



**INSTITUTIONEN FÖR FILOSOFI, LINGVISTIK
VETENSKAPSTEORI**

FÖLJSAMHET I EVIDENSRÖRELSEN

En studie av följsamhetsproblematik utifrån
exemplet Förskrivning av fysisk aktivitet

Helena Lagerlöf

Uppsats/Examensarbete:	15 hp
Program och/eller kurs:	Magisterexamensarbete i evidensbaserad, EV2130
Nivå:	Avancerad nivå
Termin/år:	Vt/2016
Handledare:	Dick Kasperowski
Examinator:	Ingemar Bohlin

Abstract

Uppsats/Examensarbete:	15 hp
Program och/eller kurs:	Magisterexamensarbete i evidensbaserad, EV2130
Nivå:	Avancerad nivå
Termin/år:	Vt/2016
Handledare:	Dick Kasperowski
Examinator:	Ingemar Bohlin
Nyckelord:	Compliance, adherence, concordance, fysisk aktivitet, interventioner

-
- Syfte:** Låg följsamhet till riktlinjer avseende fysisk aktivitet har lett till att interventioner som innebär förskrivning av fysisk aktivitet prövats. I studier och införande av dessa interventioner krävs det avvägningar avseende vårdtagares följsamhet, för att studierna ska ge rättvisande resultat, för att nå målet avseende ökad fysisk aktivitet och med samtidig respekt för individers önskemål och situation. Det övergripande syftet med denna uppsats är att öka förståelsen för följsamhetsproblematik i studier, klinik och livsmiljö, i relation till evidensbaserings formaliserade procedurer och dess ideal om bästa möjliga hälso- och sjukvård.
- Metod:** En evidensvärdering utfördes av interventionen Förskrivning av fysisk aktivitet och dess långtidseffekt på fysisk aktivitet, i enlighet med den arbetsgång som Statens beredning för medicinsk och social utvärdering (SBU) rekommenderar. Vetenskaplig litteratur om följsamhet fick sedan belysa hur följsamhetsaspekter hanterats i dessa studier och enligt evidensrörelsens formaliserade procedurer för syntetisering av studier, samt vilka effekter det får på den evidensbaserade hälso- och sjukvården.
- Resultat:** Förskrivning av fysisk aktivitet har ingen långtidseffekt på fysisk aktivitet och det vetenskapliga underlaget är begränsat, enligt resultat av SBU:s rekommenderade procedurer för evidensgradering. Följsamhet är mångfacetterat och kan beskrivas med olika begrepp. Både vilka begrepp som används och hur de definieras har förändrats genom åren. Evidensrörelsens ideal bygger i hög grad på korrekt hanterade statistiska beräkningar där följsamhet spelar en stor roll, framför allt för komplexa och mindre väl definierade interventioner som Förskrivning av fysisk aktivitet. I studier möts statistiska och etiska ideal i hantering av följsamhet vilket leder till nödvändiga avvägningar där inte alla aspekter kan tillgodoses samtidigt i studier och klinisk praktik.

Innehållsförteckning

Inledning.....	4
Bakgrund	5
Följsamhet - en förutsättning för evidensbasering	5
Förskrivning av fysisk aktivitet.....	5
Potentiella motsättningar i hantering av följsamhet i studier, klinik och livsmiljö	6
Syfte, frågeställningar och disposition	7
Delsyfte 1	7
Frågeställningar	7
Delsyfte 2	7
Frågeställningar	7
Disposition	8
Metod och material.....	9
Metod	9
Del 1.....	9
Del 2.....	10
Material	10
Evidensläget för Förskrivning av fysisk aktivitet.....	11
Inkludering av studier utifrån strukturerad forskningsfråga.....	11
Kvalitetsgranskning och sammanställning av studier	12
Resultatsammanställning och evidensgradering.....	14
Slutsats av systematisk översikt och GRADE.....	14
Grundläggande antaganden varpå slutsatsen vilar	14
Följsamhetens problematiker	18
De granskade studiernas hantering av följsamhet	18
Etiska aspekter av följsamhet.....	20
En historiskt förändrad syn på följsamhet	20
Evidensrörelsens och individens olika perspektiv på risk och nytta	21
Vad är målet och var ligger ansvaret?	22
Följsamhetens betydelse för antaganden om kausalitet.....	23
Kausalitet i evidensrörelsen.....	23
Komplexa interventioner utmanar evidensrörelsens förhållande till kausalitet	25
Alternativa analysmetoder finns men kan inte lösa problemet.....	26
Förskrivning av fysisk aktivitet synliggör brister i evidensgraderingen	27
När statistik möter etik	28
Diskussion	30
Resultat och implikationer.....	30

Metoddiskussion.....	31
Slutsatser	32
Referenslista.....	33
Bilaga 1 Kvalitetsgranskning systematisk översikt.....	37
Bilaga 2 Bedömning enligt GRADE.....	40

Inledning

Evidensrörelsen har vuxit sig stark inom hälso- och sjukvården. Därmed finns det en strävan som de flesta delar, mot en hälso- och sjukvård som bygger på vetenskaplig kunskap. Att basera beslut på vetenskap är dock inte så enkelt som det kan låta. Vissa områden är helt enkelt inte tillräckligt beforskade för att det ska finnas ett vetenskapligt underlag, och inom andra områden har forskningen visat motstridiga resultat vilket vittnar om att ett vetenskapligt resultat inte alltid är tillförlitligt.

I dag är också de flesta överens om att en evidensbaserad hälso- och sjukvård inte bör byggas enbart på vetenskap. Hänsyn måste också tas till vårdtagares önskemål och situation, samt att klinikers erfarenhet bör vägas in. Det innebär att uppgiften att arbeta evidensbaserat blir än mer komplex. Det krävs idel avvägningar när olika kunskapskällor ska integreras och ett beslut om åtgärd ska tas.

Inom evidensrörelsen har det byggts upp procedurer för hur forskningsresultat kan sammanställas och ge underlag för både beslutsfattare och kliniker. Alla formaliseringar av detta slag innebär att viss kunskap ges företräde framför annan kunskap. Det byggs upp ett ideal avseende forskningsmetodik. Men vad händer med den evidensbaserade hälso- och sjukvården när den inte kan leva upp till detta ideal eftersom försökspersoner i vetenskapliga studier och vårdtagare inte gör det som förväntas av dem? Den följsamhetsproblematik som då uppstår är i fokus för denna uppsats.

Särskilt komplex blir följsamhetsproblematiken för interventioner som i sig syftar till att öka följsamhet. Det finns riktlinjer grundade i vetenskaplig kunskap som gör gällande att varje person bör upprätthålla en viss miniminivå av fysisk aktivitet för att undvika sjukdom och förtidig död. För de som inte följer dessa riktlinjer har interventioner prövats för att öka följsamheten till riktlinjer avseende fysisk aktivitet, d.v.s. öka den fysiska aktiviteten. Ett exempel på en intervention som syftar till ökad fysisk aktivitet är att vårdgivare utför en förskrivning av fysisk aktivitet till en vårdtagare. I denna uppsats vill jag med fallet Förskrivning av fysisk aktivitet visa hur följsamhet och dess relaterade problematiker i flera avseenden spelar en viktig roll i den evidensbaserade hälso- och sjukvården.

Bakgrund

Följsamhet - en förutsättning för evidensbaserad medicin

Evidensbaserad medicin (EBM) växte fram i början av 1990-talet och konceptet har sedan spridit sig utanför medicinen till övrig vård och omsorg, och på senare år även andra välfärdsområden såsom skolans värld och politiken. Oavsett område är grundtanken i all evidensbaserad medicin att en verksamhets insatser ska baseras på ett aktuellt vetenskapligt kunskapsunderlag (Bohlin, 2011). I Sverige utför myndigheten Statens beredning för medicinsk och social utvärdering (SBU) omfattande evidensvärderingar av interventioner inom hälso- och sjukvård. Systematiska översikter av företrädesvis randomiserade kontrollerade studier (RCT:er) framhålls som bästa grund för syntetisering av kvantitativ forskning. I dessa värderas studiernas vetenskapliga kvalitet var för sig för att ge studier med mindre risk för systematiska fel ett högre genomslag i värderingen av interventionens effekt. För att sammanväga studiernas resultat gör SBU en bedömning enligt GRADE (Grading of Recommendations Assessment, Development and Evaluation) som leder fram till en värdering av en interventions effekt och det resultatets evidensstyrka på en fyrgradig skala (Statens beredning för medicinsk utvärdering, 2014, kap 10).

En ideal evidensbaserad hälso- och sjukvård innebär att forskning producerar evidensbaserade interventioner som implementeras i verksamheterna. Följsamhet spelar en avgörande roll i båda dessa steg. För att forskningen ska ge valida effektmätningar krävs att försökspersonerna är följsamma till det studieprotokoll på vilket mätningen av interventionens effekt vilar. I det andra steget krävs följsamhet i klinik och personers livsmiljö för att forskningsresultaten ska komma människor till godo och göra nytta.

Förskrivning av fysisk aktivitet

Fysisk inaktivitet är förknippad med ökad risk för flera vitt skilda sjukdomstillstånd, såsom depression, benskörhet, diabetes, cancer, hjärtkärlsjukdom och för tidig död. World Health Organization (WHO) rankar fysisk inaktivitet som den fjärde ledande riskfaktorn för förtidig död. Rekommendationen för vuxna 18-64 år är enligt WHO minst 150 minuter aerob fysisk aktivitet av måttlig intensitet eller minst 75 minuter aerob fysisk aktivitet av hög intensitet eller likvärdig kombination av dessa, samt muskelstärkande träning minst två gånger i veckan. För barn och äldre finns anpassade rekommendationer (World Health Organization [WHO], 2010, ss. 54-55). Dåvarande Statens Folkhälsoinstitut i Sverige (ersatt av Folkhälsomyndigheten 2014) har snarlika rekommendationer och beräknar att ungefär halva den svenska befolkningen når dessa mål (*FaR® : individanpassad skriftlig ordination av fysisk aktivitet*, 2011, ss.12-13). De samhällsekonomiska kostnaderna för fysisk inaktivitet har av hälsoekonomer beräknats uppgå till flera miljarder bara i Sverige varje år (Bolin & Lindgren, 2006, s. 36).

För att öka följsamheten till riktlinjer avseende fysisk aktivitet har olika metoder prövats. Upplysning och rådgivning från sjukvårdspersonal är en åtgärd som känns logisk och rimlig och har använts både mer och mindre systematiserat och även utvärderats (Anokye, Lord & Fox-Rushby, 2014). Interventioner som grundas i någon teori om beteendeförändring, exempelvis transteoretiska modellen (TTM) eller socialkognitiv teori (SCT) har också studerats (Holtzman et al., 2004). Att ordinera fysisk aktivitet genom någon form av förskrivning är en intervention som i olika varianter fått stort genomslag i framför allt Storbritannien och Norden. I Storbritannien benämns denna intervention oftast Exercise referral scheme (ERS) och i Sverige finns Fysisk aktivitet på recept (FaR®). I vetenskaplig litteratur förekommer förutom Exercise referral schemes även Exercise on prescription och Physical activity on prescription. I denna uppsats används benämningen Förskrivning av fysisk aktivitet (FFA) för att inkludera samtliga dessa interventioner.

Potentiella motsättningar i hantering av följsamhet i studier, klinik och livsmiljö

Syftet med livsstilsinterventioner som FFA är att öka följsamheten till livsstilsrekommendationer för att motverka ohälsa och sjukdom. Detta går helt i linje med evidensbaserad hälso- och sjukvård eftersom forskning entydigt visar att fysisk aktivitet förebygger ohälsa och sjukdom. Om följsamheten till riktlinjer är låg innebär det ett misslyckande för implementeringen av evidensbaserad hälso- och sjukvård. Samtidigt framhålls på många håll vikten av att lyssna till vårdtagarens önskemål. Skälen till detta är både att det förväntas öka följsamheten och framhållandet av etiska skäl där vårdtagarens autonomi betonas som ett självändamål (Felzmann, 2012). Dessa olika förhållningssätt till följsamhet kan delvis förstås genom de begrepp som används i den vetenskapliga litteraturen. Både *compliance* och *adherence* används för att beskriva följsamhet men med något olika perspektiv. *Compliance* har kritiserats för att vara ett paternalistiskt begrepp där vårdtagarperspektivet utelämnas. I definitioner av *adherence* framhålls ofta att följsamhet grundas i en överenskommelse mellan vårdgivare och vårdtagare. *Concordance* har på senare år börjat användas för att ytterligare betona vårdtagarens autonomi och vikten av dialog mellan vårdgivare och vårdtagare. Det förekommer dock olika definitioner av alla tre begreppen (Vermeire, Hearnshaw, Van Royen & Denekens, 2001, ss. 332-333).

Även i studier krävs det avvägningar relaterade till följsamhet. Att studera en intervention i en kontrollerad studie kräver följsamhet till studiens intervention för att den statistiska analysen ska ge tillförlitliga resultat. Samtidigt ger studien inte ett generaliserbart resultat om samma följsamhet inte kan upprätthållas utanför studiemiljön. Sådillvida krävs det alltid en avvägning i val av design och upplägg för studier som utvärderar effekt av interventioner som kräver ett aktivt deltagande.

Det finns alltså i den vetenskapliga litteraturen en otydlighet i vad de olika följsamhetsbegreppen står för i olika sammanhang, och det är inte självklart hur följsamhetsproblematik bäst hanteras, i personers livsmiljö, i klinik och i studier. Därmed krävs avvägningar i flera avseenden vid studerandet av FFA. Statistiska beräkningar och EBM:s grundidé att evidensbaserade metoder bör implementeras talar för att eftersträva hög följsamhet både i studier och utanför. Vårdtagarens autonomi och möjlighet att generalisera studieresultat, som också förespråkas inom EBM, kan tala emot. Hur man bäst hanterar följsamhetsproblematik avseende livsstilsrekommendationer och interventioner såsom FFA är en viktig fråga i utvecklingen av en evidensbaserad hälso- och sjukvård.

Syfte, frågeställningar och disposition

Givet de potentiella motsättningar som kan existera mellan möjligheter att producera valida forskningsresultat och implementera framforskade interventioner, med samtidig respekt för individers önskemål, är det övergripande syftet med denna uppsats att öka förståelsen för följsamhetsproblematik avseende livsstilsrekommendationer i evidensbaserad hälso- och sjukvård.

I denna text används följsamhet som ett samlingsbegrepp för alla begrepp som på något sätt avser hur en vårdtagare efterlever instruktioner, ordinationer eller överenskomna åtgärder från hälso- och sjukvården. Uppsatsen avgränsas till ett studium av följsamhetsproblematik som avser vårdtagares följsamhet och berör alltså t. ex. inte vårdgivares följsamhet till riktlinjer. I begreppet vårdtagare innefattas i denna text även alla personer som är mottagare av livsstilsrekommendationer, det vill säga aktuell målgrupp för respektive riktlinje.

Delsyfte 1

Det första delsyftet är att genom en systematisk översikt och evidensgradering sammanfatta evidensläget för FFA avseende utfallsmåttet fysisk aktivitetsnivå.

Evidensgradering grundad på en systematisk översikt valdes som metod eftersom den är internationellt framarbetad och förespråkas av bl.a. SBU och Socialstyrelsen i Sverige, samt flera internationella organisationer (Statens beredning för medicinsk utvärdering, 2014, s. 141).

Frågeställningar

Har FFA positiv effekt på fysisk aktivitetsnivå jämfört med sedvanlig vård hos fysiskt inaktiva vårdtagare i primärvården?

Vilka grundläggande antaganden vilar slutsatsen om evidensläget för FFA på?

Delsyfte 2

Det andra delsyftet är att öka förståelsen för de olika problem som kan uppstå relaterade till personers följsamhet till livsstilsrekommendationer om fysisk aktivitet och till FFA, i studier, klinik och livsmiljö. Särskilt fokus läggs vid relationen följsamhetsproblematik och evidensbaserad. För detta syfte får resultatet från det första delsyftet fungera som ett exempel på hur följsamhet har hanterats i de granskade studierna och i den systematiska översikten. Med avstamp i dessa exempel diskuteras problem, möjligheter och alternativa sätt att hantera följsamhetsproblematik.

