmålet om hvorvidt den personlige dysfunktion også udgør en afvigelse fra et begreb om mentale funktioners normalplan, og det er derfor ikke givet at den indfører nogen skelnen mellem patologiske dysfunktioner og ikke-patologiske dysfunktioner.

Samme betragtning går igen hos Wakefield og First, der mener, at for at opnå en klassifikation der ikke patologiserer normale problemer, må de diagnostiske kriterier der baserer sig på symptomer, være konstrueret således at symptomet ikke med rimelighed kan antages at være udtryk for normal funktionsevne. (Wakefield & First 2013b: 665).

6.3 DSM-5's definition af 'mental disorder'

I 2013 udkom den 5. reviderede udgave af DSM. I den forbindelse er definitionen af 'mental disorder' ændret. I den gamle udgave, DSM-IV lyder definitionen:

A mental disorder is a clinically significant behavioral or psychological syndrome or pattern that occurs in an individual and that is associated with present distress or disability or with a significantly increased risk of suffering death, pain, disability, or an important loss of freedom.

Med DSM-5 er definitionen ændret til denne formulering:

A mental disorder is a syndrome characterized by clinically significant disturbance in an individual's cognition, emotion regulation, or behavior that reflects a dysfunction in the psychological, biological, or developmental processes underlying mental functioning. (APA 2013: Use of the manual, definition of a mental disorder)

Herudover anføres det, at forventelige og kulturelt acceptable reaktioner på dødsfald og andre belastende begivenheder ikke udgør en psykisk forstyrrelse, og ligeledes at afvigende adfærd og konflikter mellem individ og samfund kun kan defineres som psykisk forstyrrelse i fald der er tale om en individuel dysfunktion. Dermed søger man altså i definitionen af psykiske forstyrrelser at imødekomme både den antipsykiatriske kritik og kritikken ang. patologisering af normal lidelse, bl.a. fremført af Nikolas Rose. Det fastslås (ifølge Wakefield & First (2013b: 664) for første gang i DSM's historie), at alle de diagnostiske klasser må leve op til definitionen af mentale forstyrrelser (APA 2013: Use of the manual, definition of a mental disorder).

Mest iøjnefaldende er det at en forøget *risiko* for død, smerte, dysfunktion eller tab af frihed ikke længere med den nye definition kan diagnosticeres som en mental forstyrrelse. Derudover går man fra at tale om et 'adfærdsmæssigt eller psykologisk syndrom eller mønster' til 'syndrom karakteriseret ved forstyrrelse i henhold til kognition, emotion eller adfærd'. Man går fra at tale om 'distress' og 'disability' til 'disturbance' og 'dysfunction'. Endelig er der i den nye definition tale om at den psykiske forstyrrelse reflekterer en dysfunktion i psykologisk, biologisk eller udviklingsmæssige processer der er underliggende for den mentale funktionsevne, der er årsag til den klinisk signifikante forstyrrelse der udgør lidelsen.

Mens det er forholdsvis ligetil at undersøge hvorvidt patienten lider af forstyrrelse af det kognitive, emotionelle eller adfærdsmæssige normalplan (der formentligt ville kunne betegnes som 'distress' eller 'disability'), er det langt sværere at undersøge hvorvidt denne forstyrrelse skyldes en underliggende dysfunktion. Definitionen omtaler dysfunktioner i både psykologiske, biologiske og udviklingsmæssige processer. Påvisningen af en dysfunktion ville således kræve fuldt kendskab og adgang til vores psykologiske og biologiske processer, hvordan de forårsager funktioner og dermed hvordan dysfunktion afviger fra disse processer. Det betyder at det skal være muligt at skelne den funktion der er i overensstemmelse med normale biologiske og psykologiske mekanismer (herunder forskellige reaktioner på forskellige typer belastning) fra de dysfunktioner der markerer unormale processer. Hvis man havde dette kendskab ville klassifikationen formentlig basere sig på netop disse processer, altså psykiatriens patogenetik. Det er imidlertid ikke muligt at teste for psykologiske og neurobiologiske dysfunktioner (Kendell & Jablensky 2003: 7). I stedet er det op til klinikerens vurdering, på baggrund af symptomer og kontekst, hvorvidt der er tale om en dysfunktion (Wakefield & First 2013b: 665). Den ændrede definition ligner altså en tilnærmelse mod en patogenetisk sygdomsmodel, hvilket også ses i udskiftningen af 'distress' og 'disability' med 'disturbance' og 'dysfunction'. 'Distress' (bekymring) som kriterium kan kun valideres på baggrund af patientens sygdomsoplevelse, mens 'disability' både kan have en objektiv karakter hvis invaliditeten kan observeres, men også kan bestå i en subjektiv erfaring af at være ude af stand til fungere normalt. Derimod retter 'disturbance' (forstyrrelse) og 'dysfunction' (dysfunktion) sig mod objektive kriterier for sygdom, ved indirekte at henvise til en eller anden form for balance, fra hvilken der kan finde en forstyrrelse sted, eller en normal funktion som dysfunktionen afviger fra.

Begrebet 'mental disorder' dækker dermed over noget andet i den nye definition end i den gamle. Mental forstyrrelse er med DSM-5 ikke længere en betegnelse for nogle fænomener som er karakteriseret ved en række symptomer og tegn og som opleves som belastende og hæmmende. Begrebet dækker over en dysfunktion i de processer som er underliggende for den mentale funktionsevne, og som giver sig til udtryk i forstyrrelser af individets kognition, emotion eller adfærd. Afvigende adfærd og forstyrret emotionel eller kognitiv funktion er ikke i sig selv udtryk for mental forstyrrelse men afhængig af tilstedeværelsen af en underliggende dysfunktion.

Den objektive markør 'underliggende dysfunktion' som afgørende for sygdom eller ej i overensstemmelse med Ross' sygdomsbegreb, så længe at den underliggende dysfunktion også for individet opleves som dysfunktion. Afvigende adfærd der skyldes underliggende dysfunktion kan

således definatorisk betegnes psykisk forstyrrelse, men kan kun forstås som sygdom i Ross' forståelse, hvis adfærden også er en funktionsforstyrrelse for det pågældende individ. Altså må sygdom erfares som noget der er hæmmende for livsmuligheder og for psykiske sygdomme hæmmende for evnen til at kommunikere meningsfuldt.

Så længe det ikke er muligt at teste for biologisk og psykologisk dysfunktion, er det dog ikke klart hvilken funktion den ændrede definition vil have. Spørgsmålet er om ændringen afspejler sig i hvordan der kan klassificeres, eller om den alene afspejler en forventning til den fremtidige forsknings afdækning af neurobiologiske, genetiske og psykologiske kausalmekanismer der kan redegøre for at sådanne dysfunktioner rent faktisk eksisterer. Kravet om underliggende dysfunktion kan ses som et klart udtryk for et biomedicinsk sygdomsbegreb, der søger at definere sygdommen som svarende til en bestemt mekanistisk funktionsfejl.

6.4 Delkonklusion

Kapitel 5 og 6 har centreret sig om arbejdsspørgsmålet:

Hvordan konceptualiseres og defineres psykisk sygdom i de klassifikationssystemer?

En række udfordringer for valid klassificering i psykiatrien har vist sig. For det første kan psykiatrien (endnu) ikke hente validitet i et empirisk kendskab til patogenese og ætiologi, altså i den kausale mekanisme. Den psykiatriske klassifikation er nominalistisk og baserer sig dermed på symptomgrupper. Symptomgrupperne mener ikke at afspejle nogen underliggende årsager eller kausalforhold. Den rette differentiering af symptomerne anses for af at være tilstrækkeligt grundlag til opnå prognostisk homogene patientgrupper.

Det har (endnu) ikke været muligt at påvise 'zones of rarity' eller på anden måde validere grænserne for hver enkelt diagnostiske klasse. Dermed tyder det på at de klasser vi benytter ikke afspejler faktiske forskelle, og den prognostiske validitet kan dermed være kompromitteret, eftersom mange patienters symptombillede vil befinde sig mellem flere diagnoser.

Endelig har den nominalistiske model været kritiseret for at "medikalisere" personlig lidelse der ikke er patologisk, ved at definere afvigelse fra sociale eller psykologiske normer som medicinske problemer. Jeg har med Alf Ross' sygdomsbegreb argumenteret for at muligheden for at afgrænse patologiske psykiske lidelser fra ikke patologiske problemer må afhænge af at der findes en norm for normalplanet for mental funktion – i forhold til hvilket dysfunktionen kan forstås. ICD-10

definerer i overensstemmelse med Ross psykisk forstyrrelse som personlig dysfunktion. Definitionen inddrager imidlertid ingen betragtninger over hvad denne dysfunktion består i og hvad den må afvige fra.

Både WHO og APA forstår psykiske forstyrrelser som dysfunktion, og er således i overensstemmelse med Ross. Da DSM-5 imidlertid definerer psykisk forstyrrelse i relation til en underliggende sygdomsfunktion, får den karakter af et sygdomsentitet-begreb. Et sådant sygdomsbegreb er ikke på baggrund af vores aktuelle viden operationelt i forhold til at drage grænser mellem patologisk og ikke-patologisk lidelse, og kan dermed ikke møde den antipsykiatriske kritik.

Da definitionerne for psykisk forstyrrelse altså ikke sikrer sig mod problemet angående grænsen mellem normal og patologisk, må de diagnostiske kriterier der baserer sig på symptomer, altså være konstrueret således at symptomerne ikke med rimelighed kan antages at være udtryk for normal funktionsevne.

Kapitel 7: Diagnosen depression

I dette kapitel vil gennemgå de diagnostiske kriterier for diagnosen depression for i det følgende kapitel at diskutere hvorvidt diagnosen sikrer prognostisk homogenitet.

7.1 De diagnostiske kriterier

I dette afsnit vil jeg gennemgå de diagnostiske kriterier for depressionsdiagnosen som de er defineret i ICD-10. Desuden ser jeg på hvilke subkategorier man arbejder med, samt øvrige værktøjer der anvendes i diagnosticeringspraksis. Dette gør jeg for bagefter at kunne diskutere validiteten af diagnosen.

Depression er i ICD-10 klassificeret som en affektiv sindslidelse og i DSM-V under mood-disorders. ICD-10 findes både i en udgave tiltænkt den kliniske praksis, kaldet *Clinical Descriptions and Diagnostic Guidelines* fra 1992, og en udgave til forskningsøjemed, kaldet *Diagnostic Criteria for Research* fra 1993. De to udgaver er fuldt kompatible og varierer primært i deres introduktioner, og altså ikke ang. de diagnostiske kriterier. Jeg bruger her *Diagnostic Criteria for Research*, da de enkelte symptomer her er klart opstillet og defineret, og da det desuden er den udgave der må forventes at anvendes i forskningsforsøg og dermed evidensproduktion.

