

**Model for udarbejdelse af fælles regionale vurderinger af
nye lægemidlers og nye indikationers kliniske merværdi**

Metodehåndbogen

Version 1.0

Model for udarbejdelse af fælles regionale vurderinger af nye lægemidlers og nye indikationers kliniske merværdi

Metodehåndbog version 1.0

Sprog: Dansk

Versionsnummer 1.0

Indhold

1.0	Introduktion	4
1.1	Om håndbogen.....	4
1.2	Definition og afgrænsning	4
1.3	Overordnet proces	5
1.4	Organisatorisk ramme	6
2	Fase I – Forberedelsesfasen	6
2.1	Varsling og foreløbig ansøgning	6
2.2	Nedsættelse og indkaldelse af fagligt forum eller fagudvalg	7
2.3	Habilitet	7
2.4	Udarbejdelse af protokol.....	7
2.4.1	Udarbejdelse af fokuserede spørgsmål og tilhørende PICO	8
2.4.2	Spørgsmål om effekten af en intervention	9
2.4.3	Typer af studiedesign.....	10
2.4.4	Afgrænsning vedrørende sprog.....	10
2.4.5	Metode for sammenligning.....	10
3	Fase II Udførelsesfasen	11
3.1	Litteratursøgning	11
3.2	Udvælgelse af evidens.....	12
3.3	Vurdering af evidens fra primære studier.....	12
3.4	Sammenfatning af evidens.....	13
3.4.1	Dataekstraktion.....	13
3.4.2	Metode for sammenligning.....	13
3.5	Gradering af evidens (GRADE)	14
3.5.1	Vurdering af tiltroen til det enkelte effektmål	15
3.5.2	Nedgradering af evidensen for effekt af interventioner	16
3.5.3	Evidensprofiler	18
3.5.4	Vurdering af tiltroen til det overordnede evidensniveau	18
3.6	Kvalitetssikring af ansøgningen	19
3.7	Fra evidens til kategori	19
3.7.1	Rationale for metodisk tilgang til vurdering af merværdi.....	21
3.7.2	Vægtning af effektmål.....	21
3.7.3	Relative effekter sammenholdt med væsentlighedskriterier.....	22
3.7.4	Vurdering af evidensens kvalitet.....	24
3.7.5	Vurdering af klinisk betydning	24
3.8	Samlet vurdering af den kliniske merværdi	24
3.9	Færdiggørelse af rapport	25
3.10	Godkendelse	25
3.11	Høring	25

4 Fase III – Beslutningsfasen	25
4.1 Endelig anbefaling om ibrugtagning som standardbehandling	25
4.2 Offentliggørelse	25

1.0 Introduktion

Danske Regioners bestyrelse har besluttet at oprette Medicinrådet, som bygger på erfaringerne fra Rådet for Anvendelse af Dyr Sygehusmedicin (RADS) og Koordinationsrådet for ibrugtagning af Sygehusmedicin (KRIS). Medicinrådets overordnede formål er:

- at sikre en hurtig og ensartet anvendelse af nye såvel som eksisterende lægemidler på tværs af sygehuse og regioner
- at stille større krav til dokumentation for, at ny og eksisterende medicin er til gavn for patienterne
- at sikre et stærkere grundlag for Amgros' prisforhandlinger og udbud.

Medicinrådet foretager overordnet set to typer af vurderinger; vurdering af nye lægemidler (eksklusiv generika og biosimilære lægemidler), og vurdering af flere lægemidler indenfor samme terapiområde.

1.1 Om håndbogen

Denne metodehåndbog omhandler alene vurdering af nye lægemidlers og nye indikationers kliniske merværdi. Metoden vedr. terapiområder er beskrevet i "Metodehåndbog for udarbejdelse af fælles regionale behandlingsvejledninger for terapiområder" som findes her. ([link](#)).

DEFACTUM har ydet konsulentbistand under udarbejdelsen af metodehåndbogen for vurdering af nye lægemidlers kliniske merværdi, og har kommenteret på udkast til metodevejledningen udarbejdet af Danske Regioner, RADS Sekretariatet og AMGROS. Modellen bygger på Sundhedsstyrelsens "Model for udarbejdelse af Nationale Kliniske Retningslinjer", Metodehåndbog version 2.1. og er tilpasset arbejdsområde - og processer i Medicinrådet.

Modellen indeholder en generisk tidsplan, vejledning til udformning af projektprotokollen, retningslinjer for indholdet af ansøgningen fra industrien, retningslinjer for Medicinrådets kvalitetssikring af industriens ansøgninger samt og vejledning til vurdering af klinisk merværdi for det nye lægemiddel eller den nye indikation.

Håndbogen skal være med til at sikre patienter, sundhedspersonale og virksomheders indsigt i Medicinrådets arbejde og metode. Herudover er håndbogen arbejdsredskab for fagudvalgsmedlemmer (kliniske eksperter) og medarbejdere i Medicinrådets Sekretariat og Rådet.

1.2 Definition og afgrænsning

Medicinrådet vurderer, hvorvidt ny markedsføringsgodkendt sygehusmedicin eller nye markedsføringsgodkendte indikationsudvidelser skal tages i brug som national standardbehandling.

Følgende nye lægemidler (generika og biosimilære produkter undtaget) kan efter ansøgning fra firmaet blive vurderet af Medicinrådet:

- Lægemidler, der kun må udleveres til sygehuse (BEGR)
- Lægemidler, som kun må udleveres til sygehuse og kun én gang efter samme recept (AP4BG)

- Lægemidler, som kun må udleveres til sygehuse eller efter ordination af nærmere bestemte speciallæger (AP4NB)
- Lægemidler, som kun må udleveres til sygehuse eller efter ordination af nærmere bestemte speciallæger, som Sundhedsstyrelsen fastsætter for hvert enkelt lægemiddel (NBS)

De fire udleveringsgrupper svarer til de grupper, som er omfattet af den prisloftaftale, som er indgået mellem Lægemiddelindustriforeningen (LIF), Sundheds- og Ældreministeriet og Danske Regioner i 2016.

At et lægemiddel anbefales til ibrugtagning som standardbehandling indebærer, at lægemidlet indføres som et alment anvendt behandlingstilbud til en patientgruppe, og at lægemidlet er umiddelbart tilgængeligt på sygehusafdelingerne.

Anbefalingerne bygger overordnet på tre trin:

1. En sundhedsfaglig vurdering og kategorisering af det nye lægemiddels/den nye indikations kliniske merværdi for patienterne.
2. En økonomisk faglig vurdering i form af en omkostningsanalyse, hvor de inkrementale omkostninger ved ibrugtagning af det nye lægemiddel/den nye indikation vurderes.
3. En forhandling, hvor vilkårene for det offentliges indkøb fastlægges.

Den sundhedsfaglige vurdering er systematisk udarbejdede vurderinger af det nye lægemiddels kliniske merværdi sammenlignet med et eller flere relevante andre lægemidler (komparatorer). Med *merværdi* menes den ekstra (inkrementale) kliniske effekt et lægemiddel kan tilbyde sammenlignet med nuværende standardbehandling. Vurderingerne indeholder således den videnskabelige og kliniske argumentation for, hvilken ekstra klinisk effekt det nye lægemiddel tilbyder, sammenholdt med komparator. Den kliniske merværdi opsummeres i en af følgende kategorier (Danske Regioners notat Model for vurdering af lægemidler dateret 04-04-2016):

- Kategori 1: Stor merværdi
- Kategori 2: Vigtig merværdi
- Kategori 3: Lille merværdi
- Kategori 4: Ingen merværdi
- Kategori 5: Negativ merværdi
- Kategori 6: Ikke-dokumenterbar merværdi

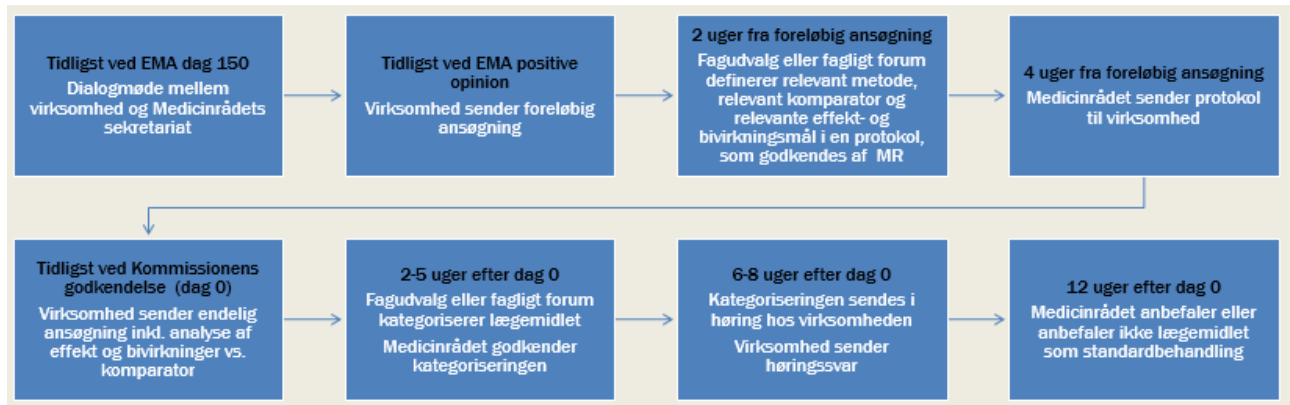
1.3 Overordnet proces

Processen for anbefalingen af et nyt lægemiddel som standardbehandling forløber overordnet i disse trin:

- Indsendelse af foreløbig ansøgning
- Udarbejdelse af projektprotokol
- Indsendelse af endelig ansøgning
- Kvalitetssikring af ansøgningen
- Kategorisering af det nye lægemiddel/den nye indikation
- Høring
- Omkostningsanalyse
- Forhandling

- Anbefaling af, hvorvidt det nye lægemiddel/den nye indikation skal tages i brug som standardbehandling eller ej
- Implementering af anbefalingen
- Monitorering af forbruget af lægemidlet

De overordnede tidsrammer fremgår af figuren herunder



Medicinrådet kan iværksætte "clock-stop" i tilfælde af komplicerende faktorer, for eksempel hvis eksterne interessenter ikke kan leve op til de anførte tidsrammer eller hvis ansøgningen er af en sådan kvalitet, at Medicinrådet selv skal udføre litteratursøgning og analyser.