Frågeställningar

Hur hanteras följsamhet i de granskade studierna?

Vilka perspektiv på följsamhet lyfts i de granskade studierna?

Vilka potentiella motsättningar finns avseende hantering av följsamhet i studier av FFA?

Vilka potentiella motsättningar finns avseende hantering av följsamhet till riktlinjer avseende fysisk aktivitet?

Hur påverkas den evidensbaserade hälso- och sjukvården av synen på följsamhet som representeras i evidensrörelsens ideal avseende analysmetoder och formaliserade procedurer för sammanställning av forskning?

Disposition

Studiens tillvägagångssätt för att besvara delsyfte 1 och 2 delges översiktligt under rubriken *Metod och material*.

Under rubriken *Evidensläget för Förskrivning av fysisk aktivitet* redovisas metoden för delsyfte 1 utförligare och dess resultat i form av en slutsats om effekt och evidensstyrka avseende effekt på fysisk aktivitet av FFA, samt därefter en redogörelse över vilka grundläggande antaganden som slutsatsen vilar på.

Under nästa huvudrubrik, *Följsamhetens problematiker*, besvaras delsyfte 2. Först ges en sammanfattning av hur följsamhet hanterats i de granskade studierna. Därefter följer en problematisering av olika följsamhetsaspekter, uppdelad enligt rubrikerna *Etiska aspekter av följsamhet* och *Följsamhetens betydelse för antaganden om kausalitet*. I sista avsnittet *När statistik möter etik* analyseras effekterna av de olika följsamhetsaspekterna tillsammans.

Under rubriken *Diskussion* diskuteras studiens resultat, dess implikationer samt uppsatsens metodologiska begränsningar. Sist dras generella slutsatser av studien som helhet.

Metod och material

Metod

Del 1.

Evidensläget för FFA jämfört med sedvanlig vård avseende utfallsmåttet fysisk aktivitet sammanfattas genom att följa arbetsgången för kvantitativa studier i SBU:s metodbok: Utvärdering av metoder i hälso- och sjukvården (Statens beredning för medicinsk utvärdering, 2014), enligt följande steg:

Steg 1. Inkludering av studier utifrån strukturerad forskningsfråga (Statens beredning för medicinsk utvärdering, 2014, kap 3, 4, 5)

Första steget syftar till att identifiera samtliga vetenskapliga artiklar av relevans för den aktuella forskningsfrågan. För detta ändamål formuleras en strukturerad forskningsfråga enligt PICO där P=Population, I=Intervention, C=Control och O=Outcome. PICO omarbetas till en söksträng som används i relevanta databaser. Irrelevanta artiklar avseende PICO exkluderas, så att samtliga artiklar som kvarstår uppfyller PICO och svarar på den aktuella forskningsfrågan.

Steg 2. Kvalitetsgranskning och sammanställning av studier (Statens beredning för medicinsk utvärdering, 2014, kap 6)

De i förra steget identifierade artiklarna kvalitetsgranskas. Eftersom enbart randomiserade kontrollerade studier inkluderas i denna studie används mall Kvalitetsgranskning för randomiserade studier av SBU (2014, bilaga 2). Eventuella studier av låg kvalitet exkluderas. Studier av hög eller medelhög kvalitet tabelleras för innehåll avseende relevant PICO.

Steg 3. Resultatsammanställning och evidensgradering (Statens beredning för medicinsk utvärdering, 2014, kap 10)

Studiernas resultat avseende effekt för aktuellt utfallsmått sammanställs och tabelleras. En evidensgradering enligt GRADE (Grading of Recommendations Assessment, Development and Evaluation) genomförs. GRADE-systemet klassificerar det vetenskapliga underlagets styrka i en evidensstyrka 1-4 där 1 är otillräckligt vetenskapligt underlag, 2 är begränsat, 3 är måttligt starkt och 4 är starkt vetenskapligt underlag. För randomiserade kontrollerade studier är utgångspunkten att evidensstyrkan är stark (4). Denna kan sänkas vid brister i studiekvalitet (*risk of bias/limitations* max -2), vid bristande överensstämmelse mellan studier (*consistency* max -2), bristande överförbarhet/relevans för aktuell forskningsfråga (*directness* max -2), bristande precision, d.v.s. osäkerhet i den sammanvägda effekten (*precision* max -2) och vid hög sannolikhet för publikationsbias (*publication bias* max -1). Publikationsbias innebär att delar av det vetenskapliga underlaget inte är tillgängligt i vetenskapliga studier, vilket kan innebära att interventionens effekt övervärderas. Detta kan misstänkas t. ex. om studierna som finns publicerade är av samma forskargrupp, eller om register över påbörjade kliniska prövningar visar att studier påbörjats men aldrig publicerats. Det går också att bedöma risken för publikationsbias genom en statistisk beräkning. För detta krävs att det finns minst fem publicerade studier. Eftersom effektmätningen i studier med hög osäkerhet sannolikhetsmässigt fördelar sig jämnt kring medelvärdet för samtliga studier, kan man misstänka publikationsbias om studier med högre osäkerhet generellt visar högre effekt (Statens beredning för medicinsk utvärdering, 2014, kap 10).

Arbetsgången enligt de tre stegen ovan, leder fram till en slutsats om vilken effekt interventionen har på det aktuella utfallsmåttet, och det vetenskapliga underlagets styrka på skalan 1-4.

Till sist beskrivs kortfattat vilka antaganden som slutsatsen grundas i. Dessa antaganden har betydelse för de följsamhetsaspekter som hanteras i del 2.

Del 2.

Perspektiv på följsamhet och hantering av följsamhet i de inkluderade studierna analyseras. En första sökning gjordes med söksträng ”compliance AND adherence AND concordance” i databaserna Pubmed och Cinahl för att identifiera artiklar där olika perspektiv och syn på följsamhet lyfts. Sökningen resulterade i 236 träffar i Pubmed och 20 träffar med limitering ”exclude medline records” i Cinahl. Många artiklar exkluderades vid läsning av abstract eftersom de inte hade forskningsfokus på följsamhet. Ytterligare sökningar gjordes för att identifiera artiklar med fokus på problematik kring låg följsamhet i studier. Denna sökning gjordes mindre systematiskt. Sökord som användes var *clinical trial*, *causal inference*, *complex intervention*, och *bias* i kombination med *patient compliance*. Snöbollstekniken användes i alla sökningar d.v.s. en studie kunde leda till en annan via dess referenser som via referenser kunde leda till ytterligare en annan. Vid läsning av studier noterades allt som ansågs kunna vara av intresse för att besvara delsyfte 2. Därefter lästes och analyserades noteringarna för att utläsa teman, mönster och motsättningar. Totalt lästes cirka 100 fulltextartiklar varav cirka 40 av dem refereras till i denna uppsats.

Material

För delsyfte 1 utgörs materialet av de systematiskt uppsökta artiklarna med relevans avseende PICO, samt SBU:s metodbok (Statens beredning för medicinsk utvärdering, 2014). Artiklarna beskrivs under rubrik *Evidensläget för Förskrivning av fysisk aktivitet*.

För delsyfte 2 utgörs materialet av artiklar från ett stort antal tidskrifter, ämnesområden från olika nationaliteter, kompletterat med böcker. Brittisk omvårdnadsforskning är mest representerad vad gäller historisk utveckling av synen på följsamhet. Men även forskning inom omvårdnad och medicin från Indien, Norden, USA och Canada förekommer. Den medicinska forskningen har inriktning mot bland annat farmakologi och psykiatri. Litteratur avseende följsamhetsproblematik i studier utgörs dels av böcker av Shadish, Cook och Campbell (2002), Cartwright och Hardie (2012), samt SBU:s metodbok (2014), dels av vetenskapliga artiklar från företrädesvis psykologisk och psykiatrisk forskning från USA. Även andra forskningsområden såsom omvårdnadsforskning från Canada och humanistisk forskning från Australien förekommer. Tidskrifterna är nästan lika många till antalet som antalet artiklar som refereras till, d.v.s. närmare 40 tidskrifter, huvuddelen antingen brittiska eller amerikanska.

Evidensläget för Förskrivning av fysisk aktivitet

Inkludering av studier utifrån strukturerad forskningsfråga

För denna studies syfte är utfallsmåttet följsamhet till rekommendationer om fysisk aktivitet, d.v.s. graden av fysisk aktivitet, av intresse. Övriga vanliga utfallsmått för FFA är olika hälsoparametrar, men här är syftet enbart att utvärdera om FFA leder till ökad följsamhet till rekommendationer om fysisk aktivitet, inte den fysiska aktivitetens effekter. De hälsorelaterade måtten hade dock kunna fungera som validering av självskattningen av den fysiska aktiviteten, eftersom de kan anses vara mindre subjektiva.

Forskningsfrågan formulerades således:

Har förskrivning av fysisk aktivitet positiv effekt på fysisk aktivitetsnivå jämfört med ingen behandling eller sedvanlig vård, hos fysiskt inaktiva vårdtagare i primärvården?

Forskningsfrågan omarbetades till PICO (P = *population*, I = *intervention*, C = *control* och O = *outcome*) för att därefter omarbetas till en söksträng.

P: Vårdtagare i primärvården som inte når målen för rekommenderad fysisk aktivitet.

I: Förskrivning av fysisk aktivitet.

C: Sedvanlig vård eller ingen behandling.

O: Fysisk aktivitetsnivå, minst 6 månaders uppföljningstid.

Söksträng: ((referral AND (schemes OR scheme)) OR prescription) AND (exercise OR (physical activity))

Limitering: Studier av en avgränsad subgrupp med allvarlig sjukdom eller funktionsnedsättning inkluderades inte.

Enligt SBU (2014, ss. 54-55) kan en tidigare utförd systematisk översikt användas om PICO i denna överensstämmer med den aktuella forskningsfrågan och är av godtagbar kvalitet. Därför gjordes en sökning med aktuell söksträng i Pubmed med filter "review", och den senast publicerade systematiska översikten med relevant PICO visade sig vara av Pavey et al (2011). Kvalitetsgranskning enligt AMSTAR (A measurement tool to assess the methodological quality of systematic reviews) som är en checklista för granskning av systematiska översikter, resulterade i betyget hög kvalitet (se Bilaga 1) och studien kunde inkluderas. Av åtta inkluderade studier i den systematiska översikten av Pavey et al (2011), exkluderades tre på grund av irrelevans för PICO, se tabell 1a.

Pavey et al. uppger att de har gjort sökningar fram till datum 2011-07-01. För att identifiera artiklar som publicerats senare gjordes en sökning med aktuell söksträng i Pubmed utan filter "review" men med filter "fulltext", "controlled trial" och publikationsdatum efter 2011-07-01. Sökningen resulterade i 170 träffar varav sju abstract bedömdes vara relevanta, men vid läsning av fulltext kunde samtliga exkluderas eftersom de inte var relevanta för PICO. Samma söksträng användes i Cinahl med limitering "exclude medline records", "fulltext" och publikationsdatum efter 2011-07-01, med 46 träffar och ett relevant abstract. Efter läsning av fulltextartikel exkluderades även denna, se tabell 1b.

Totalt inkluderades fem studier (Taylor, Doust & Webborn, 1998, Stevens, Hillsdon, Thorogood & McArdle, 1998, Harrison, Roberts & Elton, 2005, Isaacs et al., 2007, Murphy et al., 2012), samtliga från den systematiska översikten av Pavey et al.

Tabell 1a Exkluderade artiklar ur Pavey et al:s systematiska översikt

Författare (år)	Skäl för exklusion
Gusi, Reyes, Gonzalez-Guerrero, Herrera & Garcia (2008)	Ej relevant utfallsmått (ej mätt fysisk aktivitetsnivå)
Jolly et al. (2009)	Ej relevant kontrollintervention (jämför olika former av FFA)
Sorensen, Kragstrup, Skovgaard & Puggaard (2008)	Ej relevant kontrollintervention (jämför olika former av FFA)

Tabell 1b Exkluderade fulltext-artiklar Pubmed och Cinahl

Författare (år)	Skäl för exklusion
Olsson, S. J., Börjesson, M., Ekblom-Bak, E., Hemmingsson, E., Hellénus, M. L., Kallings L. V. (2015)	Ej relevant utfallsmått (ej mätt fysisk aktivitetsnivå)
Schutz, Y., Nguyen, D. M., Byrne, N. M., Hills, A. P. (2014)	Ej relevant intervention (ej förskrivning från primärvård)
Rome, Persson, Ekdahl & Gard (2009)	Ej relevant kontrollintervention (jämför olika former av FFA)
Duda, J. L., Williams G. C., Ntoumanis, N., Daley, A., Eves, F. F., et al (2014)	Ej relevant kontrollintervention (jämför olika former av FFA)
Galaviz, Levesque & Kotecha (2013)	Ej relevant utfallsmått (ingen långtidsuppföljning)
Karjalainen, J. J., Kiviniemi, A. M., Hautala, A. J., et al (2012)	Ej relevant studiedesign (ej RCT)
Gademan, M. G., Deutekom, M., Hosper, K. et al (2012)	Ej relevant studiedesign (ej RCT)
Murphy, S. M., Edwards, R. T., Williams, N., et al (2012)	Dubblett. Redan inkluderad via systematisk översikt (Pavey, 2011)

Kvalitetsgranskning och sammanställning av studier

Samtliga inkluderade studier är kvalitetsgranskade avseende *risk of bias* av Pavey et al. med resultat hög eller medelhög studiekvalitet. *Risk of bias* innebär risk för snedvridet studieresultat p.g.a. systematiska fel i studiens genomförande. Systematiska fel kan uppstå p.g.a. inkorrekt utförd randomisering och tilldelning till grupper, icke blindad utvärdering, olikheter mellan grupperna vid baslinjemätning, ej rekommenderad analysmetod eller oklar rapportering av data (Pavey et al., 2011, table 2). Studier sammanfattas med utfallsmått av intresse för aktuellt PICO i tabell 2 där även sammanfattad studiekvalitet enligt Pavey et al. framgår. Fem studier, samtliga från Storbritannien, har utvärderat effekten av FFA. Interventionen har genomförts på ett jämförbart sätt i de olika studierna, med hänvisning från vårdgivare i primärvård till 10-16 veckors subventionerad träning vid friskvårdsanläggning. En inledande konsultation och både grupp- och individuell träning har erbjudits. Utfallsmått är i samtliga fall självskattnings av fysisk aktivitetsnivå, kompletterat med andra utfallsmått som inte berörs i denna uppsats.

Tabell 2 Included studies, data relevant to PICO presented

Author Year Reference Country	Study Design	Follow -up	Population Study groups Intervention (I) vs control (C)	Patients (n)	Mean age (years)	Men /women	Outcome variables	Study quality
Taylor et al. (1998) UK	RCT	26, 37 weeks	Smokers, hypertensive or overweight I = ERS, 10 weeks, 2x30-40 min/week, moderate intensity, leisure centre C = usual care/information	142	54	53/89	Mean duration of PA of at least moderate intensity (min/week)	Moderate overall risk of bias
Stevens et al. (1998) UK	RCT	8 months	Inactive adults I = ERS, 10 weeks leisure centre C = usual care	714	59	300/414	Mean number of PA occasions	Moderate overall risk of bias
Harrison et al. (2005) UK	RCT	6, 9, 12 months	Primary care inactive patients with CHD risk factors I = ERS 12 weeks, 2 x60 min/week, individually based, leisure centre C = usual care/information	545	NR	182/363	Percentage participating in at least 90 min of moderate/vigorous PA	Moderate overall risk of bias
Isaacs et al. (2007) UK	RCT	6 months	Hypertension, obesity, raised cholesterol, type 2 diabetes I = ERS 10 weeks, 2x45 min/week, leisure centre C = usual care	632	57	211/421	Mean (SD) duration of PA of at least moderate intensity (min/week) Mean (SD) duration of total PA	Low overall risk of bias
Murphy et al. (2012) UK	RCT	12 months	Mental health issues, CHD risk factors I= ERS 16 weeks, leisure centre C= usual care/information	2160	NR	745/1415	Mean duration of total PA (min/week)	Low overall risk of bias

ERS=Exercise referral scheme, NR=Not reported, CHD=Coronary heart disease, SD=Standard deviation

Resultatsammanställning och evidensgradering

Samtliga fem studier har utvärderat långtidseffekten av FFA på fysisk aktivitetsnivå vid sex månader eller längre tid efter avslutad FFA. Två av dessa visade signifikanta skillnader (tabell 3), den ena vid sex månader (Harrison et al., 2005) och den andra vid 12 månader (Murphy et al., 2012). I studien av Harrison et al. utvärderades effekten även vid 12 månader, då utan signifikanta skillnader mellan grupperna. Den andra studien, av Murphy et al., visade i primär analys med korrigering för alla kovariater vid baslinjen, ingen signifikant skillnad. I sekundär analys där kovariaten fysisk aktivitetsnivå vid baselinjen exkluderades blev skillnaden mellan grupperna däremot signifikant. Skillnader tidigare än sex månader efter avslutad intervention var inte i fokus för denna uppgift men det kan ändå konstateras att det i flera studier fanns signifikanta korttidseffekter.