Fælles for de affektive sindslidelser er, at den fundamentale forstyrrelse er en forandring i stemningslejet sædvanligvis ledsaget af en ændring i aktivitetsniveauet. De fleste andre symptomer er forståelige i forbindelse med disse to ændringer, eller er sekundære hertil (WHO 1992: 112). I både ICD-10 og DSM-5 har man to diagnoser for depression alt efter om der er tale om gentagne depressive perioder eller blot en enkelt episode. De kaldes depressiv enkeltepisode og tilbagevendende (periodisk) depression. Kriterierne for depressiv enkeltepisode og tilbagevendende depression er ens, udover kriteriet om gentagelse. Den depressive episode er ofte relateret til stressende begivenheder eller situationer (WHO 1992: 112). Tidligere har man anvendt termen Unipolar depression for de to diagnoser, for at markere forskellen til bipolar sindslidelse, denne betegnelse er dog fjernet fra den seneste version af ICD-10 (Kragh-Sørensen & Kessing 2010: 361). Udover de to diagnoser for depression skelner man i ICD-10 mellem de affektive sindslidelser 'manisk episode', 'bipolar affektiv sindslidelse' (tidligere kaldet maniodepressivitet), 'vedblivende affektiv forstyrrelse', 'andre affektive forstyrrelser' og 'uspecificeret affektiv forstyrrelse' (WHO 1993: 13f)).

Der findes ikke et enkelt symptom der kan siges altid at være tilstede hos patienten med depression, i stedet opererer man i både ICD-10 og DSM-5 med grupper af symptomer der i forskellige kombinationer alle diagnosticeres som depression. Det betyder at depression kan fremstå klinisk meget forskelligt (Sundhedsstyrelsen 2007: 19).

Depressiv enkeltepisode og periodisk depression underinddeles i sværhedsgraderne let, moderat og svær ud fra antallet af symptomer der er tilstede. I DSM-IV anvendes diagnosen Major Depressive Disorder, som svarer til en mindst moderat depression efter ICD-10 (Kragh-Sørensen & Kessing 2010: 367).

En depressiv episode er defineret ved en række kriterier der skal være opfyldt. For det første skal der være tale om en varighed på mindst to uger. For det andet skal mindst to ud af tre følgende kernesymptomer være tilstede: 1. Nedtrykthed, 2. Nedsat lyst og interesse, 3. Nedsat energi og øget træthedsgrad. For det tredje knytter der sig en række ledsagesymptomer der er væsentlige for at bestemme sværhedsgraden af depressionen. Disse er: 1. Nedsat selvtillid eller selvfølelse, 2. Selvbeprejelser eller skyldfølelse, 3. Tanker om død eller selvmord eller suicidal adfærd, 4. Tænke- eller koncentrationsbesvær, 5. Agitation eller hæmning (øget eller nedsat psykomotorisk aktivitet), 6. Søvnforstyrrelser, 7. Appetit- eller vægtændring. (WHO 1993: 83)

Ved let depression er mindst 2 kernesymptomer tilstede mindst 2 ledsagesymptomer. Ved moderat depression er mindst 2 kernesymptomer og mindst 4 ledsagesymptomer. Ved svær depression er alle 3 kernesymptomer og mindst 5 ledsagesymptomer tilstede.

7.2 Subkategorier

Ved alle sværhedsgrader kan der desuden optræde en subtype, kaldet somatisk eller melankoliformt syndrom, hvor mindst 4 af følgende symptomer er tilstede: 1. Nedsat lyst eller interesse, 2. Svækket emotionel reaktivitet, 3. Tidlig opvågning (>2 timer før vanligt), 4. Morgenforværring, 5. Hæmning eller agitation, 6. Nedsat appetit, 7. Vægttab (>5% af vægten over en måned), 8. Nedsat libido (WHO 1992: 120). I DSM findes der en meget lig subtype under betegnelsen 'melancholia' (WHO 1993: 82).

Endvidere opererer man ved subtypen 'svær depression med psykotiske symptomer', hvor der i tillæg til kriterierne for svær depression også forekommer vrangforestillinger, hallucinationer eller stupor, der er en tilstand hvor patienten helt ophører med at kommunikere med omverdenen. Der

skelnes her desuden mellem kongruente og inkongruente psykotiske symptomer, dvs. henholdsvis vrangforestillinger og hallucinationer af affektiv karakter og uden affektiv karakter (WHO 1993: 84).

Endnu en subtype betegnes blot ”anden depressiv episode”. Denne subtype dækker over tilfælde der ikke passer beskrivelsen af depressiv enkeltepisode, men hvor det samlede diagnostiske udtryk ikke desto mindre indikerer at der er tale om depression. Det kan være tilfældet ved tilstedeværelsen af en række ikke-diagnostiske kriterier:

Examples include fluctuating mixtures of depressive symptoms (particularly those of the somatic syndrome) with non-diagnostic symptoms such as tension, worry, and distress, and mixtures of somatic depressive symptoms with persistent pain or fatigue not due to organic causes (WHO 1993: 85).

Diagnosen depression bygger altså på et specifikt kriterium angående varighed (2 uger) samt et minimum af mindst 4 af de associerede symptomer. Diagnosen kan specificeres i forhold til forløb (enkelt eller gentaget) og sværhedsgrad (antallet af associerede symptomer). Ud over disse subtyper der adskiller sig fra hinanden ved kvantitativ specifikation, opererer man med 3 kvalitativt distinkte subkategorier af depression, nemlig: svær depression med psykotiske symptomer, somatisk syndrom og anden depressiv episode, hvor der optræder symptomer der ikke normalt forbindes med depression.

7.3 Diagnostiske redskaber

I forbindelse med diagnostik anvender man ofte at semistrukturerede diagnostiske interviews eller spørgeskemaer til at spore sig ind på om ovenstående kriterier er tilstede.

I den kliniske vejledning for almen praksis for unipolar depression anbefales det at gøre brug af ICD-10 kriterierne i forbindelse med fastlæggelse af diagnosen, det fastslås endvidere at ”Depressionsdiagnosen stilles således udelukkende ud fra patientens symptomer, uden anden prioritering eller vægtning end den, patienten selv anfører, og helt uden hensyntagen til mulige årsager eller udløsende faktorer.” (Dansk Selskab for Almen Medicin 2010: 11) Som redskab i diagnosticeringen anbefaler både Dansk selskab for Almen Medicin og Sundhedsstyrelsen anvendelsen af spørgeskemaet MDI (Major Depression Inventory) hvor patienten bedes angive sværhedsgraden af 10 forskellige symptomer på en skala fra 1 til 5 (bilag 1). MDI kan også bruges til at bedømme sværhedsgraden af depressionen i form af den samlede pointscore. Begge guidelines anbefaler desuden Hamiltons depressionsskala (HAMD) (bilag 2) til yderligere vurdering af

sværhedsgraden. Et andet redskab til vurdering af sværhedsgraden er Montgomery & Åsberg Depression Rating Scale (MADRS) (bilag 3), som også nævnes i Sundhedsstyrelsens referenceprogram og som vi så anvendt i det ekseptificerede RCT. Kendskab til sværhedsgraden er vigtig både af hensyn til valg af behandling og af hensyn til at kunne følge sygdomsudviklingen (Dansk Selskab for Almen Medicin 2010: 11, Sundhedsstyrelsen 2007: 13).

Med Hamiltons depressionsskala rates 17 på forhånd definerede symptomer fra 0 til 2 for nogle symptomer og fra 0 til 4 for andre. Denne forskel gør det muligt at tillægge de mest centrale symptomer mere vægt i den samlede vurdering af diagnosens sværhedsgrad der bestemmes ud fra den samlede pointscore. Max-antallet er 52 point og en score på mellem 8 og 12 point defineres som tvivlsom depression, mellem 13 og 17 point defineres som let depression, mellem 18 og 24 som moderat depression, mellem 25 og 29 point som middelsvær depression og fra 30 og opefter som svær depression (Kragh-Sørensen & Kessing 2010: 386).

På Montgomery & Åsberg skalaen vurderes depressionen kun på 10 symptomer fra 0-6. En score på 11 eller derunder anses for remission (ingen depression), 18-21 er lig let depression, 22-29 er moderat depression og 30 og derover er svær depression (Beck et al 2005: 55).

Symptomerne der anvendes på HAMD og MADRS er er ret forskellige (se bilag), således ses det at der optræder betydeligt flere somatiske symptomer på HAMD end på MADRS.

Kapitel 8: Analyse af depressionsdiagnosen

I dette kapitel vil jeg analysere depressionsdiagnosen med henblik på at vurdere muligheden for prognostisk valide og homogene patientgrupper og dermed den diagnostiske kategori som udgangspunkt for den empiriske forskning.

Som jeg tidligere har gjort rede for er det ikke påvist ved hjælp af 'points of rarity' at diagnoserne afspejler naturligt forekommende afgrænsede patientgrupper, og mange mener at vi i stedet for at betragte psykiske lidelser som typer, der dermed former klasser, bør betragte psykiske symptomer og lidelser som kontinuerte. Hvis det er tilfældet at dette billede i højere grad svarer til virkelighedens patienter, er grænserne mellem diagnoserne mere eller mindre vilkårlige. Derfor kan det dog alligevel have en praktisk funktion at opdele lidelserne i klasser, selvom vi må være opmærksomme på at diagnoserne i højere grad vil passe på nogle patienter end andre. Det kan således være at patienter jo mere deres lidelse er knyttet til et bestemt symptom, har mere fordel af en bestemt behandling. Kendell og Jablensky taler derfor også i deres kritik af diagnosernes validitet om, at vi netop må adskille validiteten fra brugbarheden. I dette kapitel behandler jeg nogle forskellige spørgsmål der relaterer sig til diagnosernes brugbarhed.

8.1 Spørgsmålet om dysfunktion

I min analyse af de to definitioner af 'mental disorder' hos hhv. DSM-5 og ICD-10, skelnede jeg mellem 'personlig dysfunktion' om individets og eventuelt pårørendes oplevelse af nedsat funktionalitet, og 'underliggende dysfunktion' om den funktionsnedsættelse der skyldes en underliggende patologisk komponent, det være sig biologisk, psykologisk eller indlæringsmæssigt. Som jeg også gjorde rede for er begrebet 'underliggende dysfunktion' som definerende for hvad der udgør psykisk forstyrrelse, skønt et stærkt kriterium, dog ikke operationelt i praksis.

En personlig dysfunktion kan naturligvis godt være en underliggende dysfunktion, men den kan også være en ikke-patologisk og dermed "normal" psykisk reaktion, der altså ikke lever op til Ross' definition af psykisk sygdom i og med at dysfunktion er forståelig i den kontekst den optræder.

Kriterierne for depression som de beskrives i diagnosen lever op til 'personlig dysfunktion'. Kernesymptomerne 'nedtrykthed' og 'nedsat lyst og interesse' kan karakteriseres som en dysfunktion i den forstand at individet ikke oplever at være "sig selv", mens det tredje

kernesymptom 'nedsat energi og øget træthedsbarhed', må anses for en dysfunktion ift. et normalt aktivitetsniveau der sætter individet ude af stand til at gøre ting han/hun ellers ville kunne.

For alle tre kernesymptomer synes det desuden rimeligt at arbejde med en tærskel for at tale om personlig dysfunktion, i tråd med Ross' kriterium for væsentlig afvigelse fra normalplan. Ingen af kernesymptomerne kan siges i sig selv nødvendigvis at være udtryk for en personlig dysfunktion. Fx er det indlysende at 'nedtrykthed' ikke nødvendigvis er udtryk for dysfunktion. Det er derfor nødvendigt med en tærskel for hvornår nedtryktheden er 'sygelig', diagnoseværdig, eller behandlingskrævende. En sådan tærskel for personlig dysfunktion er afhængig af individets oplevelse af symptomet, men må også være afhængig af sværhedsgraden og varigheden af symptomet.