1.4 Organisatorisk ramme

Medicinrådet består af tre enheder: Rådet, Sekretariatet og fagudvalgene/faglige fora. Kommissariatet for Medicinrådet sætter følgende organisatoriske ramme for arbejdet:

- *Rådet* under Medicinrådet, der har til opgave at godkende projektprotokollen, den fælles regionale vurdering af lægemidlers merværdi samt formulere den anbefaling af, hvorvidt det nye lægemiddel/den nye indikation skal ibrugtages som standardbehandling ud fra en samlet vurdering af klinisk merværdi og omkostninger.
- Et *sekretariat* i Medicinrådet, der betjener de forskellige fagudvalg/faglige fora samt Rådet.
- *Fagudvalg/faglige fora* som er de sundhedsfaglige arbejdsgrupper bag for hver fælles regionale vurdering. Arbejdsgrupperne består af medlemmer udpeget af regioner og faglige selskaber. Herudover vil der være repræsentation af et medlem fra en patientforening.

2 Fase I – Forberedelsesfasen

2.1 Varsling og foreløbig ansøgning

Et medicinalfirma kan tage initiativ til at igangsætte en vurdering, såfremt deres nye lægemiddel forventes at opfylde de kriterier, der er nævnt i afsnit 1.2.

Processen kan begynde ved dag 150 i godkendelsesprocessen i European Medicines Agency (EMA). Dag 150 er således tidligste dato for et dialogmøde mellem firmaet og sekretariatet. På

mødet orienterer firmaet om lægemidlet, virkningsmekanismen, studierne, der ligger til grund for ansøgningen i EMA, forventet indikation, forventet tidspunkt for positive opinion, forventet udleveringsgruppe osv. Sekretariatet orienterer om Medicinrådets proces og metode vedrørende nye lægemidler.

Når EMA har anbefalet, at lægemidlet godkendes (positive opinion), kan firmaet sende en foreløbig ansøgning. Ansøgningen skal indeholde en kort beskrivelse af sygdommen og populationen samt relevante lægemiddelkarakteristika (ansøgningsskema udarbejdes senere). Hvis det nye lægemiddel forventes godkendt til flere indikationer, indsender firmaet en ansøgning per indikation. Den foreløbige ansøgning danner udgangspunkt for fagpersonernes og sekretariatets drøftelser omkring protokollen (se afsnit 2.4).

2.2 Nedsættelse og indkaldelse af fagligt forum eller fagudvalg

I forlængelse af varslingsmødet indkalder Sekretariatet de personer, der skal sikre den kliniske og patientmæssige vinkel på vurderingen.

Medicinrådet kan for eksempel beslutte at nedsætte et permanent fagligt forum på kræftområdet, som vurderer alle nye kræftlægemidler. Forummet kan for hver ny vurdering suppleres med 2-3 specialister og 1 patientrepræsentant indenfor den konkrete kræftsygdom.

For nye lægemidler til behandling af andre sygdomme end kræft indkalder Medicinrådet det eksisterende fagudvalg og supplerer med en patientrepræsentant.

For sygdomsområder, hvor der ikke findes et eksisterende fagudvalg, nedsætter Medicinrådet et nyt fagudvalg. Fagudvalgene sammensættes efter samme principper som beskrevet i metodehåndbogen vedrørende terapiområder (se link).

Der kan som udgangspunkt ikke sendes suppleanter til møder i et fagligt forum eller fagudvalg. Såfremt en repræsentant udtræder, udpeges der samtidig en ny repræsentant.

2.3 Habilitet

For at sikre gennemsigtighed og håndtere eventuelle habilitetsproblemer anvender Medicinrådet habilitetserklæringer (notat vedr. habilitet udarbejdes senere).

2.4 Udarbejdelse af protokol

Fælles regionale vurderinger af nye lægemidlers kliniske merværdi udarbejdes på baggrund af en protokol, som Rådets formandskab har godkendt.

Protokollen skal indeholde følgende elementer:

- Fokuserede spørgsmål med tilhørende udførlige PICO beskrivelser (se afsnit 2.4.1).
- En vægtning af alle effektmål ("kritiske" vs "vigtige" vs "begrænset vigtige") samt en a priori definition af, hvad der er klinisk relevante forskelle for de enkelte effektmål.
- Søgestrategi (søgetermer og informationskilder)
- Kriterier for selektion af studier (fx studiedesign, in- og eksklusionskriterier)
- En beskrivelse af hvordan data skal kvalitetsvurderes
- En beskrivelse af hvilke data, der skal ekstraheres og hvordan
- En beskrivelse af hvordan data skal opsummeres (syntetiseres)

Den godkendte protokol vil blive delt med det relevante firma umiddelbart efter godkendelse, både skriftligt og ved behov på et møde. Protokollen bliver også offentliggjort på Medicinrådets hjemmeside.

Inden første møde i forummet/fagudvalget udarbejder Sekretariatet på baggrund af den foreløbige ansøgning forslag til fokuserede spørgsmål og de øvrige elementer i protokollen. Udkastet drøftes med forummet/fagudvalget. Her er det væsentligt, at forummet/fagudvalget har mulighed for at foreslå andre klinisk relevante spørgsmål, så det sikres, at protokollen kommer til at afdække de mest relevante kliniske problemstillinger. Spørgsmålene skal optimalt formuleres, inden litteraturen på området er kendt.

Protokollen godkendes af Rådets formandskab og er "den røde tråd" igennem hele vurderingen. Protokollen udgør grundlaget for firmaets efterfølgende ansøgning, for Sekretariatets kvalitetssikring og eventuelle justering af ansøgningen, samt for den endelige anbefaling og kategorisering.

2.4.1 Udarbejdelse af fokuserede spørgsmål og tilhørende PICO

Til hvert af de fokuserede spørgsmål knytter sig en specifikation af patientgruppen, interventionen, alternativet/erne til interventionen og effektmålene. Dette afspejles i akronymet PICO (Population, Intervention, Comparison and Outcomes).

Population (patientgruppen): Definition af sygdommen/tilstanden og patientmålgruppen med relevante demografiske faktorer (f.eks. performance status).

Intervention (interventionen): Definition af interventionen, altså det nye lægemiddel. Herunder beskrives hvilke variationer (eksempelvis dosis, administrationsform o.l.), der findes af interventionen, og det angives specifikt hvilke variationer, der inkluderes i vurderingen.

Comparator (komparator): Definition af sammenligningsalternativet/erne til det nye lægemiddel. Medicinrådets endelige anbefaling er helt afhængig af, hvilken komparator, der sammenlignes med (både klinisk og omkostningsmæssigt). Derfor forbeholder Medicinrådet sig retten til at definere komparator. Her vil specialisterne vurdere af, hvilket lægemiddel, der i den daglige klinik udgør det reelle alternativ, veje tungt.

Outcome (effekt mål): Definition af de effektmål, som vurderes, at være nødvendige at vide noget om, for at opveje gavnlige mod skadelige effekter af den undersøgte intervention. Her skal indtænkes relevante effektmål for alle beslutningstagere, det vil sige patienten såvel som klinikerne.

Effektmålene kan samles under følgende overskrifter:

- Dødelighed
- Alvorlige symptomer og bivirkninger
- Helbreds-/funktionsrelateret livskvalitet
- Ikke-alvorlige symptomer og bivirkninger.

Effekt mål kan omfatte kliniske hændelser (f.eks. frakturer, slagtilfælde) og andre patientrelevante effektmål (f.eks., symptomer, livskvalitet, funktionsevne). Relevante bivirkninger skal altid inkluderes, herunder alvorlige bivirkninger (SAE, serious adverse events), f.eks. dødelighed, blødninger. Udover at gøre sig klart hvilke effektmål, der er relevante for vurderingen, bør fagudvalget også gøre sig klart, hvordan disse skal måles. I forhold til målinger af livskvalitet anvendes enten generiske eller sygdomsspecifikke redskaber, der er generelt accepterede og har høj validitet.

Derudover er det vigtigt at fastlægge hvilke tidspunkter, der er relevante at se på. Dette er vigtigt, da både effekter og skadebivirkninger kan variere betydeligt over tid, og det kan således have stor betydning, om man vælger at fokusere på effekten efter 6 uger, måneder eller år.