Någon sammanslagning av data från flera studier har inte utförts i denna uppsats. Efter kontakt med artiklarnas författare gjorde Pavey et al. (2011) en meta-analys av fyra studier (de fem inkluderade i denna studie förutom Murphy et al.) av andelen personer som uppnådde 90-150 minuter fysisk aktivitet av minst måttlig intensitet per vecka vid 6-12 månaders uppföljning. Enligt originalartiklarnas data gav meta-analysen resultatet att det förelåg en 16 procents *relativ risk* (RR =1,16 (95% konfidensintervall 1,03 till 1,30)), att uppnå 90-150 minuter fysisk aktivitet per vecka vid 6-12 månaders uppföljning med FFA jämfört med vanlig vård. Pavey et al (2011) räknade också om data för att den skulle avse *intention-to-treat*, d.v.s. en beräkning av andelen av alla som randomiserats som nådde 90-150 minuter fysisk aktivitet och inte bara de som fullföljt studien, och då var skillnaden mellan grupperna inte längre signifikant (RR=1,11 (95% konfidensintervall 0,99 till 1,25)).

Vid evidensgradering enligt GRADE gjordes avdrag på evidensstyrkan för studiekvalitet (-1) p.g.a. brister avseende blindning av utvärderare, blindning vid allokering och bortfallshantering. *Precision*, *consistency* och *directness* bedömdes tillsammans utgöra brister som motiverade en sänkning (-1). *Precision* är osäker p.g.a. breda konfidensintervall och avtagande effekt över tid. *Consistency* är något oklar, vilket beror på låg effekt och låg precision som därmed gör att studiernas resultat blir signifikanta i en del studier och icke signifikanta i andra. *Directness* påverkas av att bortfall före randomisering är stort i flera studier och att sättet att inkludera deltagare skiljer sig något mellan studierna.

Resultat av studierna var för sig samt bedömning av respektive studies *risk of bias*, *directness* och *precision* sammanfattas i tabell 3. Sammanställd GRADE-bedömning redovisas i tabell 4. Utförligare beskrivning av GRADE-bedömning framgår i Bilaga 2.

Slutsats av systematisk översikt och GRADE

Förskrivning av fysisk aktivitet har ingen effekt på fysisk aktivitetsnivå sex månader eller senare efter avslutad intervention. Det vetenskapliga underlaget bakom denna slutsats är begränsat och baseras på fem RCT ingående i en systematisk översikt. Det innebär att det skulle behövas fler studier av hög kvalitet för att med säkerhet kunna bedöma långtidseffekten av FFA på fysisk aktivitetsnivå.

Grundläggande antaganden varpå slutsatsen vilar

Den arbetsgång som följts i detta exempel och leder fram till konklusionen att FFA inte har någon effekt på fysisk aktivitetsnivå sex månader eller senare efter avslutad intervention, vilar på grundläggande antaganden om vad en effekt är och hur den mäts. Detta är av betydelse för analysen av följsamhetsaspekter som utgör det andra delsyftet för denna studie.

Evidensgraderingen utgörs av två bedömningar som i idealfallet är oberoende av varann. Dels bedöms effekten av en intervention, dels bedöms evidensstyrkan vilken ger ett mått på det vetenskapliga underlagets styrka, d.v.s. hur säker effektmätningen är. När effekten av interventionen inte är statistiskt signifikant kommer evidensstyrkan avse hur säkert det är att interventionen inte har effekt.

Bedömningen av en interventions effekt grundas ytterst i antagandet att det sanna värdet som studierna syftar till att närma sig utgörs av medelvärdet för en population. Att en stor del av populationens värden ligger långt från detta medelvärde innebär bara en ökad svårighet att bedöma medelvärdet med precision. Enstaka individers reaktion på interventionen är endast av intresse vad gäller icke önskvärda bieffekter. Inte heller har det för effektmätningen någon betydelse huruvida personerna faktiskt varit följsamma till interventionen, eftersom alla deltagare ska räknas enligt den rekommenderade analysmetoden *intention-to-treat*.

Effekten av en intervention utgörs av skillnaden mellan resultatet för de som utsätts för interventionen och de som inte utsätts. Effekten kan beräknas och presenteras på olika sätt med det görs alltid en beräkning av om skillnaden mellan grupperna sannolikhetsmässigt skulle kunna bero på slumpen. I så fall är skillnaden icke-signifikant och det innebär att slutsatsen blir att interventionen inte har någon effekt.

I värderingen av evidensstyrka görs en bedömning av risken för systematiska fel. Först bedöms risken för att systematiska fel uppstått i respektive studie som då kan ha påverkat det medelvärde som man söker. Därefter görs även en bedömning av systematiska fel vid sammanslagning av studier som kan bero på att studiepopulationerna skiljer sig åt i de olika studierna, eller att de skiljer sig mellan studierna och den population som den systematiska översikten avser, eller att det finns risk för publikationsbias.

Tabell 3 Outcome variable physical activity

Author (year)	Country	Study design	Number of patients n=	With-drawals Drop outs	Results		Risk of bias	Directness	Precision
					Intervention	Control			
Taylor et al. (1998)	UK	RCT	n = 142 I = 97 C = 45	I = 40 C = 31	Mean (SD) min/week duration of PA of at least moderate intensity of completers 26 w 183 (39) 37 w 158 (38)	Mean (SD) min/week duration of PA of at least moderate intensity of completers 26 w 206 (45) 37 w 162 (44) <i>ns between groups</i>	?	?	?
Stevens et al. (1998)	UK	RCT	n= 714 I= 363 C= 351	I= 163 C= 136	Mean PA occasions in the 4 weeks before follow-up 8 months: 5.95	Mean PA occasions in the 4 weeks before follow-up 8 months: 4,43 <i>1,52 (CI 1,14 to 1,95) between groups</i>	?	?	?
Harrison et al. (2005)	UK	RCT	n= 545 I= 275 C= 270	I= 120 C= 113	At least 90 min of moderate/vigorous PA past week of completers 6 months: 22,6% (38/168) 9 months: 24,2% (36/149) 12 months: 25,8 % (40/155)	At least 90 min of moderate/vigorous PA past week of completers 6 months: 13,6% (22/162) 9 months: 22,1% (31/140) 12 months: 20,4 % (32/157) <i>p=0,05 at 6 months ns at 9 and 12 months between groups</i>	?	?	?
Isaacs et al. (2007)	UK	RCT	n= 632 I= 317 C= 315	I= 124 C= 95	Mean (95% CI) duration of PA of at least moderate intensity min/week 6 months 65 (55 to 77) Total duration PA min/week 6 months 692 (641 to 748)	Mean (95% CI) duration of PA of at least moderate intensity min/week 6 months 58 (49 to 69) Total duration PA min/week 6 months 647 (600-699) <i>ns between groups</i>	?	?	?
Murphy et al. (2012)	UK	RCT	n= 2160 I= 1080 C= 1080	I~189 C~176	Median duration of total PA min/week at 12 months <i>Odds ratio: with all covariates included 1,18 (CI 0,99 to 1,42) ns With baseline covariate for PA omitted: 1,19 (CI 1,00 to 1,42) p<0,05</i>		+	?	?

PA=Physical activity, I= Intervention, C=Control, SD=standard deviation, ns=non significant, CI=Confidence interval

Tabell 4 Summary of findings. Long-term effects on physical activity.

Outcome variable (Number of studies) Number of patients	Design	Limitations	Consistency	Directness	Precision	Publication bias	Magnitude of effect	Relative effect	Quality of evidence grade
Physical activity (5) 4193	RCT	-1	0?	0?	0?	Unlikely	Not relevant	Non significant See text and table 3	⊕⊕○○ Low

Följsamhetens problematiker

I föregående avsnitt kunde konstateras att det verkar som att FFA inte har någon långtidseffekt på fysisk aktivitet men att evidensen för denna slutsats är begränsad. Det leder i sin tur till slutsatsen att det behöver utföras fler studier för att säkerställa den eventuella långtidseffekten av FFA. Metoden som användes för att nå dessa slutsatser kan ses som evidensbaserings gyllene standard och används av SBU och internationella aktörer (Statens beredning för medicinsk utvärdering, 2014, s. 141). I uppsatsen kommer fortsättningsvis dessa slutsatser sättas i ett större perspektiv, utan att för den skull frånga evidensbaserings mål att hälso- och sjukvården ska baseras på ett vetenskapligt underlag. Den vetenskapliga litteraturen om följsamhet används för att analysera aspekter av följsamhet i de granskade studierna och syntetiseringen av studierna. Samtidigt analyseras också den vetenskapliga litteraturen om följsamhet.

Vid läsning av vetenskaplig litteratur om följsamhet framstod två teman. Många artiklar berör de etiska aspekterna av följsamhet. Andra berör följsamhetsproblematikens betydelse för möjligheten att fastställa kausala samband. I följande redovisning kommer denna uppdelning följas.

Det första avsnittet handlar således om vårdtagares följsamhet till livsstilsrekommendationer och hur evidensrörelsens mål att implementera evidensbaserade metoder kan stå i konflikt med etiska ideal. I detta avsnitt behandlas också rationaliteten bakom antagandet att FFA leder till ökad följsamhet till riktlinjer avseende fysisk aktivitet. Detta antagande vilar på en mängd andra mer eller mindre uttalade antaganden om vad en låg följsamhet till riktlinjer beror på, var ansvaret för en låg följsamhet ligger och med vilken rätt hälso- och sjukvården kan påverka vårdtagare till ökad följsamhet.

I det andra avsnittet fokuseras på hur evidensrörelsens ideal avseende statistiska metoder och synen på kausalitet styr forskningen om FFA och följsamhet till riktlinjer avseende fysisk aktivitet.

Därefter följer ett avsnitt där dessa teman fogas samman. Här ligger fokus på att belysa hur evidensbaserings olika ideal kan komma i konflikt med varandra vad gäller interventioner som FFA. Här uppmärksammas också vilka kunskapsstyper som utelämnas i det formaliserade ideal som evidensrörelsen byggt upp.

Men först ges en kort beskrivning av i vilka avseenden och på vilket sätt de granskade studierna problematiserar eller redovisar följsamhet.

De granskade studiernas hantering av följsamhet

I ingen av de fem artiklarna förs en diskussion om orsaker till att följsamheten till riktlinjer avseende fysisk aktivitet är låg och med vilka mekanismer man antar att FFA skulle förändra detta. Det kan verka anmärkningsvärt eftersom det är själva utgångspunkten för studierna. Samtidigt är det en del av det formaliserade ideal som byggts upp inom evidensrörelsen där kvantitativa studier testar effekten av en intervention genom att jämföra den med en annan intervention. För den jämförelsen krävs inga resonemang om mekanismer. De tidskrifter som publicerar denna typ av studier har ofta begränsningar i antal ord och höga krav avseende redovisning av studiernas interna validitet, vilket gör att författarna även om de skulle vilja, svårligen kan skapa utrymme för resonemang om mekanismer. Indirekt, kan man ändå i två av studierna ana de förväntade mekanismerna av FFA. I studien av Taylor et al. mättes deltagarnas förändringsstadium enligt den transteoretiska modellen för beteendeförändring. Enligt den transteoretiska modellen är det viktigt att möta personen på den nivå som denne befinner sig för att öka motivationen (Huang et al., 2015, s. 203), men det framgår inte om eller hur detta skett i studien. I studien av Murphy et al. användes motiverande samtal. Huruvida det motiverande samtalet ska ses som en metod för att öka *adherence* till FFA eller som en del i FFA för att öka den fysiska aktiviteten, är dock oklart.

Följsamhet till fysisk aktivitet, dvs. studiernas utfall, har mätts på likartat sätt i alla studier genom att deltagarna fått uppge sin aktivitetsnivå, antingen i en dagbok eller genom en skattning i efterhand. Alla studier utom en (Stevens et al., 1998) har använt en självskattningsmetod kallad 7D-PAR som är validitets- och reliabilitetstestad med gott resultat (Isaacs et al., 2007). Det går dock inte att komma ifrån att självskattningar är subjektiva och därmed ökar risken för bedömningsbias (Statens beredning för medicinsk utvärdering, 2014, bilaga 2), vilket skapar en osäkerhet i evidensgraderingen. Studier har också visat att ett problem med följsamhetsmätningar kan vara att mätningen i sig påverkar följsamheten (Shi et al., 2010, s. 1103, Vermeire et al., 2001, s. 26). Även om dessa studier har gällt följsamhet till medicinering är det inte osannolikt att det gäller även annan följsamhet.

När det gäller följsamhet till interventionen så är skillnaderna mellan studierna stora. Fyra av fem studier har redovisat *adherence* på något sätt, en studie har inte redovisat följsamhet till interventionen över huvud taget. De olika metoderna för att mäta följsamhet får betydelse både för generalisering av studieresultat (Barbosa, Balp, Kulich, Germain & Rofail, 2012, ss. 46-47) och för sammanläggning av studieresultat i systematiska översikter (Dodd, White & Williamson, 2012, s.13). Inte minst visar de olika sätten att mäta följsamhet att interventionen har definierats på olika sätt i studierna.

I artikeln av Taylor et al. uppges att *adherence* uppgick till 9,1 av 20 föreskrivna träningstillfällen och att 28% genomförde minst 75% av de föreskrivna tillfällena. *Adherence* mättes med datoriserad registrering vid träningscentret. Signifikanta skillnader har uppmätts mellan de med hög respektive låg *adherence*, med bättre resultat för de med hög *adherence*. Dock redovisas dessa effekter bara för andra utfallsmått än fysisk aktivitet. Ett starkt samband mellan de som var s.k. *completers* - och alltså utvärderades fullt ut - och de som hade hög *adherence*, har också uppmätts. De konstaterar också att överviktiga och icke-rökare hade bättre *adherence* än icke överviktiga och rökare, vilket de menar kan få implikationer för urval av personer för FFA. De drar till sist slutsatsen att för att minska hjärt-kärl-risikfaktorerna krävs fokus på att öka *adherence* till FFA.

I studien av Steven et al. är interventionen inte definierad i antal tillfällen utan innehåller en första konsultation med rådgivning angående fysisk aktivitet och en andra konsultation efter 10 veckor för att diskutera framsteg. De har mätt att 35% av de som randomiserats till interventionsgruppen deltog vid första konsultationen och 25% vid andra konsultationen. De konstaterar också att en större andel kvinnor deltog i första konsultationen och att andra metoder för att rekrytera män behövs.

Harrison et al. uppger inte att de har mätt följsamhet över huvud taget. I ett flödesschema över studien redovisas dock att alla i både interventionsgrupp och kontrollgruppen mottog intervention respektive kontroll-intervention. Troligast är att Harrison et al. definierar interventionen som själva hänvisningen till den första konsultationen. Därav finns det heller ingen följsamhet till interventionen att mäta.