Således fungerer minimumsvarigheden på 2 uger altså som tærskel for hvornår et symptom er rimeligt at fortolke som en personlig dysfunktion. Det kan naturligvis diskuteres hvorvidt denne grænse er rimelig, men en sådan grænse vil som Ross siger, være afhængig af den kulturelle og samfundsmæssige kontekst.

Ligeledes fungerer brug af diagnoseredskaber som en tærskel for sværhedsgraden af de oplevede symptomer. Det er således muligt at udvise symptomerne på depression i antal, men i så let en grad at den samlede estimation af symptomernes sværhedsgrad, ikke kvalificerer til at der kan stilles diagnose, og derved har man indført en tærskel angående sværhedsgrad for tilstedeværelsen af personlig dysfunktion.

I forhold til Ross' sygdomsbegreb kvalificerer den ikke-underliggende dysfunktion sig ikke til sygdom. For at personlig dysfunktion kan siges at være sygdom, må der defineres en norm for hvad der forstås ved mental dysfunktion og mentalt funktionelt normalplan.

8.2 Kriteriernes karakter

Som jeg var inde på i kapitel 4 ville den ideelle klassifikation basere sig på nogle specifikke patologiske kriterier der udgjorde de nødvendige og tilstrækkelige faktorer for at diagnosen kan stilles. Dermed ville man sikre klare grænser for hvilke tilstande der skal inkluderes i diagnosen, og patientgruppen ville være homogen på de variable, man anså for de definerende karakteristika. Som jeg også gjorde rede for, er det dog vanskeligt at arbejde med et sæt af kriterier der er både nødvendige og tilstrækkelige for mange somatiske sygdomme. Dette er også tilfældet for

depression. Diagnosen er karakteriseret ved ikke at have nogen definerende kriterier, dvs. der er ikke nogle nødvendige kriterier der altid skal være til stede for at diagnosen kan stilles. I stedet er der en gruppe af 10 symptomer og et krav om at patienten udviser mindst 4. Det betyder at patienter med forskellige symptomer kan diagnosticeres ens. Ved at tillade meget forskellige kombinationer af symptomer, kommer diagnosen til at dække over mange forskellige sygdomsbilleder. For en patient kan depression fx give sig til udtryk i manglende koncentrationsevne, rastløshed og søvnforstyrrelser, og hos en anden kan den give sig til udtryk i nedsat selvfølelse, selvbefredelser, skyldfølelse og selvmordstanker. Når disse forskellige lidelser diagnosticeres under én betegnelse, er det ikke fordi de nødvendigvis menes at være ”udtryk for det samme”. Diagnosen depression er netop nominalistisk defineret og gør sig dermed ikke antagelser om lidelsernes årsager, det væsentlige for klassernes legitimitet er, jævnfør Goodwin og Guze, kommunikation og forudsigelse, med andre ord, diagnosens reliabilitet og prognostiske validitet.

Hvis vi igen skæver til den somatiske medicin, findes muligheden for forskelligartede symptombilleder som udtryk for samme diagnose også her. Men baggrunden for at betegne disse fænomener under ét er en indsigt i symptomernes underliggende årsager, typisk en patogenetisk faktor der er væsentlig for muligheden af at behandle sygdommen. Dvs. at det forhold at sygdomstilfældene ikke udgør en naturlig klasse på symptomniveau, forsvares med at de udgør en naturlig klasse på patogenetisk niveau. I den psykiatriske klassifikation er diagnoserne imidlertid ikke begrundet i en teori om sygdommens kausalitet, de er konstruerede på baggrund af observationer af symptomernes forekomst.

Alle tre kernesymptomer samt de fire ud af syv ledsagesymptomer er karakteriseret ved at referere til patientens oplevelse af sin tilstand. Psykomotorisk hæmning eller agitation er, hvis symptomet er tilstrækkeligt udtrykt, mulige for læge eller pårørende at registrere, ligesom også søvnforstyrrelser og appetit- eller vægtændring, på trods af at være afhængige af patientens opmærksomhed, oplevelse og rapportering har en mere objektiv karakter. De øvrige symptomer er ikke mulige at konstatere gennem observation, de er udelukkende af subjektiv karakter. For at kunne stille diagnosen depression er klinikerens nødt til at prøve at afkode den subjektive oplevelse som patienten udtrykker. Diagnosticeringen er altså afhængig af at patienten sætter ord på hvordan hun/han har det, hvorefter lægen afkoder patientens beretning og sammenholder den med de diagnostiske kriterier for depression. I den proces hvor patientens oplevelse af sin sindstilstand og funktionsevne, lidelsens betydning og årsag bliver til en diagnose af sværhedsgraden let, moderat

eller svær, må lægen nødvendigvis foretage en fortolkning. Lægen lægger mærke til bestemte udtalelser, tillægger dem bestemte betydninger, og stiller bestemte spørgsmål der leder patientens beretning. Det sker i overensstemmelse med anbefaling for diagnosticeringspraksis der fx kan ske på baggrund af Major Depression Index, hvor der findes et spørgsmål der svarer til hver af de 10 symptomer i diagnosen. Diagnosticeringspraksissen har altså på forhånd udpeget de væsentlige faktorer i patientens beretning og patienten kan enten bekræfte eller afkræfte lægens hypotese om en depression.

Patientens egen udlægning eller fortolkning af sin tilstand forsvinder derfor i diagnosticeringen af patientens tilstand. Gennem diagnosticeringsprocessen omformes oplevelse, holdning, erfaring og historie gennem lægens fortolkning til diagnosen depression af en bestemt sværhedsgrad baseret på hvor mange af de på forhånd bestemte kriterier, patienten møder. Patientens beretning skal indordnes under de standardiserede kriterier og der opnås dermed begrænset indsigt i patientens egentlige oplevelse af sygdom. Resultatet af den diagnostiske proces er en diagnose og et enkelt tal på MDI eller Hamiltons Depressions skala.

8.3 Depressionsskalaerne MDI, HAMD og MADRS

De tre skemaer til vurdering af sværhedsgrad er ret forskellige, Major Depression Inventory (MDI) er bygget op om de 10 ICD-kriterier med ét spørgsmål for hvert kriterie, mens Hamiltons depressionsskala (HAMD) og Montgomery-Åsberg skalaen (MADRS) også inddrager andre symptomer. MDI vægter symptomerne efter hvor meget af tiden de er tilstede, HAMD vægter dem efter hvor udtalte de er og MADRS blander de to kriterier i dens beskrivelse af symptomerne. MDI og MADRS er bygget op således at alle kriterierne vægtes ens, således at selvmordstanker vægtes lige så højt som søvnforstyrrelser og kernesymptomet nedsat stemningsleje vægtes lige så højt som nedsat eller øget appetit i vurderingen af depressionens sværhedsgrad. Dog kræves det i MDI som jo også er et diagnoseredskab (til forskel fra redskab til vurdering af sværhedsgrad), for at diagnosen depression kan stilles at kernesymptomerne er meget udtalte, dvs. tilstede over halvdelen af tiden, efter patientens vurdering.

HAMD anvender en vis differentiering mellem symptomer der anses for mere eller mindre centrale for diagnosen, således at nogle symptomer kan vægtes højere end andre. HAMD er desuden meget orienteret om psyko-somatiske symptomer, således er der 3 spørgsmål der omhandler søvn, og herudover gastrointestinale symptomer, seksuelle forstyrrelser, vægttab, og somatiske symptomer

generelt. HAMD omfatter også en række symptomer der ikke findes under ICD-kriterierne, det gælder psykiske angst-komponenter, somatiske angst-komponenter, hypokondri og manglende sygdomsindsigt. Disse symptomer tegner sig for 14 point ud af de 52 mulige og de psykosomatiske symptomer for yderligere 14 point. HAMD er ikke et diagnoseredskab, men anvendes alene til at vurdere sværhedsgrad. Diagnosen stilles på ICD-10 kriterierne der er af mere subjektiv karakter. Mens kriterierne for diagnosen altså er den subjektive oplevelse, kan sværhedsgraden af sygdommen afhænge af om den også giver sig til udtryk i somatiske, objektive eller angst-symptomer og manglende realitetssans.

Man kan forestille sig en patient der lider af alle de depressive kernesymptomer (nedsat stemningsleje, nedsat lyst og interesse, nedsat energi og øget træthedsgrad) i svær grad og er stærkt invalideret af det. Derudover lider hun også af søvnforstyrrelser af forskellig karakter. Som konsekvens af denne tilstand oplever hun også koncentrationsbesvær i middelsvær grad samt let psykomotorisk hæmning. Der er altså 3 kernesymptomer og 3 ledsagesymptomer tilstede, hvilket ifølge ICD-10 kriterierne betyder at hun har depression af let sværhedsgrad. MDI-skalaen anvendes i diagnosticeringen hvilket giver en score på 26 (ud af 40), svarende til moderat depression (symptomerne: 1, 2, 3 og 9 = 5 + symptom 7, 8b = 3, i alt 26). Efterfølgende bruges Hamiltons depressionsskala til yderligere at fastslå sværhedsgraden. Her opfylder hun kriterierne: nedsat stemningsleje (score 4), indsovningsbesvær (score 2), afbrudt søvn (score 2), tidlig morgenopvågning (score 2), arbejde og interesser (score 2) og psykomotorisk hæmning (score 2). I alt har hun en score på 14 (ud af 52) svarende til let depression.

Ser man til gengæld på MADRS kan patienten diagnosticeres med svær depression: Hun har stærkt forsinket stemningsleje, både observeret og rapporteret (score 12), udtalt søvnbesvær (score 6), koncentrationsbesvær i middelsvær grad (score 2), initiativløshed (score 2), svækket følelsesmæssigt engagement (score 4) og depressivt tankeindhold (score 4). I alt når hun op på 30, lige akkurat svarende til svær depression.

På HAMD reduceres patientens primære problem, den stærkt depressive sindstilstand, til et enkelt symptom (nedsat stemningsleje) der tæller 4 point. Til gengæld vejer hendes omfattende søvnproblemer højere (i alt 6 point). Hvis man havde nøjedes med at bruge MDI ville de depressive kernesymptomer tegne sig for 15 point og søvnproblemer for 5 point. Hvis man i stedet havde brugt MADRS havde kernesymptomerne talt for 20 point i sig selv, altså svarende til let depression, mens søvnbesvær kun tegner sig for 6 point. Denne forskel får betydning når man skal måle på i hvilken

grad en behandlingsindsats har virket. En behandling der effektivt stopper patientens søvnproblemer vil ift. HAMD have reduceret depression til en score på 8. Alene ved at behandle søvnproblemerne kan der altså næsten opnås remission, uagtet at de depressive kernesymptomer stadig er tilstede og patienten stadig oplever dårlig livskvalitet og nedsat funktionsevne. Anvendes MDI ville reduktionen af depressive symptomer kun være 5 ud af 26 ved at behandle søvnproblemerne, og på MADRS ville patienten stadig kunne diagnosticeres med depression af moderat sværhedsgrad med en behandling der kun rettede sig mod søvnproblemer. Forskellene mellem de tre skalaer betyder altså i dette tilfælde at reduktion af depressive symptomer kan måles både som en reduktion på 43% og på 19%.