Forummet/fagudvalget udvælger og vægter effektmålene ("kritisk" vs "vigtig" vs "begrænset vigtige") jf GRADE-metoden. Valgene beskrives i protokollen. Den endelige vurdering af klinisk merværdi samt den samlede kvalitetsvurdering, skal primært baseres på de kliniske effektmål.

Vigtigheden af de identificerede effektmål rangeres på en skala fra 1-9 (7-9: kritiske effektmål, 4-6: vigtige effektmål og 1-3: begrænset vigtige effektmål) for at skelne mellem de forskellige kategorier.

Resultater for alle relevante kritiske og vigtige effektmål bør inkluderes og beskrives i ansøgningen og rapporten. Derimod skal resultater for "begrænset vigtige" effektmål ikke beskrives.

Surrogatmål

Anvendelse af surrogatmål er kun relevant, når der mangler data for klinisk relevante effektmål og der samtidig er dokumenteret en klar sammenhæng mellem surrogateffekt og klinisk hændelse, altså hvor et surrogatmål forventes af prædiktere kliniske effekter baseret på epidemiologisk, patofysiologisk, terapeutisk eller anden videnskabelig evidens.

Ved anvendelse af et surrogatmål som erstatning for det kliniske effektmål, skal tiltroen til evidensen derfor altid jævnfør GRADE (se afsnit 3.5.2) nedgraderes med ét niveau, idet man sjældent kan være 100 % sikker på, at surrogateffekten reelt prædikerer den kliniske effekt.

Progressionsfri overlevelse (PFS) kan dog som udgangspunkt accepteres som et kritisk effektmål og vil ikke blive nedgraderet i de tilfælde, hvor EMA har accepteret PFS som primært effektmål i godkendelsesprocessen.

2.4.2 Vurdering af klinisk relevans

Når de "kritiske" og "vigtige" effektmål er definerede, er det vigtigt, at man for hvert effekt- og bivirkningsmål overvejer og forhåndsdefinerer, hvad der er den *mindste klinisk relevante forskel*. Forhåndsdefinitionen af klinisk relevante forskelle er vigtig for at undgå subjektive ad hoc beslutninger, der for eksempel er baseret på den tilgængelige evidens. Forskellene beskrives i protokollen.

GRADE bruger termen "Clinical Decision Threshold" for klinisk relevante forskelle, og definerer dette som den forskel, der afgør, hvorvidt man vælger – eller fravælger et lægemiddel fremfor et andet. Se afsnit 3 for en forklaring af GRADE.

For binære effektmål vil den klinisk relevante forskel ofte udtrykkes i en risikoreduktion. Grænsen for en klinisk relevant forskel vil ofte afhænge af hvilket effektmål, der er tale om. For eksempel er en mortalitetsreduktion på 10 % mere vigtig, end en reduktion på 10 % i andelen af patienter, der oplever for eksempel hudkløe.

For kontinuerlige effektmål gælder samme princip om, at den klinisk relevante forskel skal afspejle den forskel i effekt, som den velinformerede patient, eller kliniker, mener er relevant for, hvorvidt han/hun vil vælge denne behandling frem for en anden.

For kontinuerlige effektmål vil den klinisk relevant forskel ofte være udtrykt i en gennemsnitlig forskel eller proportionen af patienter, der har en fordel af behandlingen, som ligger over den klinisk relevante forskel. For eksempel kan den klinisk relevante forskel på reduktion i nakkesmerter defineres som en reduktion på 10 på en VAS skala (VAS er en visuel analog skala, der måler smerte fra 1-100). Et lægemiddel, der reducerer nakkesmerter kan for eksempel anføres som en gennemsnitlig reduktion på 30, men kan også anføres som en proportion af patienter fx "den mindste klinisk relevante forskel på 10 point blev opnået hos 40 % af patienterne".

Beslutningen om, hvad der er klinisk relevante forskelle, er ofte vanskelig. For enkelte behandlingsområder og effektmål findes validerede samlinger af "klinisk relevante forskelle", der bygger på evidens fra spørgeskemaundersøgelser i patientgruppen. For eksempel har "Cochrane back pain group" besluttet sig for klinisk relevante grænser for rygsmerter: [<http://processbook.kce.fgov.be/node/118>, 09-09-16].

Det er ikke anbefalelsesværdigt at lave systematiske og specifikke søgninger på emnet, da der ikke findes en universel accepteret og valideret metode.

Det vil dog ofte være muligt at fremsætte et ikke-valideret forslag over for forummet/fagudvalget. Dette forslag kan fremkomme på baggrund af klinisk relevante forskelle defineret i nyere eksisterende retningslinjer indenfor fagområdet. Det er vigtigt, at den endelige grænse for klinisk relevante forskelle defineres i samarbejde med klinikerne i forummet/fagudvalget.

2.4.3 Typer af studiedesign

Randomiserede studier (lodtrækningsforsøg) anses for at være det mest valide beslutningsgrundlag for at vurdere forskelle i effekt mellem et nyt lægemiddel og komparator. Som udgangspunkt anvender Medicinrådet derfor kun randomiserede studier.

Hvis EMAS's godkendelse af det nye lægemiddel/en ny indikation afviger fra dette, vil Medicinrådet overveje at inddrage andre data.

Hvis Medicinrådet har valgt en anden komparator, end den, der er anvendt i godkendelsesstudierne, kan der være behov for yderligere litteratur i form af randomiserede studier, der vurderer komparators effekt og bivirkninger i en population, der er så tæt på den her relevante population som muligt. Såfremt der ikke findes resultater fra randomiserede studier, der afspejler den relevante population, kan data fra danske nationale databaser indgå i vurderingen.

2.4.4 Afgrænsning vedrørende sprog

Der skal som udgangspunkt inkluderes litteratur på engelsk og de skandinaviske sprog.

2.4.5 Metode for sammenligning

Hvis der foreligger en direkte randomiseret sammenligning af det nye lægemiddel vs komparator, danner denne udgangspunkt for vurderingen af den kliniske merværdi. I de tilfælde, hvor Medicinrådet har valgt en anden komparator end den, firmaet har lavet sammenlignende studier overfor, skal der foretages en indirekte sammenligning. Medicinrådet

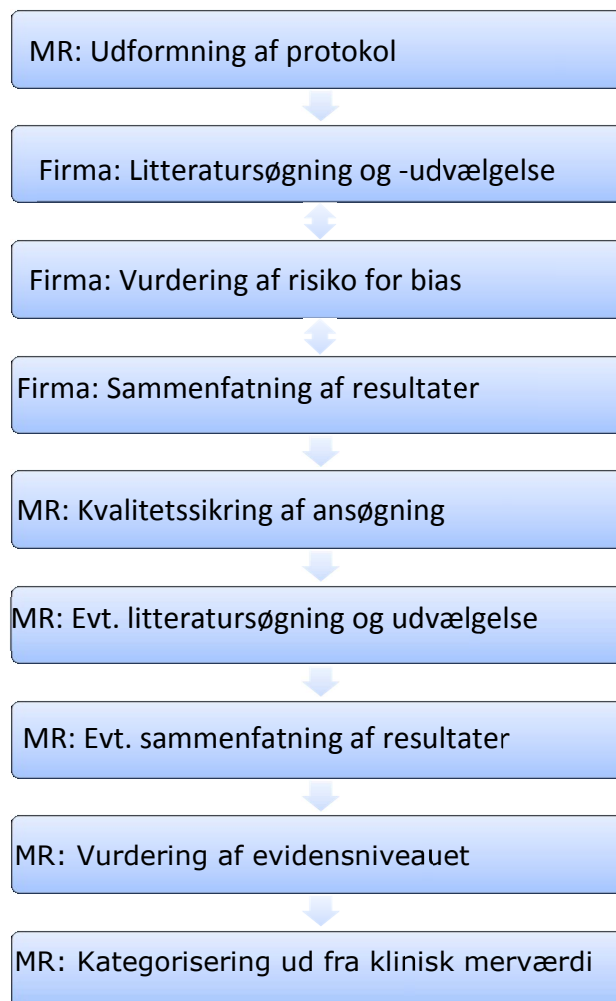
vil i protokollen beskrive, hvordan den relative effekt af det nye lægemiddel og komparator bedst vurderes i den konkrete situation (for eksempel narrativ analyse, metaanalyse, netværks-metaanalyse).

3 Fase II Udførelsesfasen

Når Europa-Kommissionen har givet markedsføringstilladelse til det nye lægemiddel, kan firmaet indsende den endelige ansøgning. Ansøgningen skal som før nævnt følge protokollen og skal indeholde en vurdering af det nye lægemiddels effekt sammenlignet med komparator, samt indeholde en vurdering af risikoen for bias i de inkluderede studier (ansøgningskema udarbejdes senere).

