

Van aanbeveling tot gevalideerde kwaliteitsindicator: een evidence-based en transparant ontwikkelp proces



CEQUID:

Cebam Quality Indicator Development

Methodologisch document

Contactinformatie

Bij vragen over dit document kan u terecht bij Cebam, het Belgisch Centrum voor Evidence-Based Medicine, cel Evaluatie.

Bijgedragen door:

dr. Marijke Avonts (cel Evaluatie Cebam)

Liesbeth Meel, MSc (cel Evaluatie Cebam)

dr. Deborah Seys (cel Evaluatie Cebam)

dr. Carolien Strouwen (coördinator cel Evaluatie Cebam)

prof. dr. Steve Van den Bulck (U Hasselt, ACHG KU Leuven)

dr. Anne-Catherine Vanhove (cel Evaluatie Cebam)

prof. dr. Patrik Vankrunkelsven (directeur Cebam)

Sofie Wouters, Msc (cel Evaluatie Cebam)

Eindredactie: Jasmine Lemmens

Contactpersoon:

dr. Carolien Strouwen, coördinator cel Evaluatie Cebam

Kapucijnenvoer 7

3000 Leuven

evaluatie@cebam.be

Versie datum: augustus 2026

Refereren naar dit document kan als volgt: Cebam. CEQUID: Cebam Quality Indicator Development. Methodologisch document [Internet]. Leuven: Cebam; 2026. Beschikbaar via: <https://www.cebam.be/>.

Copyright © 2026 Cebam vzw. Alle rechten voorbehouden.

Dit document is eigendom van Cebam en wordt beschermd door de toepasselijke auteursrechtwetgeving. Geen enkel onderdeel van deze publicatie mag worden gereproduceerd, opgeslagen in een geautomatiseerd gegevensbestand, openbaar gemaakt, verspreid of doorgegeven in welke vorm of op welke wijze dan ook zonder voorafgaande schriftelijke toestemming van Cebam, tenzij wettelijk toegestaan.

Dit document wordt uitsluitend ter beschikking gesteld voor informatieve doeleinden. Hoewel de inhoud met zorg werd samengesteld, aanvaardt Cebam geen aansprakelijkheid voor eventuele fouten, onvolledigheden of schade voortvloeiend uit het gebruik van deze informatie.

Inhoud

Contactinformatie.....	2
Inhoud	3
Samenvatting.....	5
Lijst van afkortingen.....	6
1. Doel en opzet van dit document.....	7
a) Doelgroep.....	7
b) Uitgangspunten van CEQUID	7
c) Opbouw van het document.....	8
d) Reikwijdte van de methodologie	9
e) Leeswijzer	9
2. Inleiding.....	10
a) Waarom QI ontwikkelen?.....	10
b) Internationale context en aanleiding ontwikkeling CEQUID	10
3. Conceptueel kader.....	12
a) Wat is een kwaliteitsindicator?.....	12
b) Het model van Donabedian.....	12
c) Kwaliteit volgens The Institute of Medicine	13
4. Kwaliteitscriteria voor QI.....	14
a) Waarom kwaliteitscriteria noodzakelijk zijn	14
b) Validiteit.....	14
c) Betrouwbaarheid.....	15
d) Haalbaarheid.....	15
e) Discriminerend vermogen.....	15
f) Interpreteerbaarheid.....	16
g) Aanvaardbaarheid.....	16
5. CEQUID: 8 stappen.....	17
Fase 1 - Voorbereidende fase.....	18
Fase 2 - Ontwikkelingsfase	25

Fase 3 – Praktijkttoets.....	40
6. Validatie van QI.....	47
Het AIRE-instrument.....	47
Validatiecriteria van Cebam.....	47
AIRE item 8 – Formele vaststelling via consensusprocedure.....	47
AIRE item 18 – Praktijkttoets en ongewenste neveneffecten	47
7. Algemene discussie.....	49
Inleiding.....	49
Sterktes van de CEQUID-methodologie.....	49
Integratie van bestaande conceptuele kaders.....	49
Transparantie van besluitvorming	49
Expliciete koppeling aan evidentie.....	49
Verplichte praktijktoets.....	50
Evaluatie van ongewenste effecten	50
Aandachtspunten van de CEQUID-methodologie.....	50
Hoge methodologische eisen	50
Afhankelijkheid van beschikbare richtlijnen.....	50
Consensus blijft subjectief.....	50
Contextgebondenheid	50
8. Besluit.....	51
9. Referenties	52

Samenvatting

5 UITGANGSPUNTEN

Evidentie Kwaliteitsindicatoren (QI) gebaseerd op aanbevelingen uit hoogwaardige klinische praktijkrichtlijnen	Transparantie Alle methodologische keuzes expliciet gedocumenteerd
Stakeholderbetrokkenheid Multidisciplinair, inclusief patiëntenperspectief	Praktische bruikbaarheid Haalbaar en bruikbaar in de praktijk
Kritische reflectie Mogelijke ongewenste effecten expliciet geëvalueerd	

8 STAPPEN IN 3 FASEN

Fase 1 — Voorbereiding	Fase 2 — Ontwikkeling	Fase 3 — Praktijkttoets
Stap 1 Ontwikkelgroep <i>Multidisciplinaire zorgverlener(s), methodoloog, patiëntenvertegenwoordiger</i>	Stap 3 Richtlijnen & bestaande QI <i>Systematische zoektocht; beoordeling via AGREE II en AIRE</i>	Stap 8 Bruikbaarheid en ongewenste effecten <i>Haalbaarheid, betrouwbaarheid, interpreteerbaarheid, discriminerend vermogen, gaming, risicoselectie, meetfixatie</i>
Stap 2 Afbakening bereik <i>Doelgroep, context, kwaliteitsdomeinen (IOM), beoogd gebruik</i>	Stap 4 Selectie aanbevelingen <i>Sterkte van aanbeveling, klinische relevantie, en meetbaarheid</i>	
	Stap 5 Consensusprocedure (optioneel) <i>Indien meer dan 10 aanbevelingen tweeronden Delphi met consensusbijeenkomst ter beoordeling van de aanbevelingen</i>	
	Stap 6 Matching & formulering QI <i>Koppeling aanbeveling–QI; nieuwe QI bij ontbrekende match</i>	
	Stap 7 Consensusprocedure <i>Twee rondes Delphi met consensusbijeenkomst ter beoordeling van de indicatoren</i>	

STERKTES EN AANDACHTSPUNTEN

Sterktes + Expliciete koppeling aan wetenschappelijke evidentie + Integratie van AGREE II, AIRE en RAND/UCLA + Verplichte praktijktoets en evaluatie van ongewenste effecten + Transparante documentatie en patiëntenparticipatie	Aandachtspunten – Hoge tijdsinvestering – Nood aan methodologische expertise – Professionele oordeelsvorming blijft noodzakelijk – Contextgebondenheid van QI
---	---

Conclusie

Kwaliteitsindicatoren (QI) die volledig ontwikkeld worden via CEQUID voldoen in principe aan de AIRE-criteria. Het uiteindelijke doel is om QI te bekomen waarin gebruikers vertrouwen kunnen hebben, zowel in de QI zelf als in het ontwikkelproces dat eraan voorafging.

Lijst van afkortingen

AGREE II	Appraisal of Guidelines for Research & Evaluation II
AIRE	Appraisal of Indicators through Research and Evaluation
CEQUID	Cebam Quality Indicator Development
COI	Conflict of Interest (belangenconflict)
GPS	Good Practice Statements
GRADE	Grading of Recommendations Assessment, Development and Evaluation
IOM	Institute of Medicine
KCE	Federaal Kenniscentrum voor de Gezondheidszorg
PAQS	Plateforme pour l'Amélioration continue de la Qualité des soins et de la Sécurité des patients
QI	Kwaliteitsindicator(en)
VIKZ	Vlaams Instituut voor Kwaliteit van Zorg

1. Doel en opzet van dit document

Dit document beschrijft de methodologie die ten grondslag ligt aan de CEQUID (Cebam Quality Indicator Development). CEQUID is een gestructureerde methode die binnen Cebam, het Belgisch centrum voor Evidence-Based Medicine, gevolgd wordt voor de ontwikkeling, validatie en documentatie van kwaliteitsindicatoren (QI) binnen de eerstelijnsgezondheidszorg.

De methodologie heeft als doel de ontwikkelgroep te ondersteunen bij de ontwikkeling van QI die:

- wetenschappelijk onderbouwd zijn;
- expliciet gekoppeld zijn aan richtlijnen en aanbevelingen;
- transparant en reproduceerbaar ontwikkeld worden;
- relevant zijn voor kwaliteitsverbetering;
- haalbaar zijn in de praktijk;
- kritisch geëvalueerd worden op mogelijke ongewenste effecten.

Daarnaast wil CEQUID bijdragen aan een meer uniforme en transparante aanpak van QI-ontwikkeling binnen de eerstelijnszorg.

a) Doelgroep

CEQUID is een methodologische leidraad voor alle betrokkenen in de eerstelijnsgezondheidszorg die QI willen ontwikkelen of valideren:

- Wetenschappelijke beroepsverenigingen en onderzoeksgroepen die QI voor eigen kwaliteitsprojecten ontwikkelen.
- Beleidsinstanties die kwaliteitsmeting willen inzetten voor beleidsdoeleinden.
- Onderzoeksgroepen die de methodologische kwaliteit van bestaande QI willen beoordelen via het AIRE-instrument.
- Cebam zelf als instrument voor de productie van nieuwe QI-sets.

De methodologie kan worden toegepast binnen zowel lokale als regionale of nationale kwaliteitsprojecten.

b) Uitgangspunten van CEQUID

CEQUID vertrekt vanuit vijf fundamentele uitgangspunten:

- 1) Evidentie: QI worden zoveel mogelijk gebaseerd op wetenschappelijke evidentie en aanbevelingen uit kwaliteitsvolle klinische praktijkrichtlijnen.
- 2) Transparantie: alle methodologische keuzes worden expliciet gedocumenteerd,

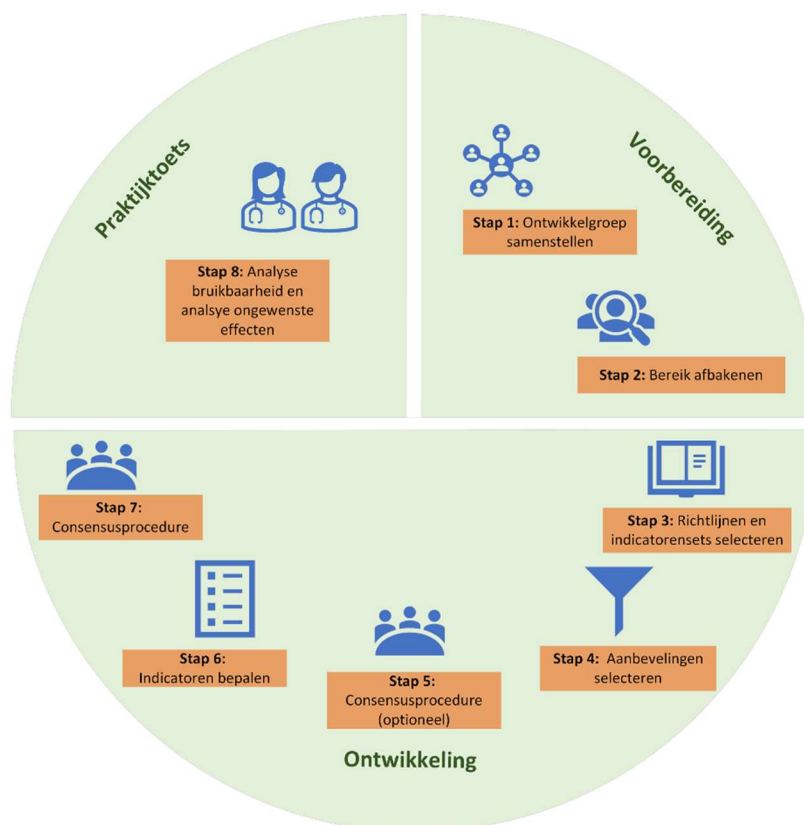
zodat het ontwikkelproces reproduceerbaar en controleerbaar blijft.

- 3) Stakeholderbetrokkenheid: relevante stakeholders, inclusief patiënten en naasten, worden waar mogelijk betrokken bij het ontwikkelproces.
- 4) Praktische bruikbaarheid: QI moeten niet alleen inhoudelijk valide zijn, maar ook bruikbaar blijken in de praktijk.
- 5) Kritische reflectie: naast mogelijke voordelen, wordt ook expliciet aandacht besteed aan beperkingen en mogelijke ongewenste effecten van QI.

c) Opbouw van het document

Eerst beschrijven we het conceptueel kader waarin CEQUID werd ontwikkeld. Vervolgens beschrijven we de CEQUID-methodologie, die bestaat uit drie opeenvolgende fasen en acht stappen (Figuur 1). Voor elke stap beschrijven we de wetenschappelijke onderbouwing, kritische reflecties, praktische uitvoering en verwachte output. De methodologie bevat drie fasen:

- fase 1: voorbereiding: samenstelling van de ontwikkelgroep en afbakening van onderwerp, doelgroep en doelstellingen;
- fase 2: ontwikkeling: selectie van richtlijnen en bestaande QI, selectie van aanbevelingen, formulering van QI en consensusvorming;
- fase 3: praktijktoets: analyse bruikbaarheid en ongewenste effecten.



Figuur 1 Overzicht van de acht stappen van CEQUID, verdeeld over drie fasen (vorbereiding, ontwikkeling en praktijktoets)

d) Reikwijdte van de methodologie

CEQUID beschrijft een methode voor de ontwikkeling en validatie van QI. Het document heeft niet als doel om implementatiestrategieën, kwaliteitsverbeteringsprogramma's of beleidskaders uit te werken. Hoewel de methodologie ontwikkeld werd met bijzondere aandacht voor de eerstelijnszorg, kunnen verschillende onderdelen ook toegepast worden in andere contexten.

e) Leeswijzer

Het document combineert wetenschappelijke onderbouwing met praktische richtlijnen voor toepassing. Elke stap binnen CEQUID bevat:

- een beschrijving van het doel;
- een overzicht van de relevante literatuur, kritische reflectie op bestaande methoden en een motivatie van de keuzes binnen CEQUID;
- praktische instructies voor uitvoering;
- een beschrijving van de verwachte output.

Op die manier beoogt CEQUID niet alleen een handleiding te zijn voor QI-ontwikkeling, maar ook een methodologisch referentiekader dat transparantie en kwaliteit binnen QI-ontwikkeling ondersteunt.

2. Inleiding

a) Waarom QI ontwikkelen?

QI worden internationaal beschouwd als een essentieel instrument voor het monitoren, evalueren en verbeteren van de kwaliteit van gezondheidszorg. Donabedian beschreef reeds dat kwaliteitsverbetering slechts mogelijk is wanneer relevante aspecten van zorg systematisch kunnen worden gemeten¹. Sindsdien hebben QI een centrale plaats verworven binnen kwaliteitsbeleid, audit- en feedbacksystemen, benchmarkinginitiatieven en gezondheidszorgonderzoek.

b) Internationale context en aanleiding ontwikkeling CEQUID

Verschillende auteurs tonen aan dat de kwaliteit van QI sterk varieert. Internationaal gebeurt de QI-ontwikkeling via uiteenlopende methodologische benaderingen^{2-6,8}. Grote verschillen worden gevonden in de selectie van richtlijnen, de vertaling van aanbevelingen naar QI en de gebruikte consensusprocedures⁴. De koppeling tussen QI en hun wetenschappelijke onderbouwing wordt niet altijd even transparant gerapporteerd⁵. Deze vaststellingen illustreren dat QI slechts zo betrouwbaar zijn als het ontwikkelproces waarop zij gebaseerd zijn.