Skalaen udpeger og sætter grænser for hvilke symptomer der er væsentlige og i hvilken grad de er væsentlige. Det betyder også at antallet af forskellige symptomer bliver af stor betydning. Det kan således let blive af større betydning at have en række symptomer af let sværhedsgrad end at have fx skyldfølelse i en sådan grad så det udgør vrangforestillinger, eller at være ude af stand til at udføre selv de mest rutineprægede aktiviteter. Behovet for numerisk at måle sværhedsgraden af depressionen er altså langt fra en neutral praksis. Med et standardspørgeskema kan man ikke sikre en nuanceret gengivelse af så forskelligartede tilstande som går under betegnelsen depression.

Efter således at være kommet frem til at diagnosen depression omfatter forskelligartede lidelser i kraft af manglen på nødvendige kriterier, og at disse tilstande også kan gengives meget forskelligt i kraft af redskaberne for diagnose, vil jeg nu undersøge subkategorierne under depression for at kunne vurdere om de giver mulighed for at opnå homogene patientgrupper til brug i RCT-forsøg.

8.4 Subkategorier

Den diagnostiske kategori depression optræder i ICD-10 som et enkelt syndrom der varierer dimensionalt over sværhedsgrad, varighed og hyppighed. Dvs. der ligger altså en formodning om at trods varians i depressionens styrke, er der tale om det samme grundlæggende syndrom, samme problem. Samtidig har klassifikationen også en vis subkategorisering i distinkte udtryk af syndromet i form af somatisk syndrom, svær depression med psykotiske symptomer og anden depression (atypiske symptomer). Der er altså varianter af sygdomsenheden, som det anses for relevant at betegne med en særlig subkategori, og dermed må antages at være kvalitativt forskellige fra depression generelt.

Somatisk syndrom

I den tidligere version af klassifikationen, ICD-9, fandtes to kvalitativt forskellige typer af depression, nemlig den endogene (indefrakommende) og den psykogene/reaktive (udefrakommende). Den endogene depression blev også sat i forbindelse med melankoli, der stadig findes som diagnose i DSM-5, og som i ICD-10 er blevet til somatisk syndrom. Den endogene depression kan altså siges at være den umotiverede, mens den psykogene opstår som reaktion på udefrakommende omstændigheder, og dermed er motiveret. Ifølge Kragh-Sørensen og Kessing har skellet mellem endogen og psykogen depression altid været uklart defineret, og de er derfor taget ud af brug i de aktuelle udgaver af ICD og DSM. Skellet var især problematisk, da der ved gentagne depressioner hos patienter først diagnosticeret med psykogen depression, ofte har skullet mindre og mindre hændelser til at udløse depressionen. Dvs. at den psykogen diagnosticerede patient gradvist blev mere og mere endogen (2010: 381). Skellet udmærker sig desuden ved at være ætiologisk, og dermed bryder det med den rent symptombaserede diagnosticering.

På trods af at det ætiologiske skel er fjernet og de to typer er samlet i én diagnose, findes skellet altså i form af muligheden for at specificere diagnosen, uanset sværhedsgrad, ved anvendelsen af betegnelsen somatisk syndrom. Den afgørende forskel er naturligvis at det somatiske syndrom defineres ift. nogle specifikke symptomer der menes at have en særlig klinisk relevans, og som af den årsag har sin egen subkategori, og ikke som tidligere på grund af årsagen.

Hvis subkategorien rent faktisk afspejler en kvalitativ forskel på to typer af depression, er det muligt at denne forskel har rødder i en specifik ætologi, eller kan udgøre et grundlag for specifik prognostik og behandlingsform. Man kunne fx forestille sig at patienter diagnosticeret med det somatiske syndrom, i højere grad responderede på medicinske eller andre fysiske behandlingsformer som elektrochok. For at kunne diagnosticeres med depression med somatisk syndrom må man leve op til kriterierne for både depression og somatisk syndrom. Kriterierne for somatisk syndrom er lidt mere specifikke end de generelle kriterier for depression. Fx findes der to distinkte symptomer for 'nedsat appetit' og 'vægttab (>5% kropsvægt)'. Når man ser på kriterierne for somatisk syndrom er der dog en del overlap mellem hvad der kvalificerer til depression uden somatisk syndrom og med somatisk syndrom.

Vi ser i henholdsvis 'depression uden somatisk syndrom' og 'depression med somatisk syndrom':
1. 'Nedsat lyst og interesse' vs. 'Nedsat lyst og interesse' 2. 'Søvnforstyrrelser' vs. 'Tidlig opvågning' (>2 timer før vanligt), 3. 'Agitation eller hæmning' vs. 'Hæmning eller Agitation,' 4.

'Appetit eller vægtændring' vs. 'Nedsat appetit', 5. 'Appetit eller vægtændring' vs. 'Vægttab (> 5% kropsvægt)'.

Det er således kun 3 kriterier nemlig morgenforværring, nedsat libido og svækket emotionel reaktion, der er specifikke for somatisk syndrom, hvoraf det sidstnævnte 'svækket emotionel reaktion' om ikke andet, så for det utrænede medicinske øje, måske kunne forveksles med 'nedsat lyst og interesse' eller 'hæmning'. I og med at kun 4 symptomer skal være tilstede for at diagnosen depression med somatisk syndrom kan stilles, kan de symptomer der kvalificerer den generelle diagnose depression, altså være de samme symptomer der kvalificerer det somatiske syndrom.

Konsekvensen af det store overlap af symptomer er, at distinktionen mellem somatisk syndrom og intet somatisk syndrom kan hvile på tilstedeværelsen af et enkelt symptom der vel at mærke ikke er et nødvendigt kriterium eller et specifikt kriterium for subkategorien. Tag fx to patienter der begge oplever symptomerne: nedsat lyst og interesse, nedsat energi og øget træthæd, nedsat selvtillid eller selvfølelse, hæmning, koncentrationsbesvær og nedsat appetit. En af dem oplever også vægttab. De kan begge diagnosticeres med moderat depression, men den ene patient kan endvidere diagnosticeres med somatisk syndrom. I stedet for vægttab, kunne den i dette tilfælde afgørende faktor for diagnosen somatisk syndrom også have været søvnforstyrrelse, svækket emotionel reaktivitet, morgenforværring eller nedsat libido.

I introduktionen til kapitlet om affektive forstyrrelser fremgår usikkerheden omkring somatisk syndroms status også: "The classification is arranged so that this somatic syndrome can be recorded by those who so wish, but can also be ignored without loss of any other information." (WHO 1992: 112)

Brugen af subkategorien somatisk syndrom er valgfri hvilket forklarer hvorfor de fleste af symptomerne er overlappende. Men derved bliver spørgsmålet, hvilken funktion subkategorien tjener, når man hverken har bestemt sig for eller imod realiteten eller anvendeligheden af dette syndrom. Subkategorien tjener ikke til at markere ætiologiske forskelle som skellet mellem endogen og psykogen depression gjorde, og den udskiller ikke syndromet som et specifikt klinisk billede, eftersom fraværet af syndromet alligevel kan indebære mange af de samme karakteristika som er definerende for syndromet. I og med at kriterierne er overlappende, virker subkategorien ikke så meget som en specifik kategori, men snarere som en dimension for depressionens somatiske udtryk, der kan variere gradvist. Der er altså muligvis basis for at skabe en patientgruppe der har en højere

grad af homogenitet og dermed måske højere prognostisk validitet, men man har ikke indført et klart skel mellem subgruppen og hovedgruppen, hvormed forskning i egentlige forskelle i behandlingsrespons og altså afdækning af den prognostiske validitet vanskeliggøres.

Subkategorien 'Anden depression'

Subkategorien 'anden depression' tjener til at udvide rammerne for hvad der kan diagnosticeres som depression. De fremhævede symptomer er karakteriseret ved at være af udpræget reaktiv karakter, i form af "tension, worry, and distress" i forbindelse med vekslende sammensætninger af depressive symptomer, særligt de somatiske, samt træthed og smerte uden biologisk årsag.

'Tension', 'worry' og 'distress' er symptomer hvis patologiske status må siges at være afhængige af konteksten de optræder i. Som almindelige psykologiske funktioner synes det her afgørende at de optræder enten umotiverede eller som urimeligt motiverede jf. Ross. Der indgår dog ingen krav i diagnosen der sikrer at symptomerne er udtryk for dysfunktion.

Subkategorien 'Svær depression med psykotiske symptomer'

Subkategorien svær depression med psykotiske symptomer er formentligt den bedst afgrænsede og mest operationelle subkategori under depression. Den er defineret ved et kriterium der er specifikt for subkategorien, og således ikke også kan tilskrives almindelig depression. Dette kriterium er tilstedeværelsen af hallucinationer, vrangforestillinger eller depressiv stupor. Symptomerne er tydelige og adskiller sig fra den overordnede depressionskategoris symptomer ved, i sig selv at være at betragte som patologiske. Dermed er det ikke blot i kraft af en række symptomer der ikke i sig selv kan betragtes som udtryk for psykisk dysfunktion at syndromet defineres. Hallucinationer, vrangforestillinger og stupor må betragtes som afvigelser fra et mentalt funktionelt normalplan defineret ved evnen til at kommunikere. Stupor er defineret som en tilstand hvor patienten slet ikke kommunikerer med omverdenen, og hallucinationer og vrangforestillinger hindrer netop meningsfuld kommunikation, da det viser sig i en uenighed om hvad der udgør virkeligheden.

8.5 Delkonklusion

Kapitel 7 og 8 har centreret sig om arbejdsspørgsmålet:

Hvordan er diagnosen depression defineret og i hvilken grad sikrer den prognostisk validitet?

Diagnosen er ikke i stand til at skelne patologisk lidelse, dvs. sygdom, fra ikke-patologisk lidelse, hvis vi forholder os til Ross' sygdomsbegreb. Diagnosen identificerer personlig dysfunktion og lever således op til ICD-10's egen definition af mental forstyrrelse, men da de fleste af symptomerne ikke i sig selv kan anses for patologiske og diagnosen ikke forholder sig til en norm for mental funktionel normalplan, findes der ikke nogen ramme for hvornår symptomerne med rimelighed kan antages at være udtryk for dysfunktion. Dermed slår diagnosen potentielt sygdom og ikke patologiske symptom mønstre sammen i kategorien depression.

Diagnosen depression har ingen specifikke kriterier, men er defineret ved en gruppe symptomer. Dette giver et utal af kombinationsmuligheder der lever op til de diagnostiske kriterier. Diagnosen dækker således over meget forskellige kliniske sygdomsbilleder.

De tre kernesymptomer for depression samt 4 ud af 7 ledsagesymptomer er karakteriseret ved kun at være subjektivt tilgængelige. Det betyder at diagnosticering af depression er en fortolkende praksis, hvor patientens sygdomsoplevelse søges forstået af lægen. Lægen må dog fortolke patientens udtalelser i forhold til de fastlagte symptomer der er definerende for diagnosen. Denne fortolkning oversættes videre til en numerisk værdi på en depressionsskala i vurderingen af depressionens sværhedsgrad.

Skalaen der benyttes til at bestemme sværhedsgraden er af stor betydning for hvilke symptomer der vægtes, og samme tilstand kan derfor diagnosticeres forskelligt alt efter skala. Skalaerne kan således også have betydning i RCT der tester behandlingseffekter.