Medicinrådet anvender GRADE (Grading of Recommendations Assessment, Development and Evaluation system), der er et internationalt brugt systematisk og transparent system til at vurdere kvaliteten af evidens og til at give anbefalinger. Den følgende tekst vil kort beskrive systemet. For mere detaljeret information henvises til GRADE arbejdsgruppens hjemmeside:
<http://www.gradeworkinggroup.org>

GRADE-processen starter med formuleringen af de fokuserede spørgsmål, som foregår i protokolfasen. I vurderingen af nye lægemidler er dette Medicinrådets opgave. Herefter foretages en systematisk søgning efter litteratur. Dette er firmaets opgave. Efter at litteraturen er fundet, og man har fastlagt evidensgrundlaget for hvert fokuseret spørgsmål (firmaets opgave), giver GRADE eksplicite kriterier, hvormed man kan vurdere kvaliteten af evidensen ud fra fire kriterier fra meget lav til høj. Kriterierne inkluderer risiko for bias, inkonsistens, unøjagtighed, hvor indirekte evidensen er samt risikoen for publikationsbias. Firmaet vurderer risiko for bias i ansøgningen. Denne proces sikrer, at der systematisk identificeres, vurderes og sammenfattes viden om interventionerne. Hele ansøgningen kvalitetssikres af Medicinrådet, der har ret til at udføre andre litteratursøgninger og -udvælgelser samt analyser, såfremt Medicinrådet ikke er enig i firmaets tilgang. Firmaets ansøgning offentliggøres som en del af det endelige beslutningsgrundlag.



3.1 Litteratursøgning

Litteraturen skal søges i henhold til beskrivelsen i protokollen.

De fælles regionale vurderinger af nye lægemidlers kliniske merværdi bygger i udgangspunktet på de randomiserede data i EMAs Scientific Discussion vedrørende det nye lægemiddel. Andre relevante data, som firmaet er i besiddelse af, kan evt. bidrage til den samlede vurdering. Alle resultater vedrørende komparator skal være publiceret i enten peer-review'ede artikler eller i EMAs Scientific Discussions (da EMAs proces anerkendes som en form for peer review).

Eventuel litteratursøgning efter primærlitteratur foretages i OVID Medline. Ansøgningen skal indeholde en søgeprotokol. En søgeprotokol har til formål at skabe overblik og transparens i processen med at indsamle evidens. Ligeledes er søgeprotokollen med til at sikre konsistens ved opfølgning eller gentagelse af søgningen.

Søgeprotokollen skal indeholde følgende elementer:

- Fokuserede spørgsmål
- In- og eksklusionskriterier (defineret på forhånd)
- Informationskilder (databaser, internetsider, registre, etc.)
- Søgestrategi for hver enkelt informationskilde
- Eventuelle restriktioner (f.eks. publikationstype)
- Strategi for gennemgang og udvælgelse af den fundne litteratur

Skabelon til søgeprotokol findes i ansøgningsskemaet.

Hele processen udføres af firmaet og valideres/kvalitetssikres af Sekretariatet i samarbejde med fagudvalget/det faglige forum.

3.2 Udvalgelse af evidens

Efter hver søgning skal den fundne litteratur sorteres og vurderes. Det skal vurderes, om den fundne litteratur er af tilstrækkelig høj kvalitet og om den indeholder de elementer, der skal bruges for at kunne vurdere den kliniske merværdi. Til at bevare et løbende overblik over hvilken evidens, der er fundet til at besvare de forskellige fokuserede spørgsmål, findes en tabel i ansøgningsskemaet.

3.3 Vurdering af evidens fra primære studier

Individuelle studiers risiko for bias vurderes ved hjælp af checklister. Til vurdering af bias i randomiserede skal Cochranes Risk of Bias Tool benyttes (http://handbook.cochrane.org/chapter_8/8_assessing_risk_of_bias_in_included_studies.htm).

Firmaet anfører sin vurdering af risikoen for bias i tabellen i ansøgningsskemaet. Vurderingen valideres/kvalitetssikres af Sekretariatet i samarbejde med fagudvalget/det faglige forum.

Videre læsning:

Lundh A, Gøtzsche PC. Recommendations by Cochrane Review Groups for assessment of the risk of bias in studies. BMC Med Res Methodol. 2008 Apr 21;8:22.

3.4 Sammenfatning af evidens

Ved hjælp af litteratursøgnings- og sorteringsprocessen beskrevet i de foregående afsnit, identificeres de studier, der danner evidensgrundlaget for vurderingerne. I dette afsnit beskrives hvorledes evidensen sammenfattes.

Hele processen udføres af firmaet og valideres/kvalitetssikres af Sekretariatet i samarbejde med fagudvalget/det faglige forum.

3.4.1 Dataekstraktion

Dataekstraktion

Ved dataekstraktion fremlægges relevante oplysninger fra inkluderede studier. Processen bør ske systematisk, konsekvent og stringent for hvert enkelt studie indeholdende samme typer af oplysninger for de enkelte studier. Dataekstraktionstabellerne i ansøgningsskemaet skal anvendes. Det er vigtigt, at beskrivelserne ikke efterlader tvivl om karakteristika for patientgruppe, intervention og komparator, herunder kontekst.

3.4.2 Metode for sammenligning

Direkte sammenligninger fra randomiserede studier er at foretrække, når effekten af forskellige interventioner vurderes i forhold til hinanden. Desværre foreligger de ønskede head-to-head data ikke altid. Indirekte sammenligninger er sammenligninger, der er lavet mellem interventioner, der ikke er sammenlignet direkte med hinanden. Indirekte sammenligninger fra randomiserede studier er ikke randomiserede sammenligninger, men er i realiteten observationelle resultater på tværs af studier, og bør derfor nøje vurderes i forhold til bias (for eksempel i form af confounding).

I situationer, hvor der både eksisterer indirekte og direkte evidens for en sammenligning, bør der - med mindre den direkte evidens er behæftet med meget store mangler i studiedesign - primært lægges vægt på resultaterne fra de direkte sammenligninger.

Resultaterne fra de primære studier kan præsenteres deskriptivt (narrativt) eller om muligt i en meta-analyse (kvantitativ opsummering). En meta-analyse er en metode, som gør det muligt at sammenfatte resultater fra flere randomiserede kontrollerede studier, så man opnår et bedre datagrundlag for at vælge den bedste behandling. Det er ikke altid tilrådeligt at lave en meta-analyse. Studier kan være for uensartede til at det giver mening at sammenfatte dem i en meta-analyse. Dette kan skyldes at for eksempel populationerne eller interventionerne i de enkelte studier er for forskellige. Der kan også være tale om at studierne design er for forskellige. I sådanne tilfælde frarådes det også at lave en simpel optælling over, hvor mange studier der støtter en given intervention og hvor mange der fraråder den, da denne form for analyse antager, at studier er fuldstændigt ens med hensyn til deres størrelse (antal patienter eller events) og deres metodologiske kvalitet. Det er derfor nødvendigt, mere detaljeret, at beskrive hvad de enkelte studier fandt med hensyn de inkluderede effektmål, hvor store de er og hvad deres metodologiske kvalitet er. Et eksempel på en narrativ analyse findes her ([LINK](#))

Metaanalysen tager udgangspunkt i enkeltstudiers effektestimater og konfidensintervaller. I selve analysen foretages en vægtning af disse estimater, der primært er baseret på de absolutte tal fra studierne (antal hændelser og deltagere i hvert enkelt studie). Meta-analysen beregner et samlet effektestimat for hvert effektmål. Med det samlede estimat etableres et konfidensinterval. Den grafiske fremstilling af en meta-analyse kaldes et "forest plot".

Det anbefalede output af en meta-analyse er relative mål fx RR, OR eller HR, fordi relative mål har den fordel, at de som regel er stabile på tværs af populationer med forskellige baseline værdier.

Absolutte mål har den fordel at de er meget anvendelige for klinisk beslutningstagning, og det vil ofte være nødvendigt at estimere absolutte mål for at vurdere, hvorvidt forskellen er klinisk relevant (se afsnit om klinisk relevante forskelle). Den valide metode til at estimere den absolutte risikoreduktion er at gange den relative risiko fra meta-analysen med det mest plausible estimat af baselinerisikoen i kontrolgruppen. Da baselinerisiko i praksis ofte er svært at estimere, er det en mulighed at tage denne risiko fra resultaterne i kontrolgruppen. Hvis data vurderes at være tilstrækkeligt homogene på tværs af studier, er det muligt at lave meta-analyser.

Principperne for meta-analyser følges som beskrevet i:
Cochrane Handbook for Systematic Reviews of Interventions

Den statistiske fremstilling af indirekte sammenligninger kan foretages i form af fx en netværks metaanalyse. Når Medicinrådet i protokollen beskriver, hvilken analysemetode der skal benyttes, er beslutningen truffet i tæt samarbejde med en biostatistiker. Tilsvarende vil Sekretariatets kvalitetssikring af de indsendte analyser ske i tæt samarbejde med en biostatistiker.

Medicinrådet vil i de tilfælde, hvor der ikke findes head-to-head data typisk bede firmaet om at udføre indirekte analyser i narrativ form. Medicinrådet foretager selv statistiske bearbejdnings på indirekte sammenligninger i de tilfælde, hvor det skønnes nødvendigt.