CEQUID werd ontwikkeld als een methodologisch kader voor de ontwikkeling en validatie van QI binnen de Belgische eerstelijnszorg. De methodologie vertrekt vanuit de vaststelling dat QI-ontwikkeling vaak gefragmenteerd verloopt^{2,4-6}. Ontwikkelgroepen moeten verschillende instrumenten combineren zonder duidelijke richtlijnen over hoe die samen gebruikt kunnen worden.

Een breed aanvaard principe binnen QI-ontwikkeling is dat QI niet zonder meer kunnen worden overgenomen uit andere landen of zorgsystemen. Internationale ervaringen tonen aan dat succesvolle implementatie van QI sterk afhankelijk is van een formeel ontwikkelkader, betrokkenheid van relevante stakeholders en een duidelijke koppeling met nationale richtlijnen⁷. De toepasbaarheid van een QI wordt mede bepaald door contextuele factoren, zoals de organisatie en financiering van de gezondheidszorg, de rol van zorgverleners, de beschikbaarheid van gegevensbronnen en de kenmerken van de doelpopulatie. Internationale ervaringen tonen aan dat succesvolle implementatie van QI afhankelijk is van een duidelijke aansluiting bij nationale richtlijnen, betrokkenheid van relevante stakeholders en een formeel kader voor QI-ontwikkeling en -beheer.

Daarnaast blijkt uit de literatuur dat er verschillende modellen bestaan voor de relatie tussen richtlijnontwikkeling en QI-ontwikkeling. Hoewel een nauwe afstemming tussen beide processen wordt aanbevolen, bestaat er geen consensus over de optimale organisatievorm.

Sommige auteurs pleiten voor maximaal hergebruik van bestaande QI-sets om efficiëntie en internationale vergelijkbaarheid te bevorderen, terwijl anderen benadrukken dat contextuele verschillen de validiteit, haalbaarheid en acceptatie van QI aanzienlijk kunnen beïnvloeden. Die spanning tussen standaardisatie en contextualisering maakt een expliciete beoordeling van de lokale toepasbaarheid van QI noodzakelijk^{2,4-6}.

Langendam et al. concludeerden dat QI-ontwikkeling internationaal gekenmerkt wordt door aanzienlijke methodologische heterogeniteit en dat een volledig geïntegreerd conceptueel kader voor richtlijn- en QI-ontwikkeling grotendeels ontbreekt. CEQUID beoogt die kloof gedeeltelijk te overbruggen door een expliciete, transparante en reproduceerbare methodologie te beschrijven⁶.

De methodologie bouwt voort op bestaande internationale kaders:

- AGREE II van Brouwers et al. voor richtlijnbeoordeling⁸;
- AIRE van de Koning et al. voor beoordeling van indicatorontwikkeling⁹;
- RAND/UCLA-methode voor consensusvorming^{10,11}.

CEQUID combineert die bestaande bouwstenen in één geïntegreerd ontwikkelproces en vult ze aan met een bijkomende praktijktoets. De methodologie sluit bovendien aan bij eerdere Belgische initiatieven voor evidence-based indicatorontwikkeling en bouwt verder op bestaande ervaring van het Vlaams Instituut voor Kwaliteit van Zorg (VIKZ)¹². Binnen België bestaan er nog andere organisaties die QI ontwikkelen, zoals onder andere het Federaal Kenniscentrum voor de Gezondheidszorg (KCE)¹³ en het Plateforme pour l'Amélioration continue de la Qualité des soins et de la Sécurité des patients (PAQS). De meerwaarde van CEQUID ligt in de integratie van bestaande methodologische kaders binnen één coherent ontwikkelproces en een praktijktoets. Binnen CEQUID vertrekken we daarom van bestaande evidentie en beschikbare indicatoren, maar we gaan steeds na in welke mate die aansluiten bij de Belgische eerstelijnscontext. Daarbij besteden we bijzondere aandacht aan de koppeling met (inter)nationale richtlijnen, de beschikbare data-infrastructuur, de organisatie van zorg en de haalbaarheid van implementatie.

3. Conceptueel kader

a) Wat is een kwaliteitsindicator?

QI vormen de hoeksteen van kwaliteitsmeting in de gezondheidszorg. Hoewel er verschillende definities bestaan, delen die een aantal gemeenschappelijke kenmerken. Campbell et al. definiëren QI als meetbare elementen van de klinische praktijk waarvoor evidentie of consensus bestaat dat zij gebruikt kunnen worden om de kwaliteit van zorg te beoordelen². Braspenning et al. beschrijven QI als operationele vertalingen van kwaliteitsdoelstellingen die toelaten om de mate waarin gewenste zorg wordt geleverd systematisch te evalueren¹⁴.

Binnen CEQUID wordt een QI gedefinieerd als: 'Een meetbaar aspect van zorg waarvoor wetenschappelijke evidentie of formele consensus bestaat dat verbetering op dit aspect samenhangt met verbetering van de kwaliteit van zorg.' Die definitie benadrukt dat QI geen doel op zich zijn, maar instrumenten die kwaliteitsverbetering ondersteunen.

b) Het model van Donabedian

Het meest invloedrijke theoretische kader voor kwaliteitsmeting werd ontwikkeld door Donabedian¹. Hij onderscheidde drie complementaire perspectieven op kwaliteit van zorg:

- structuurindicatoren: organisatorische voorwaarden waarbinnen zorg wordt geleverd (bv. het percentage zorgverleners [bv. kinesitherapeuten] dat jaarlijks geaccrediteerde bijscholingen volgt volgens een vast opleidingsplan);
- procesindicatoren: meten wat zorgverleners daadwerkelijk doen (bv. het percentage diabetespatiënten met een griepvaccinatie in het voorbije jaar);
- uitkomstindicatoren: beschrijven de gevolgen van zorg voor patiënten (bv. het percentage diabetespatiënten met Hemoglobine A1c < 7% (53 mmol/mol)).

Deze driedeling vormt nog steeds de basis van vrijwel alle systemen voor kwaliteitsmeting en leveren elk complementaire informatie over de kwaliteit van zorg.

CEQUID kiest niet voor één exclusief type indicator, maar legt de focus eerder op procesindicatoren. Die keuze sluit aan bij Mant¹⁵, Braspenning et al.¹⁴ en Campbell et al.², die benadrukken dat procesindicatoren, doordat ze rechtstreeks gekoppeld zijn aan het handelen van zorgverleners, daardoor beter bruikbaar zijn voor kwaliteitsverbetering. Want ze zijn :

- beter beïnvloedbaar;
- gevoeliger voor kwaliteitsverschillen;
- eenvoudiger interpreteerbaar;
- doorgaans betrouwbaarder meetbaar.

Deze argumenten worden verder ondersteund door Abel et al., die in een grootschalige studie binnen de eerstelijnszorg aantoonde dat uitkomstindicatoren op praktijkniveau vaak onvoldoende betrouwbaar zijn voor vergelijking tussen praktijken¹⁶. Uitkomstindicatoren worden echter niet uitgesloten. Wanneer voldoende risicocorrectie mogelijk is en de relatie met kwaliteit voldoende plausibel is, kunnen zij een waardevolle aanvulling vormen.

c) Kwaliteit volgens The Institute of Medicine

Naast Donabedian¹ vormt het kwaliteitskader van het Institute of Medicine (IOM)¹⁷ een tweede belangrijk theoretisch fundament van CEQUID. In *Crossing the Quality Chasm* beschreef het IOM zes fundamentele dimensies van kwaliteitsvolle zorg:

1. Veiligheid (Safety): het voorkomen van schade bij zowel cliënten/patiënten als medewerkers.
2. Tijdigheid (Timely): het leveren van zorg binnen een passende termijn om schadelijke vertragingen te voorkomen.
3. Doeltreffendheid (Effective): het aanbieden van zorg of behandelingen waarvan wetenschappelijk is aangetoond dat ze patiënten helpen, en het vermijden van behandelingen die geen bewezen meerwaarde hebben.
4. Doelmatigheid (Efficient): het voorkomen van verspilling van middelen die geen waarde toevoegen voor de patiënt of het beoogde resultaat, zoals personeelsinzet, geld, materialen en tijd.
5. Gelijkwaardigheid (Equitable): het bieden van kwalitatief hoogwaardige zorg voor alle patiënten, ongeacht persoonlijke, sociale of geografische kenmerken.
6. Patiëntgerichtheid (Patient-Centered): het leveren van zorg die respectvol is en aansluit bij de waarden, wensen en behoeften van de patiënt.

Sommige dimensies, zoals veiligheid en doeltreffendheid, lenen zich goed voor QI-ontwikkeling. Patiëntgerichtheid, gelijkwaardigheid en continuïteit van zorg blijken vaak moeilijker te operationaliseren. Daarnaast kan een sterke focus op bepaalde kwaliteitsdomeinen ertoe leiden dat andere domeinen minder aandacht krijgen. Smith beschreef dat risico als tunnelvisie of meetfixatie, waarbij aandacht verschuift naar wat gemeten wordt ten koste van andere relevante aspecten van kwaliteit¹⁸.

Binnen CEQUID worden de IOM-dimensies niet gebruikt als een verplicht classificatiesysteem, maar als een reflectiekader. De ontwikkelgroep wordt aangemoedigd om expliciet na te denken over welke kwaliteitsdimensies relevant zijn voor het onderwerp en welke mogelijk onderbelicht blijven.

4. Kwaliteitscriteria voor QI

a) Waarom kwaliteitscriteria noodzakelijk zijn

De ontwikkeling van QI heeft slechts meerwaarde wanneer de uiteindelijke QI voldoende kwaliteit bezitten. Het bestaan van een richtlijn of aanbeveling impliceert niet automatisch dat daarvoor een bruikbare QI kan worden ontwikkeld. Campbell et al. benadrukken dat QI niet alleen klinisch relevant moeten zijn, maar ook valide, betrouwbaar, haalbaar en interpreteerbaar². Ook de Koning et al. beschrijven binnen het AIRE-instrument verschillende kwaliteitsdimensies die essentieel zijn voor de beoordeling van indicatoren⁹.

CEQUID beschouwt kwaliteitscriteria niet als een afzonderlijke validatiestap, maar als een rode draad doorheen het volledige ontwikkelproces. Van richtlijnselectie tot praktijk gaan we telkens na in welke mate QI aan die criteria voldoen. We beoordelen de QI idealiter op validiteit, betrouwbaarheid, haalbaarheid, discriminerend vermogen, interpreteerbaarheid en aanvaardbaarheid.

b) Validiteit

CEQUID erkent dat geen enkele afzonderlijke validatiestap volstaat om de validiteit van QI volledig aan te tonen. Daarom kiezen we voor een meerlagige benadering waarbij inhoudsvaliditeit, face validity, construct validity, haalbaarheid en betrouwbaarheid in opeenvolgende stappen worden geëvalueerd. Tijdens de QI-ontwikkeling wordt er ingezet op:

- expliciete koppeling aan richtlijnen: vergroot de kans dat de QI relevante aspecten van kwaliteitsvolle zorg meet (inhoudsvaliditeit);
- systematische selectie van aanbevelingen: versterkt de inhoudsvaliditeit door enkel representatieve aanbevelingen te selecteren;
- formele consensusprocedure: experts beoordelen of QI logisch, relevant en geschikt zijn;
 - o Daarvoor bekijken we of de QI een goede maat voor kwaliteit lijkt (face validity) en of de QI de relevante aspecten van het kwaliteitsdomein afdekt (inhoudsvaliditeit).
- praktijktoets: levert informatie over haalbaarheid, betrouwbaarheid, interpreteerbaarheid en discriminerend vermogen.
 - o Wanneer QI tijdens de praktijktoets daadwerkelijk verschillen in zorgkwaliteit detecteren zoals theoretisch verwacht, draagt deze stap ook bij aan de construct validity.
 - o Indien vergelijking mogelijk is met een externe gouden standaard, kan bovendien informatie worden verzameld over criterion validity, al is dat in de praktijk van QI vaak moeilijk realiseerbaar.

c) Betrouwbaarheid

Betrouwbaarheid verwijst naar de mate waarin een QI consistente resultaten oplevert wanneer dezelfde situatie opnieuw wordt beoordeeld. Mokkink et al. beschrijven betrouwbaarheid als een noodzakelijke voorwaarde voor betekenisvolle metingen¹⁹. Een QI die verschillende resultaten oplevert afhankelijk van beoordelaar, tijdstip of registratiepraktijk kan moeilijk gebruikt worden voor kwaliteitsverbetering.

CEQUID beschouwt betrouwbaarheid als een essentieel criterium voor implementatie. Daarom wordt tijdens een praktijktoets expliciet nagegaan of:

- de definities eenduidig zijn;
- er interpretatieverschillen optreden;
- de resultaten reproduceerbaar zijn.

d) Haalbaarheid

Haalbaarheid verwijst naar de praktische mogelijkheid om een indicator te meten. Braspenning et al. benadrukken dat indicatoren slechts bruikbaar zijn wanneer de noodzakelijke gegevens op een haalbare manier kunnen worden verzameld¹⁴.

Belangrijke aspecten zijn:

- beschikbaarheid van gegevens;
- registratielast;
- technische uitvoerbaarheid;
- benodigde middelen.

Er bestaat een terugkerende spanning tussen wetenschappelijke kwaliteit en praktische uitvoerbaarheid. Uitstekend onderbouwde QI kunnen onbruikbaar zijn wanneer de benodigde gegevens niet beschikbaar zijn, terwijl eenvoudig meetbare QI niet noodzakelijk de meest relevante zijn. Binnen CEQUID is haalbaarheid een essentieel kwaliteitscriterium. Een QI die niet haalbaar meetbaar is, kan geen duurzame bijdrage leveren aan kwaliteitsverbetering. Daarom vormt de praktijktoets een verplicht onderdeel van de methodologie.

e) Discriminerend vermogen

Een QI moet relevante verschillen in kwaliteit van zorg tussen zorgverleners, organisaties of tijdsperiodes zichtbaar kunnen maken. Sommige QI vertonen een plafondeffect waarbij vrijwel alle zorgverleners maximale scores behalen. Dergelijke QI kunnen nog steeds nuttig zijn voor monitoring, maar hebben vaak beperkte meerwaarde voor kwaliteitsverbetering. De literatuur biedt geen duidelijke grenswaarden voor wat als voldoende discriminerend vermogen moet worden beschouwd.

CEQUID beschouwt discriminerend vermogen als een belangrijke eigenschap voor QI die gebruikt worden voor benchmarking of kwaliteitsverbetering. Tijdens de praktijktoets wordt expliciet nagegaan of QI voldoende variatie vertonen.

f) Interpreteerbaarheid

Interpreteerbaarheid verwijst naar de mate waarin gebruikers begrijpen:

- wat een QI meet;
- hoe resultaten moeten worden geïnterpreteerd;
- welke beperkingen bestaan.

Grol en Wensing benadrukken dat kwaliteitsinformatie slechts bruikbaar is wanneer gebruikers de betekenis ervan correct begrijpen²⁰. Interpreteerbaarheid wordt vaak onderschat binnen QI-ontwikkeling. QI kunnen technisch correct zijn, maar verkeerd geïnterpreteerd worden door gebruikers, beleidsmakers of het publiek. Dat risico neemt toe wanneer QI gebruikt worden buiten de context waarvoor zij oorspronkelijk ontwikkeld werden.