Subkategorien svær depression med psykotiske symptomer er veldefineret i kraft af specifikke patologiske kriterier. De øvrige subkategorier, 'somatisk syndrom' og 'anden depression', er ikke tilstrækkeligt selvstændigt defineret ift. den generelle diagnose, til at det er muligt i forskningen at undersøge hvorvidt der er tale om en specifik ætiologi eller behandlingsrespons.

Kapitel 9: Diskussion

Jeg har gennem specialet arbejdet med mine tre arbejdsspørgsmål. Jeg vil i dette afsluttende kapitel forsøge at samle trådene fra mine analysedele og diskutere implikationerne af et RCT-baseret evidensbegreb til brug i vidensproduktion om depressionsbehandling.

Jeg starter i en central delkonklusion fra kapitel 4: En forudsætning for relevansen af RCT-genereret evidens er en homogen patientpopulation mht. behandlingsrespons. Det er det jeg også betegner prognostisk validitet. Da RCT oftest baserer sig på den diagnostiske klasse generelt, er det diagnosens prognostiske validitet der er afgørende for hvorvidt man kan generalisere evidensens resultater til at gælde patienter med depression som samlet gruppe. Spørgsmålet om prognostisk validitet har været en rettesnor gennem de følgende kapitler og jeg starter derfor med at samle op på de foreløbige resultater af undersøgelsen.

9.1 Homogenitet og prognostisk validitet

At den prognostiske validitet af patientpopulationen der deler diagnosen depression, ikke er høj, indikeres i eksemplet på RCT og review, i og med at patienterne netop ikke reagerer ens på behandlingen. I RCT'et af Wade et al. mødte 45% af forsøgspopulationen i indsatsgruppen ikke succeskriteriet på minimum en halvering af score på Montgomery Åsbergskalaen og i reviewet oplevede gennemsnitligt 40% ingen respons. 75-79% af de patienter der oplevede effekt af behandlingen kunne lige så godt have modtaget en placebo-indsats, hvilket reducerer den reelle effekt af behandlingen til 13% af forsøgspopulationen. Disse resultater er ikke enestående (Gøtzsche 2014: 8). Ifølge Sundhedsstyrelsen og flere kritikere viser studierne endog ofte bedre resultater end der kan opnås i den kliniske praksis, som følge af inklusions- og eksklusionskriterier og generelt kompromitteret ekstern validitet (afsnit 4.2). Det er dog ikke muligt at slutte direkte den relativt lille effekt til patientgrundlagets manglende homogenitet. For at undersøge den prognostiske validitet af den diagnostiske klasse har jeg derfor set på grundlaget for psykiatrisk klassifikation. Mens man i den somatiske medicin sikrer prognostisk validitet ved at basere sig på patogenetisk definerede diagnoser, er dette endnu ikke en mulighed i psykiatrien. Den psykiatriske klassifikation kan dog forsøge at sikre sig at diagnosernes kriterier opdeler patienterne således at det svarer til faktisk forekommende distinkte patientgrupper gennem påvisningen af 'zones of rarity'. Hvis den diagnostiske klasse netop omfatter hvad der svarer til en naturligt forekommende gruppe af sygdomstilfælde, er det sandsynligt at diagnosen afspejler nogle reelle underliggende forskelle der

forårsager de forskellige psykiske symptombilleder. Dermed har man et godt udgangspunkt for forskning i behandling, da man ved at patienternes tilstand sandsynligvis vil kunne bedres ved at identificere denne mekanisme og målrette behandlingen denne. Det er imidlertid endnu ikke lykkedes at påvise nogen 'zone of rarity' mellem depressionsdiagnosen og andre diagnostiske klasser eller mellem depression og ingen depression. Det indikerer at diagnosen ikke svarer til en distinkt sygdomsmekanisme eller 'sygdomsenhed'. Det at der ikke kan påvises 'zones of rarity' har også konsekvenser på symptomplan. I og med at mange sygdomstilfælde ikke entydigt befinder sig indenfor én klasse, men viser symptomer fra flere, vil det være umuligt på baggrund af de eksisterende diagnostiske klasser at opnå homogenitet på symptomniveau. Desuden giver de diagnostiske kriterier for depression mulighed for vidt forskellige sygdomsbilleder. Det betyder at patienter diagnosticeret med depression ikke nødvendigvis har ret meget til fælles med andre med samme diagnose, mens nogle til gengæld vil have en del tilfælles med patienter med andre diagnoser. Endelig koncentrerer diagnosen sig om bestemte symptomer, nemlig de der beskrives i de diagnostiske kriterier. Patienten kan således også have andre symptomer, der ikke tegner et symptombillede der passer til en diagnose. Dermed dækker diagnosen ikke nødvendigvis over patientens samlede symptombillede, hvilket giver mulighed for yderligere variation. På denne baggrund tegner mulighederne for en prognostisk homogen patientpopulation på baggrund af diagnosen depression ikke godt, da diagnosen ikke ser ud til at afspejle en specifik patogenetisk klasse og samtidig ikke sikrer en symptommæssig homogen patientgruppe. Endelig kan det bemærkes at depression heller ikke på det ætiologiske niveau, altså den initierende årsag til at sygdomsprocessen sættes i gang, er homogen. Dette spørgsmål har ikke indgået i undersøgelsen, men det er velkendt at folk udvikler depression på baggrund af en række forskellige situationer, af forskellige årsager og nogen gange tilsyneladende uden nogen form for identificerbar årsag.

Et problem for skabelsen af homogene patientgrupper er at diagnosticeringen af depression nødvendigvis hviler på en fortolkning af subjektive symptomer. De diagnostiske kriterier er afgørende for præcis hvilke patienter der kan få diagnosen depression, samt hvilken sværhedsgrad tilstanden vurderes at have. Homogeniteten afhænger således af kriterierne. De diagnostiske kriterier er fastlagte på forhånd og retter sig mod de sygdomselementer man anser for definerende for sygdommen. Depression adskiller sig dog radikalt fra den somatiske sygdom i det forhold at *kernesymptomerne* kun er subjektivt tilgængelige. Kernesymptomerne beskriver det som forstås som definerende for sygdommen, mens de øvrige følgesymptomer ifølge ICD-10 forstås som følge af kernesymptomerne. I diagnosticeringen af en patient med depression, kræves det derfor også at

det er belyst at minimum to ud af tre af kernesymptomerne (nedtrykthed, nedsat lyst og interesse og nedsat energi og øget trætheds) er tilstede, da disse symptomer drejer sig om det der primært forstås ved den depressive tilstand, nemlig det nedsatte stemningsleje. En depressionsdiagnose uden nedsat stemningsleje ville mangle begrebsvaliditet, dvs. manglende overensstemmelse mellem begrebet depressiv og diagnosen depression (afsnit 5.3). Det er ikke muligt for lægen objektivt at måle det nedsatte stemningsleje, og for så vidt heller ikke for patienten. Når diagnosen skal stilles får patienten spørgsmålet ”hvor ofte indenfor de sidste 14 dage har du følt dig nedtrykt?” (jf. MDI, bilag 1). Men for patienten vil et sådant symptom rimeligvis heller ikke have karakter af en erfaring der kan gengives objektivt og således indpasses svarmulighederne ”hele tiden”, ”det meste af tiden” eller ”lidt over halvdelen af tiden”. Et præcist svar på et sådant spørgsmål kræver at patienten kan genkalde sig sit stemningslejes op og nedture gennem 14 dage, men patienten har jo ikke nødvendigvis været konstant optaget af at registrere disse variationer og hvis han havde (fordi han vidste at han skulle besvare et spørgeskema), er det ikke utænkeligt, at denne opmærksomhed i sig selv kunne have en effekt på stemningslejet. Hvad jeg prøver at nå frem til er, at det er meget vanskeligt at gengive subjektive oplevelser i objektive mål. Ikke desto mindre er det de subjektive symptomer der er de væsentligste i dannelsen af en patientgruppe der gerne skulle være homogen for at vi kan bruge den i behandlingssammenhæng. I dannelsen af denne patientgruppe må man forlade sig på patientens fortolkning og måde at udtrykke sin oplevelse af at være deprimeret, og på lægens fortolkning af patientens subjektive tilkendegivelse af sin sygdomsoplevelse. Det gør nødvendigvis diagnosen mere usikker at beskrivelsen af tilstanden hviler på fortolkning af subjektive oplevelser, oversat til objektive kriterier, end hvis diagnosen havde hvilet på observation. Og det betyder at der kommer større usikkerhed omkring homogeniteten af den diagnostiske patientgruppe end det er tilfældet for somatiske sygdomme. Dette er en principiel forskel mellem somatiske og psykiske sygdomme, der betyder at kravet om homogene diagnostiske patientgrupper i evidensproduktionen, stiller psykiatrien og diagnosen depression dårligere når effekten af behandlingsindsatser skal undersøges.

9.2 Objektivisering af det subjektive

For at lappe på det problem jeg har beskrevet ovenfor, kan man forsøge at måle depression på mere objektive kriterier. Det ses mest tydeligt i Hamiltons Depressionsskala, der anvender flere objektive kriterier for depression end subjektive.

EBM er afhængig af at kunne måle effekter i forhold til standardiserede målestokke. Oplevelsen af depression oversættes til standardsymptomer der placeres på en skala for omfang og vægt, hvorefter de lægges sammen og angives som et enkelt tal.

Grunden til at RCT nyder den øverste placering i evidenshierarkiet er, at den menes bedst til at udelukke muligheden af usikkerheder, dvs. bias. Men den bevægelse der sker i fortolkningsprocessen fra oplevelse til tal er på ingen måde bias-fri. Den baserer sig netop på bestemte kriterier og bestemte måder at måle disse kriterier. Som jeg har vist i et eksempel i kapitel 8, vil resultatet af behandlingsindsatsens effekt, være stærkt afhængig af hvilken skala man anvender til at måle sværhedsgraden. Det vil sige at behandlingen vurderes effektiv eller ikke-effektiv alt efter hvilke depressionssymptomer skalaen måler på og hvordan den vægter dem. Behandlingen vurderes derimod ikke på baggrund af hvilke symptomer der er væsentlige for patienten og hvordan patienten vægter disse. Dermed indebærer de processer EBM er afhængig af, nemlig kvantificeringen af den subjektive oplevelse, en standardisering som flytter udgangspunktet for behandling fra den individuelle oplevelse af sygdom til nogle standardiserede kriterier for sygdom. Dermed sker der en reduktion, i og med at det subjektive kun forstås i forhold til standardiserede kriterier. Ved kun at måle på standardiserede objektiverede symptomer overser man jo netop at den lidelse man forsøger at behandle i sin essens er en subjektiv oplevelse – det fænomen der forstås ved depression er et subjektivt fænomen. I Hamiltons Depressionsskala, som anvendes i det eksemplificerede review, tager man som sagt i vidt omfang udgangspunkt i objektive tegn på depression, der angår funktionelle egenskaber der kan identificeres objektivt. Der spørges således ikke til patientens lyst og interesse i sine daglige gøremål, men til patientens evne til udføre disse aktiviteter. Det er primært følgesymptomerne frem for kernesymptomerne som man måler sværhedsgrad og behandlingseffekt på, frem for de oplevelsesmæssige kernesymptomer som man ellers primært forstår som definerende for depression. Det bliver altså en reduktion fra subjektivt erfarede symptomer der er bærere af individuel meningstillæggelse, til objektive symptomer der vægtes i forhold til standardiserede mål. Det tal på Hamiltons depressionsskala som angiver depressionens sværhedsgrad kunne ligeså godt dække over en anden symptomsammensætning og en anden sygdomsoplevelser. Depression kan eksempelvis være primært karakteriseret ved den manglende lyst og interesse, hæmning, koncentrationsbesvær og nedsat appetit – altså en generelt apatisk tilstand, eller den kan være præget af skyldfølelse, nedsat selvfølelse, agitation og søvnforstyrrelser. På en skala kan de to tilfælde imidlertid oversættes til det samme tal. I reduktionen fra det subjektive til det objektive foregår altså et meningstab med det formål at kunne

måle effekter som det foreskrives i EBM. Når man foretager RCT-forsøg er det de standardiserede og i høj grad objektive symptomer der er bestemmende for behandlingsindsatsens effektivitet. Man kan mistænke at det netop er derfor, at muligheden opstår for at behandlingsindsatser (antidepressiva) der medfører en række emotionelle bivirkninger som for eksempel emotionel følelseløshed, dalende empati, og færre positive tanker (Read et al. 2014) findes effektive mod depression. Disse emotionelle bivirkning ligger tæt op ad kernesymptomerne for depression, men behandlingen kan alligevel siges at behandle depression, hvis man vel at mærke vurderer behandlingseffekten på nogle helt andre parametre som fx søvnbesvær, maveproblemer, hæmning og vægttab.