Videre læsning:

Cochrane Handbook Systematic Reviews of Interventions <http://handbook.cochrane.org/>
EUnetHTAs vejledning: COMPARATORS & COMPARISONS: Direct and indirect comparisons

3.5 Gradering af evidens (GRADE)

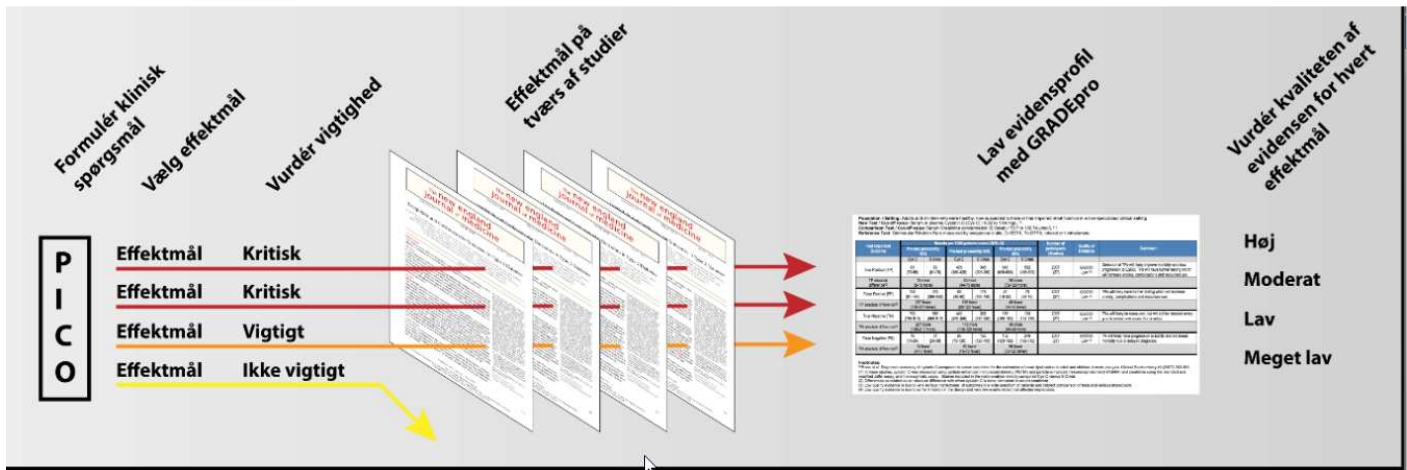
Når man systematisk har gennemgået den identificerede litteratur, der relaterer sig til de fokuserede spørgsmål, skal man vurdere kvaliteten af evidensen.

GRADE (Grading of Recommendations Assessment, Development and Evaluation system) er et internationalt brugt systematisk og transparent system til at vurdere kvaliteten af evidens og til at give anbefalinger.

GRADE systemet adskiller sig fra andre evidensvurderingssystemer på den måde, at man i GRADE vurderer kvaliteten af evidensen per effektmål og på tværs af studier.

I ansøgningen fra firmaet skal firmaet give sit bud på risikoen for bias i de inkluderede studier. De øvrige trin i GRADE-processen varetager Sekretariatet i tæt samarbejde med klinikerne i fagudvalget/det faglige forum.

En skematisk proces over GRADE processen se fig. 1



Den følgende tekst vil gennemgå systemet i nærmere detaljer. For mere information henvises til GRADE working groups hjemmeside, hvor der blandt andet er links til en artikelserie i Journal of Clinical Epidemiologi, som gennemgår de forskellige aspekter af GRADE processen (se www.gradeworkinggroup.org).

GRADE-processen begynder med fokuserede spørgsmål, hvor også alle de kritiske og vigtige effektmål er specificerede. Efter at evidensen er identificeret, giver GRADE eksplicite kriterier, hvormed man kan vurdere kvaliteten af evidensen ud fra fire kriterier fra meget lav til høj.

Kriterierne inkluderer: risiko for bias, inkonsistens, unøjagtighed, hvor indirekte evidensen er samt risikoen for publikations bias.

Graderingen af evidensen sker i to etaper. Først graderes tiltroen til estimerne af effekten for de enkelte effektmål. Herefter graderes det overordnede niveau af evidens for det fokuserede spørgsmål.

3.5.1 Vurdering af tiltroen til det enkelte effektmål

For hvert fokuseret spørgsmål foretager Medicinrådets Sekretariat en separat vurdering af evidensen. Såfremt der foreligger data for flere relevante effektmål (f.eks. død, myokardieinfarkt, apopleksi, blødning) lægges der særligt vægt på de for patienterne mest kritiske effektmål.

Eksempelvis er død og myokardieinfarkt mere patientrelevante end kolesterolniveau eller carotis intima tykkelse. Der bør altid medtages minimum én skadesvirkning i vurderingen.

Før den endelige vurdering af evidensens kvalitet pr. klinisk spørgsmål, foretages en evidensvurdering pr. effektmål. I GRADE opererer man med fire niveauer for kvaliteten af evidensen:

Høj (⊕⊕⊕⊕)

Vi er meget sikre på, at den sande effekt ligger tæt på den estimerede effekt.

Moderat (⊕⊕⊕○)

Vi er moderat sikre på den estimerede effekt. Den sande effekt ligger sandsynligvis tæt på denne, men der er en mulighed for, at den er væsentligt anderledes.

Lav (⊕⊕○○)

Vi har begrænset tiltro til den estimerede effekt. Den sande effekt kan være væsentligt anderledes end den estimerede effekt.

Meget lav (⊕○○○)

Vi har meget ringe tiltro til den estimerede effekt. Den sande effekt vil sandsynligvis være væsentligt anderledes end den estimerede effekt.

For behandlingsinterventioner har man som udgangspunkt høj tiltro til evidens fra randomiserede forsøg (lodtrækningsforsøg), mens observationelle studier (databasestudier) som udgangspunkt er af lav kvalitet.

Tiltroen til evidens fra randomiserede forsøg kan nedgraderes efter en gennemgang af studierne risiko for bias, konsistensen af resultaterne på tværs af studier, den samlede præcision og størrelsen af effektestimater, hvor direkte de inkluderede studier svarer på det fokuserede spørgsmål samt risikoen for publikationsbias.

Table x ned- og opgradering af evidens for effekt af interventioner (GRADE)

Evidensniveau udgangspunkt	Studiedesign	Nedgradering	Opgradering	Endeligt evidensniveau
Høj	Randomiserede forsøg	<ul style="list-style-type: none">• Risiko for bias• Inkonsistens• Indirekte evidens• Unøjagtighed• Publikationsbias	<ul style="list-style-type: none">• Effektstørrelse• Dosisrespons• Konfounding	Høj
				Moderat
Lav	Observationelle studier			Lav
				Meget lav

3.5.2 Nedgradering af evidensen for effekt af interventioner

Evidensen kan nedgraderes et til to evidensniveauer for hvert af følgende domæner:

- Risiko for bias
- Inkonsistens
- Indirekte evidens
- Unøjagtighed
- Publikationsbias

For hvert domæne kan en nedgradering foretages. Er problemet mindre nedgraderes et niveau (eksempelvis fra høj til moderat) og er problemet stort nedgraderes det to niveauer (eksempelvis fra høj til lav). Nedenfor gennemgås kort de forskellige domæner, for en nærmere beskrivelse se de forskellige artikler i GRADE serien i Journal of Clinical Epidemiology.

Risiko for bias

Findes der retningslinjer eller systematiske oversigtsartikler, hvor der er foretaget en relevant vurdering af bias i primærstudierne, kan disse overtages. Ellers kan man bruge Cochrane's *Risk of Bias Tool* til at vurdere randomiserede forsøg. Eksempelvis vil et randomiseret forsøg med manglende blinding, risiko for selektionsbias grundet manglende skjult allokering og stort frafald nedgraderes to niveauer, hvorimod et randomiseret forsøg alene med moderat frafald vil nedgraderes et niveau. Selektiv rapportering af effektmål (kun de mest positive rapporteres) kan også være en bias, som bør vurderes.

Inkonsistens

Hvis der er inkonsistens i resultaterne fra forskellige studier nedgraderes evidensen. Inkonsistens kan skyldes forskelle i populationer, interventioner, behandling i kontrolgruppen eller definition af effekter. Er der en god forklaring, der opfylder subgruppe kriterierne, herunder at de stemmer med få og a priori definerede hypoteser, så nedgraderer man ikke, men vil snarere differentiere de anbefalinger, man vil komme med til forskellige subgrupper.

Indirekte evidens

Relaterer evidensen sig ikke direkte til ens kliniske spørgsmål foretages en nedgradering. Indirekte evidens kan have to baggrunde:

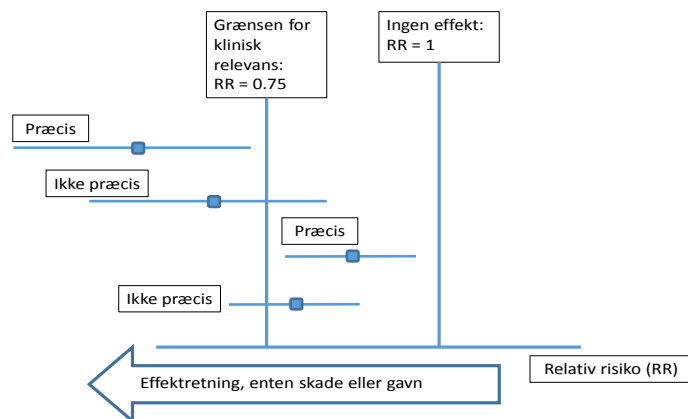
1. man ønsker at sammenligne to behandlinger over for hinanden, men de enkelte behandlinger er kun sammenlignet overfor placebo.
2. Der er forskelle i population, intervention, kontrolgruppe eller den måde effekterne er målt på mellem det kliniske spørgsmål og de tilgrundliggende studier. Eksempler: Spørgsmålet relaterer sig til ældre med multiple komorbiditeter, men evidensen stammer fra midaldrende relativt raske personer. Spørgsmålet relaterer sig til kirurgisk behandling på et provinshospital i et udviklingsland, men evidensen relaterer sig til behandling på et højtspecialiseret universitetshospital i et industrialiseret land. Spørgsmålet relaterer sig til 40 mg simvastatin behandling i kontrolgruppen, men evidensen stammer fra studier med 10 mg simvastatin. Spørgsmålet relaterer sig til diabetiske komplikationer, men evidensen stammer fra effekt på blodsukker.

