

Noter til SfR checkliste 2

Randomiserede kontrollerede undersøgelser

1. **Sektion 1** afdækker den interne validitet af artiklen. Det vurderes, om undersøgelsen er udført grundigt, og om undersøgelsens udfald skyldes den behandling, der undersøges. Hvert spørgsmål afdækker et område af metodologien, som har væsentlig betydning for en undersøgelses konklusioner.

1.1 Er der en velafgrænset og klinisk relevant problemstilling?

Medmindre en klar og velafgrænset problemstilling er specificeret i artiklen, kan det være vanskeligt at bedømme, i hvor høj grad kriterierne er opfyldt, eller hvor relevant den er i forhold til de problemstillinger, I selv har valgt til referenceprogrammet.

1.2 Blev forsøgspersonerne randomiseret?

Det er grundlæggende, at der er foretaget en randomisering af forsøgspersonerne til fx aktiv behandling eller placebo. **Hvis randomisering ikke er angivet, skal artiklen forkastes.** Hvis beskrivelsen af randomiseringen er utilstrækkelig, eller hvis den anvendte randomiseringsmetode ikke er korrekt (fx randomisering vha. dato eller alternering mellem to grupper), eller i øvrigt ufuldkommen, skal undersøgelsen reduceres i kvalitetsniveau.

1.3 Var behandlings- og kontrolgrupperne ens ved undersøgelsens start?

Patienter, der inkluderes i en undersøgelse, bør være så ens som mulig, for at eliminere bias. Der bør i undersøgelsen angives signifikante forskelle i gruppernes sammensætning mht. køn, alder, sygdomsstadie (hvis relevant), social status, etnisk oprindelse eller eventuelle konkurrerende lidelser. Nogle af disse faktorer kan dog være indeholdt i inklusions- og eksklusionskriterierne. Er disse faktorer ikke afdækket, eller anvendes der usammenlignelige grupper, skal undersøgelsen nedgraderes i evidensniveau.

1.4 Var randomiseringen (allokeringen) skjult

Undersøgelser har vist, at hvis blinding og randomisering kan gennemskues, kan behandlingseffekten overvurderes med op til 40 %. Centraliseret randomisering (allokering), computer-randomiseringssystemer eller anvendelse af ensartede mærkede medicinpakninger anses alle for at være effektive metoder til at skjule randomiseringen og tages som udtryk for en velgennemført undersøgelse. Blindingen bør opretholdes gennem hele undersøgelsen til og med den statistiske analyse. Hvis randomiseringen er nem at gennemskue, må undersøgelsen tildeles et lavere evidensniveau, **og kan eventuelt forkastes, hvis randomiseringen ikke er skjult.**

1.5 Blev forsøgspersonerne, behandler og forsker blindet?

Blinding kan udføres på tre niveauer. Ved enkelt blinding er patienten uvidende om den aktuelle behandling. Ved dobbelt blinding er både patient og behandler uvidende om den aktuelle behandling. Ved trippel blinding er patient, behandler og forskeren

(investigator) uvidende om behandlingen. Generelt gælder det, at jo højere blindingsniveau, jo mindre er risikoen for bias.

1.6 Er alle relevante resultater målt standardiseret, gyldigt og pålideligt?

Hvis der er udeladt væsentlige resultater, eller hvis de ikke er vurderet tilstrækkeligt i artiklen, bør artiklen tildeles et lavere evidensniveau. Den bør også nedgraderes, hvis de anvendte målemetoder er utilstrækkeligt validerede eller er anvendt inkonsekvent (f.eks. ved spørgeskemaer, laboratorieanalyser osv.). Er der ofte en større eller mindre måleusikkerhed).

1.7 Bortset fra den undersøgte behandling, blev grupperne så behandlet ens?

Hvis nogle patienter har fået yderligere behandling, også selvom det kun drejer sig om vejledning og rådgivning, er denne behandling en potentiel confounder, som kan påvirke behandlingsresultatet. **Hvis grupperne ikke blev behandlet ens, skal undersøgelsen forkastes, medmindre der ikke foreligger anden evidens på dette område.**

1.8 Hvor mange af de rekrutterede personer blev medtaget i de endelige analyser?

Det skal give anledning til bekymring, hvis antallet af patienter, der udgår af studiet, er meget højt. Generelt anses en frafaldsrate (dropout) på 20 % som værende acceptabelt, men der er store variationer. Man må dels vurdere, *hvorfor* patienterne udgik, dels *hvor mange* der udgik. Frafallraten er forventeligt højere i længerevarende studier end i korte studier. En høj frafaldsrate vil oftere medføre en nedgradering af evidensniveau, snarere end forkastelse.

1.9 Blev alle forsøgspersonerne analyseret i henhold til randomiseringen?

Det forekommer sjældent, at alle patienter randomiseret til aktiv behandling, modtager behandlingen i hele forsøgsperioden. Det er ligeledes sjældent at alle patienter i kontrolgruppen ikke modtager behandling i samme periode. Desuden kan patienter afslå behandling i forløbet, ligesom der kan opstå bivirkninger eller kontraindikationer, som fører til skift til den anden gruppe. Hvis grupperne fortsat skal være sammenlignelige efter randomiseringen, *skal* resultaterne analyseres i henhold til den gruppe, patienterne oprindeligt var randomiseret til, uafhængigt af den behandling de faktisk fik. (Dette princip benævnes "*intention to treat-analyse*"). **Hvis det klart fremgår, at analyserne ikke var efter "*intention to treat-princippet*", bør artiklen forkastes.** Hvis der ikke foreligger yderligere evidens på det aktuelle område, bør artiklen indgå, men vurderes som en ikke-randomiseret kohorteundersøgelse (SfR checkliste 3).

1.10 Er resultaterne homogene mellem de forskellige undersøgelsessteder (multicenterundersøgelser)?

I multicenterundersøgelser øges pålideligheden til undersøgelsen, hvis der er opnået enslydende resultater på de forskellige deltagende centre.

Sektion 2 forholder sig til den generelle metodologiske kvalitet af RCT'en, baseret på svarene i sektion 1, graderet i tre grader:

- ++ Alle eller de fleste** kriterier er opfyldt.
Kriterier, der ikke er opfyldt, vil **meget sjældent ændre** undersøgelsens konklusioner.
- + Nogle** kriterier er opfyldt.
De kriterier, der ikke er opfyldt eller grundigt beskrevet, vil **sjældent ændre** studiets konklusioner.
- ÷ Få eller ingen** kriterier er opfyldt.
Undersøgelsens konklusioner kan **meget vel ændres**.

Denne graduering, sammenholdt med typen af undersøgelsen, afgør det endelige evidensniveau.

Formålet med de tre andre spørgsmål er at sammenfatte jeres syn på RCT'ens kvalitet og dens anvendelighed på netop jeres referenceprograms patientmålgruppe.

Sektion 3 sammenfatter nøglepunkter om artiklen, som skal anvendes, når der senere i arbejdsprocessen skal udarbejdes konkrete anbefalinger til referenceprogrammet.

28.01.2004