CEQUID vereist daarom gestandaardiseerde indicatorfiches waarin expliciet beschreven wordt:

- wat de QI meet;
- hoe resultaten geïnterpreteerd moeten worden;
- welke beperkingen gelden.

g) Aanvaardbaarheid

QI hebben enkel impact wanneer gebruikers bereid zijn ermee te werken². Aanvaardbaarheid omvat onder meer:

- klinische relevantie;
- organisatorische relevantie;
- patiëntrelevantie.

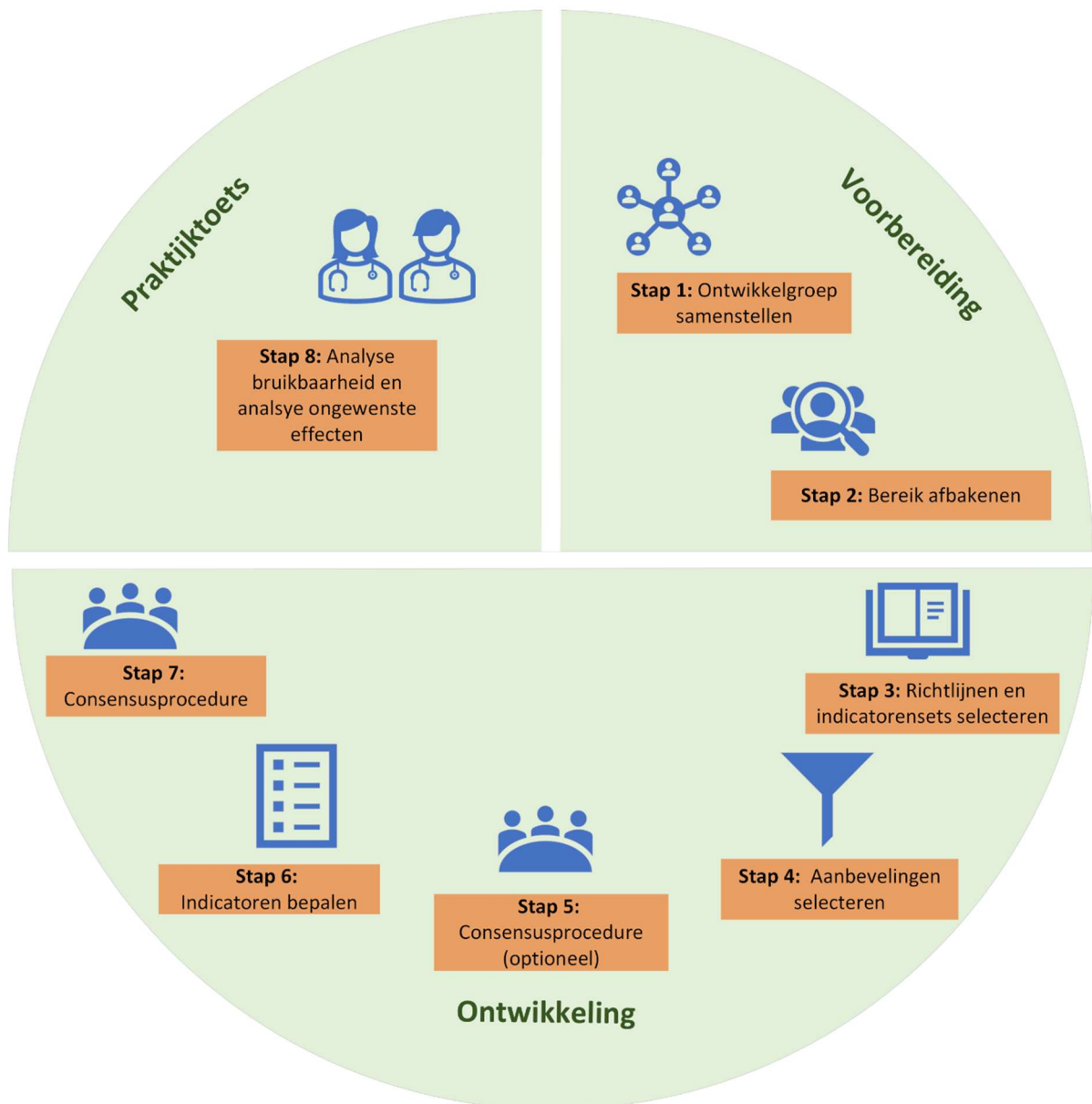
Aanvaardbaarheid wordt vaak beïnvloed door factoren die buiten de indicator zelf liggen, zoals:

- beschikbare middelen;
- registratielast;
- beleidscontext;
- ervaring met kwaliteitsmeting.

Binnen CEQUID worden stakeholders en patiënten vroeg in het ontwikkelproces betrokken om de aanvaardbaarheid van QI te verhogen.

5. CEQUID: 8 stappen

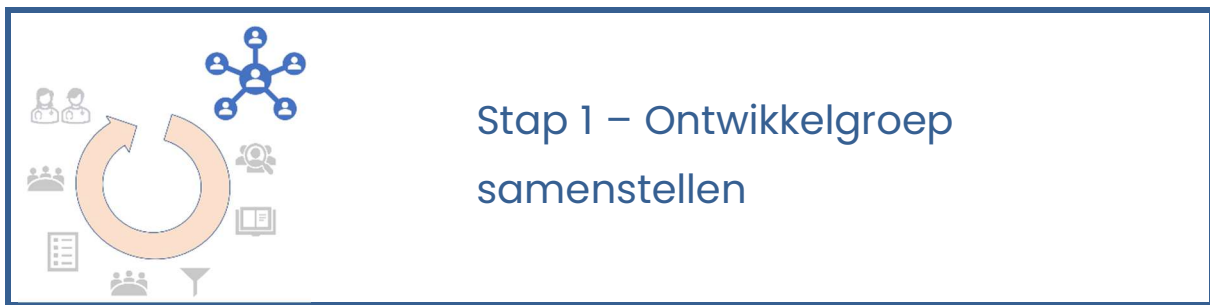
CEQUID omvat acht stappen, verdeeld over drie fasen (Figuur 1). Het uiteindelijke doel van CEQUID is om QI te ontwikkelen waarin gebruikers vertrouwen mogen hebben, zowel in de QI zelf als in het ontwikkelproces dat eraan voorafging.



Figuur 2 Overzicht van de acht stappen van CEQUID, verdeeld over drie fasen (voorbereiding, ontwikkeling en praktijktoets)

Fase 1 – Voorbereidende fase

De voorbereidende fase legt de fundamenten voor het volledige QI-ontwikkelp proces. Methodologische keuzes die vroeg in het ontwikkelingsproces worden gemaakt, hebben een belangrijke invloed op de kwaliteit, validiteit en bruikbaarheid van de uiteindelijke QI-set^{2,4,6}. Campbell et al. beschrijven QI-ontwikkeling als een proces waarin de samenstelling van de ontwikkelgroep, de afbakening van het onderwerp en de betrokkenheid van stakeholders bepalend zijn voor de validiteit en aanvaardbaarheid van de uiteindelijke QI². Binnen CEQUID wordt deze fase gezien als een essentieel kwaliteitselement.



Doel

Het samenstellen van een multidisciplinaire ontwikkelgroep die beschikt over de noodzakelijke methodologische en klinische expertise voor de ontwikkeling van QI, aangevuld met patiëntenvertegenwoordiging om de relevantie en het maatschappelijke draagvlak ervan te waarborgen.

Aanpak

De CEQUID vertrekt vanuit de veronderstelling dat het onderwerp waarvoor QI ontwikkeld worden en de projectleider gekend zijn. De projectleider coördineert het hele proces. De projectleider stelt een ontwikkelgroep samen die bestaat uit methodologen en adviserende stakeholders. Die geven input doorheen alle fasen. De ontwikkelgroep voert de processtappen uit, bewaakt de methodologische kwaliteit en legt verantwoording af over elke beslissing.

Waarom is deze stap noodzakelijk?

QI-ontwikkeling vereist verschillende vormen van expertise. Enerzijds is inhoudelijke expertise noodzakelijk om de klinische relevantie van aanbevelingen en QI te beoordelen, en ook om belangrijke contextelementen rond het onderwerp aan te dragen. Anderzijds is methodologische expertise nodig om richtlijnen kritisch te evalueren, consensusprocedures te begeleiden en validatiestappen uit te voeren.

Brouwers et al. benadrukken binnen AGREE II dat stakeholderbetrokkenheid een essentieel kwaliteitscriterium vormt voor richtlijnontwikkeling. Een vergelijkbaar principe geldt voor QI-ontwikkeling⁸. Dudley et al. beschrijven dat QI meer kans hebben op succesvolle implementatie wanneer relevante stakeholders al tijdens de ontwikkelingsfase worden betrokken³. Kötter et al. wijzen er bovendien op dat de samenstelling van de ontwikkelgroep een directe invloed kan hebben op de uiteindelijke QI-selectie⁴.

Wetenschappelijke onderbouwing, kritische reflectie en keuze binnen

CEQUID

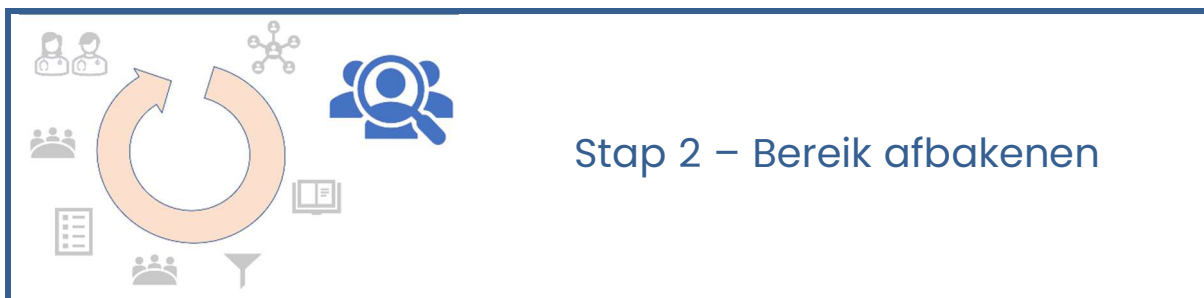
Thema	🔗 Wetenschappelijke onderbouwing	⚠️ Kritische reflectie	✓ Keuze binnen CEQUID
Multidisciplinaire expertise	QI-ontwikkeling vereist zowel wetenschappelijke als praktijkgerichte expertise ^{2,4,8,21} .	Hoewel multidisciplinaire ontwikkelgroepen breed worden aanbevolen, bestaat weinig empirisch onderzoek naar de optimale samenstelling ervan. De literatuur biedt geen duidelijke antwoorden over ideale groepsgrootte of minimale disciplines.	De ontwikkelgroep is multidisciplinair samengesteld met minstens 1 methodoloog, 1 of meerdere klinische experts met domeinexpertise en toekomstige gebruikers. Stakeholders en patiëntenvertegenwoordigers worden betrokken. De samenstelling wordt expliciet gemotiveerd en gerapporteerd.
Stakeholders-betrokkenheid	Betrokkenheid van relevante stakeholders is een essentieel kwaliteitscriterium. Zij dragen bij aan identificatie van prioriteiten, beoordeling van relevantie en haalbaarheid ⁸ .	Er bestaat een spanningsveld tussen representativiteit en efficiëntie. Grotere groepen brengen meer perspectieven samen, maar kunnen besluitvorming vertragen. Kleinere groepen werken efficiënter, maar lopen een risico op eenzijdige besluitvorming ^{4,10} .	De stakeholderanalyse wordt afgestemd op het onderwerp en de context, en keuzes voor stakeholdergroepen worden expliciet verantwoord.
Patiënten-participatie	Patiëntenparticipatie komt binnen indicatorontwikkeling nog relatief beperkt voor ⁴ . Patiënten beschrijven andere kwaliteitsprioriteiten dan zorgverleners en onderzoekers, wat kan bijdragen aan meer patiëntgerichte QI ²² . Patiënten worden nog relatief beperkt betrokken bij QI-ontwikkeling ²³ .	Hoewel de meerwaarde van patiëntenparticipatie breed wordt erkend, bestaat weinig consensus over de optimale vorm van betrokkenheid. Bergholtz et al. waarschuwen voor tokenisme, waarbij patiënten formeel betrokken worden zonder wezenlijke invloed op de besluitvorming ²² .	Patiëntenparticipatie wordt beschouwd als een potentiële bron van aanvullende expertise. Wanneer patiënten niet worden opgenomen in de ontwikkelgroep, wordt deze keuze expliciet verantwoord.
Belangenconflicten	Transparantie rond belangenconflicten is essentieel voor de geloofwaardigheid van ontwikkelprocessen ⁸ .	De aanwezigheid van specifieke stakeholders kan onbedoeld leiden tot dominantie van bepaalde perspectieven of belangen. Volledige neutraliteit is daarom niet haalbaar. Transparante rapportering van belangenconflicten vermindert dit risico, maar kan het niet volledig uitsluiten ^{8,22} .	Potentiële belangenconflicten worden systematisch geregistreerd en gerapporteerd. Alle leden van de ontwikkelgroep ondertekenen een verklaring van belangen (COI-verklaring), conform de aanbevelingen van AGREE II.

Output

Deze stap resulteert in:

- een multidisciplinair samengestelde ontwikkelgroep;
- een overzicht van expertisegebieden;
- een stakeholderanalyse;
- een overzicht van belangenverklaringen;
- een motivatie van de gekozen samenstelling.

In [Stap 2](#) – Afbakening van bereik, doelgroep en doelstellingen wordt onder meer de keuze voor kwaliteitsdomeinen, doelgroep afbakening en het spanningsveld tussen brede versus smalle QI-sets methodologisch verder uitgewerkt.



Doel

Het expliciet definiëren van het onderwerp, de doelgroep, de relevante kwaliteitsdomeinen vastleggen en het beoogde gebruik van de QI.

Aanpak

De ontwikkelgroep bakent in deze stap vier elementen expliciet af, die samen het kader vormen voor de volledige ontwikkelfase.

- 1) *Beoogde gebruikers*: voor wie zijn de QI bedoeld (bv huisartsen, verpleegkundigen, apothekers, kinesitherapeuten, beleidsmakers)?
- 2) *Bereik van het onderwerp*: welke patiëntenpopulatie, zorgfase en zorgaspecten worden gedekt?
- 3) *Doel van de QI*: interne kwaliteitsverbetering, externe rapportering, onderzoek? Door wie gaan de indicatoren gebruikt worden?
- 4) *Betrokken kwaliteitsdomeinen*: selectie uit de zes IOM domeinen van kwaliteit (veiligheid, doeltreffendheid, patiëntgerichtheid, tijdigheid, doelmatigheid en gelijkwaardigheid)¹⁷

Waarom deze stap noodzakelijk is

De ontwikkeling van kwaliteitsindicatoren begint met het bepalen van wat precies gemeten moet worden en waarom.

Campbell et al. benadrukken dat QI steeds context gebonden zijn. Een QI die relevant is voor kwaliteitsverbetering binnen één context is niet noodzakelijk bruikbaar binnen een andere context². Brouwers et al. beschouwen een duidelijke afbakening van scope en doel als een fundamenteel kwaliteitscriterium binnen richtlijnontwikkeling. Een vergelijkbare redenering geldt voor QI-ontwikkeling⁸. QI die expliciet gekoppeld zijn aan een of meerdere kwaliteitsdomeinen kennen een hogere face validity en acceptatie bij zorgverleners^{4,17}.

Wanneer doelstellingen onvoldoende expliciet worden geformuleerd, ontstaat het risico dat:

- QI verschillende doelen tegelijk proberen te dienen;
- relevante kwaliteitsdomeinen over het hoofd worden gezien;
- gebruikers resultaten verkeerd interpreteren;
- implementatieproblemen ontstaan.