Link

[http://www.jclinepi.com/article/S0895-4356\(11\)00183-1/fulltext](http://www.jclinepi.com/article/S0895-4356(11)00183-1/fulltext)

Unøjagtighed

Hvis effektestimateret er unøjagtigt, dvs. konfidensintervallet er bredt, foretages en nedgradering. Det vurderes, om konfidensintervallet overlapper den mindste klinisk relevante forskel, og om man ville komme med forskellige anbefalinger i hver sin ende af konfidensintervallet.



Publikationsbias

Hvis der er tegn på publikationsbias (manglende publicering af hele studier) eller selektiv rapportering af effekter (kun de mest positive rapporteres) foretages en nedgradering. Det ses eksempelvis, hvis der kun er publiceret studier, der anvender metoder til at måle effektmål, som ikke svarer til, hvad man vil forvente at anvende (eksempelvis brug af en atypisk depressionsskala).

3.5.3 Evidensprofiler

Formålet med præsentationen af evidensprofilerne i ansøgningskemaet er at give et overblik over studierne karakteristika og punkttestimater.

3.5.4 Vurdering af tiltroen til det overordnede evidensniveau

Når tiltroen til de enkelte effekttestimater er vurderet på effektmål-niveau, skal den samlede tiltro til evidensen vurderes på tværs af effektmål. Dette afgøres af de effektmål, der er defineret som kritiske for udarbejdelsen af vurderingen, og det overordnede evidensniveau fastsættes, som udgangspunkt, efter det lavest vurderede kritiske effektmål.

Når vurderingen af klinisk merværdi skal formuleres, tager Medicinrådet udgangspunkt i tiltroen til hvorvidt effekttestimaterne er pålideligt underbyggede og kan støtte en given kategorisering.

Når nye lægemidler vurderes, vil der ofte kun være ganske få eller måske kun et enkelt RCT-studie til rådighed. Set i lyset af, at en given positiv effekt beskrevet i ét studie, i mange tilfælde, ikke kan genfindes i efterfølgende studier, vil man (med rette) som bedømmer kunne føle sig utryk ved at betegne ét RCT-studie som "høj kvalitet af evidens". På den anden side vil det ikke være hensigtsmæssigt at nedgradere kvaliteten af evidens automatisk i de tilfælde, hvor der kun findes et studie. Et enkelt stort veltilrettelagt og veludført multicenter-RCT kan levere høj kvalitet af evidens. I tilfælde, hvor der kun findes et studie, anbefales særlig grundig granskning af alle relevante domæner (risiko for bias, indirekte evidens, unøjagtighed og publikationsbias)

Videre læsning:

GRADE arbejdsgruppen (www.gradeworkinggroup.org)

Artikelserie i Journal of Clinical Epidemiology om GRADE
(http://www.gradeworkinggroup.org/publications/JCE_series.htm)

3.6 Kvalitetssikring af ansøgningen

Sekretariatet kvalitetssikrer i tæt samarbejde med fagudvalget/det faglige forum firmaets ansøgning efter samme anvisning, som er beskrevet i afsnittet ovenfor. Derudover foretages det meste af GRADE-vurderingen i Sekretariatet/fagudvalget/forummet (undtaget vurderingen af risiko for bias, hvor firmaet også udarbejder en analyse). Såfremt firmaets og Medicinrådets analyser varierer, er det Medicinrådets analyser, der kommer til at ligge til grund for vurderingen. Både firmaets ansøgning og Medicinrådets analyser offentliggøres.

3.7 Fra evidens til kategori

Kategorisering af merværdi er en central proces i forbindelse med vurdering af nye lægemidler. Med kategoriseringen af merværdi ønskes det at tydeliggøre niveauet for de samlede effekter og bivirkninger af en given behandling på patient-relevante effektmål. Kategoriseringen af merværdi kan desuden bidrage med vigtige oplysninger til patienter og læger, om de forventede fordele af en ny behandling.

De endelige kategorier er som nævnt i afsnit 1.2 følgende:

- Kategori 1: Stor merværdi
- Kategori 2: Vigtig merværdi
- Kategori 3: Lille merværdi
- Kategori 4: Ingen merværdi
- Kategori 5: Negativ merværdi
- Kategori 6: Ikke-dokumenterbar merværdi.

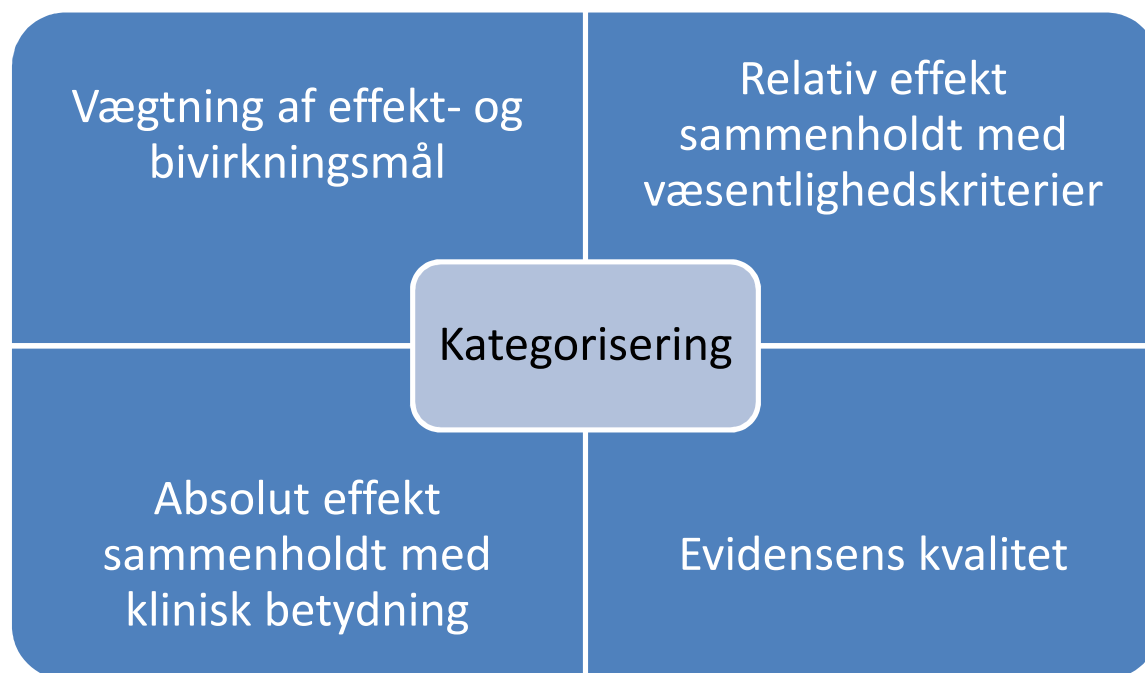
I det følgende beskrives den metodetilgang, som anvendes ved kategorisering, og som sikrer, at vurderinger af klinisk merværdi gennemføres systematisk og transparent. Denne proces sikrer et pålideligt vurderingsgrundlag for kategoriseringer og dermed kommende anbefalinger.

Medicinrådet og Sekretariatet vil på baggrund af indkomne ansøgninger, foreliggende studier og andet materiale foretage en lægefaglig og statistisk vurdering af det nye lægemiddel og indplacere det i en af seks kategorier for merværdi (jævnfør Danske Regioners notat: Metode for vurdering af lægemidler).

Det er ikke muligt at anvende en stringent algoritme for denne proces, og det skal derfor understreges, at processen indebærer både kvalitative og kvantitative trin.

Overordnet bygger vurderingen på følgende fire elementer:

- Vægtningen af effektmål
- Vurderingen af lægemidlets relative effekt sammenholdt med internationalt anerkendte væsentlighedskriterier
- Vurderingen af lægemidlets absolutte effekt sammenholdt med de klinisk relevante forskelle
- Vurderingen af evidensens kvalitet



Merværdi kan som omtalt beskrives som den ekstra kliniske effekt et lægemiddel tilbyder i forhold til nuværende behandling med hensyn til livsforlængelse (mortalitet), symptomer, bivirkninger og livskvalitet. Merværdien vurderes med udgangspunkt i de GRADE-vurderede effekt- og risikoforhold. For at der kan blive tale om en klinisk merværdi, skal effekten af et nyt lægemiddel være signifikant større end for sammenligningsgrundlaget. Dette er tilfældet for de tre første kategorier (kategori 1-3).

Ved kategori 4-6 vurderes det, at det nye lægemiddel ikke tilfører klinisk merværdi. Her kan merværdien kategoriseres som:

Kategori 4 - "ingen merværdi", hvor den kliniske merværdi beviseligt ikke er større end værdien ved det eksisterende lægemiddel (komparator).

Kategori 5 - "negativ merværdi", hvor det nye lægemiddels kliniske merværdi er signifikant lavere end den kliniske værdi af det eksisterende lægemiddel (komparator).