Studien av Isaacs et al. är publicerad i Health Technology Assessment och omfattar inte mindre än 184 sidor totalt inklusive bilagor. Denna studie har jämfört FFA med dels bara rådgivning men också med en intervention som innebar ledarledda promenader. Bakgrunden till studien är mer gediget beskriven än i andra artiklar, men består framför allt i en beskrivning av tidigare studier av FFA. Slutsatsen de drar av studien är att alla tre interventioner hade effekt på fysisk aktivitetsnivå, utan signifikanta skillnader mellan grupperna. *Adherence* till interventionen mättes på två sätt, dels genom dagböcker, dels genom registrering vid aktiviteterna. Dessa mätningar var inte helt samstämmiga men det konstaterades att 42% av deltagarna nådde 75-100% *adherence*.

Murphy et al. uppger att 43,8% fullföljde det 16 veckor långa programmet. Subgruppsanalyser i denna studie visade att *adherence* hade stor betydelse för utfallet. Bland de med hög följsamhet var samtliga utfall signifikant bättre.

Resultatet av studierna är helt avhängigt synen på följsamhet. I två av artiklarna (Harrison et al., 2005, Isaacs et al., 2007) presenteras resultatet av studierna som en skillnad mellan grupper av de som fullföljt studien. Huruvida de varit följsamma till behandlingen är däremot inte klart, men bortfallet är

alltså inte inkluderat i studien, och det är kanske sannolikt att de personer som utgör bortfallet också haft en lägre följsamhet till interventionen. Detta sätt att hantera bortfall är inte enligt rekommendation enligt SBU. I den systematiska översikten av Pavey et al (2011) hanterades detta genom att räkna om resultaten av studierna, vilket fick till följd att det som i studierna presenterades som ett positivt resultat för FFA blev i den systematiska översikten inte längre en signifikant skillnad mellan grupperna.

I de granskade studierna har redovisningen av följsamhet till interventionen alltså hanterats på olika sätt. Det som ändå framgår tydligt av den redovisning som gjorts är att följsamheten är låg när de förskrivna träningsstillfällena räknas som en del av interventionen. Sannolikt är följsamheten i studierna, om den skulle syntetiseras, klart lägre än 50%. Det som däremot inte framgår tydligt i någon av artiklarna är hur de etiska aspekterna hanterats. Utgångspunkten är ju att vårdgivarna står inför uppgiften att försöka påverka vårdtagarna, tillika försökspersonerna, till en förändrad livsstil vilket innebär etiska avvägningar.

Etiska aspekter av följsamhet

En historiskt förändrad syn på följsamhet

Compliance är det äldsta av de tre mest frekvent använda begreppen för att beskriva följsamhet. Termen härstammar från en tid då det togs för självklart att målet var att vårdtagaren skulle följa vårdgivarens rekommendation. En av de mest citerade definitionerna av *compliance* lyder ”*the extent to which a person’s behavior (in terms of taking medications, following diets, or executing lifestyle changes) coincides with medical or health advice*” (Haynes, Taylor & Sackett, 1979, ss. 1-2). *Compliance* eller snarare *noncompliance* uppmärksammades mest som ett problem i avseendet att vårdtagarna inte följde läkarens råd eller ordination i tillräckligt hög utsträckning, och därför inte fick den bästa möjliga behandlingen (Marston, 1970). *Noncompliance* klassificerades 1973 av North American Nursing Diagnosis Association (NANDA) som en omvårdnadsdiagnos, och blev i och med det en diagnos som sjuksköterskor kan ställa (Russell, Daly, Hughes & Hoog Co, 2003, s. 283). Några år senare började diagnosen *noncompliance* kritiserars (Stanitis & Ryan, 1982), likväl som begreppet *compliance* kom att kritiserars av allt fler för sin paternalistiska utgångspunkt där vårdtagarens röst inte får gehör (Donovan & Blake, 1992, Lutfey & Wishner, 1999).

Som en följd av denna kritik började begreppet *adherence* att användas mer frekvent. *Adherence* till långvarig behandling har av WHO (2003, s. 17) definierats som ”*the extent to which a person’s behaviour – taking medication, following a diet, and/or executing lifestyle changes, corresponds with agreed recommendations from a health care provider*”. Denna definition lyfter till skillnad från definitioner av *compliance* att det är fråga om en överenskommelse mellan vårdtagare och vårdgivare, och WHO betonar i rapporten att *adherence* inte ska förväxlas med *compliance*. Detta till trots har både *adherence* och *compliance* kritiserats för att vara paternalistiska och sakna vårdtagarperspektivet (Russell et al., 2003, Bissell, May & Noyce, 2004), och Bissonette (2008) fann att de båda begreppen ofta används som synonymier. *Noncompliance* kvarstår som diagnos i NANDA (Herdman, Kamitsuru & International, 2015) men det kan noteras att begreppet *nonadherence* används i dess diagnosbeskrivning.

Både *compliance* och *adherence* negligerar den normativa frågan huruvida det är rätt eller fel att följa vårdgivarens rekommendation, menar Chakrabarti (2014, s.34). *Informed adherence* har ibland använts som en kompromiss mellan *informed consent* och *adherence*, för att ytterligare betona att rekommendationen från vårdgivaren föregås av en diskussion mellan vårdtagare och vårdgivare där vårdtagaren står inför ett val att antingen följa vårdgivarens rekommendation eller inte (Chakrabarti, 2014, s.33). I definitioner av *concordance*, som har likställts med patientcentrerad vård (Bell, Airaksinen, Lyles, Chen & Aslani, 2007, s.710), förordas diskussionen inte bara i syfte att öka vårdtagarens autonomi utan också för att relationen mellan vårdgivare och vårdtagare i sig är viktig (Felzmann, 2012). *Concordance* handlar till skillnad från *adherence* och *compliance* alltså inte om

vårdtagarens slutliga beteende utan om dennes upplevelse av konsultationen, menar Hobden (2006, s. 258). Snowden (2014, s. 57) gjorde en begreppsanalys av *concordance* och kom fram till följande definition: "The process of enlightened communication between the person and the healthcare professional leading to an agreed treatment and ongoing assessment of this as the optimal course". Han konstaterade också att det finns skillnader mellan olika yrkesgruppers uppfattning om innebörden i *concordance*.

Den historiska utvecklingen av synen på följsamhet grundas till stor del i ett etiskt dilemma där vårdgivarens kunskap om bästa möjliga hälso- och sjukvårdsintervention vägs mot vårdtagarens rätt att välja. Begreppens förekomst i den vetenskapliga litteraturen, där *adherence* i hög grad ersatt *compliance*, symboliserar en utveckling där vikten av vårdtagarens integritet och fria val har uppvärderats. Denna utveckling har skett parallellt med att vårdtagarens önskemål och situation värderas högre i de senare definitionerna av EBM. De flesta verkar i dag anse att om möjligt bör vårdtagaren vara delaktig i beslut om behandling. Dock finns det situationer då vårdtagaren inte bedöms ha förmågan att själv fatta långsiktigt riktiga beslut och i dessa fall kan paternalism och tvångsvård försvaras. I dessa fall görs en bedömning av risk och nytta för vårdtagaren (Felzmann, 2012, s. 407).

Idag förekommer alla tre begreppen i vetenskaplig litteratur, med skiftande definitioner. Det finns förespråkare för *concordance* som menar att utvecklingen att frångå *compliance* som följsamhetsbegrepp symboliserar ett paradigmskifte (Lutfey & Wishner, 1999, s.635), medan andra har tonat ner skillnaderna (Spiro, 2001, s. 5).

Evidensrörelsens och individens olika perspektiv på risk och nytta

Flera studier tyder på att hälften av alla vårdtagare med långvarig behandling för fysiskt eller psykiskt sjukdomstillstånd inte är följsamma till ordinerad medicinering (World Health Organisation, 2003, s. 7). I en studie av patienter med diabetes typ 2 var det fem till sex gånger vanligare att patienterna upplevde problem med följsamhet till rekommenderad diet och fysisk aktivitet än till medicinering (Moreau et al., 2009, s. 151). För patienter med kardiovaskulär sjukdom har dock följsamheten till mediciner och annan behandling visats vara jämförbar (Klotsche et al., 2011, s. 90). WHO (2010, ss. 54-55) rekommenderar minst 150 minuter aerob fysisk aktivitet av måttlig intensitet eller minst 75 minuter aerob fysisk aktivitet av hög intensitet eller likvärdig kombination av dessa, samt muskelstärkande träning minst två gånger i veckan. Dåvarande Statens Folkhälsoinstitut (ersatt av Folkhälsomyndigheten 2014) uppskattar att ungefär halva den svenska befolkningen uppnår den svenska rekommendationen om minst 30 min måttlig fysisk aktivitet dagligen (*FaR® : individanpassad skriftlig ordination av fysisk aktivitet*, 2011, ss. 12-13). I England har 20-40% av befolkningen beräknats nå målet om fem tillfällen med 30 minuter fysisk aktivitet i veckan (Health and Safety Executive, 2009).

För rekommendationer om fysisk aktivitet finns det således ett stort glapp mellan å ena sidan evidensrörelsens strävan att låta vetenskaplig kunskap om effekterna av fysisk aktivitet och inaktivitet styra, och å andra sidan vårdtagarnas följsamhet till dessa rekommendationer. Det väcker frågan om vad detta glapp orsakas av. Ett sätt att studera saken är att kartlägga orsaker till låg följsamhet, vilket gjorts men då framför allt vad gäller följsamhet till medicinering. Faktorer av betydelse för följsamhet delas av WHO (2003, ss. 27-31) in i fem interagerande domäner: socio-ekonomiska förhållanden, hälso- och sjukvårdssystem, patientrelaterade, terapielaterade och sjukdomsrelaterade faktorer. I en systematisk översikt av 51 tidigare systematiska översikter av följsamhet till medicinering summerades inte mindre än 771 determinanter för följsamhet (Kardas, Lewek & Matyjaszczyk, 2013, s.1), där WHO:s alla fem domäner var väl representerade. Även om denna studie gjordes på följsamhet till medicinering är det sannolikt att följsamhet till livsstilsrekommendationer har minst lika många determinanter. Många vårdtagare får dessutom flera rekommendationer och/eller ordinationer som kräver följsamhet på flera sätt. När det gäller diabetes t.ex. så får varje vårdtagare individanpassade rekommendationer som varierar över tid och spänner över flera områden, från mediciner till fotvård

och fysisk aktivitet (Hearnshaw & Lindenmeyer, 2006, s. 726). I en kvalitativ studie av 1358 personer som varit icke-följsamma till FFA i Sverige konstaterades att sjukdom och smärta var vanliga orsaker bland äldre personer och ekonomiska faktorer och tidsbrist var vanligast angivna orsaker bland de yngre. Det fanns även skillnader i orsaker till icke följsamhet beroende på diagnos och kön (Leijon, Faskunger, Bendtsen, Festin & Nilsen, 2011, s. 237).

Den låga följsamheten skulle också kunna förklaras med att evidensrörelsens perspektiv skiljer sig från den individuella vårdtagarens perspektiv vad gäller risk och nytta med fysisk (in-)aktivitet. Det är förvisso vetenskapligt säkerställt att fysisk inaktivitet ökar risken både för ett antal sjukdomar och för att livet förkortas. Men risken är långsiktig och det handlar just om en risk vilket innebär att *sannolikheten* för sjukdom och förkortat liv ökar vid fysisk inaktivitet. På gruppnivå kommer den ökade sannolikheten innebära att fler drabbas av sjukdom och förtidig död, men många individer kan trots fysisk inaktivitet leva ett långt och friskt liv. Eftersom man inte på förhand vet vilka som kommer drabbas måste den evidensbaserade hälso- och sjukvården eftersträva att så många människor som möjligt lever efter de rekommendationer som gör att så få människor som möjligt drabbas. Att så få som möjligt drabbas innebär mindre lidande totalt sett och det innebär också en effektivare resursanvändning för hälso- och sjukvården, vilket gynnar alla medborgare. Det innebär att bedömningen av risk/nytta inte kan begränsas till en individ utan också måste göras på gruppnivå. Det är dock inte självklart att bedömningen av risk/nytta görs på samma sätt av den enskilda individen, vilket kan yttra sig som bristande följsamhet. Samtidigt är det inte alltid så att förskrivning av läkemedel och andra interventioner är nödvändiga eller ens rätt i den meningen att de leder till avsedd effekt. Att inte följa rekommendationen kan i dessa fall spara resurser och minimera biverkningarna (Felzmann, 2012, s. 406). När det gäller fysisk aktivitet är dock denna risk liten eftersom biverkningarna sällan är allvarliga och samhällskostnaderna är låga.

Bernardini (2004, s. 223) åskådliggör med exemplet blodtryckssänkande medicinering hur vårdgivare och vårdtagare väger in olika aspekter i värderingen av risk och nytta. Vid milt förhöjt blodtryck har blodtryckssänkande mediciner visat sig ha god effekt på gruppnivå genom en stor minskning av antalet allvarliga utfall. Detta är grunden för vårdgivarens ordination. Däremot är risken för varje individ med milt förhöjt blodtryck relativt låg, samtidigt som risken för biverkningar av medicinering är betydligt högre. Både den något minskade risken för allvarligt utfall och risken för eller redan upplevda biverkningar kan värderas olika av vårdtagare, beroende på kulturella och personliga preferenser. En del personer värderar varje möjlighet att förlänga sitt liv högt medan det för andra kan finnas ett stort motstånd mot medicinering över huvud taget. Avvägningen kommer således resultera i olika beslut där några vårdtagare är följsamma och några är det inte. Den vårdtagare som inte följer vårdgivarens ordination ser sig dock sällan som icke följsam utan anser sig ha tagit ett rationellt beslut för att undvika biverkningar, att testa effekten av en medicin, eller att undvika stigmatisering kring sjukdom eller ohälsa, konstaterar Bernardini (2004, s. 223). Ett liknande resonemang kan föras för fysisk aktivitet. Det är vetenskapligt fastställt att fysisk aktivitet har god effekt på gruppnivå och det är därför rimligt att den evidensbaserade hälso- och sjukvården eftersträvar en generellt ökad fysisk aktivitet hos de som ligger i riskzonen eftersom det kan rädda liv och lidande. För den enskilda individen är det dock andra aspekter som vägs in. Även om fysisk aktivitet inte har så många potentiella biverkningar så är det en åtgärd som kräver mycket av vårdtagaren vilket kan leda till mer eller mindre medvetna val av vårdtagare att inte följa riktlinjer avseende fysisk aktivitet.

Vad är målet och var ligger ansvaret?

I arbetet med att öka följsamheten till livsstilsrekommendationer är frågan också om det ytterst handlar om vilken attityd och vilka metoder som har störst chans att öka följsamheten, eller om egenvärdet i att vårdgivaren bemöter vårdtagaren med respekt och väger in dennes situation och önskemål, är viktigare. I det senare fallet är det inte givet att målsättningen i varje individuellt fall är ökad följsamhet. Kaufman och Birks (2011, ss. 13-15) menar att det inte finns några belegg för det ibland underliggande antagandet att *concordance* skulle leda till ökad *adherence*. Samtidigt är det heller inte självklart att ett delat beslutstagande alltid välkomnas av vårdtagare. En kvalitativ studie (Du Pasquier

& Aslani, 2008, s. 846) av vårdtagares syn på *concordance* och *adherence* i primärvården visade att de var positiva till ökad information från läkare men tveksamma till att själva vara delaktiga i beslut. Denna studie gällde dock mediciner och huruvida det är överförbart på andra åtgärder är osäkert. Det har också mycket riktigt påpekats att hela systemet för hantering av mediciner grundas i en paternalistisk syn där vårdtagaren inte själv kan välja de mediciner som önskas (Horne R, Weinman J, Barber N, Elliott R & M., 2005, s. 117). Så i hur hög grad som hälso- och sjukvården faktiskt kan grundas i *concordance* är inte heller självklart.