Een duidelijke afbakening vormt daarom de basis voor alle verdere beslissingen binnen CEQUID.

Wetenschappelijke onderbouwing, kritische reflectie en keuze binnen CEQUID

Thema	🔗 Wetenschappelijke onderbouwing	⚠️ Kritische reflectie	✓ Keuze binnen CEQUID
Afbakening van de doelgroep	QI worden steeds ontwikkeld voor een specifieke populatie. Een duidelijke definitie van de doelgroep verhoogt de validiteit en interpreteerbaarheid van QI ² .	Een brede doelgroep verhoogt de toepasbaarheid, maar kan specificiteit verminderen. Een te nauwe doelgroep verhoogt de relevantie, maar beperkt generaliseerbaarheid ^{2,24} .	CEQUID vereist een expliciete beschrijving, inclusief patiëntengroep, leeftijdscategorie, zorgsetting, tijdsframe en betrokken zorgverleners. Alle keuzes worden gemotiveerd en gedocumenteerd.
Afbakening van de context	QI die ontwikkeld worden binnen één zorgsysteem, zijn niet noodzakelijk overdraagbaar naar andere settings. QI zijn contextgebonden. Factoren zoals organisatie van zorg, financieringsstructuur en gegevensbeschikbaarheid bepalen of een QI bruikbaar is binnen een bepaalde setting ²⁴ .	Een te brede contextdefinitie kan contextverschillen maskeren, een te nauwe afbakening beperkt bruikbaarheid.	CEQUID vereist dat expliciet wordt beschreven voor welke context QI ontwikkeld worden en welke contexten buiten de scope vallen. Daardoor worden de grenzen van de toepasbaarheid van de QI transparant gemaakt.
Definiëren van het doel van QI	Er zijn diverse toepassingen: kwaliteitsverbetering, audit en feedback, benchmarking, beleidsmonitoring en onderzoek ^{2,25,26} .	QI voor interne kwaliteitsverbetering zijn niet noodzakelijk geschikt voor benchmarking of financieringsbeslissingen, wat kan leiden tot ongewenste effecten zoals gaming ¹⁸ .	CEQUID vereist dat het primaire doel van de QI-set vooraf wordt vastgelegd. Daardoor kan de ontwikkelgroep de QI afstemmen op het beoogde gebruik en kunnen beoogde gebruikers de resultaten correct interpreteren.
Scope van het project	De literatuur beschrijft geen empirisch onderbouwde optimale afbakening van de scope.	Een brede scope laat toe om meerdere kwaliteitsaspecten tegelijk te evalueren, maar verhoogt de complexiteit, de benodigde middelen en het risico op een heterogene QI-set. Een te smalle scope kan daarentegen belangrijke kwaliteitsaspecten uitsluiten ^{2,4,6} .	CEQUID vereist dat de scope expliciet wordt beschreven en verantwoord.
Kwaliteitsdomeinen	Het Institute of Medicine beschrijft zes fundamentele dimensies van kwaliteitsvolle zorg: veiligheid, doeltreffendheid, patiëntgerichtheid, tijdigheid, doelmatigheid en gelijkwaardigheid ¹⁷ .	Sommige kwaliteitsdomeinen zijn gemakkelijker meetbaar dan andere. Daardoor ontstaat het risico dat ontwikkelgroepen focussen op wat gemakkelijk meetbaar is in plaats van op wat maatschappelijk belangrijk is.	CEQUID gebruikt de IOM-domeinen als reflectiekader, zonder te vereisen dat alle domeinen vertegenwoordigd zijn. Gemotiveerde keuzes worden gedocumenteerd.

Output

Deze stap resulteert in:

- een duidelijk omschreven onderwerp;
- een afgebakende doelgroep;
- een beschrijving van de context;
- een expliciete doelstelling voor de QI-set;
- een overzicht van relevante kwaliteitsdomeinen;
- een motivatie van de gemaakte keuzes.

Deze output vormt de basis voor de selectie van richtlijnen en bestaande QI in de volgende fase.

Fase 2 – Ontwikkelingsfase

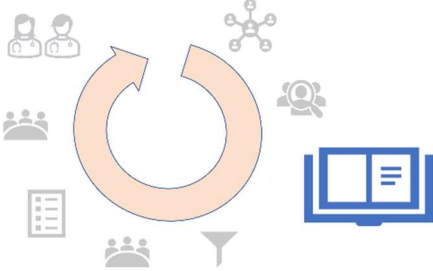
Na de voorbereidende fase (fase 1) beschikt de ontwikkelgroep over:

- een duidelijk afgebakend onderwerp;
- een gedefinieerde doelgroep;
- een omschreven context;
- een multidisciplinair samengestelde ontwikkelgroep.

De ontwikkelingsfase vormt de wetenschappelijke kern van CEQUID. In deze fase worden aanbevelingen uit klinische praktijkrichtlijnen systematisch vertaald naar mogelijke QI.

Campbell et al. beschrijven deze fase als de cruciale schakel tussen wetenschappelijke evidentie en kwaliteitsmeting². Tijdens deze fase worden de grootste methodologische verschillen tussen QI-ontwikkelingsmethoden zichtbaar⁴. Binnen CEQUID wordt daarom bijzondere aandacht besteed aan:

- expliciete koppeling aan wetenschappelijke evidentie;
- transparantie;
- reproduceerbaarheid;
- expliciete besluitvorming;
- documentatie van alle keuzes.



Stap 3 – Richtlijnen en bestaande QI selecteren

Doel

Het identificeren van de best beschikbare wetenschappelijke evidentie en bestaande QI als basis voor de ontwikkeling van een nieuwe QI-set.

Aanpak

De ontwikkelgroep voert twee parallelle zoekacties uit:

- 1) Een systematische search naar relevante klinische praktijkrichtlijnen, beoordeeld via het AGREE II-instrument. Daarvoor geldt:
 - Als de richtlijnen door Cebam werden gevalideerd, voert de ontwikkelgroep een relevantietoets uit.
 - Als de richtlijnen niet door Cebam werden gevalideerd, voert de ontwikkelgroep een AGREE II-beoordeling uit, of minimaal een verkorte AGREE-screening (mini-AGREE)⁸.
- 2) Een gelijkaardige search naar bestaande QI-sets voor het betrokken onderwerp, beoordeeld via het AIRE-instrument⁹.
 - Als de QI gebaseerd zijn op een organisatie met Cebam-accreditatie, evalueert de ontwikkelgroep alle AIRE-items met uitzondering van items 9, 10 en 11.
 - Ook als de QI gebaseerd zijn op een richtlijn waarvoor een AGREE II-beoordeling werd toegepast, evalueert de ontwikkelgroep alle AIRE-items met uitzondering van items 9, 10 en 11.
 - Als de QI gebaseerd zijn op een richtlijn of andere bron zonder AGREE II-beoordeling, doorloopt de ontwikkelgroep het volledige AIRE-instrument.

Waarom is deze stap noodzakelijk?

De kwaliteit van een QI wordt in belangrijke mate bepaald door de kwaliteit van de bronnen waarop die gebaseerd is. Als de onderliggende evidentiebasis onvoldoende duidelijk wordt beschreven, wordt het moeilijk om de wetenschappelijke onderbouwing van QI te beoordelen. Bovendien bestaan er aanzienlijke verschillen tussen QI-ontwikkelingsmethoden wat betreft de selectie en beoordeling van richtlijnen, waardoor QI rond hetzelfde onderwerp kunnen steunen op uiteenlopende niveaus van evidentie. De beperkte transparantie waarmee richtlijnselectie vaak wordt gerapporteerd, wordt daarbij beschouwd als een van de belangrijkste methodologische tekortkomingen binnen bestaande QI-ontwikkelingsprojecten^{2,4,24}. Een systematische review van Langendam et al. bevestigde bovendien dat er internationaal een grote methodologische heterogeniteit bestaat in de

ontwikkeling van richtlijn gebaseerde QI. De auteurs identificeerden aanzienlijke verschillen in de selectie van aanbevelingen, het gebruik van consensusmethoden, de operationalisering van QI en de mate waarin validatiestappen worden beschreven. Die bevinding onderstreept de noodzaak van een transparante en reproduceerbare ontwikkelmethodologie⁶.

Binnen CEQUID vormt een systematische selectie en kwaliteitsbeoordeling van richtlijnen en bestaande indicatoren daarom een essentiële eerste stap van de ontwikkelingsfase.

Wetenschappelijke onderbouwing, kritische reflectie en keuze binnen CEQUID

Thema	🔗 Wetenschappelijke onderbouwing	⚠️ Kritische reflectie	✓ Keuze binnen CEQUID
Klinische praktijkrichtlijnen als vertrekpunt	Richtlijnen bieden een synthese van best beschikbare wetenschappelijke evidentie. Internationale QI-ontwikkelingsmethoden gebruiken richtlijnen bovendien het vaakst als primaire bron ^{2,4} .	QI-ontwikkeling kan plaatsvinden in domeinen met beperkte evidentie. Strikte afhankelijkheid van richtlijnen kan daardoor leiden tot ondervetegenwoordiging van bepaalde kwaliteitsaspecten ⁴ .	CEQUID kiest bewust voor een richtlijn-gebaseerde benadering. Kennishiaten worden expliciet gedocumenteerd zodat zichtbaar blijft waar QI-ontwikkeling steunt op beperkte evidentie.
Beoordeling van richtlijnkwiteit	AGREE II beoordeelt richtlijnen systematisch op scope en doel, stakeholderbetrokkenheid, methodologische nauwkeurigheid, helderheid van presentatie en redactionele onafhankelijkheid ⁸ .	AGREE II werkt bewust zonder vaste afkapwaarden, waardoor professionele oordeelsvorming noodzakelijk blijft ⁸ .	CEQUID vereist een systematische kwaliteitsbeoordeling met vastgelegde selectiecriteria, gedocumenteerde beoordelingen en expliciete motivatie van uitsluiting.
Gebruik van bestaande QI	AIRE ondersteunt de systematische beoordeling van de methodologische kwaliteit. Hergebruik van bestaande QI vermijdt dubbel werk en bevordert internationale vergelijkbaarheid ^{9,8} .	QI zijn contextspecifiek en niet noodzakelijk overdraagbaar. Methodologische tekortkomingen kunnen worden overgenomen bij onkritisch hergebruik ⁵ .	Bestaande QI worden systematisch beoordeeld op methodologische kwaliteit, relevantie, actualiteit en toepasbaarheid voor zij verder worden gebruikt.
Transparantie van selectieprocedures	Transparante beschrijving van de selectie van richtlijnen en QI verhoogt de reproduceerbaarheid en geloofwaardigheid ^{4,5} .	Zelfs met gestandaardiseerde instrumenten kunnen interpretaties van kwaliteitscriteria verschillen. Volledige objectiviteit is niet haalbaar.	CEQUID vereist volledige documentatie van zoekstrategieën, selectiecriteria, kwaliteitsbeoordelingen en redenen voor inclusie of exclusie.

Output

Deze stap resulteert in:

- een overzicht van geselecteerde richtlijnen met AGREE II-beoordelingen;
- een overzicht van bestaande QI met AIRE-beoordelingen;
- volledige zoekstrategieën;
- motivatie van inclusie- en exclusiebeslissingen.



Stap 4 – Aanbevelingen selecteren

Deze output vormt de basis voor de selectie van aanbevelingen in [stap 4](#).

Doel

Het identificeren van aanbevelingen die voldoende wetenschappelijk onderbouwd, klinisch relevant en meetbaar zijn om als basis te dienen voor kwaliteitsindicatoren.

Aanpak

De ontwikkelgroep maakt een gestructureerd overzicht van de aanbevelingen uit de geselecteerde richtlijnen, gegroepeerd per zorgaspect. Vervolgens beoordeelt zij welke aanbevelingen bruikbaar zijn voor de formulering van een QI, op basis van drie criteria:

1. Soort aanbeveling: enkel sterke aanbevelingen worden meegenomen:
 - a. Aanbevelingen met een GRADE 1
 - b. Aanbevelingen op basis van goede praktijk (Good Practice Statements)
 - c. Aanbeveling op basis van consensus, als ze geformuleerd zijn als sterke aanbeveling en mits kritische reflectie bij de onderbouwende rationale
2. Relevantie ten aanzien van het bereik en de kwaliteitsdomeinen uit [stap 2](#).
3. Specificiteit en meetbaarheid: de aanbeveling moet voldoende specifiek en meetbaar kunnen worden omgezet in een QI met een teller en een noemer. Aanbevelingen die te breed (op basis van [stap 2](#)), te algemeen of inherent kwalitatief van aard zijn, worden niet geselecteerd.

De ontwikkelgroep labelt de bewijskracht van de aanbeveling consistent conform het graderingssysteem dat in de geselecteerde richtlijnen wordt gehanteerd (bv GRADE²⁷). Als de ontwikkelgroep meer dan 10 aanbevelingen selecteert, past ze een RAND/UCLA-methode^{10,11} toe voor prioritering ([Stap 5](#)).

Waarom is deze stap noodzakelijk?

Niet elke aanbeveling uit een klinische praktijkrichtlijn is geschikt om te worden omgezet naar een QI. Sommige aanbevelingen zijn:

- onvoldoende specifiek;
- onvoldoende meetbaar;
- zwakke aanbevelingen, die voor hun toepassing te sterk afhankelijk zijn van de context of van de casus van de patiënt, of waarbij een alternatieve aanpak even goed te verantwoorden is;
- weinig relevant voor kwaliteitsverbetering;
- moeilijk operationaliseerbaar.

Die aanbevelingen worden niet geselecteerd.

Campbell et al. benadrukken dat indicatorontwikkeling een expliciete vertaalslag vereist van aanbevelingen naar meetbare kwaliteitsaspecten². Langendam et al. beklemtonen dat er aandacht nodig is voor het onderscheid tussen de zekerheid van het onderliggende bewijs en de sterkte van een aanbeveling. Ontwikkelgroepen hebben begeleiding nodig wanneer ze deze concepten toepassen in de ontwikkeling van QI⁶. Kötter et al. beschrijven deze stap als een van de meest bepalende fasen binnen het volledige ontwikkelproces. QI gebaseerd op aanbevelingen van hogere evidentieniveaus behalen consistent hogere validiteitsscores bij consensuspanels. Indicatoren gebaseerd op consensus-only aanbevelingen kregen significant lagere scores op validiteit en klinische relevantie⁴. In veel QI-ontwikkelingsmethodes wordt een consensusprocedure gebruikt bij:

- selectie van aanbevelingen;
- prioritering van aanbevelingen;
- selectie van kandidaat-indicatoren⁴.

De kwaliteit van de uiteindelijke QI-sets wordt daarom in belangrijke mate bepaald door de kwaliteit van de geselecteerde aanbevelingen.