Kategori 6 - "ikke-dokumenterbar merværdi", hvor en eventuel klinisk merværdi ikke kan kvantificeres, fordi det videnskabelige datagrundlag ikke tillader dette (pga. sparsom evidens eller meget lav kvalitet af evidens), eller hvor den kliniske merværdi kun er dokumenteret via surrogatmål. Her kan protokolleret ibrugtagning overvejes.

3.7.1 Rationale for metodisk tilgang til vurdering af merværdi

I dette afsnit beskrives den metodiske tilgang til vurdering af merværdi og de definitioner og kriterier, der ligger til grund for en kategorisering samt evidens- og metodegrundlag, som ligger til grund for udarbejdelse af de væsentlighedskriterier, der er centrale for kategorisering af et nyt lægemiddel. I nedenstående afsnit beskrives blandt andet hvordan væsentlighedskriterier og vægtningen (relevansen) af effektmål hænger sammen.

De nye lægemidler, der tilfører klinisk merværdi, opdeles i kategori 1-3, og det er særligt disse tre kategorier som skal afgrænses metodisk og indenfor hvilke størrelsen af merværdi skal beskrives. I appendiks A er metodetilgangen til den indledende operationalisering for kategorisering af klinisk merværdi beskrevet i detaljer.

3.7.2 Vægtning af effektmål

Som beskrevet ovenfor er relevante effektmål vægtet som henholdsvis "kritiske", "vigtige" og "ubegrænset vigtige" i protokollen. De kritiske vægter højest i vurderingen og de vigtige næsthøjest. De begrænset vigtige indgår ikke. Vægtningen vil variere mellem sygdomsområder (fx vil forlænget overlevelse være kritisk i nogle sammenhænge, og livskvalitet i andre, mens afhjælpning af ikke- alvorlige men meget hyppige bivirkninger, vil være kritisk indenfor nogle sygdomsområder)

3.7.3 Relative effekter sammenholdt med væsentlighedskriterier

I denne del af vurderingen tages afsæt i estimat og konfidensinterval. Som det ses i tabellerne herunder, skal konfidensintervallet på estimatet indenfor et givent effektmål ligge helt under et vist væsentlighedskriterium, for at kunne indplacere effekten i en foreløbig kategori. Det er således afgørende, at den øvre grænse for konfidensintervallet er mindre end væsentlighedskriteriet (Ref: General Methods - Version 4.2. IQWIG. April 2015).

Dermed inddrages ikke kun effekttestimatet, men også til en vis grad sandsynligheden for den givne effekt. Hermed tilvejebringes et mere validt grundlag for kategorisering af merværdi. Væsentlighedskriterier er som nævnt individuelle for de enkelte effektmål/grupper af effektmål.

Væsentlighedskriterierne for de forskellige effektmål og de foreløbige merværdikategorier er vist i nedenstående tabeller.

Tabel x: Levetidsforlængelse. Statistiske væsentlighedskriterier for relative effektmål. Den foreløbige definition af merværdien er her alene baseret på kvantitative egenskaber.

Foreløbig kategori	Levetidsforlængelse[#] vurderet pba. relativ risiko, odds ratio eller hazard ratio (væsentlighedskriteriet er øvre grænse for konfidensintervallet)
Foreløbig kategori 1: Stor merværdi	Stor øgning i livsforlængelse (signifikant) baseret på et væsentlighedskriterium = 0,85[#]
Foreløbig kategori 2: Vigtig merværdi	Moderat (signifikant) øget livsforlængelse baseret på et væsentlighedskriterium = 0,95[#]
Foreløbig kategori 3: Lille merværdi	Enhver øgning (signifikant) i livsforlængelse baseret på et væsentlighedskriterium = 1,00[#]

[#] Væsentlighedskriterium som ikke må overskrides af øvre grænse for konfidensintervallet.

Med det sædvanlige signifikansniveau på 5 %, er enhver statistisk signifikant stigning i overlevelsestid mindst klassificeret som "lille merværdi", altså en tærskelværdi/øvre grænse for konfidensintervallet < 1. Hvis merværdien derimod skal kategoriseres i kategori 1, må den øvre grænse for konfidensintervallet på estimatet ikke overskride 0,85.

Tabel x: Alvorlige symptomer og bivirkninger. Statistiske væsentlighedskriterier for relative effektmål. Definitionen af den foreløbige merværdi er her alene baseret på kvantitative egenskaber.

Foreløbig kategori	Alvorlige symptomer og bivirkninger^{&} vurderet pba. relativ risiko, odds ratio eller hazard ratio (væsentlighedskriteriet er øvre grænse for konfidensintervallet)
Foreløbig kategori 1: Stor merværdi	Længerevarende fravær baseret på et væsentlighedskriterium = 0,75[#] (risiko \geq 5 %) ^b
Foreløbig kategori 2: Vigtig merværdi	Reduktion eller relevant undvigelse baseret på et væsentlighedskriterium = 0,90[#]
Foreløbig kategori 3: Lille merværdi	Enhver reduktion baseret på et væsentlighedskriterium = 1,00[#]

[&] For alle patientrapporterede effektmål forudsættes anvendelse af valideret og etableret måleskala.

[#] Væsentlighedskriterium som ikke må overskrides af øvre grænse for konfidensintervallet.

^b Risiko skal ved baseline være mindst 5 % for mindst én af de sammenlignede grupper hvilket inddrager relevansen af risikoen i vurderingen.

I forhold til alvorlige symptomer og bivirkninger vil enhver statistisk signifikant reduktion også mindst repræsentere en "lille merværdi" på baggrund af alvorlighedsgraden af effektmålet. Væsentlighedskriterierne for kategori 1 og 2 er derimod lavere end for levetidsforlængelse. For at opnå en stor merværdi – kategori 1 - indenfor disse effektmål er der samtidig opstillet et kriterium om, at risikoen for den undersøgte hændelse skal være mindst 5 % i mindst en af de sammenlignede grupper. Hermed får incidensen på befolkningsniveau betydning for kategoriseringen og kriteriet skærper de kvantitative krav til en høj kategorisering.

I de tilfælde hvor helbreds-/funktionsrelateret livskvalitet er vægtet som et kritisk eller vigtigt effektmål, indgår det i den samlede vurdering af klinisk merværdi. Det vurderes hvorvidt de dokumenterede ændringer i livskvalitet ved den ny behandling er stor, vigtig eller lille. Stor klinisk merværdi - kategori 1 - kræver en dokumenteret stor positiv forbedring. Ingen forbedring eller forværring af livskvaliteten tæller ned i den samlede vurdering af merværdi.

Tabel x: Ikke- alvorlige symptomer og bivirkninger. Statistiske væsentlighedskriterier for relative effektmål. Den foreløbige merværdi er her alene baseret på kvantitative egenskaber.

Foreløbig kategori	Ikke- alvorlige symptomer og bivirkninger vurderet pba. relativ risiko, odds ratio eller hazard ratio (væsentlighedskriterium er øvre grænse for konfidensintervallet)
Foreløbig kategori 1: Stor merværdi	<i>Ikke muligt at opnå kategori 1</i>
Foreløbig kategori 2: Vigtig merværdi	Reduktion eller relevant undvigelse baseret på et væsentlighedskriterium = 0,80[#]
Foreløbig kategori 3: Lille merværdi	Reduktion eller relevant undvigelse baseret på et væsentlighedskriterium = 0,90[#]

Væsentlighedskriterium som ikke må overskrides af øvre grænse for konfidensintervallet

For ikke- alvorlige symptomer og bivirkninger ligger væsentlighedskriterierne lavere end blot signifikant forskel i kategorien "lille merværdi" grundet alvorlighedsgraden af effektmålene. Således starter væsentlighedskriteriet her på 0,9. Denne kategorisering medfører at statistisk signifikante effekter over væsentlighedskriteriet på 0,9 skal kategoriseres som "ingen merværdi". Grundet alvorlighedsgraden af effektmålene er det heller ikke muligt at kategorisere effektforhold indenfor disse effektmål som "stor merværdi".

3.7.4 Vurdering af klinisk betydning

I protokollen er den mindste klinisk betydende forskel defineret. Resultaterne fra analyserne sammenholdes med disse definitioner og kan – hvis de viser sig ikke at være klinisk betydende – medføre en nedgradering af den foreløbige kategori.

3.7.5 Vurdering af evidensens kvalitet

Evidensens kvalitet vurderes pr. effektmål og pr. klinisk spørgsmål som beskrevet i afsnit 3.5. I den foreløbige vurdering af merværdi kan der nedjusteres, såfremt kvaliteten af evidensen er lav eller meget lav.

3.8 Samlet vurdering af den kliniske merværdi

Den samlede kategorisering af klinisk merværdi ift. det vurderede lægemiddel inddrager således systematisk:

- Vægtningen af effektmål
- Vurderingen af lægemidlets relative effekt sammenholdt med internationalt anerkendte væsentlighedskriterier
- Vurderingen af evidensens kvalitet
- Vurderingen af lægemidlets absolutte effekt sammenholdt med de klinisk relevante forskelle.

Det er ikke muligt at fremlægge en eksplicit fremgangsmåde, som kan benyttes i forbindelse med alle kategoriseringer. Det er væsentligt at understrege, at vurderingsprocessen indeholde både kvantitative og kvalitative trin. Medicinrådet ønsker at fremlægge de kriterier og forudsætninger, som kategoriseringerne bygger på, på transparent vis. Særligt skal balancen

mellem de væsentligste fordele og risici beskrives og kommenteres, således at det bliver synligt, hvorledes vægtningen af disse er foregået. I forbindelse med op- og nedkategoriseringer skal det især tydeliggøres, hvilke kriterier der har påvirket kategoriseringen og i hvor høj grad.