9.3 Manglende kohærens mellem sygdomsniveauer

Med diagnosticering på baggrund af symptomer opnås der mulighed for at opnå en vis homogenitet, selv om den aktuelle depressionsdiagnose omfatter forholdsvist forskellige symptombilleder. Men spørgsmålet er om homogenitet med hensyn til symptombillede er det afgørende for den prognostiske validitet. Et stort problem for RCT-baseret evidens i psykiatrien er at der ikke er opnået den sygdomsindsigt der er nødvendig som fundament for prognostisk valide diagnostiske kategorier. Det kausale net der skal forklare hvorfor sygdommen udvikles er simpelthen slet ikke afdækket.

Hvis man skal overføre de erfaringer der gælder klassifikation i den somatiske medicin til psykiatrien og til depression, så er der ikke grundlag for at antage at homogene grupper på symptomniveau også er homogene grupper på patogenetisk og ætiologisk niveau. Samme symptom kan sagtens være udtryk for forskellige underliggende sygdomsmekanismer og samme sygdomsmekanisme kan give sig til udtryk i forskellige symptomer. Ud fra et nominalistisk udgangspunkt og 'the minimal medical model', er det imidlertid slet ikke interessant at beskæftige sig med årsager til sygdom. Det vil sige at en sådan klassifikation også egner sig bedst til at informere behandlingsindsatser der virker på symptomer.

Hvis man således laver en behandlingsindsats der skal virke på sygdommens patogenetiske niveau (hvis indsatsen fx har til formål at rette op på en skæv serotoninbalance som altså udgør en formodet underliggende dysfunktion der manifesterer sig i depressionssymptomer), men samtidig baserer sin patientpopulation på en klassifikation der udelukkende fokuserer på at skabe patientpopulationer der er homogene på symptomniveau, så er der ikke tilstrækkeligt grundlag til at

antage at denne population er homogen på det relevante niveau, altså patogenetisk niveau, og dermed er prognostisk valid. Selvom man ved, at en skæv serotoninbalance forårsager depression, er det ikke serotoninbalancen patienterne er diagnosticeret på baggrund af, og man kan derfor ikke vide om symptomerne i de specifikke tilfælde er forårsaget af noget andet.

Hvis man sammenligner med den somatiske medicin kan man ikke slutte direkte fra symptomet mavesmerte til det patogenetiske niveau mavesår, ligesom man ikke kan slutte fra symptomet feber til nogen bestemt sygdom. Symptomerne indikerer mulige årsagsmekanismer, der efterfølgende kan testes for, ved fx laboratorietests (fx ved infektioner), scanning (fx ved læsioner, knoglebrud, cyster osv.), eller direkte observation (fx ved hals eller ørebetændelse). Alt efter graden af sygdomsindsigt, dvs. i hvilket omfang det kausale netværk af årsagsfaktorer er afdækket, kan man bruge symptomerne som indikator på årsagsfaktorerne. Depression og psykiatrien generelt er karakteriseret ved ikke at have en ret høj grad af sygdomsindsigt. Selvom neurovidenskaben er i stand til at observere flere og flere tilstande, er det stadig ikke muligt i at kortlægge det samlede kausalnetværk der forårsager sygdommen. Nikolas Rose illustrerer denne pointe ved henvisning til forskningen i Alzheimers. De seneste år har man haft stor succes med at påvise hvordan neurologiske faktorer og dispositioner har betydning for udvikling af Alzheimers. Ikke desto mindre er der ikke tale om nogen simpel relation mellem disse neurologiske fund og den faktiske udvikling af sygdommen. Det er ikke muligt på baggrund af en hjernescanning at vurdere om personen rent faktisk lider af demens. De neurologiske fund man lægger til grund for sygdommen viser sig ikke kun i forbindelse med symptomer på Alzheimers, og man kan derfor ikke slutte fra hjernens tilstand til den faktiske kognitive kapacitet, eller fra symptomerne på demens til bestemte neurologiske tilstande (Rose 2010: 76f). Biologien er således ikke alene determinerende for den kognitive funktion og kan således kun redegøre for nogle faktorer i årsagskomplekset, og de neurologiske fund udgør således ikke tilstrækkeligt grundlag til at skabe en homogen population.

På samme vis er det, selv hvis der kan påvises en sammenhæng mellem depressive tilstande og serotoninbalancen, ikke muligt at slutte direkte fra depressionssymptomer til skæv serotoninbalance uden at man rent faktisk har testet for denne.

I kraft af den manglende indsigt i sygdomsmekanismen ved man ikke reelt hvilke faktorer der sikrer prognostisk validitet i forhold til behandling på patogenetisk niveau, og som dermed er de relevante at måle homogeniteten på. Ved at basere studier af behandlinger der virker på sygdomsmekanisme-niveau på en diagnostisk gruppe der er defineret i kraft af sin homogenitet på symptomniveau, gør

man sig den antagelse at det patogenetiske niveau og symptomniveauet er kohærente. Det er der ikke tilstrækkeligt belæg for.

9.4 Subgrupper og generaliseringsspørgsmålet

Hvis det ikke er muligt at opstille diagnostiske kriterier der sikrer en prognostisk homogen patientpopulation, så er diagnosen ikke et tilstrækkeligt grundlag at basere behandlingspraksis på. For at forbedre behandlingspraksis for alle de patienter der ikke falder ind under de 13% der har gavn af SSRI, kan man i evidensproduktionen forsøge at gøre sig erfaringer om de betingelser der kan have indflydelse på behandlingseffekten.

Jeg har derfor diskuteret brugen af subgrupper i evidensproduktionen. I det eksemplificerede RCT og review undersøger man ikke forskelle i behandlingseffekt på patientsubgrupper, men i Sundhedsstyrelsens anbefalinger er retningslinjerne fordelt på de subgrupper som også optræder i diagnosen, altså let, moderat, svær og psykotisk, og desuden særlige anbefalinger i forbindelse med depression hos gravide, ammende, ældre, patienter med demens, patienter med somatisk comorbiditet og patienter med personlighedsforstyrrelse. Heraf kan man se at der i et vist omfang findes evidens omhandlende patientsubgrupper, om end den stadig er sparsom (Sundhedsstyrelsen 2007: 52ff).

I EBM antager man at depression kan oversættes til objektive symptomer og måles herefter, og i de faktiske Cochrane reviews foretager man kun undersøgelser af objektive faktorer som betydende for behandlingseffekten (antal symptomer, alder, fødsel og comorbiditet). Valget af faktorer kan dermed afspejle en naturalistisk-objektivistisk sygdomsopfattelse hvor sygdomme virker ind på en standardiseret krop (den postnatale krop, den somatisk syge krop og den aldrende krop), der som sådan er adskilt fra individet. Den relativistisk-subjektivistiske sygdomsopfattelse ville derimod lægge mere vægt på individets egen måde at forholde sig til sygdommen som betydende for behandlingers effekt og hensigtsmæssighed generelt. Jeg har i kapitel 4 nævnt, at i og med at depression ofte er forbundet med patientens specifikke livssituation, erfaringer eller hændelser, er sociale faktorer oplagte også at undersøge i relation til behandlingseffekter.

Praktisk er det imidlertid en uhensigtsmæssig tilgang at begynde at udføre de meget omfattende RCT for at teste for alle relevante subgrupper, for problemet er jo, at man netop ikke ved, hvilke subgrupper der ville være de mest relevante at betragte for at opnå prognostisk homogene patientgrupper.

Hvis man skal tage den relativistisk-subjektivistiske position seriøst handler det dog ikke kun om at undersøge sociodemografiske kontekstuelle variable i forhold til sociale årsagsfaktorer. Det handler også om den måde individet forholder sig til sin sygdom som afgørende i forhold til behandlingsindsatsers resultat. Det kunne fx dreje sig om patientens tidligere erfaringer med sygdomsforløb og behandlingsmetoder, tillidsforholdet til lægen, patientens fortolkning og forståelse af depressionens årsager, motivation for behandlingstilgang, kulturel baggrund og værdier og overbevisninger. Disse faktorer er relateret til de såkaldte non-specifikke behandlingselementer der netop har vist sig at være af større betydning for depressionsbehandling end de specifikke behandlingsindsatser, med en placeborespons på 42% i RCT-eksemplet. Denne form for subjektive kontekstmæssige faktorer er karakteriseret ved at være svære at opstille kvantitativt, da de er udtryk for patientens egen fortolkning og meningstillæggelse. Det er således ikke kun et spørgsmål om hvilke objektive forhold der kan karakterisere patientens liv, men også hvilken mening patienten tillægger disse forhold. Viden af denne karakter vil derfor sandsynligvis ikke komme fra RCT, da det vil være nødvendigt at anlægge en langt mere forståelsesorienteret og dermed kvalitativ tilgang i forskningstilgangen der involverer patienten som subjekt.

Hvis man ser på depression som fænomen, der i kraft af per definition at være en subjektiv oplevelse af sygdom, er afhængig af personens individuelle måde at håndtere og forstå sin tilstand på, bliver det altså nødvendigt at undersøge individets subjektivitet som ramme for kausale mekanismer i form af vilje, værdier, overbevisninger osv. Det kausalbegreb som der arbejdes med i RCT-baseret evidens, samvariansen mellem flere faktorer, er derfor ikke tilstrækkeligt til at skabe viden der kan fungere som behandlingsgrundlag for den enkelte patient.

Kapitel 10: Konklusion

Jeg vil afslutningsvist forsøge at konkludere på undersøgelsen i en besvarelse af min problemformulering:

Hvilken betydning har det for den viden som depressionsbehandlingen i den kliniske praksis baserer sig på, at den primært er tilvejebragt på baggrund af et evidensbegreb som er udviklet indenfor den somatisk medicinske videnskab, og som baserer sig på randomiserede kontrollerede forsøg?