Wetenschappelijke onderbouwing, kritische reflectie en keuze binnen CEQUID

Thema	🔗 Wetenschappelijke onderbouwing	⚠️ Kritische reflectie	✓ Keuze binnen CEQUID
Aanbevelingen als basis voor QI	QI zijn operationele vertalingen van aanbevelingen. Wanneer aanbevelingen wetenschappelijk onderbouwd zijn, kunnen zij een sterke basis vormen voor kwaliteitsmeting ^{7,8} .	Niet alle aanbevelingen zijn even geschikt. Aanbevelingen kunnen te algemeen, onvoldoende specifiek of onvoldoende onderbouwd zijn om te worden omgezet naar valide en meetbare QI. Professionele oordeelsvorming blijft noodzakelijk ^{2,4} .	CEQUID documenteert systematisch de koppeling tussen aanbeveling en indicator, zodat de wetenschappelijke onderbouwing transparant blijft.
Graad van evidentie en aanbevelingssterkte	Volgens de GRADE Working Group moet een onderscheid worden gemaakt tussen de kwaliteit van bewijs en de sterkte van aanbevelingen. Beide elementen vallen niet noodzakelijk samen ²⁷ .	Gebruik van aanbevelingssterkte als enig selectiecriteria kan leiden tot uitsluiting van relevante aanbevelingen. Omgekeerd kan opname van zwak onderbouwde aanbevelingen de robuustheid verminderen.	CEQUID beoordeelt zowel de kwaliteit van bewijs als de sterkte van aanbevelingen. Beide worden expliciet gedocumenteerd.
Klinische relevantie	QI focussen idealiter op aspecten van zorg met een aantoonbare impact op patiëntuitkomsten, waarvoor kwaliteitsverschillen mogelijk zijn ² .	Klinische relevantie blijft gedeeltelijk afhankelijk van context en perspectief, wat tot verschillen in beoordeling kan leiden.	CEQUID beoordeelt aanbevelingen op klinische impact, verbeterpotentieel en relevantie voor patiënten, zorgverleners en andere stakeholders.
Meetbaarheid	Voor QI-ontwikkeling moet een aanbeveling kunnen worden vertaald naar een teller, een	Aanbevelingen rond communicatie, gedeelde besluitvorming en persoonsgerichte zorg zijn klinisch	Meetbaarheid is een belangrijk, maar niet exclusief criterium.

Thema	🔗 Wetenschappelijke onderbouwing	⚠️ Kritische reflectie	✓ Keuze binnen CEQID
	noemer, operationele definities en concrete meetcriteria ⁴ .	relevant, maar moeilijk te operationaliseren. Daardoor bestaat risico op meetfixatie ⁵ .	Uitgesloten aanbevelingen worden systematisch gedocumenteerd, zodat onmeetbare kwaliteitsaspecten zichtbaar blijven.
Evidentie versus consensus	In sommige zorgdomeinen is hoogwaardige evidentie beperkt beschikbaar. Consensusaanbevelingen kunnen een rol spelen binnen richtlijnen en kwaliteitsverbetering ^{2,28} . Blozik et al. pleiten voor strikte minimumdrempels voor lage kwaliteitsrichtlijnen, omdat zwak onderbouwde QI het vertrouwen in de volledige QI-set kunnen ondermijnen ²⁸ .	Het omzetten van consensusaanbevelingen naar QI verhoogt het risico dat onzekerheden worden vertaald naar formele kwaliteitsmetingen. Er ontstaat een fundamenteel spanningsveld tussen wetenschappelijke robuustheid en volledigheid. QI afgeleid van lagere kwaliteitsrichtlijnen kunnen waardevolle meting mogelijk maken mits beperkingen transparant worden gerapporteerd ⁴ .	CEQID sluit consensusaanbevelingen niet uit. Sterk geformuleerde aanbevelingen worden meegenomen, zwak geformuleerde niet. De evidentiebasis en onzekerheden worden steeds gedocumenteerd.
Verschillen tussen richtlijnen	Aanbevelingen rond hetzelfde onderwerp kunnen tussen richtlijnen verschillen in sterkte, evidentieniveau of interpretatie ^{6,27} .	Ontwikkelgroepen moeten keuzes maken tussen soms tegenstrijdige aanbevelingen. Professionele oordeelsvorming blijft noodzakelijk.	CEQID vereist dat verschillen tussen richtlijnen expliciet worden gedocumenteerd en besproken, met transparante motivatie van de uiteindelijke selectie.
Prioritering van aanbevelingen	Delphi en RAND/UCLA laten toe aanbevelingen gestructureerd te beoordelen op relevantie, impact, verbeterpotentieel en meetbaarheid. Anonieme beoordelingsrondes beperken groepsdruk ^{4,10} .	Delphi is tijdsintensief; consensus valt niet noodzakelijk samen met wetenschappelijke correctheid. Er bestaat geen internationale consensus over optimale panelgrootte of consensusdrempels ⁶ .	Bij minder dan 10 aanbevelingen beslist de ontwikkelgroep rechtstreeks. Bij meer dan 10 aanbevelingen wordt een RAND/UCLA-methode ^{10,11} aanbevolen met de leden van de ontwikkelgroep (Stap 5).

Output

Deze stap resulteert in:

- een overzicht van alle geëvalueerde aanbevelingen;
- een lijst van geselecteerde aanbevelingen;
- een overzicht van uitgesloten aanbevelingen;
- een motivatie van alle selectiebeslissingen;
- een overzicht van geïdentificeerde kennishiaten.

Deze output vormt de basis voor de formulering en matching van QI in [stap 5](#). In die stap worden de geselecteerde aanbevelingen op basis van evidentie, consensus, relevantie en meetbaarheid samengebracht met de bijbehorende kwaliteitsindicatoren.



Stap 5 – Consensusprocedure (optioneel)

Let op: Deze stap is enkel van toepassing als er 10 of meer aanbevelingen uit [stap 4](#) geselecteerd werden.

Doel

Wanneer er 10 of meer aanbevelingen voorliggen, voert de ontwikkelgroep een consensusprocedure uit. Doel is om collectief vast te stellen welke aanbevelingen werkelijk belangrijk zijn voor de definitieve QI-set, rekening houdend met de relevantie ten aanzien van het bereik en de kwaliteitsdomeinen die in [stap 2](#) bepaald werden.

Aanpak

De ontwikkelgroep is het meest geschikte forum voor deze beslissing omdat:

- ze betrokken was bij [stap 2](#) en alle aanbevelingen grondig kent;
- ze volledig zicht heeft op welke aanbevelingen binnen het bereik vallen;
- ze de samenhang tussen aanbevelingen begrijpt;
- de consensusprocedure sneller opgestart kan worden.

De consensusprocedure verloopt via de RAND/UCLA Appropriateness Method^{10,11}, een combinatie van individuele anonieme scoring en gestructureerde groepsdiscussie. De procedure bestaat uit 4 fasen:

- Fase 1: individuele beoordeling. Panelleden beoordelen elke aanbeveling onafhankelijk van elkaar via een online vragenlijst (Delphi ronde 1). Voor elke aanbeveling kennen ze een score toe op een negenpuntsschaal (1 = zeer laag, 9 = zeer hoog) om de geschiktheid van de aanbeveling te beoordelen, met een focus op de **dimensies relevantie en haalbaarheid**. Daarnaast prioriteren de panelleden de aanbevelingen voor elk onderdeel van het zorgproces (Top 3). Als er voor een onderdeel van het zorgproces minder dan 3 aanbevelingen zijn, wordt de prioritering van dat onderdeel niet bevestigd.
- Fase 2: Feedback. Na verwerking van de eerste beoordelingsronde ontvangt elk panellid een individueel feedbackrapport met de eigen scores en een overzicht van de scoreverdeling binnen het panel. Dat wordt besproken tijdens fase 3.
- Fase 3: Consensusbespreking. De panelleden bespreken de resultaten van de eerste beoordelingsronde tijdens een online consensusprocedure. Daarbij focussen ze op aanbevelingen waarvoor nog onduidelijkheid of discussie bestaat²⁹.

- Fase 4: Finale beoordeling. Na de consensusbespreking volgt een laatste beoordelingsronde waarin panelleden de finale aanbevelingen een laatste keer onafhankelijk beoordelen (Delphi ronde 2).

Waarom is deze stap noodzakelijk?

Wanneer er meer dan 10 aanbevelingen geselecteerd worden uit [stap 4](#), zou dat zonder deze stap kunnen leiden tot ontwikkeling van QI voor elk van die aanbevelingen. Dat is praktisch niet haalbaar en inefficiënt. Door deze stap van consensusbeoordeling in te voegen, beperken we ons tot echt relevante aanbevelingen. We kiezen hier voor kwaliteit boven kwantiteit, want niet alle aanbevelingen zijn even cruciaal voor kwaliteitsverbetering. Door deze consensusstap concentreren we ons op de meest impactvolle aanbevelingen, zodat we diepere en betere QI kunnen ontwikkelen, die ook bruikbaar zijn. Ook na deze stap kunnen er nog meer dan 10 aanbevelingen overblijven als het panel die allemaal relevant vindt om mee te nemen naar de volgende fase. Idealiter wordt er naar het eindresultaat toe gefocust op een set van een 5-tal QI. Dat is gemakkelijker voor zorgverleners om toe te passen, en zorgt voor minder administratieve last, betere naleving en meer impact op kwaliteitsverbetering.

Door de ontwikkelgroep zelf die selectie te laten maken, ontstaat er:

- eigenaarschap over de gekozen aanbevelingen;
- duidelijkheid over prioriteiten;
- gezamenlijke betrokkenheid bij de uiteindelijke QI-set.

Output

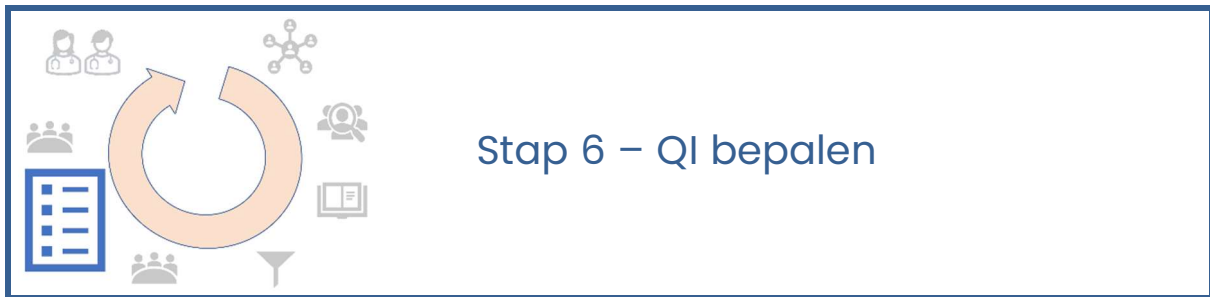
Deze stap resulteert in:

- duidelijk inzicht in welke aanbevelingen prioritair zijn;
- individuele scores en groepsscores voor alle aanbevelingen;
- documentatie van de 2 Delphirondes;
- motivatie van inclusie- en exclusiebeslissingen;
- overzicht van punten van consensus en punten van onenigheid.

De geselecteerde aanbevelingen worden in de volgende stap gebruikt om QI te bepalen.

Wetenschappelijke onderbouwing

Voor de wetenschappelijke onderbouwing bij deze stap verwijzen we naar de gelijknamige titel bij [stap 7](#).



Doel

Het systematisch koppelen van geselecteerde aanbevelingen aan bestaande QI en, waar nodig, het ontwikkelen van nieuwe QI die wetenschappelijk onderbouwd, meetbaar en toepasbaar zijn binnen de beoogde context.

Aanpak

In [stap 5](#) worden de geselecteerde aanbevelingen en de bestaande QI naast elkaar gelegd in een gestructureerde matching-oefening. Dat leidt tot 3 mogelijke uitkomsten:

- Koppeling: bestaande QI uit (inter)nationale sets worden gekoppeld aan de relevante aanbevelingen en waar nodig aangepast voor de Belgische context (teller, noemer, in-/exclusiecriteria).
- Nieuwe formulering: voor aanbevelingen zonder bestaande matchende QI formuleert de ontwikkelgroep een nieuwe QI conform de AIRE-criteria (teller, noemer, in-/exclusiecriteria, dataverenisten).
- Geen indicator mogelijk: aanbevelingen waarvoor geen goed formuleerbare QI mogelijk is, worden gedocumenteerd als 'niet meetbaar' en opgenomen in de rapportering.
- Indicatoren die niet aan een aanbeveling gekoppeld kunnen worden: als deze relevant zijn voor het afgebakende bereik ([stap 2](#)) voegt de ontwikkelgroep een toelichting toe over het wetenschappelijke bewijs, zoals gedocumenteerd in de beschikbare onderbouwing bij de QI.

Waarom is deze stap noodzakelijk?

Na selectie van aanbevelingen moet worden bepaald hoe die aanbevelingen kunnen worden vertaald naar concrete QI. Die vertaalslag is niet vanzelfsprekend. Deze matchingoefening vermijdt dubbel werk en verankert de evidentiebasis van de QI in bestaand internationaal werk. Het hergebruik van bestaande QI kan de acceptatie bij zorgverleners verhogen, mits contextaanpassing².

Campbell et al.¹ benadrukken dat aanbevelingen en QI verschillende concepten zijn. Aanbevelingen beschrijven wat wenselijke zorg is, terwijl QI beschrijven hoe die zorg gemeten kan worden².

Niet elke aanbeveling leidt tot één QI. Omgekeerd kan één QI meerdere aanbevelingen weerspiegelen. Kötter et al. beschrijven deze stap als een van de minst gestandaardiseerde onderdelen van QI-ontwikkeling. Veel publicaties rapporteren onvoldoende hoe aanbevelingen precies werden vertaald naar indicatoren⁴.

Binnen CEQUID wordt daarom bijzondere aandacht besteed aan transparantie en reproduceerbaarheid van deze vertaalslag.

Wetenschappelijke onderbouwing, kritische reflectie en keuze binnen CEQUID

Thema	🔗 Wetenschappelijke onderbouwing	⚠️ Kritische reflectie	✓ Keuze binnen CEQUID
Hergebruik van bestaande QI	Bestaande QI bieden voordelen zoals internationale vergelijkbaarheid en vermindering van ontwikkeltijd ^{2,9} .	QI zijn context-specifiek en worden soms overgenomen zonder kritische evaluatie van actuele relevantie of methodologische kwaliteit ⁴ .	CEQUID hanteert een getrapte aanpak: eerst wordt nagegaan of kwalitatief hoogwaardige QI beschikbaar zijn, gevolgd door beoordeling op kwaliteit, actualiteit en toepasbaarheid.
Contextuele validatie	QI kunnen niet los worden gezien van de context waarin zij werden ontwikkeld. Factoren zoals zorgorganisatie, financiering en registratiesystemen beïnvloeden de toepasbaarheid ⁵ .	Hergebruik kan leiden tot onbedoelde overdracht van aannames die niet aansluiten bij de lokale context.	Voor elke voorlopige QI beoordeelt CEQUID expliciet de overdraagbaarheid naar de beoogde context.
Operationalisering van aanbevelingen	Operationalisering omvat het vertalen van aanbevelingen naar meetbare QI. Daarvoor moeten minimaal een teller, noemer, inclusiecriteria, exclusiecriteria en gegevensbron worden vastgelegd ^{2,4,24} .	Eén aanbeveling kan leiden tot meerdere potentiële QI; de uiteindelijke keuze blijft gedeeltelijk afhankelijk van professionele oordeelsvorming.	Wanneer er geen geschikte QI beschikbaar is, ontwikkelt CEQUID een nieuwe op basis van systematische operationalisering. Alle keuzes worden expliciet vastgelegd.
Relatie met kwaliteitscriteria	Tijdens QI-ontwikkeling wordt nagegaan in welke mate een voorlopige QI voldoet aan kwaliteitscriteria zoals validiteit, betrouwbaarheid, haalbaarheid en interpreteerbaarheid. Die beoordeling voorkomt dat aanbevelingen automatisch worden omgezet naar QI zonder methodologische evaluatie ^{2,9} .	Een QI die sterk scoort op één kwaliteitscriterium scoort niet noodzakelijk goed op andere criteria. Een zeer valide QI kan bijvoorbeeld moeilijk meetbaar zijn, terwijl een eenvoudig meetbare indicator beperkt kan bijdragen aan kwaliteitsverbetering. Er bestaat geen universeel optimale balans tussen deze criteria.	CEQUID beoordeelt elke voorlopige QI systematisch op validiteit, betrouwbaarheid, haalbaarheid en interpreteerbaarheid. De resultaten van die beoordeling worden gedocumenteerd en meegenomen in de verdere selectie- en validatieprocedure.
Eenvoud versus nauwkeurigheid	QI-ontwikkeling vereist een afweging: eenvoudige QI zijn gemakkelijker meetbaar, complexere QI weerspiegelen de klinische werkelijkheid beter ^{2,30} .	Te sterke vereenvoudiging leidt tot verlies van klinische informatie; overmatige complexiteit vermindert haalbaarheid.	CEQUID streeft naar QI die voldoende eenvoudig zijn voor praktische toepassing, zonder essentiële klinische aspecten te verliezen.
Risico op verlies van nuance	Richtlijnen bevatten vaak uitzonderingen en contextuele afwegingen die niet volledig in een QI kunnen worden opgenomen ^{2,30} .	Operationalisering gaat altijd gepaard met vereenvoudiging, wat kan leiden tot verlies van klinische nuance.	CEQUID erkent dat operationalisering altijd vereenvoudiging impliceert. Alle keuzes en aannames worden expliciet gedocumenteerd.
Onafhankelijkheid	De formulering wordt door minstens twee personen onafhankelijk beoordeeld ⁴ . De tweede beoordelaar hoeft geen	Volledige onafhankelijkheid is in de praktijk niet altijd haalbaar. Beoordelaars beschikken vaak al over voorkennis van het	Binnen de ontwikkelgroep wordt gekeken hoe de onafhankelijkheid gerealiseerd kan worden door