Förskrivning av fysisk aktivitet kan också ses som ett sätt för hälso- och sjukvården att ta sitt ansvar för att öka *adherence* till riktlinjer avseende fysisk aktivitet, eftersom *adherence* till skillnad från *compliance* ses som ett delat ansvar mellan vårdtagare och vårdgivare. Istället för att konstatera att en stor del av befolkningen är *noncompliant* och lägga ansvaret för detta på de enskilda individerna, så tar hälso- och sjukvården sitt ansvar och utformar en intervention för att främja implementeringen. Det som däremot är mindre väl motiverat är hur denna intervention ska fungera, det vill säga med vilka mekanismer som FFA antas leda till långsiktigt ökad fysisk aktivitet. Även om interventionen beskrivs snarlikt i de granskade artiklarna så finns det en viktig skillnad. I några av artiklarna nämns modeller för beteendeförändring som en ingrediens i interventionen. Hur viktiga dessa är som en del av interventionen är dock oklart eftersom de i andra studier inte nämns överhuvudtaget, trots att interventionerna benämns och i övrigt beskrivs mycket lika. Eftersom forskning visat att följsamhetsproblematik har ett mycket stort antal determinanter är det rimligt att tro att det krävs individanpassade åtgärder för att öka följsamhet. Det är också ett rimligt antagande att den låga följsamheten till fysisk aktivitet till viss del kan förklaras som bristande motivation. Interventioner som syftar till ökad motivation och stöd för beteendeförändring är då logiska. Det är kanske inte lika logiskt att en förskrivning i sig skulle ha effekt, men det är ett sätt för hälso- och sjukvården att stärka det hälsofrämjande och sjukdomsförebyggande perspektivet (Folkhälsomyndigheten). Förskrivningen ger tyngd åt rekommendationen och visar att det är en åtgärd som är lika viktig som att vårdtagaren tar sina mediciner. Å andra sidan är vårdtagarens respekt för förskrivning av mediciner uppenbarligen inte heller särskilt hög. Enligt WHO (2003, s. 21) har beräkningar gjorts att bara hälften av alla med kronisk sjukdom medicinerar enligt ordination i industriländer. I utvecklingsländer bedöms siffran vara avsevärt lägre ändå.

Följsamhetens betydelse för antaganden om kausalitet

Kausalitet i evidensrörelsen

Inom EBM är RCT:er och systematiska översikter av företrädesvis RCT:er de högst värderade källorna till kunskap om en interventions effekt (Statens beredning för medicinsk utvärdering, 2014, s. 43). Dessa tillvägagångssätt för att försöka fastställa kausala samband grundas i en kontrafaktisk definition av kausalitet. Det innebär att effekten av en intervention definieras som den hypotetiska skillnaden mellan scenariot om personen utsätts respektive inte utsätts för en intervention. Eftersom det är omöjligt att studera två alternativa scenarier för en person studeras den kausala effekten som medelvärde för en grupp som utsätts för interventionen, vilket jämförs med en kontrollgrupps medelvärde. Kontrollgruppens intervention får utgöra en approximation av det kontrafaktiska utfallet. En förutsättning för att den kontrollerade studien ska ge rättvisande resultat är att grupperna är lika i alla avseenden av betydelse för utfallet, förutom interventionen. Randomiseringen nyttjar slumpen för att uppnå detta syfte (Shadish et al., 2002 s. 1-13).

Resultatet av en RCT kan räknas fram på olika sätt beroende på hur forskningsfrågan är ställd. Den analysmetod som förordas inom EBM är *intention-to-treat* (ITT) -analysen (Statens beredning för medicinsk utvärdering, 2014, Bilaga 2:8). Denna analysmetod utförs genom att samtliga deltagare som randomiseras innefattas i analysen, oavsett om de varit följsamma till interventionen och oavsett om de varit tillgängliga för utvärdering. *Intention-to-treat*-analysen svarar således på frågan om vilken den genomsnittliga effekten blir för samtliga studiedeltagare, där man kan anta att några deltagare har genomfört interventionen exakt enligt studieprotokollet, några har delvis följt interventionen och några

har sannolikt inte tagit del av interventionen över huvud taget. Likaledes beräknas medelvärdet på alla deltagare oavsett om de funnits tillgängliga för utvärdering, vilket innebär att de saknade värdena måste ersättas på något sätt. En s.k. imputeringsmetod gör detta möjligt, där man på ett standardiserat sätt ersätter de saknade värdena. Detta kan ske på olika sätt, t. ex. med aktuellt baslinjevärde, vilket innebär att bortfallet förutsätts ha oförändrade värden efter interventionen (Statens beredning för medicinsk utvärdering, 2014, Bilaga 2:9).

Att alla randomiserade deltagare innefattas i analysen innebär att ITT-analysen som metod tar full hänsyn till randomiseringen och därigenom ger förutsättningar för hög intern validitet (Sagarin et al., 2014, ss. 320, 328). Intern validitet värderas högt inom EBM och det är såtillvida inte förvånande att ITT-analysen rekommenderas. Att ITT-analysen förordas innebär också att man inom EBM prioriterar forskningsfrågor som avser att bedöma effekten av en policy, d.v.s. effekten av att hänvisa en grupp personer till en intervention. För denna forskningsfråga bryr man sig således inte om ifall personerna verkligen mottar interventionen. Denna forskningsfråga är ofta den mest intressanta för en beslutsfattare som ska avgöra om ett brett införande av en intervention bör finansieras. En alternativ forskningsfråga, som man inte får svar på med en ITT-analys, är vilken effekt de personer skulle få som skulle vara följsamma till interventionen, d.v.s. vilken effekten är av att motta interventionen. Denna fråga är ofta mer intressant för en enskild vårdtagare och i det aktuella fallet även för dennes vårdgivare. För en vårdgivare som funderar på att byta inriktning på behandling för samtliga sina vårdtagare är fortfarande policy-frågan mer intressant (Shrier et al., 2014, s. 29).

Ett annat sätt att beskriva vad ITT-analysen mäter är att den ger en beräkning av kombinationen interventionens kausala effekt och implementeringen av interventionen. I detta avseende kan ITT-analysen sägas mäta *effectiveness* snarare än *efficacy*, där *effectiveness* avser effekt i en så klinisknära miljö som möjligt med de begränsningar som där finns avseende bl. a. följsamhet. *Efficacy* däremot mäter interventionens kausala effekt vilken mäts uteslutande för de som varit följsamma och mottagit interventionen fullständigt (Dodd et al., 2012, s. 2). Den effekten är därmed oftast omöjlig att uppnå på gruppnivå i kliniken. När ITT-analysen ger svaret att effekten av en intervention är låg innebär det att *effectiveness* är låg, men inte nödvändigtvis att *efficacy* är låg, eftersom det kan ha varit implementeringen som inte lyckats. *Intention-to-treat*-analyser ger bara en rättvisande bild av den kausala effekten av interventionen om följsamheten är fullständig, annars underestimerar de oftast men de kan även överestimera och under vissa omständigheter till och med visa en förändring i fel riktning, eftersom det är möjligt att effekten för de följsamma och de icke följsamma skiljer sig åt på ett systematiskt sätt (Sagarin et al., 2014, s. 320). Att ITT-analys vid låg följsamhet ger en skattning av interventionens *effectiveness*, vilar dock på ett viktigt antagande. En förutsättning är nämligen att följsamheten i verkligheten är den samma som i studien, vilket i praktiken sällan kan garanteras (Sagarin et al., 2014, s. 320). De tillfrågades preferenser kan påverka vilka som väljer att delta i studien vilket innebär att de som randomiseras är en subgrupp av populationen med en eventuell annan följsamhet (Sidani, Miranda, Epstein & Fox, 2009, s. 56). Att så ibland troligen är fallet styrks av att följsamheten ofta är högre i studier än i verkligheten (Kane, Kishimoto & Correll, 2013, s. 223).

Som Howick, Glasziou och Aronson (2013, s. 277) påpekar så ger kliniska studier, och i ännu högre grad meta-analyser, en bild av genomsnittlig effekt som inte behöver vara representativ för en enda av de ingående individerna. För interventioner som FFA där följsamheten är låg är detta påpekande mycket relevant. För forskningsfrågan vilken effekten är av en policy, behöver det inte vara ett problem eftersom det är just den genomsnittliga effekten som då är av intresse. Frågan är dock om läsare alltid tolkar studiernas resultat utifrån denna forskningsfråga och med ovanstående resonemang i beaktande. Eftersom effekten mätt med ITT-analys är låg samtidigt som följsamheten är låg skulle man kunna hävda att det troligen är implementeringen snarare än själva interventionen som inte fungerat. Å andra sidan är det sannolikt så att följsamheten till interventionen är låg delvis p.g.a. att interventionen är som den är. Om interventionen är sådan att ingen vill följa den säger det inte bara något om implementeringen utan också om interventionen. När det gäller FFA är det alltså inte

självlklart hur effekten av implementering och intervention kan särskiljas och därmed blir inte heller skiljelinjen mellan *efficacy* och *effectiveness* helt tydlig.

Komplexa interventioner utmanar evidensrörelsens förhållande till kausalitet

För att mäta effekten av en intervention i en kontrollerad studie krävs det att interventionen är definierad. Annars vet man inte vad man mäter och det blir svårt att säkerställa att alla andra variabler är lika mellan grupperna. När det gäller FFA är det oklart om träningsstillfällena ska ses som en del av interventionen, eller om de är en följd av en lyckad intervention som istället definieras som informationen som delges personen vid konsultationerna. I de granskade studierna vittnar följsamhetsmätningarna om att det finns olika uppfattningar om detta. De bakomliggande teorierna om orsaker till låg följsamhet till livsstilsrekommendationen fysisk aktivitet får betydelse i detta sammanhang. Handlar det om att personen behöver upplysning eller utbildning för att förstå konsekvenserna av fysisk aktivitet och inaktivitet? I så fall kan hänvisningen och konsultationerna ses som intervention och den fysiska aktiviteten vid träningscentret är en del av utfallet ökad fysisk aktivitet. Eller handlar det om att hjälpa personen införa och stödja en ny vana? I så fall kan den fysiska aktiviteten under interventionstiden ses som en del av interventionen. Eller ska man rentav se det som att det är själva hänvisningen som är interventionen, eftersom möjligheten till fysisk aktivitet finns där sedan tidigare, och att få konsultation vid ett träningsställe finns det också möjlighet till även när man inte blir hänvisad från primärvården. Men i så fall blir skillnaden mellan intervention och kontroll så hårfin att studien framstår som meningslös att genomföra. I de flesta av studierna har nämligen information om fysisk aktivitet och om träningsställen lämnats till kontrollgruppen. Det förefaller alltså finnas en oenighet mellan de granskade studierna om hur interventionen är definierad, vilket avslöjas av deras olika sätt att mäta följsamhet. Detta är något förvånande eftersom interventionen i stort beskrivs lika. Ytterst grundas detta i en oklarhet om interventionens mekanistiska kärna, d.v.s. med vilka mekanismer FFA antas fungera. Mot denna bakgrund kan FFA ses vara en komplex intervention, d.v.s. en intervention som omfattar flera separata delar som verkar nödvändiga för att interventionen ska fungera samtidigt som den mekanistiska kärnan i interventionen är svår att specificera (Medical Research Council, 2000, s. 2).

Evidensbaseringsens procedurer grundas i en tilltro till att randomiseringen med hjälp av slumpen ska göra två grupper så lika i alla avseenden att utfallet med säkerhet kan hänföras till interventionen. När det gäller komplexa interventioner uppstår dock problem vad gäller möjligheten att i studier fastställa kausala samband som kan överföras till andra förhållanden. En fördel med randomiseringen, som brukar framhållas, är att det inte krävs att man på förhand vet vilka faktorer som är av betydelse för utfallet eftersom slumpen gör grupperna lika även avseende icke kända faktorer. Dock finns det sammanhang då dessa faktorer ändå kan få en avgörande betydelse. En anledning till detta är att ett utfall ofta styrs av s.k. INUS-villkor snarare än regelrätta orsaksfaktorer. Ett INUS-villkor är en *Insufficient but Necessary part of an Unnecessary but Sufficient condition*, d.v.s. en otillräcklig men nödvändig del av ett icke nödvändigt men tillräckligt villkor. INUS-villkoren innebär att kontextuella faktorer kan ha stor betydelse vilket kan göra det närmast omöjligt att överföra resultat av en studie från en miljö till en annan (Cartwright & Hardie, 2012, ss. 63-64). Med en löst definierad intervention så som FFA och en rad kontextuella faktorer av betydelse är det inte osannolikt att överförbarheten är begränsad. En kvalitativ studie av FFA har t.ex. visat att träningscentrets öppettider kan ha betydelse för utfallet (Wormald & Ingle, 2004). Cartwright & Hardie (2012) menar därför att kliniska studier aldrig kan ersätta en förståelse för en interventions verk samma mekanismer. Även om studier skulle visa att FFA har fungerat krävs det en förståelse för *hur* interventionen har fungerat och med vilka stödfaktorer, för att veta vad som skulle krävas för att den skulle fungera även i ett annat sammanhang.

Rickles (2009) förklarar hur även den interna validiteten kan hotas av att förhållanden om orsak och verkan sällan är okänsliga för andra faktorer. När det gäller enkla interventioner kan randomiseringen fylla sitt syfte, att göra grupperna tillräckligt lika avseende mätbara och icke mätbara variabler av betydelse för utfallet. För komplexa interventioner däremot, där dessa variabler av betydelse för utfallet kan vara beroende av varandra, kan en liten skillnad mellan grupperna ge stora effekter på

utfallet och det är därmed osäkert om randomiseringen fyller sitt syfte, hävdar Rickles (2009, s. 87). En grundläggande förutsättning för att kunna identifiera kausala samband genom kontrollerade studier är att ett potentiellt utfall är oberoende av fördelningen av deltagare till grupperna. Detta antagande brukar benämnas SUTVA, ”stable unit treatment value assumption” (Marcus, Stuart, Wang, Shadish & Steiner, 2012, s. 2). I icke blindade randomiserade studier kan deltagarnas preferenser äventyra SUTVA. Randomiseringen gör att det bildas två subgrupper i varje grupp, baserat på om de hamnade i den grupp de önskade eller inte. Vårdtagarens preferenser avseende behandlingsmetod har i flera studier visats ha effekt på engagemang i behandlingen, följsamhet och utfall. Om den ena behandlingen är attraktiv för en större andel deltagare är det ofrånkomligt att det uppstår skillnader mellan grupperna eftersom den behandling som är mest attraktiv kommer ha större andel som fick önskad behandling (Sidani et al., 2009, ss. 56-57, Macias et al., 2009, s. 2). Rickles (2009) menar alltså dessutom att för komplexa interventioner kan denna skillnad i preferenser vara mycket liten mellan grupperna och ändå få en stor betydelse för utfallet, eftersom det kan finnas andra skillnader mellan grupperna som samverkar med skillnaden i preferenser.

I de granskade studierna är generellt sett bortfallet före randomisering stort, bortfallet under studiens gång är stort och följsamheten är låg (bortsett från studien av Harrison et al. som definierat interventionen annorlunda). Bortfall, som utgörs av de som inte utvärderats, och följsamhet, som mäter genomförande av interventionen, är i teorin oberoende av varandra. För en komplex intervention som FFA är det dock inte omöjligt att det finns gemensamma faktorer som påverkar de båda. I samtliga granskade studier var bortfallet före randomisering stort. Vad det beror på berörs inte i artiklarna, men det är sannolikt så att en anledning till bortfall är att personen som tillfrågas att delta i studien saknar motivation att genomföra interventionen, vilket förmodligen också är en vanlig orsak till att följsamheten blir låg för de som ändå väljer att delta i studien. För en person som inte varit följsam till interventionen är det kanske heller inte lika lockande att delta i utvärderingen. Detta samband har också Taylor et al. (1998, s. 598) kunnat notera i sin studie där s.k. *completers* i högre grad var följsamma till interventionen genom att de deltog i fler träningsstillfällen. Sambandet var möjligt att upptäcka eftersom *adherence* registrerades på träningsstället och alltså kunde mätas även för de som inte var tillgängliga vid uppföljningen. När bortfallet är stort och det synes finnas ett samband mellan bortfall och andra faktorer som dessutom kan skilja sig åt mellan grupperna bidrar detta till en ökad osäkerhet i tolkningen av resultatet. Ytterst innebär detta att det kan ifrågasättas om de grundläggande antaganden som evidensbaseringsens procedurer bygger på kan garanteras, när det är komplexa interventioner som ska utvärderas.