I min undersøgelse af RCT-baseret evidens som metodisk tilgang til at skabe viden om depressionsbehandling, har jeg fundet at der er grundlæggende forskelle mellem somatisk medicinske sygdomme og depression, der har betydning for relevansen og hensigtsmæssigheden af primært at basere behandlingspraksis på viden der er tilvejebragt gennem denne metodik.

Jeg har fundet at en forudsætning for at RCT-baseret evidens skal kunne bruges som grundlag for valg af behandling for den enkelte patient er, at den baserer sig på prognostisk homogene patientgrupper. Hvis ikke det er tilfældet, vil det ikke være muligt at vide, hvilke patienter der med fordel kan behandles med den givne indsats, og hvilke der ikke vil opleve den ønskede effekt.

For somatiske sygdomme er det generelt muligt at opnå forholdsvis prognostisk homogene patientgrupper, mens der er flere problemer forbundet med at skabe en prognostisk homogen patientgruppe for diagnosen depression. Jeg mener at dette kan tilskrives fire fundamentale forskelle mellem somatiske sygdomme generelt og depression.

For det første har det i psykiatrien vist sig svært at drage grænser mellem forskellige sygdomme og mellem sygdom og normalitet. For somatiske sygdomme består inklusionskriterierne for den diagnostiske gruppe i den observerede tilstedeværelse af en bestemt fysisk afvigelse fra et funktionelt normalplan. For depression er de kernesymptomer der er definerende for sygdommen ikke i sig selv udtryk for afvigelse fra normale mentale funktioner. Da man ikke arbejder med nogen norm for mentale funktioners normalplan, men alene baserer sig på tilstedeværelsen af symptomer, bliver grænsen for hvad der kan betragtes som dysfunktion og dermed en patologisk tilstand, flydende. Endvidere er det ikke påvist at diagnosernes grænser svarer til forskellige typer af patienter. Dvs. der er lige så mange patienter der oplever præcis de symptomer der er knyttet til depressionsdiagnosen som der er patienter der oplever et miks af symptomer fra forskellige diagnoser. Da de diagnostiske kriterier desuden giver mulighed for stor diversitet, betyder det at nogle af patienterne indenfor diagnosen vil have få symptomer tilfælles, mens de måske vil have

mere tilfælles med patienter der er placeret under en anden diagnose. Det flydende spektrum af symptomer gør altså at diagnostiske grupper ikke repræsenterer én tilstand men relaterede tilstande, og de grænser som diagnosen drager afspejler ikke naturligt forekommende grænser, men konstruerede grænser.

For det andet er somatiske sygdomme defineret omkring kriterier der kan identificeres ved observation. I kraft af sygdommens fysiske karakter, er det muligt at måle dem objektivt. De definerende kriterier for depression er derimod kun subjektivt tilgængelige. Det betyder at dannelsen af patientgruppen hviler på en fortolkning i stedet for observation. Den subjektive oplevelse er ikke mulig at oversætte til objektive symptomer uden at der sker et meningstab. Fortolkningen sker endvidere i forhold til nogle fastlagte rammer for, hvilke symptomer der er væsentlige og hvilke der ikke er væsentlige for diagnosen. Desuden er RCT-forsøget afhængig af at effekten kan måles numerisk, og depressionens sværhedsgrad angives derfor på baggrund af standardiserede skalaer. Fortolkningen fra sygdommen som oplevelse til sygdommen som tal sker altså igennem nogle rammer der øver stærk indflydelse på processen. Denne proces medfører en langt højere usikkerhed for dannelsen af homogene patientgrupper for diagnosen depression end for somatiske sygdomme generelt.

Den tredje forskel der gør sig gældende, er forskellen mellem den somatiske medicins sygdomsindsigt og sygdomsindsigten i depression. I den somatiske medicin baserer man diagnoserne på sygdomsmekanismen der er del af det kausalkompleks der forårsager sygdommen. Somatiske sygdomme er dermed klassificeret på baggrund af netop det kriterium som er genstand for behandlingsindsatsen. Selvom individuel variation kan have betydning for behandlingens effekt, må det siges at være en temmelig god indikator for patientgruppens prognostiske homogenitet, at inklusionskriteriet er tilstedeværelsen af en bestemt sygdomsmekanisme. Den samme sygdomsindsigt findes ikke for depression, og det betyder at man ikke reelt ved, hvilke faktorer der er de væsentlige for at opnå en patientpopulation der kan forventes at respondere sammenligneligt på en given behandling. Depression diagnosticeres på baggrund af symptomer. Når man undersøger effekten af antidepressiva som behandlingseffekt, er der derimod tale om en indsats der retter sig mod sygdommens patogenetiske niveau. Der er dog ikke, på baggrund af erfaringen fra den somatiske medicin, tilstrækkeligt grundlag for at antage at en patientpopulation der er homogen på symptomniveau, også udgør en homogen gruppe på patogenetisk niveau og det er derfor ikke muligt at slutte direkte fra kriterierne for depression til en specifik årsag.

For det fjerde er somatiske sygdomme og depression fundamentalt set meget forskellige fænomener. Somatiske sygdomme er generelt og først problemer der angår den fysiske verden. En behandlingsindsats må altså primært måles på dens evne til at bringe en kropslig mekanisme i overensstemmelse med et kropsligt funktionelt normalplan. Disse kan måles og observeres. For depression er det problem der behandles *oplevelsen* af ikke at fungere. Når man objektiviserer oplevelsen af depression i form af standardiserede og målbare symptomer som behandlingseffekten kan måles på, så overser man at den sygdom man forsøger at behandle kun kan erfares subjektivt. Det er vist (ved RCT-forsøg og Cochrane reviews) at patienter med depression har en meget høj placeborespons. De behandlingselementer der betegnes placebo, er netop kendetegnet ved at angå oplevelsen af behandlingen. I og med at depression i sin grundsubstans er et subjektivt fænomen er det nødvendigt at undersøge hvordan den subjektive erfaring af sygdommen formes, dvs. de mekanismer der danner ramme for individets oplevelse. Disse er netop individuelle, og kan fx være afhængige af den betydning som individet tillægger depressionen. En patientgruppe der kun er baseret på standardiserede symptomer, vil altså ikke tage hensyn til de individuelle forskelle der kan have betydning for behandlingsrespons.

Når jeg dermed skal besvare min problemformulering, bliver det klart at den viden man opnår om depressionsbehandling gennem RCT-baseret evidens er karakteriseret ved at generalisere over en patientpopulation, der ikke kan antages at have den samme grad af prognostisk validitet som somatiske sygdomme. Derved fortæller evidensen os noget om den generelle effekt af behandlingsindsatser for depression, men da diagnosen ikke kan forventes at tegne en tilstrækkelig prognostisk homogen gruppe, er der betydelige usikkerheder forbundet ved at generalisere resultaterne og anvende dem i behandlingsvalget for den enkelte patient, da de kan dække over store individuelle forskelle. Konsekvensen kan derfor blive at mange patienter kan behandles med indsatser der ikke har nogen effekt, er overflødige eller direkte skadelige.

Evidensbaseret medicin søger at forholde sig til sygdomme som objektive fænomener, men derved overses det at psykiatriens genstandsfelt er af en anden karakter. Evidensbaseret medicin anlægger en mekanistisk og standardiserende tilgang til forskning i behandlingseffekter. Denne tilgang har sit teoretiske fundament i den somatiske medicin. Men i anvendelsen af samme metodologi indenfor psykiatrien forbigås væsentlige aspekter af psykiatriens genstandsfelt og interesse, idet man i anvendelsen af metodologien, ikke forholder sig til at psykiatri nødvendigvis er en fortolkende praksis der forholder sig til subjektive fænomener.

Litteraturliste

American Psychiatric Association (2013): *Diagnostic and Statistical Manual of Mental Disorders* (5. Udgave; DSM-5) APA. Washington.

Arroll, B., Elley, C.R., Fishman, T., Goodyear-Smith, F.A., Kenealy, T., Blashki, G., Kerse, N. & MacGillivray, S. (2014): *Antidepressants versus placebo for depression in primary care (Review)*. The Cochrane Collaboration. John Wiley & Sons, Ltd.

Bastian, Hilda, Glasziou, Paul & Chalmers, Iain (2010): *Seventy-five trials and eleven systematic reviews a day: How will we ever keep up?* PLOS Medicine, vol. 7, nr. 9.

Bhaskar, Roy (2008): *A Realist Theory of Science* (2. Udgave). Routledge. New York

Bluhm, Robyn & Borgerson, Kirstin (2011): *Evidence-Based Medicine*, i Gifford, Fred (ed.): *Philosophy of Medicine*. North Holland, Elsevier. Great Britain.

Bluhm, Robyn (2007): *Clinical trials as nomological machines: Implications for evidence-based medicine*, i Kincaid, H. & McKittrick, J. (ed): *Establishing Medical Reality*. Springer. Dordrecht, Holland.

Cartwright, Nancy (2007): *Are RCTs the Gold Standard?* BioSocieties, vol. 2, nr. 1.

The Cochrane Collaboration 1: Dansk brugervejledning: Hvad er et randomiseret kontrolleret forsøg? Lokaliseret 4/3 2015: <http://www.cochrane.dk/clib/rct.htm>

The Cochrane Collaboration 2: Dansk brugervejledning: Sådan læser man en oversigt. Lokaliseret 4/3 2015: <http://www.cochrane.dk/clib/laese.htm#hvad>

The Cochrane Collaboration 3: Dansk brugervejledning: Hvad er Cochranesamarbejdet? Lokaliseret 4/3 2015: <http://www.cochrane.dk/clib/omcc.htm>

Cohen, David & Jacobs, David. H. (2007): *Randomized controlled trials of antidepressants: clinically and scientifically irrelevant*. Debates in Neuroscience, vol. 1.

Cooper, Rachel (2013): *Diagnosing the Diagnostic and Statistical Manual of Mental Disorders : Fifth Edition*. Karnac Books. UK.

Danmarks Apotekerforening (2013): *Analyse. Kvinder i Syddanmark bruger oftest antidepressiv medicin*. Lokaliseret 30.10 2014: http://www.apotekerforeningen.dk/~media/Apotekerforeningen/analysersundhed/10042013_antidepressivmedicin.ashx

Dansk Selskab for Almen Medicin (2010): *Klinisk vejledning for almen praksis. Unipolar depression. Diagnostik og behandling*. Lægeforeningens forlag. København

Danske Regioner (2009): *Psykisk sygdom og arbejdsmarkedet. Debatoplæg*. Lokaliseret 28.10 2014: <http://www.regioner.dk/~media/Filer/Social%20og%20psykiatri/PsykiskSygdomTilNettet2.ashx>

Danmarks Radio, P1 Debat, 17. Januar 2014. Lokaliseret 11.11 2014: <http://depressionsforeningen.dk/p1-debat-anti-depressiver-loesningen-eller-problemet/>

Ehrenberg, Alain (2010): *Det udmattede selv. Depression og samfund*. Informations forlag, 1. Udgave, 1. oplag. Polen.

Evidence-Based Medicine Working Group (1992): *Evidence-based Medicine. A new approach to teaching the practice of medicine*. Journal of the American medical association. Vol. 268, nr. 17.