Thema	🔗 Wetenschappelijke onderbouwing	⚠ Kritische reflectie	✓ Keuze binnen CEQUID
	externe expert te zijn, maar moet wel onafhankelijk zijn van het initiële formuleringsproces ¹⁹ .	onderwerp, de richtlijn of eerdere discussies. Een bijkomend spanningsveld is dat personen die niet betrokken waren bij eerdere discussies onafhankelijker kunnen oordelen, maar mogelijk minder inzicht hebben in de achterliggende keuzes en context ^{4,8} .	bijvoorbeeld wisselende taakverdeling binnen de ontwikkelgroep: wie formuleert, beoordeelt niet. Een lid van de ontwikkelgroep heeft kennis van de discussies en kan daardoor minder onafhankelijk beoordelen. Hoe dat verloopt wordt expliciet gedocumenteerd.

Output

Deze stap resulteert in:

- een overzicht van bestaande QI die werden geëvalueerd;
- een overzicht van nieuwe QI;
- een overzicht van aanbevelingen waarvoor het niet mogelijk is om een goed formuleerbare QI te ontwikkelen.
- voorlopige QI-fiches;
- documentatie van alle operationaliseringskeuzes;
- een voorlopige QI-set.

Deze voorlopige QI-set vormt de input voor de consensusprocedure in [stap 7](#).



Doel

Het systematisch beoordelen van de klinische relevantie, inhoudelijke validiteit, prioriteit, meetbaarheid, aanvaardbaarheid en inschatting van mogelijke ongewenste effecten van de voorlopige QI door middel van een transparante en reproduceerbare consensusprocedure met een bredere stakeholdergroep.

Aanpak

De consensusprocedure verloopt via de RAND/UCLA Appropriateness Method^{10,11}, een combinatie van individuele anonieme scoring en gestructureerde groepsdiscussie. De procedure bestaat uit 4 fasen:

- Fase 1: individuele beoordeling. Panelleden beoordelen elke QI onafhankelijk van elkaar via een online vragenlijst (Delphi ronde 1). Voor elke QI wordt een score toegekend op een negenpuntsschaal (1 = zeer laag, 9 = zeer hoog) om de geschiktheid van de QI te beoordelen, nu ook met een focus op de **dimensie meetbaarheid**. Daarnaast prioriteren de panelleden de QI voor elk onderdeel van het zorgproces (Top 3). Indien er voor een onderdeel van het zorgproces minder dan 3 QI zijn, wordt de prioritering van dat onderdeel niet bevestigd.
- Fase 2: Feedback. Na verwerking van de eerste beoordelingsronde ontvangt elk panellid een individueel feedbackrapport met de eigen scores en een overzicht van de scoreverdeling binnen het panel. Dat wordt besproken tijdens fase 3.
- Fase 3: Consensusbespreking. De resultaten van de eerste beoordelingsronde worden besproken tijdens een online consensusprocedure. Daarbij wordt gefocust op QI waarvoor nog onduidelijkheid of discussie bestaat²⁹.
- Fase 4: Finale beoordeling. Na de consensusbespreking volgt een laatste beoordelingsronde waarin panelleden de finale QI-set een laatste keer onafhankelijk beoordelen (Delphi ronde 2). In deze fase lezen alle panelleden de definitieve formulering ter goedkeuring en geven een bijkomende prioritering voor de implementatie mee.

De samenstelling van het panel werd al voorbereid bij de stakeholderanalyse in [stap 1](#). Minimaal zijn de volgende profielen vereist:

- methodologen;
- klinische experts relevant voor het onderwerp;
- vertegenwoordigers van de doelgroep;
- patiëntenvertegenwoordigers.

Bijkomend moet er ook bekeken worden of er andere relevante stakeholders betrokken moeten worden, zoals bv. beleidsmakers, zorgorganisaties of beroepsverenigingen.

De samenstelling, de gehanteerde schaal, de consensusdrempel, het aantal rondes en de procedure als er geen consensus wordt gevonden worden vooraf vastgelegd en volledig gerapporteerd in het methodologisch rapport.

Waarom is deze stap noodzakelijk?

Zelfs wanneer QI gebaseerd zijn op hoogwaardige richtlijnen en zorgvuldig werden geoperationaliseerd, blijven belangrijke beoordelvragen bestaan. Wetenschappelijke evidentie alleen biedt niet altijd een antwoord op vragen zoals:

- Is deze QI voldoende relevant voor kwaliteitsverbetering?
- Is deze QI prioritair ten opzichte van andere indicatoren?
- Is deze QI aanvaardbaar voor toekomstige gebruikers?
- Is deze QI voldoende meetbaar in de praktijk?

De RAND/UCLA Appropriateness Method werd ontwikkeld precies om situaties te ondersteunen waarin wetenschappelijke evidentie moet worden gecombineerd met professionele expertise³¹. Campbell et al. benadrukken dat QI-ontwikkeling niet uitsluitend evidence-based kan zijn. Interpretatie van evidentie blijft noodzakelijk². Consensusvorming vormt daarom een essentiële brug tussen wetenschappelijke kennis en praktische toepasbaarheid.

Wetenschappelijke onderbouwing, kritische reflectie en keuze binnen

CEQUID

Thema	🔍 Wetenschappelijke onderbouwing	⚠️ Kritische reflectie	✓ Keuze binnen CEQUID
Consensus als aanvulling op evidentie	QI-ontwikkeling combineert evidentie en professionele expertise. Consensusvorming ondersteunt op gestructureerde manier beoordeling van prioriteit, haalbaarheid en relevantie ² .	Consensusvorming is afhankelijk van de samenstelling van het panel en de gekozen methodologie. Consensus mag niet geïnterpreteerd worden als bewijs van juistheid ^{4,10,11} .	CEQUID beschouwt consensus als aanvulling op evidentie, niet als vervanging. Consensusmethoden brengen evidentie, praktijkervaring en haalbaarheid systematisch samen.
RAND/UCLA Appropriateness Method	De RAND/UCLA Appropriateness Method ^{10,11} omvat individuele beoordeling, geaggregeerde feedback, groepsdiscussie en een herbeoordeling. De methode creëert expliciet ruimte voor discussie over interpretatieverschillen en meningsverschillen ^{4,11,31} .	Er bestaat geen internationale consensus over optimaal aantal rondes, panelgrootte of drempels ^{4,10} .	CEQUID beveelt de RAND/UCLA Appropriateness Method ^{10,11} aan als transparante en reproduceerbare consensusprocedure.
Belgische ervaring	Ook binnen de Belgische context werd een RAND/UCLA-gebaseerde methode succesvol toegepast voor de ontwikkeling	Er bestaat slechts beperkte empirische literatuur die verschillende consensusmethoden	CEQUID bouwt verder op internationale én Belgische ervaring met

Thema	🔗 Wetenschappelijke onderbouwing	⚠️ Kritische reflectie	✓ Keuze binnen CEQUID
	van kwaliteitsindicatoren in de eerstelijnszorg. Die ervaring ondersteunt de praktische toepasbaarheid van gestructureerde consensusmethoden binnen de context waarvoor CEQUID ontwikkeld werd ^{21,32} .	rechtstreeks met elkaar vergelijkt. Daardoor blijft het moeilijk om uitspraken te doen over de superioriteit van een bepaalde methode binnen specifieke zorgcontexten.	consensusmethoden en kiest voor een pragmatische aanpak die aansluit bij bestaande expertise en praktijkervaring.
Panelgrootte	Kleinere panels bevorderen efficiëntie; grotere panels bieden meer perspectieven en representativiteit ¹⁰ .	Er bestaat geen duidelijk optimaal aantal deelnemers in de literatuur ¹⁰ .	CEQUID adviseert een panelsamenstelling van 9 tot 18 stakeholders met verschillende expertises ¹¹ . De keuze om het aantal panelleden te beperken, gebeurt vanuit de RAND/UCLA Appropriateness Method ^{10,11} . Dit maakt de procedure gemakkelijker werkbaar. De samenstelling van het panel wordt expliciet gemotiveerd.
Consensusdrempels	Consensusdrempels variëren sterk tussen studies. In de literatuur worden drempels van 70%, 75%, 80% en 90% gerapporteerd, wat illustreert dat consensuscriteria gedeeltelijk arbitrair blijven ¹⁰ .	Het ontbreken van uniforme drempels bemoeilijkt vergelijking tussen studies. Deze keuze beïnvloedt welke QI uiteindelijk worden geselecteerd.	CEQUID vereist dat consensuscriteria vooraf worden vastgelegd en transparant worden gerapporteerd.
Patiëntenparticipatie	Patiënten stellen andere kwaliteitsprioriteiten dan zorgverleners. QI zonder patiëntenperspectief ondermeten daarom systematisch patiëntgerichtheid ²² .	Bergholtz et al. wijzen op het risico van 'tokenism' en de grote variatie in terminologie ('patiëntenbetrokkenheid', 'co-design', 'ervaringsdeskundigheid') die betekenisvolle vergelijking bemoeilijkt ²² .	De rol van patiëntenvertegenwoordigers, de manier waarop zij worden ondersteund en hun inbreng worden gewogen en vooraf schriftelijk vastgelegd in een participatieprotocol.
Belangenconflicten	Transparantie rond belangenconflicten is essentieel voor de geloofwaardigheid van ontwikkelprocessen ⁸ .	De aanwezigheid van specifieke stakeholders kan onbedoeld leiden tot dominantie van bepaalde perspectieven of belangen. Volledige neutraliteit is daarom niet haalbaar. Transparante rapportering van belangenconflicten vermindert dit risico, maar kan het niet volledig uitsluiten ^{8,22} .	Mogelijke belangenconflicten worden systematisch geregistreerd en gerapporteerd. Alle leden van het consensuspanel ondertekenen een verklaring van belangen (COI-verklaring), conform de aanbevelingen van AGREE II.

Output

Deze stap resulteert in:

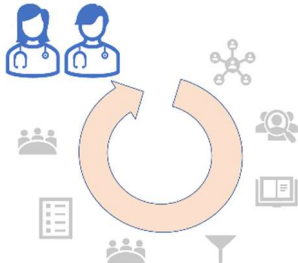
- een inhoudelijk gevalideerde voorlopige QI-set en QI-fiches;
- individuele en geaggregeerde scores van de panelleden over de QI;
- documentatie van alle beoordelingsrondes;

- motivatie van inclusie- en exclusiebeslissingen;
- beschrijving van het consensuspanel;
- beschrijving van de gevolgde procedure.

De aldus inhoudelijk gevalideerde QI-set vormt de input voor de praktijktoets in fase 3. Na de 3^e fase zijn de QI ontwikkeld volgens de CEQUID-methode. De tweedeling tussen QI-set met en zonder praktijktoets sluit aan bij het concept van 'progressive endorsement'. De QI-sets zonder praktijktoets krijgen een lagere bewijsstatus^{3,33} en zijn geschikt voor interne kwaliteitsverbetering, maar niet voor externe benchmarking.

Fase 3 – Praktijkttoets

De ontwikkelingsfase resulteert in een set voorlopige QI, waarvoor de inhoudelijke validiteit en relevantie werden beoordeeld door experts en stakeholders. Die consensusvalidatie is een belangrijke stap, maar laat nog geen uitspraken toe over de praktische toepasbaarheid. Om na te gaan of indicatoren daadwerkelijk meetbaar, interpreteerbaar en bruikbaar zijn, worden zij in fase 3 onderworpen aan een testprocedure.



Stap 8 – Bruikbaarheid en ongewenste effecten

Doel

Deze stap gaat na of de ontwikkelde QI praktisch bruikbaar zijn. Haalbaarheid, betrouwbaarheid, validiteit, interpreteerbaarheid, aanvaardbaarheid en discriminerend vermogen van de ontwikkelde QI zonder praktijktoets uit [stap 7](#) worden geëvalueerd in een reële context.

Daarnaast identificeert, beoordeelt en documenteert de ontwikkelgroep in deze stap op systematische wijze de mogelijke ongewenste effecten van QI vóór brede implementatie.

Aanpak

Aan de hand van een praktijktoets wordt binnen de Belgische praktijkcontext getoetst of de ontwikkelde QI haalbaar, betrouwbaar en interpreteerbaar zijn en een voldoende discriminerend vermogen hebben. De praktijktoets is een verplicht onderdeel van een volwaardige CEQUID-procedure en omvat 4 deeltappen:

- 1) Operationalisering: bepalen hoe de QI worden berekend, welke gegevensbronnen beschikbaar zijn en welke teller- en noemerdefinities operationeel haalbaar zijn.
- 2) Haalbaarheidstoets: kritische evaluatie van het dataverzamelingsproces, inclusief de inspanningsvereisten voor de (registrerende) zorgverlener.
- 3) Kwantitatieve evaluatie: kritische evaluatie van de eerste resultaten van de QI in een testgroep, ten aanzien van hun betrouwbaarheid en discriminerend vermogen.
- 4) Kwalitatieve evaluatie: interviews of focusgroepen met de testgroep om interpretatie, aanvaardbaarheid en ongewenste neveneffecten te toetsen. Een gestandaardiseerde checklist voor ongewenste effecten¹⁸ is standaard onderdeel van deze evaluatie.

Na de praktijktoets brengt de ontwikkelgroep alle verzamelde gegevens samen en beoordeelt ze. Zij neemt de volgende beslissingen:

- Welke QI zijn geschikt voor opname in de finale QI-set? Die QI hebben geen substantiële barrières, risico's of problemen.
- Welke QI moeten worden bijgesteld? Moeten er bv. aanpassingen gebeuren aan de teller, noemer of exclusiecriteria op basis van de praktijktoetsbevindingen?
- Welke QI worden voorlopig niet opgenomen? Redenen daarvoor kunnen zijn: dataproblemen, ongewenste neveneffecten of onvoldoende discriminerend vermogen. De reden wordt expliciet gedocumenteerd.

Waarom is deze stap noodzakelijk?