Ytterligare ett problem för evidensbaseringsens metoder är att det verkar finnas ett samband mellan låg följsamhet och ökad dödlighet, oavsett om det är följsamhet till en förmodat verksam behandling eller till placebo. Detta har tolkats som att god följsamhet fungerar som ett surrogatmått för ett allmänt hälsosamt beteende, vilket brukar benämnas *healty adherer*-effekten (Simpson et al., 2006, s. 5). Om andelen följsamma skiljer sig mellan grupperna riskerar i så fall resultatet för gruppen med flest följsamma att övervärderas. I studier som jämför en aktiv intervention med en kontrollintervention som inte kräver någon följsamhet kan man inte veta hur grupperna skiljer sig åt i detta avseende.

Alternativa analysmetoder finns men kan inte lösa problemet

Shrier (2014, s. 29) har identifierat fyra olika typer av forskningsfrågor som alla kräver olika analysmetoder för att ge ett valitt resultat, där som tidigare nämnts, frågan om *effectiveness* mätt med ITT-analys är den bäst ansedda inom EBM. Men när *effectiveness* visar sig vara låg kan man fortfarande vara intresserad av vilken effekt interventionen har för de som är följsamma. I dessa lägen kan man göra så kallade post hoc analyser för att studera *efficacy*. Traditionellt mäts *efficacy* med *per protocol* (PP) eller *as treated* (AT) -analyser, men dessa analyser tar inte hänsyn till att det finns personer som tar den aktiva behandlingen oavsett vad de tilldelas i randomiseringen likväl som det finns personer som inte tar den aktiva behandlingen oavsett vad de tilldelas. För dessa personer saknas den kontrafaktiska kontrollen vilket äventyrar den interna validiteten och riskerar att leda till ett missvisande resultat. Detta är också en av anledningarna till att dessa analysmetoder generellt sett inte

rekommenderas inom EBM (Statens beredning för medicinsk utvärdering, 2014, Bilaga 2:8). På senare år har nya analysmetoder börjat användas för att estimeras *efficacy*, under benämningen *complier average causal effect* (CACE). Det finns flera varianter på beräkning av CACE men de har det gemensamt att de avser beräkna effekten för de som är följsamma till avsedd intervention eller kontrollintervention, oavsett vilken grupp de hamnar i. Det innebär en förbättrad intern validitet jämfört med PP- och AT-analyser (Shrier et al., 2014, s. 29). Även metoder för att göra beräkningar av CACE när följsamheten inte kan kategoriseras dikotomt har föreslagits (Becque, White & Haggard, 2015).

Det finns även alternativa analysmetoder av RCT för att ta hänsyn till deltagarnas preferenser. Gemensamt för dessa är att de försöker särskilja effekten av deltagarnas preferenser och behandlingens effekt. Detta kan åstadkommas t. ex. genom att randomisera i två omgångar (Marcus et al., 2012, s. 15, Younge, Kouwenhoven-Pasmooij, Freak-Poli, Roos-Hesselink & Hunink, 2015, ss. 2013-2016, Sidani et al., 2009, s. 63).

Dessa alternativa analysmetoder borde kunna bidra till säkrare effektmätningar som ett komplement till ITT-analysen, men de verkar inte ha fått så stort genomslag inom EBM. De är i vilket fall inte nämnda eller rekommenderade i SBU:s metodbok. Metoderna löser heller inte problemet att den mekanistiska kärnan i FFA är oklar vilket innebär att det fortfarande med resultat av dessa analysmetoder skulle vara svårt att dra slutsatser om hur FFA kan användas, för vilka vårdtagare och under vilka omständigheter.

Förskrivning av fysisk aktivitet synliggör brister i evidensgraderingen

När en intervention visar effekt i flera studier med ITT-analys är följsamheten som validitetshot ett mindre problem eftersom effekten uppenbarligen är tillräckligt stor trots att inte alla genomfört interventionen. När däremot effekten är låg eller motstridig i studierna får följsamhet och bortfall större betydelse. Kvalitetsgranskningen kommer då resultera i en lägre evidensstyrka likväl som ITT-analys ger förhållandevis låg effekt. Det innebär att slutsatsen enligt GRADE blir att det krävs fler studier för att kunna uttala sig om interventionens effekt. Slutsatsen dras oavsett om den låga evidensgraderingen beror på en dåligt genomförd studie eller på att interventionen och studiepopulationen är så beskaffade att det blir stort bortfall och låg följsamhet. Det innebär också att om vi antar att det finns interventioner utan effekt som p.g.a. hur de är beskaffade aldrig kan uppnå hög följsamhet och lågt bortfall, så kan det aldrig bekräftas med GRADE att dessa har låg effekt. Slutsatsen kommer alltid bli att evidensstyrkan är för låg för att säkerställa om interventionen har effekt. Detta måste ses som en svaghet i GRADE och ytterst även EBM eftersom systemet i så fall uppmuntrar till att utföra studier av samma fenomen i det oändliga.

Man kan också konstatera att bedömningen av om följsamheten är godtagbar som är en av de frågor som ligger till grund för evidensgraderingen, inte är så lätt att göra som det kan tyckas. Både definitioner av följsamhet och mätmetoder kan skilja sig mellan studierna. När det dessutom är oklart vad som innefattas i interventionen och vad följsamheten ska mätas till blir det än osäkrare. Dodd et al. (2012, s. 7) fann i en systematisk översikt av 66 slumpvis valda studier, att bara 7 (11%) hade en fullgod rapportering avseende följsamhet till långvarig behandling. Rapporteringen var bättre för kortvarig (36%) och enstaka (83%) behandling. Om vi därtill lägger att det ofta sker ett bortfall redan före randomiseringen som kan påverka den senare följsamhetsmätningen, så blir det än mer komplicerat att bedöma följsamhet.

Enligt kategoriseringen av forskningsfrågor av Shrier (2014, s. 29) kan det konstateras att evidensgraderingar baserade på systematisk översikt och GRADE alltid svarar på samma forskningsfråga, vilken effekt en policy-ändring, t.ex. ett brett införande av interventionen FFA, skulle ge. När det gäller interventioner med låg följsamhet kommer de med största sannolikhet ge en lägre effekt än vad den kausala effekten, *efficacy*, är för de personer som är följsamma. Studierna mäter alltså inte *efficacy*, men man kan också starkt ifrågasätta om de mäter *effectiveness*. Även om

analysmetoden ITT mäter *effectiveness* så gör bortfallet före randomisering och studiens tillrättaläggande ofta att det inte heller blir en sann mätning av *effectiveness*. När det gäller FFA där följsamheten är låg innebär detta att evidensgraderingen ger viss vägledning till beslutsfattare huruvida interventionen ska införas brett, men för en kliniker som står med en vårdtagare i behov av ökad fysisk aktivitet ger evidensgraderingen liten vägledning. Det genomsnittliga värdet för effekten av FFA säger mycket lite om vilken effekten är för respektive vårdtagare. Det som däremot kan vägleda en kliniker är vetenskapen att få personer är följsamma till interventionen FFA. Om man som kliniker tror sig ha förmåga att välja ut de personer som kommer vara följsamma alternativt motivera personer så att de blir följsamma, kan det finnas anledning att fortsätta använda interventioner så som FFA, trots att effekten bedömts vara låg i systematiska översikter och enligt GRADE.

Sammanfattningsvis så grundas evidensrörelsens procedurella metoder för att fastställa en interventions effekt i flera antaganden, som inte är självklara. En interventions effekt kan definieras på andra sätt och det går att ställa andra forskningsfrågor. Samtidigt finns det inom evidensrörelsens metoder en stor osäkerhet kring tolkning av studier där följsamhet spelar en viktig roll. Man kan fråga sig om det i dessa lägen är rimligt att ändå försöka använda dessa metoder så gott det går, eller om forskningen hellre bör söka andra metoder.

När statistik möter etik

Livsstilsrekommendationerna avseende fysisk aktivitet utgår från medicinska aspekter och grundas i ett vetenskapligt underlag som visat att fysisk aktivitet ger goda hälsoeffekter. I de tidigaste definitionerna av EBM skulle detta underlag kanske räcka för att ställa krav på vårdtagare som inte uppnår målen avseende fysisk aktivitet att öka sin fysiska aktivitet. Dagens evidensrörelse värderar vårdtagarens situation och önskemål vilka mycket väl kan strida mot de medicinska aspekterna. Det kan finnas anledningar till att en person gör valet att leva med de hälsorisker som en lägre fysisk aktivitetsnivå innebär. Den vetenskapliga litteraturen där *concordance* och *adherence* ofta förespråkas framför *compliance* kan ses som en utveckling där man i dag mer än tidigare betonar individens rätt att av personliga skäl frångå rekommendationer och ordinationer. Samtidigt kan en systematisk översikt ge vid handen att interventionen FFA inte har effekt och det är högst troligt att slutsatsen varit en annan om fler personer varit följsamma till interventionen i de studier som slutsatsen grundas på. Här finns en paradox inbyggd eftersom följsamhet ses som ett fritt val i det ena avseendet medan följsamhet är en förutsättning för interventionens effekt i det andra avseendet. Det grundas ytterst i att man betonar individanpassning i den evidensbaserade hälso- och sjukvården i det ena avseendet, men när det gäller bedömning av en interventions effekt är utrymme för individanpassning mycket begränsat.

När man inom evidensrörelsen och forskningen överlag alltmer betonar vikten av att anpassa vården efter individen och dess önskemål och sociala kontext, har inte samma glidning skett med hantering av följsamhet i studier enligt EBM. *Intention-to-treat*-analys av RCT som SBU förordar grundas i begreppet *compliance*, eftersom vårdtagaren blir tilldelad och förväntas vara följsam till en intervention. Det är i ett avseende fullt logiskt eftersom statistikens lagar inte ändras av förändringar i normer. Men det är också ett faktum att vårdtagarens följsamhet påverkas av dennes preferenser för behandlingen, vilket i sin tur kan påverkas av huruvida vårdtagaren varit delaktig i beslut om behandlingsmetod. Att vårdtagaren ska få vara med och bestämma handlar således inte bara om de etiska aspekterna av att öka vårdtagarens självbestämmande utan också om att optimera effekten av behandlingen. Resultatet av en studie som inte tar hänsyn till vårdtagarens önskemål kan med andra ord inte hur som helst generaliseras till en kontext där vårdtagarens önskemål ska vara styrande. Det innebär också att en traditionell RCT som anses vara gyllene standard inom EBM, i detta avseende strider mot en av EBMs övriga grundpelare, nämligen att vårdtagaren ska vara delaktig i valet av behandling. Det finns dock alternativa analysmetoder där man försöker balansera så att man får bästa statistiska underlag utan att begränsa vårdtagarens valfrihet mer än absolut nödvändigt. Dessa har dock

inte slagit igenom. Ett problem med dessa är att de är tids- och kostnadskrävande vilket också måste vägas in när studiedesign väljs.

Följsamhet mäts i de flesta kvantitativa studier på något sätt, men det är ofta oklart hur följsamheten är definierad och därmed hanteras. I studier eftersträvas ofta hög följsamhet för att kunna göra säkrare statistiska mätningar, men huruvida studiedeltagarna uppmuntras och i så fall i vilken grad, till så hög följsamhet som möjlighet framgår inte alltid. McKay et al. (2015, s. 382) som forskar på idrottsskador menar att *efficacy*-studier kräver *compliance* medan pragmatiska *effectiveness*-studier istället bör mäta och tolka *adherence*. Det är en intressant tanke men man kan fråga sig vad det skulle innebära i praktiken. Om det skulle innebära att deltagaren lämnas större utrymme att själv bestämma följsamhet vid mätning av *adherence*, så undrar man vilka metoder som skulle användas för att hålla en hög *compliance*. När sedan utfallsmåttet är fysisk aktivitetsnivå är det svårt att se hur man skulle kunna hålla en strikt skillnad mellan *compliance* och *adherence*. Vad innebär t.ex. *compliance* till ett motiverande samtal? Det är i alla fall uppenbart svårt att mäta följsamhet till interventioner som kräver ett kognitivt engagemang.

För interventioner som inbegriper en motiverande åtgärd är följsamhet till intervention och följsamhet till utfall inte helt oberoende. Med FFA försöker man först nå en följsamhet till FFA som man sedan hoppas ska ge en följsamhet till fysisk aktivitet. Det finns ingen absolut gräns mellan att motivera personen till att komma på ett träningstillfälle under interventionen och att fortsätta vara fysiskt aktiv efteråt. Man kan lätt få uppfattningen att evidensrörelsen vill låtsas att det alltid går att uppnå följsamhet till interventionen bara kvaliteten på studien är tillräckligt hög. När själva ursprunget till studien grundas i en problematik som bygger på att följsamheten är låg blir det märkligt. Här finns en grundläggande skillnad vad gäller följsamhet till en förhållandevis enkel intervention som medicinering och en komplex intervention som FFA. Om en vårdtagare i en studie av blodtryckssänkande medicinering inte är följsam till interventionen, säger det inget om den blodtryckssänkande medicinens kausala effekt. Den kausala effekten av en medicin är inte beroende av att vårdtagaren vill ta medicinen. Däremot i en studie av en intervention som syftar till att öka motivationen för fysisk aktivitet, säger det förmodligen något om interventionens effekt om vårdtagaren inte är följsam till interventionen. Om vårdtagaren inte ens vill delta i interventionen är det troligt att interventionen heller inte skulle ha någon effekt eftersom den kräver ett aktivt deltagande. Samtidigt innebär det också att de som är följsamma till den interventionen kanske skulle kunna förmås att öka sin fysiska aktivitet på många andra sätt eftersom de är öppna för att göra en ansträngning för att öka sin fysiska aktivitet. Effekten av interventionen blir därmed minst sagt svårtydd.

När det gäller komplexa interventioner som dessutom är löst definierade finns det många felkällor i syntetiseringen av studier. Till stor del handlar det om att följsamheten till interventionen inte kan säkerställas eftersom följsamheten är själva ursprungsproblemet. Men även om man visste att syntetiseringen gav ett statistiskt korrekt svar kan man ifrågasätta värdet av den informationen, i alla fall för en vårdgivare som står inför uppgiften att förbättra en vårdtagares hälsa genom ökad fysisk aktivitet, eftersom det medelvärde som ges kan ligga långt från individernas värden. Framför allt framstår det tydligt och klart att denna typ av studier inte kan ge en fullständig bild av en intervention.

Diskussion

Resultat och implikationer

Det som framstår tydligt efter granskning av studier av FFA är hur många osäkerheter det finns. Interventionen är otvetydigt definierad och det är oklart vilka problem den avser lösa och på vilket sätt. Resultatet av studierna är i många fall oklart. Syntetiseringen av studier grundas i en procedur som ska säkerställa hög validitet, men resultatet av denna studie visar att det målet inte är lätt att uppnå för interventioner som FFA.

Den ideala vetenskapliga kumulativiteten kan beskrivas med en metafor av Thomas Kuhn (1996). Han menar att normalvetenskapen (Kuhns egen benämning) fungerar som ett stort pussel där varje ny kunskap utgör en pusselbit som kan passas in i den redan existerande kunskapens påbörjade pussel. Men Kuhn (1996) visar också hur pusslen bytts ut under historiens gång genom att de teoretiska och metodologiska antaganden som normalvetenskapen vilar på har förändrats. Därför måste vetenskapens kumulativitet anses vara begränsad till rådande paradig. Dessutom finns det inom framför allt andra områden än naturvetenskapen, kontroverser och konkurrerande uppfattningar som synliggör att vetenskap inte alltid växer som normalvetenskap inom ett paradig (Kuhn, 1996, s. x). Den systematiska översikten och evidensgraderingen av FFA gör gällande att det är möjligt att bygga vetenskap kumulativt inom detta område. Samtidigt är det högst tveksamt om studier av FFA kan antas uppfylla kriterierna för normalvetenskap avseende teoretiska och metodologiska antaganden, vilka Kuhn menar är en förutsättning för den kumulativt växande normalvetenskapen.