Gannik, Dorte E. (2000): *The social construction of diagnosis*, i: Launsø, L. & Gannik, D. E.: *Disease, Knowledge, and Society*. Samfundslitteratur. Gylling

Gannik, Dorte E. (2005): *Social sygdomsteori – et situationelt perspektiv*. Forlaget Samfundslitteratur. Gylling

Gannik, Dorte E. (2008): *Virkelighed, teorier og metode i sygdomsforskningen*. Tidsskrift for forskning og sygdom og samfund, vol. 5, nr. 9.

Ghaemi, S. Nassir (2003): *The Concepts of Psychiatry: A Pluralistic Approach to the Mind and Mental Illness*. John Hopkins University Press. Baltimore, MD.

Goodwin, D.W. & Guze, S.B.: *Preface to the first edition*, in North, C.S. & Yutzy, S.H. (ed.) (2010): *Goodwin and Guze's Psychiatric Diagnosis*. 6th edition. Oxford University Press. New York.

Gomory, Tomi (2013): *The Limits of Evidence-Based Medicine and Its Application to Mental Health Evidence-Based Practice: Part One*. Ethical Human Psychology and Psychiatry, vol. 15, nr. 1.

Gupta, Mona (2007): *Does Evidence-Based Medicine Apply to Psychiatry?* Theoretical medicine and bioethics, vol. 28.

Gøtzsche, Peter (2007): *Rational Diagnosis and Treatment. Evidence-based Clinical Decision-making*. 4th edition. Wiley. Padstow, Cornwall, UK.

Gøtzsche, Peter (2013): *Dødelig medicin og organiseret kriminalitet. Hvordan medicinalindustrien har korruperet sundhedsvæsenet*. People's Press. København.

Gøtzsche, Peter (2014): *Why I think antidepressants cause more harm than good*. Lancet Psychiatry, vol. 1.

Harari, Edwin (2001): *Whose evidence? Lessons from the philosophy of science and the epistemology of medicine*. The Australian and New Zealand Journal of Psychiatry, vol. 35, nr. 6.

Hansen, Hanne Foss (2014): *Organisation of evidence-based knowledge production: Evidence hierarchies and evidence typologies*. Scandinavian Journal of Public Health. Vol. 42, nr. 13.

Healy, David (2007): *The New Anecdotes*. Ethical Human Psychology and Psychiatry, vol. 9, nr. 3.

Higgins, J.P.T. & Green, S. eds. (2011): *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. Tilgængelig via: www.cochrane-handbook.org.

Hjørland, Birger (2011): *Evidence-Based Practice: An Analysis Based on the Philosophy of Science*. Journal of the American Society for Information Science and Technology, vol. 62, nr. 7.

- Horwitz, Ralph I., Singer, Burton H., Makuch, Robert W. & Viscoli, Catherine M. (1996): *Can Treatment That Is Helpful on Average be Harmful to Some Patients? A Study of the Conflicting Information Needs of Clinical Inquiry and Drug Regulation*. Journal of Clinical Epidemiology, vol. 49, nr. 4.
- Howick, J. H. (2011): *Philosophy of evidence-based medicine*. Wiley-Blackwell. Hoboken, NJ. ProQuest ebrary.
- Institut for rationel farmakoterapi (2010): *Den nationale rekommandationsliste, antidepressiva*. http://www.irf.dk/dk/rekommandationsliste/baggrundsnotater/nervesystemet_analgetika_og_psykofarmaka/atc-gruppe_n06a_96_antidepressiva_.htm
- Kendell, Robert (2008): *Five Criteria for an Improved Taxonomy of Mental Disorders*, i Helzer, J. & Hudziak, J.(red.): *Defining Psychopathology in the 21st Century*. American Psychiatric Publishing. Tilgængelig på EBL Reader.
- Kendell, R. & Jablensky, A. (2003): *Distinguishing Between the Validity and Utility of Psychiatric Diagnoses*. The American Journal of Psychiatry, Vol 160, nr. 1.
- Kragh-Sørensen, & Kessing, (2010): *Affektive sindslidelser*, i Mors, O., Kragh-Sørensen, P. & Parnas, J. (2010): *Klinisk Psykiatri*. 3. Udgave. Munksgaard. Viborg.
- Kuhn, Thomas (1973): *Videnskabens revolutioner*. (2. Udgave) Fremads samfundsvidenskabelige serie. Odense
- Launsø, Leila & Gannik, Dorte Effersøe (2000): *Disease, Knowledge, and Society*. Samfundslitteratur. Gylling
- Leibovivi, L. (2001): *Effects of remote, retroactive intercessor y prayer on outcomes in patients with bloodstream infection: randomised controlled trial*. BMJ, Vol. 323,
- Maier, Thomas (2006): *Evidence-based psychiatry: understanding the limitations of a method*. Journal of evaluation in clinical practice, vol. 12, nr. 3.
- Moncrieff, J., Wessely, S. & Hardy, R. (2012): *Active placebos versus antidepressants for depression (review)*. The Cochrane Collaboration. John Wiley & Sons, Ltd.
- Murphy, Dominic (2013): *Philosophy of Psychiatry*, *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), Edward N. Zalta (ed.), Lokaliseret 24.11 2014 via: <http://plato.stanford.edu/entries/psychiatry/>
- North, C.S. & Yutzy, S.H. (ed.) (2010): *Goodwin and Guze's Psychiatric Diagnosis*. 6th edition. Oxford University Press. New York.
- Mors, O., Kragh-Sørensen, P. & Parnas, J. (2010): *Klinisk psykiatri*. 3. Udgave. Munksgaard. Viborg.
- Parker, G., Graham, R., Hadzi-pavlovic, D, Friend, P., Synnott, H. & Barrett, M. (2013): *Does testing for bimodality clarify whether the bipolar disorders are categorically or dimensionally different to unipolar depressive disorders?* Journal of Affective Disorders, vol. 137.

- Pedersen, Stig Andur (1990): *Causal and diagnostic reasoning*. Tekster fra IMFUFA, tekst nr. 193. Roskilde Universitetscenter.
- Pedersen, Stig Andur (2006): *Ross og sygdomsbegrebet*, i Holtermann, Jacob v. H. & Ryberg, Jesper (red.): *Alf Ross. Kritiske gensyn*. Jurist- og Økonomforbundet. **BY**
- Psykiastrifonden (2001): *Fakta og Depression*. Psykiastrifonden. København.
- Read, J., Cartwright, C. & Gibson, K. (2014): *Adverse emotionall and interpersonal effects reported by 1829 New Zealanders while taking antidepressants*. Psychiatry Research, vol. 216.
- Rieper, Olaf & Hansen, Hanne Foss (2007): *Metodedebatten om evidens*. AKF. København.
- Rose, Nikolas (2006): *Disorders Without Borders? The Expanding Scope of Psychiatric Practice*. Biosocieties, vol. 1, nr.4.
- Rose, Nikolas (2010): *Normality and pathology in a biomedical age 1*. The Sociological Review, vol. 57, nr. 2.
- Ross, Alf (1979): *Sygdomsbegrebet*. Bibliotek for læger, vol. 171.
- Ross, Alf (1980): *Det psykopatologiske sygdomsbegreb*. Bibliotek for læger, vol. 172.
- Sackett, D., Rosenberg, W., Gray, J., Muir, R. & Richardson, W. (1996): *Evidence-Based Medicine. What it is and what it isn't*. BMJ, vol. 12.
- Sehon, S. & Stanley, D. (2003): *A philosophical analysis of the evidence-based medicine debate*. BMC Health Services Research, vol. 3, nr. 14.
- Statens Serum Institut (2012): *Forbruget af antidepressiva 2001-2011*. Statens Seruminstitut, Dataleverancer og lægemiddelstatistik. Lokaliseret 30.10 2014:
- Straus, S., Richardson, W.S., Glasziou, P. & Haynes, R.B. (2005): *Evidence-based medicine: How to practice and teach EBM* (3rd ed.). Churchill Livingstone. Edinburgh.
- Sundhedsstyrelsen, Sekretariatet for Referenceprogrammer (2007): *Referenceprogram for unipolar depression hos voksne*. Sundhedsstyrelsen. København.
http://www.ssi.dk/~media/Indhold/DK%20-%20dansk/Aktuelt/Nyheder/AD_rapport_generel_2012_ver6.ashx
- Sundhedsstyrelsen, Dansk Selskab for Almen Medicin (2013): *Målbeskrivelse for speciallægeuddannelsen i almen medicin*. Sundhedsstyrelsen. København. Lokaliseret 1/3 2015:
<https://sundhedsstyrelsen.dk/~media/8F03E1CF1AA14EA58F160E5CA26E0DDC.ashx>
- Sørensen, Thorkild (2004): *Epidemiologiens bidrag til viden og evidens i forebyggelsen*, i Bruun, J.J., Hanak, M.L., Koefoed, B.G. (red.): *Sundhedsstyrelsen. Viden og evidens i forebyggelsen*. Sundhedsstyrelsen. København.
- Thomas, P., Bracken, P. & Timimi, S. (2012): *The anomalies of evidence-based medicine in psychiatry: time to rethink the basis of mental health practice*. Mental Health Review Journal, vol17, nr. 3.

Thornton, Tim (2007): *Essential Philosophy of Psychiatry*. Oxford University Press. Gosport, Hampshire.

Thrysoe, Willy (2007): *Menneskets seksualitet*. Forlaget Hovedland. Beder

Varga, Somogy (2013): *Classification, Validity and Psychiatric Kinds*, i Stoyanov, Drozdostoj (ed.): *Psychopathology. Theory, Perspectives and Future Approaches*. Nova Science Publishers, Inc. New York.

Wade, A., Lemming, O.M., Hedegaard, K.B (2002): *Escitalopram 10 mg/day is effective and well tolerated in a placebo-controlled study in depression in primary care*. *International clinical psychopharmacology*, vol. 17, nr. 3.

Wakefield, Jerome C. & First, Michael B. (2013a): *Diagnostic Validity and the Definition of Mental Disorder : A Program for Conceptually Advancing Psychiatry*. *The Canadian Journal of Psychiatry*, vol. 58, nr. 12

Wakefield, Jerome C. & First, Michael B. (2013b): *Diagnostic Criteria as Dysfunction Indicators : Bridging the Chasm Between the Definition of Mental Disorder and Diagnostic Criteria for Specific Disorders*. *The Canadian Journal of Psychiatry*, vol. 58, nr.12

WHO (1992): *The ICD-10 Classification of Mental and Behavioral Disorders: Clinical descriptions and diagnostic guideline*. World Health Organization. Geneva.

WHO (2010): *ICD-10 Version:2010*, located 15.9 2014.

<http://apps.who.int/classifications/icd10/browse/2010/en#/F30-F39>

Williams, Ben A. (2010): *The Perils of Evidence-Based Medicine*. *Perspectives in Biology and Medicine*, vol. 53, nr. 1.

Worrall, John (2010): *Evidence: Philosophy of Science Meets Medicine*. *Journal of Evaluation in Clinical Practice*, vol. 16, nr. 2.

Wulff, Henrik R., Pedersen, Stig Andur & Rosenberg, Raben (1986): *Medicinsk filosofi*. Munksgaard. Viborg.

Wulff, Henrik R. & Gøtzsche, Peter (1997): *Rationel Klinik. Evidensbaserede diagnostiske og terapeutiske beslutninger*. Munksgaard. København.