Testen van praktische bruikbaarheid

Veel QI-ontwikkelmethoden eindigen na de consensusprocedure in [stap 7](#). Daardoor worden QI soms geïmplementeerd zonder dat hun praktische bruikbaarheid vooraf werd onderzocht. Een formele praktijktoets wordt slechts beperkt beschreven binnen internationale methoden voor QI-ontwikkeling. Daarnaast rapporteren veel publicaties onvoldoende over de prestaties van QI buiten de ontwikkelomgeving, waardoor het moeilijk blijft om hun praktische bruikbaarheid en generaliseerbaarheid te beoordelen^{4,6,24}. Internationale ontwikkelprogramma's beschouwen piloottesting als een essentiële stap om de haalbaarheid, gegevensbeschikbaarheid, betrouwbaarheid en praktische toepasbaarheid van QI te evalueren vóór implementatie³⁴. Consensus over relevantie betekent immers niet automatisch dat een QI:

- meetbaar is;
- interpreteerbaar is;
- voldoende gegevens beschikbaar heeft;
- praktisch uitvoerbaar is;
- kwaliteitsverschillen kan detecteren.

Daarom beschouwt CEQUID de praktijktoets als een essentiële validatiestap.

Ongewenste effecten voorkomen

QI worden doorgaans ontwikkeld met als doel kwaliteitsverbetering te ondersteunen. Toch tonen verschillende auteurs aan dat kwaliteitsmeting ook ongewenste effecten kan veroorzaken. Smith beschreef reeds dat prestatie-QI gedragsveranderingen kunnen uitlokken die niet noodzakelijk leiden tot betere zorg¹⁸. Orueta et al. bevestigden later dat QI soms aanleiding geven tot ongewenste neveneffecten, zoals risicoselectie, meetfixatie of gaming³⁵.

Internationale kwaliteitskaders benadrukken bovendien dat kwaliteitsmetingen niet geïsoleerd geïnterpreteerd mogen worden, maar steeds bekeken moeten worden binnen een bredere set van complementaire QI en kwaliteitsinformatie. Afzonderlijke indicatoren meten immers slechts een beperkt aspect van zorgkwaliteit en kunnen daardoor een onvolledig beeld geven³⁶.

Hoewel dergelijke effecten vaak pas zichtbaar worden tijdens implementatie, kunnen sommige risico's al tijdens de ontwikkelingsfase worden geïdentificeerd.

CEQUID besteedt daarom expliciet aandacht aan mogelijke ongewenste effecten voor het QI definitief aanbeveelt. Op die manier verhoogt de methodologische robuustheid en implementatiezekerheid.

Wetenschappelijke onderbouwing, kritische reflectie en keuze binnen CEQID

Thema	🔗 Wetenschappelijke onderbouwing	⚠️ Kritische reflectie	✓ Keuze binnen CEQID
Haalbaarheid	Haalbaarheid is een van de belangrijkste voorwaarden voor succesvolle implementatie. Aanbevolen wordt om al tijdens de ontwikkeling aandacht te besteden aan gegevensbeschikbaarheid, registratielast en technische uitvoerbaarheid ¹⁴ .	Een spanningsveld betreft de vraag of QI moeten worden aangepast wanneer zij moeilijk meetbaar blijken. Aanpassing verhoogt haalbaarheid, maar kan belangrijke kwaliteitsaspecten doen verloren gaan.	CEQID beoordeelt systematisch de beschikbaarheid van gegevens, de registratielast en de technische uitvoerbaarheid. Haalbaarheid wordt meegenomen in de eindbeoordeling, maar vormt niet het enige criterium.
Betrouwbaarheid	Betrouwbare QI moeten consistente resultaten opleveren. Definities moeten eenduidig zijn en verschillende beoordelaars moeten tot vergelijkbare resultaten kunnen komen ¹⁹ .	Zelfs goed gedefinieerde QI kunnen in de praktijk verschillend worden geïnterpreteerd. Variaties in registratiepraktijken beïnvloeden de betrouwbaarheid.	CEQID evalueert of QI consistent kunnen worden gemeten en of definities voldoende duidelijk zijn.
Discriminerend vermogen	Een QI moet relevante kwaliteitsverschillen zichtbaar kunnen maken. Tijdens de praktijktoets wordt aandacht besteed aan scoreverdeling, plafondeffecten (een groot aantal deelnemers in de populatie behaalt een score dicht tegen de maximumscore, wat weinig ruimte geeft om verschillen in de populatie zichtbaar te maken) en bodemeffecten (idem aan plafondeffecten, maar dan omwille van een groot aantal scores dicht tegen de minimumscore) ¹⁶ .	Discriminerend vermogen kan sterk afhankelijk zijn van de onderzochte populatie of context.	CEQID onderzoekt de spreiding van scores in een basismeting en beoordeelt of de QI voldoende onderscheid maken binnen de onderzochte populatie.
Interpreteerbaarheid	Gebruikers moeten begrijpen wat een QI meet en hoe resultaten geïnterpreteerd moeten worden. De praktijktoets biedt de kans om problemen vroegtijdig te identificeren ²⁰ .	Interpretatieproblemen worden vaak pas zichtbaar wanneer QI daadwerkelijk worden toegepast ³⁴ .	CEQID verzamelt gebruikersfeedback over interpretatie en rapportering. Waar nodig worden definities of fiches aangepast.
Aanvaardbaarheid	QI die als onduidelijk, onrechtvaardig of administratief belastend worden ervaren, lopen een groter risico op weerstand ² .	Aanvaardbaarheid verschilt tussen stakeholders. Volledige consensus is niet altijd haalbaar.	CEQID betreft toekomstige gebruikers om aanvaardbaarheid te evalueren. Feedback over meerwaarde, duidelijkheid en registratielast wordt meegenomen.
Noodzaak van praktijktoets	Een praktijktoets levert informatie over praktische bruikbaarheid die niet via literatuuronderzoek of consensusprocedure kan worden verkregen. Langendam et al. stelden vast dat evaluatie en testing minder consistent worden beschreven ⁶ .	Voorstanders van een pragmatische aanpak argumenteren dat consensusvalidatie soms voldoende kan zijn en dat een praktijktoets tijdrovend en duur is.	CEQID beschouwt de praktijktoets als een noodzakelijke stap. QI die geen volledige praktijktoets hebben doorlopen, worden gelabeld als 'QI zonder praktijktoets'. Ze zijn geschikt voor interne kwaliteitsverbetering, maar

Thema	🔗 Wetenschappelijke onderbouwing	⚠ Kritische reflectie	✓ Keuze binnen CEQID
			niet voor externe benchmarking.
Verbeterpotentieel	Het effect van QI is sterk afhankelijk van de context: zorgverleners met lage basisprestatie profiteren meer ²⁴ .	Dit impliceert dat de meerwaarde van QI niet uniform is en dat doelgerichte implementatiestrategieën noodzakelijk zijn.	CEQID besteedt aandacht aan de heterogeniteit van de testgroep.
Beperkte generaliseerbaarheid	Resultaten van praktijktoetsen zijn steeds afhankelijk van de context waarin zij worden uitgevoerd. Factoren zoals organisatiegrootte, digitaliseringsgraad en zorgcontext kunnen de prestaties van QI beïnvloeden.	Een indicator die goed presteert in een grote groepspraktijk of academische setting zal niet noodzakelijk dezelfde prestaties vertonen in kleinere organisaties of andere zorgcontexten. Daardoor blijft volledige generaliseerbaarheid moeilijk.	CEQID documenteert de kenmerken van de testomgeving expliciet, zodat gebruikers de overdraagbaarheid van de resultaten beter kunnen beoordelen.
Wanneer is een QI voldoende getest?	De literatuur biedt weinig duidelijke richtlijnen over minimale steekproefgroottes, duur van praktijktoetsen of het aantal deelnemende organisaties.	Daardoor blijft professionele oordeelsvorming noodzakelijk bij de interpretatie van testresultaten en de beslissing om een QI al dan niet verder te implementeren.	CEQID legt geen vaste drempels op voor de praktijktoets, maar vereist dat de omvang van de test, de deelnemers en de beperkingen van de resultaten transparant worden gerapporteerd. Wanneer een formele praktijktoets niet mogelijk is, worden QI expliciet gerapporteerd als QI zonder praktijktoets.
Gaming	Gaming wordt beschreven als het bewust aanpassen van gedrag om betere QI-resultaten te behalen, zonder dat de werkelijke kwaliteit van zorg verbetert. Voorbeelden zijn registreren zonder daadwerkelijk handelen, selectieve registratie en aanpassing van registratiecriteria ¹⁸ .	De empirische evidentie over de frequentie en impact van gaming blijft beperkt. Veel publicaties zijn gebaseerd op casestudies. Het blijft moeilijk om vooraf in te schatten welke QI het grootste risico vormen ^{18,37} .	CEQID vereist dat mogelijke gamingrisico's tijdens de ontwikkelingsfase worden besproken en dat mitigerende maatregelen worden gedocumenteerd.
Risicoselectie	QI kunnen onbedoeld stimuleren dat complexe patiënten worden vermeden wanneer risicocorrectie ontbreekt. Dit is relevant bij benchmarking en externe evaluatie ³⁵ .	Niet alle vormen van risicoselectie zijn gemakkelijk detecteerbaar. Volledig corrigeren voor patiëntcomplexiteit is vaak moeilijk.	CEQID vereist dat wordt nagegaan of patiëntkenmerken of risicocorrectie noodzakelijk zijn voor correcte interpretatie van een QI.
Meetfixatie	Meetfixatie verwijst naar het verschijnsel waarbij aandacht verschuift naar gemeten aspecten van zorg, terwijl niet-gemeten aspecten minder aandacht krijgen. ¹⁸	Het spanningsveld tussen meten en verbeteren kan niet volledig worden opgelost. Het risico op meetfixatie is inherent aan QI.	CEQID gebruikt de IOM-kwaliteitsdomeinen als reflectiekader. Ontwikkelgroepen motiveren expliciet welke domeinen wel en niet worden meegenomen.
Registratielast	Kwaliteitsmetingen kunnen leiden tot een toename van administratieve belasting. Een hoge registratielast beïnvloedt aanvaarbaarheid en implementatie negatief ¹⁴ .	Het streven naar nauwkeurige meting kan botsen met de wens om registratielast te beperken.	CEQID beoordeelt expliciet de registratielast. Waar mogelijk wordt gebruikgemaakt van routinematig geregistreerde gegevens.

Thema	🔗 Wetenschappelijke onderbouwing	⚠ Kritische reflectie	✓ Keuze binnen CEQUID
Onbedoelde gedragsverandering	QI beïnvloeden het gedrag van zorgverleners en organisaties. Naast beoogde effecten kunnen zij leiden tot verschuiving van prioriteiten of overbehandeling ³⁵ .	Niet alle ongewenste effecten zijn vooraf voorspelbaar; sommige worden pas zichtbaar na langdurig gebruik.	CEQUID beschouwt identificatie van mogelijke ongewenste effecten als vast onderdeel. Ontwikkelgroepen bespreken systematisch welke gedragsveranderingen een indicator kan uitlokken.
Systematische beoordeling van ongewenste effecten	Internationaal bestaan er verschillende beschrijvingen van ongewenste effecten van QI, maar er is momenteel geen breed aanvaard instrument beschikbaar dat vergelijkbaar is met AGREE II voor richtlijnen of AIRE voor indicatoren ⁹ .	Het ontbreken van een gestandaardiseerde beoordelingsmethode maakt dat de identificatie van risico's grotendeels afhankelijk blijft van professionele oordeelsvorming. Daardoor kunnen verschillende ontwikkelgroepen tot verschillende inschattingen komen.	CEQUID erkent dat ongewenste effecten niet volledig kunnen worden uitgesloten. Daarom worden mogelijke risico's systematisch besproken, gedocumenteerd en meegenomen in de besluitvorming.

Output

Deze stap resulteert in:

- een praktijktoetsrapport;
 - o Dit rapport bevat ook een overzicht van alle QI die niet geselecteerd worden.
- evaluatie van haalbaarheid;
 - o Voor alle QI wordt weergegeven of die in de huidige context al dan niet haalbaar zijn om te meten.
- evaluatie van betrouwbaarheid;
- evaluatie van discriminerend vermogen;
- gebruikersfeedback;
- een overzicht van mogelijke ongewenste effecten van QI en maatregelen om die te beperken;
- een methodologisch verantwoordingsdocument;
- QI-fiches per QI (teller, noemer, in-/exclusiecriteria, level van evidentie van de aanbeveling waarop de QI gebaseerd is en praktijktoetsresultaten);
- een definitieve QI-set.

De verplichting tot praktijktoets vormt een van de belangrijkste verschillen tussen CEQUID en verschillende bestaande QI-ontwikkelingsmethoden. Deze keuze verhoogt de methodologische robuustheid, transparantie en implementatiekansen. Tegelijk verhoogt zij ook de ontwikkelingsduur, benodigde middelen en organisatorische complexiteit.

CEQUID vertrekt vanuit de overtuiging dat QI die gebruikt worden voor kwaliteitsverbetering, benchmarking of beleidsdoeleinden voldoende robuust moeten zijn om die bijkomende investering te rechtvaardigen.

Na voltooiing van [stap 8](#) kan een QI beschouwd worden als volledig ontwikkeld volgens de

CEQUID-methodologie.

6. Validatie van QI

Het AIRE-instrument

AIRE (Appraisal of Indicators through Research and Evaluation) is een gevalideerd instrument voor de beoordeling van de kwaliteit en methodologische onderbouwing van bestaande QI-sets. Het instrument bevat 20 stellingen, ingedeeld in 4 blokken:

- Blok 1 (items 1-5): doel, relevantie en organisatorisch verband
- Blok 2 (items 6-8): betrokkenheid van belanghebbenden
- Blok 3 (items 9-11): wetenschappelijk bewijs
- Blok 4 (items 12-20): verdere onderbouwing, formulering en gebruik

AIRE werd ontwikkeld door de Koning et al. als een equivalent van AGREE voor QI en biedt een gestructureerde methode om de transparantie, validiteit en kwaliteit van QI-ontwikkeling te beoordelen⁹. Validatieonderzoek door Van den Bulck et al. toonde aan dat AIRE in staat is onderscheid te maken tussen QI-sets van hoge en lage methodologische kwaliteit en dat hogere AIRE-scores samenhangen met een grotere acceptatie van QI binnen consensusprocedures³². Daardoor wordt AIRE internationaal beschouwd als een belangrijk instrument voor de kritische evaluatie en selectie van bestaande kwaliteitsindicatoren.

Validatiecriteria van Cebam

Wanneer we QI ontwikkelen volgens de volledige CEQUID-methodologie, streven we maximaal na dat ze voldoen aan de kwaliteitscriteria van het AIRE-instrument. CEQUID sluit inhoudelijk nauw aan bij AIRE, maar legt op twee domeinen strengere methodologische eisen op.

AIRE item 8 – Formele vaststelling via consensusprocedure

Voor de formele vaststelling van QI vereist AIRE dat relevante stakeholders betrokken worden bij de ontwikkeling en beoordeling van QI. CEQUID gaat daarin verder door een expliciete en vooraf gedefinieerde formele consensusprocedure te vereisen, gebaseerd op de RAND/UCLA Appropriateness Method^{10,11}. Daarbij beoordelen panelleden QI eerst onafhankelijk van elkaar voordat de resultaten tijdens een gestructureerde consensusprocedure worden besproken. Die aanpak vermindert het risico op groepsdenken en dominantie door individuele experts en verhoogt de transparantie van de besluitvorming. Onderzoek van Kötter et al. toont aan dat dergelijke iteratieve consensusprocedures bijdragen aan een meer gestructureerde en reproduceerbare selectie van QI⁴.