Cartwright och Hardie (2012, ss. 17-18) använder en liknande metafor som Kuhn, men som ännu tydligare pekar på den vetenskapliga kumulativitetens skörhet. De liknar slutsatsen om en interventions effekt vid den högsta stenen på en pyramid, där varje argument utgör en byggsten som vilar på andra byggstenar av argument. Bilden tydliggör att det räcker att en enda av pyramidens byggstenar är en icke verifierad hypotes, för att slutsatsen inte ska vara starkare än den hypotesens berättigande. Det innebär alltså att en icke verifierad hypotes som senare visar sig vara felaktig kan riva hela pyramiden av argument och falsifiera slutsatsen. Evidensrörelsens formaliserade procedurer har för avsikt att ge bästa möjliga vetenskapliga underlag för slutsatser om interventioners effekt, men samtidigt bygger de på en formalisering i flera led med många antaganden längs vägen. Ett misstag eller en enda oklarhet på vägen kan göra att slutprodukten blir missvisande eller direkt felaktig, och att målet därmed inte nås.

Evidensbaseringsmetoder RCT:er och metaanalyser bestämmer effekten av en intervention som medelvärde för en grupp. Detta sätt att mäta effekt grundas i ett antagande om att det finns något gemensamt för hur personerna i gruppen reagerar, eller möjligen det mer pragmatiska och något cyniska antagandet att effekten av en intervention bara är intressant om den gäller för tillräckligt många personer. Om man drar det resonemanget till sin spets så innebär det att det inte finns något utrymme för originalitet inom EBM. Statens beredning för medicinsk och social utvärdering (SBU) (2014, s. 52) konstaterar att biverkningar ofta inte utvärderas bäst i RCT:er eftersom de kan vara sällsynta och då inte säkert ger något genomslag i en begränsad kontrollerad studie. De upptäcks ofta bättre i observationsstudier som bygger på ett större underlag. I detta hänseende medger man alltså att människor reagerar olika och att de negativa effekter som uppstår för enskilda personer också behöver uppmärksammas. Positiva effekter som uppstår för enskilda personer verkar dock inte finnas något utrymme för inom EBM. Vad är det som säger att det bara är de negativa effekterna som kan uppstå hos enskilda individer? Borde inte den evidensbaserade hälso- och sjukvården se även till de människor som i något avseende är annorlunda, inte bara genom att undvika biverkningar?

Om kunskap definieras som resultatet av flera RCT:er får det också långtgående konsekvenser för vilka forskningsfrågor som kan ställas. Andra studiedesigner kan lyfta andra betydelser av FFA som inte framgår i RCT:er, och de kan också utvärdera helt andra interventioner. T.ex. menar Taylor et al.

(1998, s. 599) att FFA kan fylla en viktig funktion genom att myndigheter visar att man menar allvar med att fysisk aktivitet är viktigt. Denna effekt kan svårigen studeras i en RCT. När interventioner som kan testas i RCT:er visar sig ha tveksam effekt såsom FFA gjort, finns det också större anledning att vidga perspektivet och söka annan kunskap. Glass et al. (2013, s. 10) pekar på omöjligheten i att studera folkhälsointerventioner med RCT:er. De åskådliggör detta med exemplet rökavvänjning där effekten av ökad skatt på tobak är svårt att studera i RCT:er medan rökavvänjningsprogram går utmärkt att studera. Liknande förhållanden kan finnas mellan FFA och en samhällsplanering som gynnar fysisk aktivitet. Lambert (2006) har genom en historisk genomgång visat hur evidensrörelsen i takt med att kritik uppstått successivt plockat in ytterligare perspektiv och vidgat begreppet EBM, så att vårdtagares önskemål och situation samt klinikers kunskap värdesätts. På senare tid har även det vetenskapliga underlaget breddats till att kunna innefatta kvalitativa forskningsresultat. Slutsatserna av denna studies genomgång av följsamhet och FFA stödjer behovet av en sådan utveckling där kunskap inte definieras så att forskningen begränsas till vissa metodiker.

Metoddiskussion

Huvudsyftet för denna uppsats har varit att öka förståelsen för följsamhetsproblematik. En avgränsning gjordes i ett av delsyftena att fokusera på effekter av FFA på fysisk aktivitet jämfört med sedvanlig vård eller ingen behandling. Därigenom uteslöts studier av FFA jämfört med annan intervention för att öka fysisk aktivitet. Eftersom FFA är relativt löst definierat och det finns en otydlighet i vad som är en FFA-intervention och vad som ska definieras som en annan behandling, gjordes valet att enbart inkludera studier som värderar FFA jämfört med sedvanlig vård eller ingen behandling för att göra skiljelinjen så tydlig som möjligt. Även studier som jämfört så kallad hög dos och låg dos av FFA uteslöts av samma skäl, allt för att öka förståelsen för följsamhetsproblematik till livsstilsrekommendationer. Syftet var inte att vägleda kliniker eller beslutsfattare i vilken form av FFA eller ens vilken form av intervention för att öka fysisk aktivitet som har bäst effekt. Övriga metoder är troligen dessutom än mindre definierade och studierna kommer därmed vara heterogena och svåra att syntetisera. För att enstaka studier ska bedömas ge mer än ”otillräckligt vetenskapligt underlag” och lägsta gradering, krävs starka skäl, enligt SBU (2014, s. 151), något som i dessa fall svårigen kan uppnås.

När målsättningen är att upptäcka alla studier av relevans för ett visst område finns självklart en risk att man inte lyckas nå målet, utan missar studier. De studier som inkluderades för granskning i denna studie var samtliga inkluderade i en tidigare systematisk översikt vars sökning var mycket rigorös. Sökningen som jag gjorde efter studier som publicerats 2011 eller senare var inte lika grundlig. Det går inte att utesluta att det finns artiklar att finna i andra databaser eller med andra sökord. Av praktiska skäl gjordes en limitering till elektroniskt tillgängliga fulltext-artiklar, vilket givetvis också kan ha resulterat i att relevanta studier valdes bort. Att söka igenom alla tänkbara databaser med fler sökord, samt beställa artiklar som inte är tillgängliga elektroniskt, bedömdes ta för lång tid i anspråk i denna uppsats där huvudsyftet inte var att bedöma evidensen för FFA.

När det gäller sökningen av studier för det andra delsyftet, är risken ännu större att relevanta studier har missats. Problematikeringen i denna studie gör inga anspråk på att vara fullkomlig. Både de statistiska och de etiska aspekterna av följsamhet hade med säkerhet kunnat belysas bättre med ett större material, men den begränsade sökningen som genomfördes för denna uppsats har ändå kunnat ge material för att visa på grundläggande problematiker när det gäller svårigheterna att evidensbasera FFA.

Slutsatser

Synen på följsamhet till medicinering och livsstilsrekommendationer har förändrats under de senaste decennierna, från det relativt enkla sättet att se på följsamhet, definierat som *compliance*, till att begreppen *adherence* och *concordance* allt mer tagits i bruk. Med dessa begrepp blir följsamhet något mycket komplext utan rätt och fel och med ett delat ansvar mellan vårdgivare och vårdtagare. Samtidigt vilar de forskningsmetoder som förespråkas inom EBM fortfarande på begreppet *compliance*. Detta resulterar i en paradox och ställer till problem i utförande och tolkning av studier av interventioner som normalt sett inte ger hög följsamhet och där följsamheten inte enkelt kan definieras.

Förskrivning av fysisk aktivitet är en intervention med oklar definition, oklara verkningsmekanismer och oklar effekt. Det är tveksamt om RCT är en lämplig utvärderingsmetod av interventioner för att öka fysisk aktivitet hos befolkningen. Framför allt är det tveksamt om ytterligare RCT:er bör utföras av FFA. Samtidigt är det inom evidensrörelsen den gängse slutsatsen att flera välgjorda studier behövs, när evidensgraderingen som i detta fall resulterar i att det vetenskapliga underlaget bedöms vara begränsat. De orsaker som finns till att personer inte är fysiskt aktiva i den grad som rekommenderas varierar och därmed varierar också lösningen. Mot denna bakgrund kan man fråga sig om det ens är intressant att studera huruvida det ger en signifikant skillnad mellan grupper för en intervention som innebär att alla inaktiva ges en förskrivning av fysisk aktivitet.

Trots dessa grundläggande problematiker vill jag framhålla att denna uppsats i all sin kritik till evidensbaserad ända inte vill hävda att alternativet utan evidensbaserad skulle vara bättre. Evidensbaserad är en strävan att bygga hälso- och sjukvård på ett vetenskapligt underlag vilket aldrig kan vara fel. Däremot kan man diskutera metoderna som används, och det är tydligt att det finns förbättringspotential. Men det står också klart att många av de problem som lyfts i denna uppsats inte har någon enkel lösning och kanske ingen lösning över huvud taget. Likväl är det viktigt att vara medveten om de problem och begränsningar som evidensrörelsen står inför, när man drar slutsatser av dess resultat.

Referenslista

- Anokye, N. K., Lord, J. & Fox-Rushby, J. (2014). Is brief advice in primary care a cost-effective way to promote physical activity? *Br J Sports Med*, 48, 202-6.
- Barbosa, C. D., Balp, M.-M., Kulich, K., Germain, N. & Rofail, D. (2012). A literature review to explore the link between treatment satisfaction and adherence, compliance, and persistence. *Patient preference and adherence*, 6, 39.
- Becque, T., White, I. R. & Haggard, M. (2015). A causal model for longitudinal randomised trials with time-dependent non-compliance. *Stat Med*, 34, 2019-34.
- Bell, J. S., Airaksinen, M. S., Lyles, A., Chen, T. F. & Aslani, P. (2007). Concordance is not synonymous with compliance or adherence. *Br J Clin Pharmacol*, 64, 710-1; author reply 711-3.
- Bernardini, J. (2004). Ethical issues of compliance/adherence in the treatment of hypertension. *Advances in Chronic Kidney Disease*, 11, 222-227.
- Bissell, P., May, C. R. & Noyce, P. R. (2004). From compliance to concordance: barriers to accomplishing a re-framed model of health care interactions. *Soc Sci Med*, 58, 851-62.
- Bissonnette, J. M. (2008). Adherence: a concept analysis. *J Adv Nurs*, 63, 634-43.
- Bohlin, I. (2011). Evidensbaserat beslutsfattande i ett vetenskapligt samhälle. Om evidensrörelsens ursprung, utbredning och gränser. I: Bohlin, I. & Sager, M. (red.) *Evidensens många ansikten : evidensbaserad praktik i praktiken*. Lund: Arkiv förlag.
- Bolin, K. & Lindgren, B. (2006). *Fysisk inaktivitet - produktionsbortfall och sjukdomskostnader*. Stockholm: FRISAM.
- Cartwright, N. & Hardie, J. (2012). *Evidence-based policy : a practical guide to doing it better*. New York: Oxford University Press.
- Chakrabarti, S. (2014). What's in a name? Compliance, adherence and concordance in chronic psychiatric disorders. *World J Psychiatry*, 4, 30-6.
- Dodd, S., White, I. R. & Williamson, P. (2012). Nonadherence to treatment protocol in published randomised controlled trials: a review. *Trials*, 13, 84.
- Donovan, J. L. & Blake, D. R. (1992). Patient non-compliance: deviance or reasoned decision-making? *Soc Sci Med*, 34, 507-13.
- Du Pasquier, S. & Aslani, P. (2008). Concordance-based adherence support service delivery: consumer perspectives. *Pharm World Sci*, 30, 846-53.
- FaR® : *individ Anpassad skriftlig ordination av fysisk aktivitet*. (2011). Östersund: Statens folkhälsoinstitut.
- Felzmann, H. (2012). Adherence, compliance, and concordance: an ethical perspective. *Nurse Prescribing*, 10, 406-411.
- Folkhälsomyndigheten. *Hälso- och sjukvårdens roll*, [Elektronisk]. Tillgänglig: <https://www.folkhalsomyndigheten.se/far/halso-och-sjukvardens-roll/>.
- Galaviz, K., Levesque, L. & Kotecha, J. (2013). Evaluating the effectiveness of a physical activity referral scheme among women. *J Prim Care Community Health*, 4, 167-71.
- Glass, T. A., Goodman, S. N., Hernan, M. A. & Samet, J. M. (2013). Causal inference in public health. *Annu Rev Public Health*, 34, 61-75.
- Gusi, N., Reyes, M. C., Gonzalez-Guerrero, J. L., Herrera, E. & Garcia, J. M. (2008). Cost-utility of a walking programme for moderately depressed, obese, or overweight elderly women in primary care: a randomised controlled trial. *BMC public health*, 8, 231-231.
- Harrison, R. A., Roberts, C. & Elton, P. J. (2005). Does primary care referral to an exercise programme increase physical activity 1 year later? A randomized controlled trial. *Journal of Public Health*, 27, 25-32.
- Haynes, R., Taylor, D. & Sackett, D. (1979). *Compliance in health care*. Baltimore: John Hopkins university press.
- Health and Safety Executive (2009). *Health Survey for England – 2008. Physical activity and fitness. Summary of key findings*. London: Department of Health.

- Hearnshaw, H. & Lindenmeyer, A. (2006). What do we mean by adherence to treatment and advice for living with diabetes? A review of the literature on definitions and measurements. *Diabet Med*, 23, 720-8.
- Herdman, T. H., Kamitsuru, S. & International, N. (2015). *Omvårdnadsdiagnoser: definitioner och klassifikation 2015-2017*. Lund: Studentlitteratur.
- Hobden, A. (2006). Concordance: a widely used term, but what does it mean? *Br J Community Nurs*, 11, 257-60.
- Holtzman, J., Schmitz, K., Babes, G., Kane, R. L., Duval, S., Wilt, T. J., MacDonald, R. M. & Rutks, I. (2004). Effectiveness of behavioral interventions to modify physical activity behaviors in general populations and cancer patients and survivors. *Evid Rep Technol Assess (Summ)*, 1-8.
- Horne R, Weinman J, Barber N, Elliott R & M., M. (2005). Concordance, adherence and compliance in medicine taking. Report for the National Co-ordinating Centre for NHS Service Delivery and Organisation R & D.
- Howick, J., Glasziou, P. & Aronson, J. K. (2013). Problems with using mechanisms to solve the problem of extrapolation. *Theor Med Bioeth*, 34, 275-91.
- Huang, H. Y., Lin, Y. S., Chuang, Y. C., Lin, W. H., Kuo, L. Y., Chen, J. C., Hsu, C. L., Chen, B. Y., Tsai, H. Y., Cheng, F. H. & Tsai, M. W. (2015). Application of the Transtheoretical Model to Exercise Behavior and Physical Activity in Patients after Open Heart Surgery. *Zhonghua Minguo Xin Zang Xue Hui Za Zhi*, 31, 202-8.
- Isaacs, A. J., Critchley, J. A., Tai, S. S., Buckingham, K., Westley, D., Harridge, S. D., Smith, C. & Gottlieb, J. M. (2007). Exercise Evaluation Randomised Trial (EXERT): a randomised trial comparing GP referral for leisure centre-based exercise, community-based walking and advice only. *Health Technol Assess*, 11, 1-165, iii-iv.
- Jolly, K., Duda, J. L., Daley, A., Eves, F. F., Mutrie, N., Ntoumanis, N., Rouse, P. C., Lodhia, R. & Williams, G. C. (2009). Evaluation of a standard provision versus an autonomy promotive exercise referral programme: rationale and study design. *BMC Public Health*, 9, 176.
- Kane, J. M., Kishimoto, T. & Correll, C. U. (2013). Non-adherence to medication in patients with psychotic disorders: epidemiology, contributing factors and management strategies. *World Psychiatry*, 12, 216-26.
- Kardas, P., Lewek, P. & Matyjaszczyk, M. (2013). Determinants of patient adherence: a review of systematic reviews. *Front Pharmacol*, 4, 91.
- Kaufman, G. & Birks, Y. (2011). Helping patients adhere to their medicines: Part 1: when describing a patient's ability or willingness to take their medicines correctly, do you use the terms compliance, adherence or concordance? These three commonly used words describe subtle, yet significant, differences in the patient-professional relationship. *Practice Nurse*, 41, 13.
- Klotsche, J., Leistner, D. M., Pieper, L., Pittrow, D., Zeiher, A. M. & Wittchen, H. U. (2011). The DETECT adherence score – structure and psychometric exploration of a novel approach to measure adherence to drug and non-drug interventions in primary care. *International Journal of Methods in Psychiatric Research*, 20, 82-92.
- Kuhn, T. S. (1996). *The structure of scientific revolutions*. 3rd ed. Chicago, Ill: University of Chicago Press.
- Lambert, H. (2006). Accounting for EBM: Notions of evidence in medicine. *Social Science & Medicine*, 62, 2633-2645.
- Leijon, M. E., Faskunger, J., Bendtsen, P., Festin, K. & Nilsen, P. (2011). Who is not adhering to physical activity referrals, and why? *Scand J Prim Health Care*, 29, 234-40.
- Lutfey, K. E. & Wishner, W. J. (1999). Beyond "compliance" is "adherence". Improving the prospect of diabetes care. *Diabetes Care*, 22, 635-9.
- Macias, C., Gold, P. B., Hargreaves, W. A., Aronson, E., Bickman, L., Barreira, P. J., Jones, D. R., Rodican, C. F. & Fisher, W. H. (2009). Preference in random assignment: implications for the interpretation of randomized trials. *Adm Policy Ment Health*, 36, 331-42.
- Marcus, S. M., Stuart, E. A., Wang, P., Shadish, W. R. & Steiner, P. M. (2012). Estimating the causal effect of randomization versus treatment preference in a doubly randomized preference trial. *Psychol Methods*, 17, 244-54.