3.9 Færdiggørelse af rapport

Det er Sekretariatet, som samler og sammenfatter den egentlige løbende tekst i den endelige rapport vedrørende den fælles regionale vurdering af klinisk merværdi af nye lægemidler. Fagudvalget/det faglige forum kommenterer, redigerer og godkender den færdige rapport og kategorisering, som herefter skal godkendes af Rådet. En skabelon for rapporten kan ses her ([link](#)).

3.10 Godkendelse

Der afholdes et månedligt Medicinrådsmøde. Rapporten vedrørende det nye lægemiddel sendes med den øvrige dagsorden til Rådets medlemmer før mødet. Medicinrådet tager stilling til rapporten på mødet.

3.11 Høring

Den foreløbige kategorisering og rapport vil efter Medicinrådets møde blive sendt til firmaet. Herefter har firmaet mulighed for at fremsende høringssvar. Både den oprindelige rapport inklusiv kategorisering og firmaets høringssvar offentliggøres, når hele processen er afsluttet.

Sekretariatet behandler høringssvarene i dialog med fagudvalget/det faglige forum. Hvis høringssvarene giver anledning til en ændring af den oprindelige kategori, skal denne godkendes af Medicinrådet. Ændringer i rapporten, som ikke påvirker den endelige kategori, kan godkendes af Medicinrådet pr. e-mail.

4 Fase III – Beslutningsfasen

Sideløbende med de ovenfor beskrevne processer, har Amgros udarbejdet en omkostningsanalyse og forhandlet vilkår for det offentlige indkøb af lægemidlet med firmaet.

4.1 Endelig anbefaling om ibrugtagning som standardbehandling

Når Medicinrådet formulerer den endelige anbefaling om ibrugtagning, sker det efter en samlet vurdering af den kliniske merværdi og de samlede omkostninger.

4.2 Offentliggørelse

Efter Medicinrådet har truffet sin beslutning, offentliggøres den endelige rapport vedrørende klinisk merværdi, den eventuelt oprindelige rapport og høringssvar, samt industriens ansøgninger (konfidentielle data undtaget) og protokollen.

Appendiks A: Operationalisering for kategorisering af *merværdi*

De lægemidler der tilfører merværdi opdeles i tre kategorier:

- Kategori 1: Stor merværdi sammenlignet med standardbehandling
- Kategori 2: Vigtig merværdi sammenlignet med standardbehandling
- Kategori 3: Lille merværdi sammenlignet med standardbehandling

Da de fleste metodiske diskussioner netop omhandler disse tre kategorier må især kategorierne 1 til 3 afgrænses og beskrives for at vurdere størrelsen af merværdi. Med hensyn til vurdering af kategori 1 til 3 skal blandt andet kvalitative egenskaber (vægtning af effektmål) og eksplicit kvantitative egenskaber (for eksempel "større" vs. "moderat" stigning i livsforlængelse) vurderes. I tabel 1 vises eksempler på en første operationalisering af de tre kategoriers merværdi.

Udgangspunktet for vurdering af merværdi er patient-relevante effektmål samt et kvalitetsvurderet evidensgrundlag via GRADE.

Tabel 1. Første operationalisering af kategori 1-3 for merværdi.

Kategori	Stor merværdi (kategori 1)	Eksempler på forbedringer kan være sygdomsremission, markant stigning i overlevelsetiden, langvarigt fravær af alvorlige sygdomssymptomer, eller udtalt fravær af alvorlige bivirkninger.
	Vigtig merværdi (kategori 2)	Eksempler på forbedringer er mærkbar lindring af sygdomstilstanden, en moderat stigning i overlevelsetiden, lindring af alvorlige symptomer, og fravær af alvorlige bivirkninger og væsentligt fravær af andre bivirkninger.
	Lille merværdi (kategori 3)	Eksempler på forbedringer er reduktion i ikke alvorlige symptomer og / eller fravær af bivirkninger.

Med henblik på yderligere operationalisering af merværdi skal der kigges på effektmål. Det skal indenfor grupper af effektmål vurderes, eller angives, hvad der bevirker at en merværdi kategoriseres som fx stor, jf. tabel 2. Denne gruppering gennemføres udelukkende med henblik på opstilling af væsentlighedskriterier indenfor hver kategori. Jo alvorligere effekter, jo højere væsentlighedskriterier (og dermed lettere at opnå større merværdi).

Tabel 2. Beskrivelse af merværdi indenfor grupper af effektmål.

		Effektmål			
		Levetidsforlængelse	Symptomer (morbidity)	Helbredsrelateret livskvalitet	Bivirkninger
Kategori	Stor merværdi Vedvarende og stor forbedring i effektforhold, som ikke tidligere er opnået i forhold til relevant sammenligningsgrundlag	Stor forbedring i overlevelsestid	Langvarigt fravær af alvorlige (eller svære) symptomer (eller senkomplikationer)	Stor forbedring i livskvalitet	Udtalt fravær af alvorlige (eller svære) bivirkninger
	Vigtig merværdi Markant forbedring i effektforhold, som ikke tidligere er opnået i forhold til relevant sammenligningsgrundlag.	Moderat forbedring i overlevelsestid	Lindring af alvorlige (eller svære) symptomer (eller senkomplikationer) Væsentlig reduktion i ikke-alvorlige (eller ukomplicerede) symptomer (eller senkomplikationer)	Vigtig forbedring i livskvalitet	Relevant undvigelse af alvorlige (eller svære) bivirkninger. Væsentlig undvigelse af andre (ikke-alvorlige eller ukomplicerede) bivirkninger
	Lille merværdi Moderat og ikke kun marginal forbedring i effektforhold, som ikke tidligere er opnået i forhold til relevant sammenligningsgrundlag.	Enhver forbedring i overlevelsestid	Enhver reduktion af alvorlige (eller svære) symptomer (eller senkomplikationer) Reduktion i ikke-alvorlige (eller ukomplicerede) symptomer (eller senkomplikationer)	Relevant forbedring i livskvalitet	Enhver statistisk signifikant reduktion i alvorlige (eller svære) bivirkninger Relevant undvigelse af andre (ikke-alvorlige eller ukomplicerede) bivirkninger

I tabel 3 fortsættes kategorisering med udgangspunkt i overordnede vægtninger af effektmål, hvor blandt andet alvorlige symptomer og bivirkninger grupperes samlet. Den overordnede vægtning af effektmål har betydning for hvordan det givne lægemiddel indplaceres i én af de seks kategorier. Effektmål vægtes altså efter, hvor vigtige de er for beslutningstagning omkring anvendelse af nye lægemidler. Det er vigtigt at bemærke, at denne vægtning påvirker frembringelsen af væsentlighedskriterier, men at vægtningen af effektmål ift. det enkelte lægemiddel godt kan se anderledes ud og bør foretages på tværs af alle effektmål.

Tabel 3. Matrix af kategori for merværdi samt vægtede effektmål

	Levetidsforlængelse	Alvorlige symptomer og bivirkninger	Helbredsrelateret livskvalitet	Ikke-alvorlige symptomer og bivirkninger
Kategori 1: Stor merværdi	Stor øgning i livsforlængelse	Længerevarende frihed	Stor forbedring	Ikke relevant
Kategori 2: Vigtig merværdi	Moderat øget livsforlængelse	Reduktion eller relevant undvigelse	Vigtig forbedring	Vigtig undvigelse
Kategori 3: Lille merværdi	Enhver øgning i livsforlængelse	Enhver reduktion	Relevant forbedring	Relevant undvigelse

Det kan fremhæves, at "ikke- alvorlige symptomer og bivirkninger" kun kan kategoriseres i kategori 2 og 3 ("vigtig" og "lille" merværdi) og altså ikke på noget tidspunkt kan kategoriseres i kategori 1. Dette skyldes alene, at mindre væsentlige effektforhold ikke kan bidrage med stor merværdi ind i sundhedsvæsenet.

Den videre operationalisering vedrører opgørelse af effektstørrelser, hvor relative effektmål altid bør foretrækkes frem for absolutte effektmål, da generaliserbarheden af absolutte mål til andre populationer er begrænset. Relative mål derimod har den fordel, at de som regel er stabile på tværs af populationer med forskellige baseline værdier, hvilket er nyttigt, når resultater fra forskellige studier skal kombineres i en metaanalyse. Ved absolutte effektmål er det desuden ikke muligt at opstille generelle væsentlighedskriterier, da udgangspunktet (forhåndsrisikoen) ikke er kendt. Det vil som nævnt altid være relevant at præsentere symptomforekomst i patientgruppen. Dette muliggør større sammenligningsgrundlag indenfor og mellem de enkelte vurderinger. Typisk anvendes relativ risiko (RR), odds ratio (OR) eller hazard ratio (HR) med henblik på at estimere merværdi. Relativ risiko (RR), som ofte benyttes, er et forholdstal, hvormed sandsynligheden for en hændelse i en gruppe i forhold til en anden gruppe angives. Således angiver en RR på 0,5 en halvering af risiko ift. anvendelse af andet lægemiddel. Såfremt evidensgrundlaget ikke muliggør fremstilling af relative effektmål kan absolutte effektmål beskrives.