AIRE item 18 – Praktijkttoets en ongewenste neveneffecten

Daarnaast besteedt CEQUID meer aandacht aan de mogelijke ongewenste effecten van QI. Hoewel AIRE aandacht heeft voor de praktische toepasbaarheid van QI, vereist CEQUID expliciet dat de ontwikkelgroep nagaat of QI aanleiding kunnen geven tot ongewenste

effecten zoals meetfixatie, tunnelvisie, misinterpretatie, gaming of risicoselectie. Indien dergelijke risico's worden geïdentificeerd, beveelt CEQUID aan om aanvullende QI of monitoringmechanismen te voorzien die die effecten kunnen detecteren. Die benadering sluit aan bij het werk van Smith, die verschillende vormen van ongewenste effecten van prestatiemeting beschreef¹⁸, en bij latere bevindingen van Orueta et al., die aantoonde dat problemen zoals meetfixatie, risicoselectie en attributiebias belangrijke aandachtspunten blijven bij het gebruik van QI³⁵. Ook internationale organisaties benadrukken dat QI bij voorkeur worden geïnterpreteerd binnen een bredere QI-set en niet als geïsoleerde QI om ongewenste effecten zoals meetfixatie en tunnelvisie te beperken. Het combineren van complementaire QI kan bijdragen aan een evenwichtigere beoordeling van zorgkwaliteit en helpen om ongewenste effecten te detecteren^{36,38}.

Door deze aanvullende vereisten biedt CEQUID een uitbreiding op AIRE waarbij niet alleen de methodologische kwaliteit van QI centraal staat, maar ook de kwaliteit van het besluitvormingsproces en de beheersing van mogelijke ongewenste effecten van kwaliteitsmeting.

7. Algemene discussie

Inleiding

De ontwikkeling van QI heeft de afgelopen decennia een belangrijke plaats ingenomen binnen kwaliteitsverbetering, gezondheidszorgbeleid en evaluatie van zorgprocessen. Ondanks deze brede toepassing bestaat er internationaal nog steeds aanzienlijke variatie in de wijze waarop QI worden ontwikkeld, gevalideerd en geïmplementeerd.

QI-ontwikkeling wordt vaak gekenmerkt door heterogene methodologische keuzes, beperkte transparantie en variabele rapportering van validatiestappen^{2,4,24}.

CEQID werd ontwikkeld als antwoord op die uitdagingen en beoogt een transparant, reproduceerbaar en wetenschappelijk onderbouwd ontwikkelproces aan te bieden voor kwaliteitsindicatoren binnen de eerstelijnszorg.

Sterktes van de CEQID-methodologie

Integratie van bestaande conceptuele kaders

CEQID vertrekt niet vanuit een volledig nieuw conceptueel kader, maar bouwt voort op internationaal aanvaarde instrumenten:

- AGREE II van Brouwers et al. voor richtlijnbeoordeling⁸;
- AIRE van de Koning et al. voor beoordeling van indicatorontwikkeling⁹;
- RAND/UCLA Appropriateness Method voor consensusvorming^{10,11}.

Hierdoor sluit de methodologie aan bij bestaande internationale standaarden.

Transparantie van besluitvorming

Kötter et al. beschrijven beperkte transparantie als een van de meest voorkomende tekortkomingen binnen indicatorontwikkeling⁴. CEQID probeert dat te ondervangen door systematische documentatie van:

- richtlijnselectie;
- aanbevelingsselectie;
- operationalisering;
- consensusvorming;
- praktijktoets;
- risicoanalyse.

Hierdoor wordt het ontwikkelproces beter reproduceerbaar.

Expliciete koppeling aan evidentie

Campbell et al. benadrukken dat kwaliteitsindicatoren idealiter rechtstreeks gekoppeld zijn

aan wetenschappelijke evidentie². De richtlijngebaseerde benadering versterkt inhoudelijke validiteit, wetenschappelijke geloofwaardigheid en traceerbaarheid.

Verplichte praktijktoets

Een onderscheidende eigenschap van CEQUID is de expliciete vereiste tot praktijktoets. In de literatuur wordt die stap beperkt beschreven^{4,24}. Door een praktijktoets verplicht te maken wordt niet alleen gekeken naar inhoudelijke validiteit, maar ook naar haalbaarheid, betrouwbaarheid, interpreteerbaarheid en gebruikservaringen.

Evaluatie van ongewenste effecten

Veel QI-ontwikkelingsmethoden focussen uitsluitend op mogelijke voordelen van QI. CEQUID erkent expliciet dat QI ook negatieve effecten kunnen veroorzaken. Door een systematische risicoanalyse op te nemen, probeert CEQUID die effecten zichtbaar te maken voordat brede implementatie plaatsvindt.

Aandachtspunten van de CEQUID-methodologie

Hoge methodologische eisen

De uitgebreide aanpak van CEQUID verhoogt de methodologische robuustheid, maar brengt ook nadelen met zich mee. De methodologie vereist tijd, expertise, middelen en organisatorische ondersteuning. De volledige toepassing van de CEQUID-methodologie kan daardoor uitdagend zijn.

Afhankelijkheid van beschikbare richtlijnen

CEQUID vertrekt vanuit het principe dat QI zoveel mogelijk gebaseerd moeten zijn op richtlijnen. Die keuze verhoogt de wetenschappelijke onderbouwing, maar creëert ook een afhankelijkheid van beschikbare richtlijnen. Voor sommige zorgdomeinen ontbreken hoogwaardige richtlijnen of zijn aanbevelingen voornamelijk gebaseerd op consensus. Daardoor kunnen relevante kwaliteitsaspecten ondervertegenwoordigd blijven.

Consensus blijft subjectief

Ondanks het gebruik van gestructureerde consensusmethoden, blijft professionele oordeelsvorming een belangrijk onderdeel van QI-ontwikkeling. De uiteindelijke QI-set wordt mede beïnvloed door panelsamenstelling, expertise van de deelnemers en interpretatie van de evidentie. Volledige objectiviteit is daarom niet haalbaar.

Contextgebondenheid

QI die ontwikkeld zijn binnen de Belgische eerstelijnszorg zijn niet noodzakelijk rechtstreeks overdraagbaar naar andere settings of zorgsystemen²⁴.

8. Besluit

De ambitie van CEQUID is niet om bestaande methodologieën te vervangen, maar om hun sterktes te combineren binnen één geïntegreerd ontwikkelproces. Door expliciet aandacht te besteden aan wetenschappelijke evidentie, methodologische transparantie, stakeholderbetrokkenheid, patiëntenparticipatie, praktijktoets en ongewenste effecten beoogt CEQUID een kader aan te bieden dat zowel wetenschappelijk robuust als praktisch toepasbaar is.

De systematische review van Langendam et al. toonde aan dat QI-ontwikkeling internationaal gekenmerkt wordt door aanzienlijke methodologische heterogeniteit⁶. CEQUID beoogt niet om die variatie weg te nemen, maar wel om QI-ontwikkeling te structureren via een transparante en expliciet gedocumenteerde methodologie.

De uiteindelijke doelstelling is niet om zoveel mogelijk QI te ontwikkelen, maar wel om QI te ontwikkelen waarin gebruikers vertrouwen mogen hebben, zowel in de QI zelf als in het proces dat eraan voorafging.

9. Referenties

1. Donabedian A. Evaluating the quality of medical care. 1966. *Milbank Q.* 1966;83(4):691-729. doi:10.1111/j.1468-0009.2005.00397.x
2. Campbell SM, Braspenning J, Hutchinson A, Marshall M. Research methods used in developing and applying quality indicators in primary care. *Qual Saf Health Care.* 2002;11(4):358-364. doi:10.1136/qhc.11.4.358
3. Dudley L, Mamdoo P, Naidoo S, Muzigaba M. Towards a harmonised framework for developing quality of care indicators for global health: a scoping review of existing conceptual and methodological practices. *BMJ Health Care Inform.* 2022;29(1):e100469. doi:10.1136/bmjhci-2021-100469
4. Kötter T, Blozik E, Scherer M. Methods for the guideline-based development of quality indicators--a systematic review. *Implement Sci IS.* 2012;7:21. doi:10.1186/1748-5908-7-21
5. Becker M, Breuing J, Nothacker M, et al. Guideline-based quality indicators-a systematic comparison of German and international clinical practice guidelines. *Implement Sci IS.* 2019;14(1):71. doi:10.1186/s13012-019-0918-y
6. Langendam MW, Piggott T, Nothacker M, et al. Approaches of integrating the development of guidelines and quality indicators: a systematic review. *BMC Health Serv Res.* 2020;20:875. doi:10.1186/s12913-020-05665-w
7. Nothacker M, Bolster M, Steudtner M, et al. International experiences in the development and implementation of guideline-based quality indicators: a qualitative study. *BMJ Open.* 2021;11(1):e039770. doi:10.1136/bmjopen-2020-039770
8. Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *CMAJ Can Med Assoc J J Assoc Medicales Can.* 2010;182(18):E839-842. doi:10.1503/cmaj.090449
9. de Koning J, Smulders A, Klazinga NS. Appraisal of Indicators through Research and Evaluation (AIRE) Instrument. *Amsterdam: Academic Medical Center.* Published online 2006.
10. Jandhyala R. Delphi, non-RAND modified Delphi, RAND/UCLA appropriateness method and a novel group awareness and consensus methodology for consensus measurement: a systematic literature review. *Curr Med Res Opin.* 2020;36(11):1873-1887. doi:10.1080/03007995.2020.1816946
11. Fitch K, ed. *The Rand/UCLA Appropriateness Method User's Manual.* Rand; 2001.
12. Plessers M, Ghekiere A, De Wachter ; Dirk, Deneckere S, Ramaekers D. Het ontwikkelen van evidence-based indicatoren van kwaliteit van zorg in Vlaanderen: een methodologie [Internet]. 2019. Available from: www.zorgkwaliteit.be.

13. Vlayen, Joan, Van De Water, Geert, Camberlin, Cécile, et al. *Klinische kwaliteitsindicatoren. Objective Elements – Communication (OEC)*. Federaal Kenniscentrum voor de Gezondheidszorg (KCE); 2006. doi:DOI : 10.57598/R41A
14. Braspenning J, Hermens R, Calsbeek H, Campbell S, van der Wees P, Grol R. Indicators for Quality and Safety of Care. In: *Improving Patient Care*. John Wiley & Sons, Ltd; 2020:131-154. doi:10.1002/9781119488620.ch7
15. Mant J. Process versus outcome indicators in the assessment of quality of health care. *Int J Qual Health Care J Int Soc Qual Health Care*. 2001;13(6):475-480. doi:10.1093/intqhc/13.6.475
16. Abel G, Saunders CL, Mendonca SC, Gildea C, McPhail S, Lyratzopoulos G. Variation and statistical reliability of publicly reported primary care diagnostic activity indicators for cancer: a cross-sectional ecological study of routine data. *BMJ Qual Saf*. 2018;27(1):21-30. doi:10.1136/bmjqs-2017-006607
17. Institute of Medicine (US) Committee on Quality of Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century*. National Academies Press (US); 2001. Accessed May 28, 2026. <http://www.ncbi.nlm.nih.gov/books/NBK222274/>
18. Smith P. On the unintended consequences of publishing performance data in the public sector. *Int J Public Adm*. 1995;18(2-3):277-310. doi:10.1080/01900699508525011
19. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737-745. doi:10.1016/j.jclinepi.2010.02.006
20. Grol R, Wensing M, Grimshaw J. *Improving Patient Care: The Implementation of Change in Health Care*. 3rd ed. John Wiley & Sons, Incorporated; 2020.
21. Grypdonck L, Aertgeerts B, Luyten F, et al. Development of quality indicators for an integrated approach of knee osteoarthritis. *J Rheumatol*. 2014;41(6):1155-1162. doi:10.3899/jrheum.130680
22. Bergholtz J, Wolf A, Crine V, Cleeve H, Santana MJ, Björkman I. Patient and public involvement in healthcare: a systematic mapping review of systematic reviews – identification of current research and possible directions for future research. *BMJ Open*. 2024;14(9):e083215. doi:10.1136/bmjopen-2023-083215
23. Kötter T, Schaefer FA, Scherer M, Blozik E. Involving patients in quality indicator development – a systematic review. *Patient Prefer Adherence*. 2013;7:259-268. doi:10.2147/PPA.S39803

24. Nothacker M, Stokes T, Shaw B, et al. Reporting standards for guideline-based performance measures. *Implement Sci IS*. 2016;11:6. doi:10.1186/s13012-015-0369-z
25. Ivers N, Yogasingam S, Lacroix M, et al. Audit and feedback: effects on professional practice. *Cochrane Database Syst Rev*. 2025;3(3):CD000259. doi:10.1002/14651858.CD000259.pub4
26. Jamtvedt G, Flottorp S, Ivers N. Audit and Feedback as a Quality Strategy. In: *Improving Healthcare Quality in Europe: Characteristics, Effectiveness and Implementation of Different Strategies [Internet]*. European Observatory on Health Systems and Policies; 2019. Accessed May 28, 2026. <https://www.ncbi.nlm.nih.gov/books/NBK549284/>
27. Grade. <https://www.gradeworkinggroup.org/>
28. Blozik E, Nothacker M, Bunk T, Szecsenyi J, Ollenschläger G, Scherer M. Simultaneous development of guidelines and quality indicators -- how do guideline groups act? A worldwide survey. *Int J Health Care Qual Assur*. 2012;25(8):712-729. doi:10.1108/09526861211270659
29. Grant S, Khodyakov D. Proposal for a critical appraisal tool for studies using the Delphi method (DCAT). *BMJ*. 2025;391:e084509. doi:10.1136/bmj-2025-084509
30. Mainz J. Defining and classifying clinical indicators for quality improvement. *Int J Qual Health Care J Int Soc Qual Health Care*. 2003;15(6):523-530. doi:10.1093/intqhc/mzg081
31. Brook RH, Chassin MR, Fink A, Solomon DH, Kosecoff J, Park RE. A method for the detailed assessment of the appropriateness of medical technologies. *Int J Technol Assess Health Care*. 1986;2(1):53-63. doi:10.1017/s0266462300002774
32. Van den Bulck SA, Vankrunkelsven P, Goderis G, et al. Development of quality indicators for type 2 diabetes, extractable from the electronic health record of the general physician. A rand-modified Delphi method. *Prim Care Diabetes*. 2020;14(1):75-84. doi:10.1016/j.pcd.2019.05.002
33. National Quality Forum. Measure Evaluation Criteria. Published online 2012.
34. Campbell SM, Kontopantelis E, Hannon K, Burke M, Barber A, Lester HE. Framework and indicator testing protocol for developing and piloting quality indicators for the UK quality and outcomes framework. *BMC Fam Pract*. 2011;12:85. doi:10.1186/1471-2296-12-85
35. Orueta JF, García-Alvarez A, Grandes G, Nuño-Solinís R. The Origin of Variation in Primary Care Process and Outcome Indicators. *Medicine (Baltimore)*. 2015;94(31):e1314. doi:10.1097/MD.0000000000001314

36. Quality in the new health system – maintaining and improving quality from April 2013. GOV.UK. Accessed June 3, 2026. <https://www.gov.uk/government/publications/quality-in-the-new-health-system-maintaining-and-improving-quality-from-april-2013>
37. Mannion R, Braithwaite J. Unintended consequences of performance measurement in healthcare: 20 salutary lessons from the English National Health Service. *Intern Med J.* 2012;42(5):569–574. doi:10.1111/j.1445-5994.2012.02766.x
38. Romano P, Hussey PS, Ritley D. Selecting Quality and Resource Use Measures: A Decision Guide for Community Quality Collaboratives. Published online January 1, 2010. Accessed June 2, 2026. https://www.rand.org/pubs/external_publications/EP201000121.html