- Marston, M. V. (1970). Compliance with medical regimens: a review of the literature. *Nurs Res*, 19, 312-23.
- McKay, C. D. & Verhagen, E. (2015). 'Compliance' versus 'adherence' in sport injury prevention: why definition matters. *British Journal of Sports Medicine*, bjsports-2015-095192.
- Medical Research Council (2000). A Framework for development and evaluation of RCTs for complex interventions to improve health.
- Moreau, A., Aroles, V., Souweine, G., Flori, M., Erpeldinger, S., Figon, S., Imbert, P., Denis, A., Iwaz, J., Riou, J. P. & Ploin, D. (2009). Patient versus general practitioner perception of problems with treatment adherence in type 2 diabetes: from adherence to concordance. *Eur J Gen Pract*, 15, 147-53.
- Murphy, S. M., Edwards, R. T., Williams, N., Raisanen, L., Moore, G., Linck, P., Hounsome, N., Din, N. U. & Moore, L. (2012). An evaluation of the effectiveness and cost effectiveness of the National Exercise Referral Scheme in Wales, UK: a randomised controlled trial of a public health policy initiative. *J Epidemiol Community Health*, 66, 745-53.
- Pavey, T. G., Taylor, A. H., Fox, K. R., Hillsdon, M., Anokye, N., Campbell, J. L., Foster, C., Green, C., Moxham, T., Mutrie, N., Searle, J., Trueman, P. & Taylor, R. S. (2011). Effect of exercise referral schemes in primary care on physical activity and improving health outcomes: systematic review and meta-analysis. *Bmj*, 343, d6462.
- Rickles, D. (2009). Causality in complex interventions. *Med Health Care Philos*, 12, 77-90.
- Rome, A., Persson, U., Ekdahl, C. & Gard, G. (2009). Physical activity on prescription (PAP): costs and consequences of a randomized, controlled trial in primary healthcare. *Scand J Prim Health Care*, 27, 216-22.
- Russell, S., Daly, J., Hughes, E. & Hoog Co, C. (2003). Nurses and 'difficult' patients: negotiating non-compliance. *J Adv Nurs*, 43, 281-7.
- Sagarin, B. J., West, S. G., Ratnikov, A., Homan, W. K., Ritchie, T. D. & Hansen, E. J. (2014). Treatment noncompliance in randomized experiments: statistical approaches and design issues. *Psychol Methods*, 19, 317-33.
- Shadish, W. R., Cook, T. D. & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Shi, L., Liu, J., Koleva, Y., Fonseca, V., Kalsekar, A. & Pawaskar, M. (2010). Concordance of adherence measurement using self-reported adherence questionnaires and medication monitoring devices. *Pharmacoeconomics*, 28, 1097-107.
- Shrier, I., Steele, R. J., Verhagen, E., Herbert, R., Riddell, C. A. & Kaufman, J. S. (2014). Beyond intention to treat: what is the right question? *Clin Trials*, 11, 28-37.
- Sidani, S., Miranda, J., Epstein, D. & Fox, M. (2009). Influence of treatment preferences on validity: a review. *Can J Nurs Res*, 41, 52-67.
- Simpson, S. H., Eurich, D. T., Majumdar, S. R., Padwal, R. S., Tsuyuki, R. T., Varney, J. & Johnson, J. A. (2006). A meta-analysis of the association between adherence to drug therapy and mortality. *Bmj*, 333, 15.
- Snowden, A., Martin, C., Mathers, B. & Donnell, A. (2014). Concordance: a concept analysis. *J Adv Nurs*, 70, 46-59.
- Sorensen, J. B., Kragstrup, J., Skovgaard, T. & Puggaard, L. (2008). Exercise on prescription: a randomized study on the effect of counseling vs counseling and supervised exercise. *Scand J Med Sci Sports*, 18, 288-97.
- Spiro, H. (2001). Compliance, Adherence, and Hope. *Journal of Clinical Gastroenterology*, 32, 5-5.
- Stanitis, M. A. & Ryan, J. (1982). Noncompliance: an unacceptable diagnosis? *American Journal of Nursing*, 82, 941-942 2p.
- Statens beredning för medicinsk utvärdering (2014). *Utvärdering av metoder i hälso- och sjukvården: en handbok*. Stockholm: Statens beredning för medicinsk utvärdering (SBU).
- Stevens, W., Hillsdon, M., Thorogood, M. & McArdle, D. (1998). Cost-effectiveness of a primary care based physical activity intervention in 45-74 year old men and women: A randomised controlled trial. *British Journal of Sports Medicine*, 32, 236-241.

- Taylor, A. H., Doust, J. & Webborn, N. (1998). Randomised controlled trial to examine the effects of a GP exercise referral programme in Hailsham, East Sussex, on modifiable coronary heart disease risk factors. *J Epidemiol Community Health*, 52, 595-601.
- Vermeire, E., Hearnshaw, H., Van Royen, P. & Denekens, J. (2001). Patient adherence to treatment: three decades of research. A comprehensive review. *Journal of Clinical Pharmacy and Therapeutics*, 26, 331-342.
- World Health Organisation (2003). Adherence to long-term therapies. *Evidence for action*. Geneva: World Health Organisation.
- World Health Organization [WHO] (2010). Global recommendations on physical activity.
- Wormald, H. & Ingle, L. (2004). GP exercise referral schemes: improving the patient's experience. *Health Education Journal*, 63, 362-373.
- Younge, J. O., Kouwenhoven-Pasmooij, T. A., Freak-Poli, R., Roos-Hesselink, J. W. & Hunink, M. M. (2015). Randomized study designs for lifestyle interventions: a tutorial. *Int J Epidemiol*, 44, 2006-19.

Bilaga 1 Kvalitetsgranskning systematisk översikt

Mall för kvalitetsgranskning av systematiska översikter enligt AMSTAR [1,2] version 2012:1.

AMSTAR ger en beskrivning av hur författarna har genomfört en systematisk översikt och om översikten uppfyller grundläggande kvalitetskrav.

Författare: Pavey et al. År: 2011 Artikelnummer: doi: 10.1136/bmj.d6462

	Ja	Nej	Kan inte svara	Ej tillämpl
<p>1. Redovisas en förutbestämd metod för genomförandet? Forskningsfrågan och inklusionskriterierna ska vara fastställda innan översikten genomförs.</p>	X			
<p>2. Gjordes studieurval och dataextraktion av två oberoende granskare? Minst två oberoende granskare ska ha utfört dataextraktionen, och ett konsensusförfarande bör vara definierat för att lösa oenigheter.</p>	X			
<p>3. Var litteratursökningen av tillfredsställande omfattning? Sökningen bör göras i minst två elektroniska databaser. Översikten ska ange de årtal och databaser som ingår (t ex Central, Embase och Medline). Ämnesord ("keywords") och/eller MeSH-termer ska anges och i tillämpliga fall sökstrategin. Alla sökningar bör kompletteras med genomgång av översiktsartiklar, läroböcker, aktuella innehålls förteckningar, ämnesspecifika databaser och register eller rådfrågning av experter, samt av referenslistorna i de framtagna studierna.</p>	X			
<p>4. Användes studiernas publikationsform som ett inklusions-/exklusionskriterium? Författarna bör ange om alla typer av publikationer omfattades av litteratursökningen. Om litteratur har exkluderats pga publikationsform (t ex "grå litteratur") eller pga språk, etc ska detta anges.</p>	X			
<p>5. Finns förteckningar över inkluderade och exkluderade studier? En förteckning över medtagna respektive uteslutna studier bör finnas i rapporten.</p>	X			

	Ja	Nej	Kan inte svara	Ej tillämpl
<p>6. Har de inkluderade studiernas karakteristika och resultat redovisats?</p> <p>Kända faktorer hos deltagarna i de utvärderade studierna ("patient characteristics"), såsom ålder, etnicitet, kön, relevanta socioekonomiska data, sjukdomstillstånd, varaktighet, svårighetsgrad och andra sjukdomar, bör anges i rapporten. Uppgifter om deltagarna, åtgärd/behandling och utfall i studierna bör presenteras i sammanfattad form, t ex i en tabell.</p>	X			
<p>7. Har den vetenskapliga kvaliteten hos de ingående studierna utvärderats och dokumenterats?</p> <p>Förutbestämda metoder för kvalitetsvärderingen ska anges. För effektstudier bör exempelvis framgå om författarna valt att bara ta med randomiserade, dubbelblindade studier med kontrollgrupper som får placebo. För andra studietyper gäller andra ställningstaganden.</p>	X			
<p>8. Har vederbörlig hänsyn tagits till de inkluderade studiernas vetenskapliga kvalitet vid formulering av slutsatserna?</p> <p>Utvärderingen av metodologisk stringens och vetenskaplig kvalitet ska framgå i översiktens analys och dess slutsatser, och tydligt anges vid utformning av rekommendationer.</p>	X			
<p>9. Användes lämpliga metoder för sammanvägning av studiernas resultat?</p> <p>Lämpligheten i att lägga samman resultaten från de olika studierna bör säkerställas genom bedömning av de ingående studiernas homogenitet (dvs Chi-2-test för beräkning av homogenitet, I²). Om heterogenitet finns bör man använda en modell som tar hänsyn till slump effekter ("random effects model") och/eller överväga om det ur klinisk synpunkt är lämpligt att slå ihop resultaten.</p>	X			
<p>10. Har sannolikheten för publikationsbias* bedömts?</p> <p>En bedömning av publikationsbias bör omfatta en kombination av grafiska hjälpmedel (t ex med "funnel plot" eller andra tester) och/eller statistiska metoder (t ex Eggers regressionsanalys).</p>		X		
<p>11. Är eventuella intressekonflikter angivna?</p> <p>Eventuella sponsorer och bidragsgivare bör tillkännages både i den systematiska översikten och i de ingående studierna.</p>	X			

* SBU:s kommentar: Publikationsbias leder till snedvriden publikation, t ex att positiva resultat publiceras oftare än negativa resultat.

Kommentarer:

Att sannolikheten för publikationsbias inte bedömts motiveras i artikeln med att underlaget var för litet.

Huvudsaklig brist i denna studie är att man inte utfört någon sammanvägd bedömning av evidensstyrkan. Studiekvaliteten är bedömd och resultatet baseras på studier av hög eller medelhög kvalitet. Men bedömning av precision, consistency och directness för det sammanvägda resultatet har inte utförts. Det är heller inget som bedöms med denna mall. Däremot har jag gjort den bedömningen för studierna i uppsatsen.

Referenser

1. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology* 2007,7:10.
2. Shea BJ, Bouter LM, Peterson J, Boers M, Andersson N, Ortiz Z, et al. External validation of a measurement tool to assess systematic reviews (AMSTAR). 2007. *PLoS ONE* 2:e1350.

Bilaga 2 Bedömning enligt GRADE

Work sheet –Certainty of evidence (GRADE)

Disease/ disorder: Fysisk inaktivitet
Intervention/ method vs. control: FFA vs. sedvanlig vård eller ingen behandling
Outcome variable: Graden av fysisk aktivitet
Included studies RCT, No 5 SR, No..... Cohort studies, No.....
Number of patients: 4193

Assessment of: risk of bias (“internal validity”) external validity” and precision

+ = No or minor problems ? = Some problems - = Major problems

Study limitations

Study	Directness	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other bias	Summary of study limitations	Precision
Taylor	?	+	-	?	+	+		?	?
Stevens	?	+	-	?	?	?		?	?
Harrison	?	?	-	?	?	?		?	?
Isaacs	?	?	-	?	+	+		?	?
Murphy	?	+	-	+	+	+		+	?

Study limitations (Randomisation, blinding, follow-up, drop-out/withdrawals, intention-to-treat)	Mark with cross-sign	
No serious limitations	0	
Some limitations (<i>but not enough to downgrade</i>)	0?	
Serious limitations (<i>downgrade one step</i>)	-1	x
Very serious limitations (<i>downgrade two steps</i>)	-2	
Comment limitations or reasons to downgrade: Oklar blindning av allokering i två studier, deltagare och		

personal ej blindad i någon studie, utvärderare en blindade i alla utom en studie, stort bortfall och låg följsamhet i alla studier, ej ITT-analys i två studier.

Consistency (Estimate of relative effect, same magnitude and direction across studies? overlapping confidence intervals?)	Mark with cross-sign	
✗ Based on meta-analysis? Statistical analysis of heterogeneity: ✗ χ^2 <input type="checkbox"/> I^2		
No serious inconsistency	0	
Some inconsistency (<i>but not enough to downgrade</i>)	0?	x
Serious inconsistency (<i>downgrade one step</i>)	-1	
Very serious inconsistency (<i>downgrade two steps</i>)	-2	
Comment limitations or reasons to downgrade: Signifikanta skillnader i några studier, ingen skillnad i andra.		

Directness (study population – external validity, specificity of intervention, relevance of the comparator to the intervention, clinical setting, adequate time of follow-up)	Mark with cross-sign	
No uncertainty	0	
Some uncertainty (<i>but not enough to downgrade</i>)	0?	x
rious indirectness (<i>downgrade one step</i>)	-1	
Very serious indirectness (<i>downgrade two steps</i>)	-2	
Comment limitations or reasons to downgrade: Stort bortfall före randomisering, oklart definierad intervention.		

Precision (Few events, wide confidence intervals that also include possible unfavourable effects)	Mark with cross-sign	
No imprecision	0	
Uncertain precision (<i>but not enough to downgrade</i>)	0?	x
Serious imprecision (<i>downgrade one step</i>)	-1	
Very serious imprecision (<i>downgrade two steps</i>)	-2	
Comment limitations or reasons to downgrade: Breda konfidensintervall.		

Publication bias (Few and small studies from one research group or company which all show the same type of results, well.known unpublished studies)	Mark with cross-sign	
Unlikely	0	x
Uncertainty (<i>but not enough to downgrade</i>)	0?	
Likely (<i>downgrade one step</i>)	-1	
Very likely (<i>downgrade two steps</i>)	-2	
Comment limitations or reasons to downgrade: Ej undersökt, men eftersom resultatet av studierna visar låg effekt är risken troligen mindre.		

Magnitude of effect	Mark with cross-sign
----------------------------	----------------------

Not relevant	0	x
Large effect (RR<0,5 or >2) (<i>upgrade one step</i>)	+1	
Very large effect (RR<0,2 or >5) (<i>upgrade two steps</i>)	+2	
Comment limitations or reasons to downgrade:		

Comments on other important aspects of the level of evidence (clear dose-response gradient that may allow upgrading?, confounders that clearly reduce the magnitude of the effect?)		
Is the sum of uncertainties (?) enough to motivate downgrading with one step?	Mark with cross-sign	
Yes	- 1	x
No	0	

Certainty of evidence	Mark with plus-signs	
High	⊕⊕⊕⊕	
Moderate	⊕⊕⊕○	
Low	⊕⊕○○	x
Very low	⊕○○○	

Namn: Helena Lagerlöf

Datum: 2016-04-12