
**De ontwikkeling van een leidraad voor effectevaluatie
in het Vlaamse preventieve gezondheidsbeleid**

Eva Rens
Elien Van Hout
Evelien Coppens

LUCAS KU Leuven

Steunpunt Welzijn, Volksgezondheid en Gezin
Minderbroedersstraat 8 – B-3000 Leuven
Tel 0032 16 37 34 32
E-mail: swvg@kuleuven.be
Website: <http://www.steunpuntwvg.be>

**STEUNPUNT
WELZIJN
VOLKSGEZONDHEID &
GEZIN**

Publicatie nr. 2026/07
SWVG-Rapport 64

Titel: De ontwikkeling van een leidraad voor effectevaluatie in het Vlaamse preventieve gezondheidsbeleid

Onderzoekers: dr. Eva Rens, dr. Elien Van Hout

Promotor: dr. Evelien Coppens

LUCAS KU Leuven

Dit rapport kwam tot stand met de steun van de Vlaamse Overheid. In deze tekst komen onderzoeksresultaten van de auteur(s) naar voor en niet die van de Vlaamse Overheid. De Vlaamse Overheid kan niet aansprakelijk gesteld worden voor het gebruik dat kan worden gemaakt van de meegedeelde gegevens.

Niets uit deze uitgave mag worden verveelvoudigd en/of openbaar gemaakt zonder uitdrukkelijk te verwijzen naar de bron.

No material may be made public without an explicit reference to the source.

Het consortium SWVG

<i>Deelnemende instellingen - Dagelijks bestuur</i>		
KU Leuven		
· LUCAS, Centrum voor zorgonderzoek en consultancy		Prof. dr. Anja Declercq
· Instituut voor Sociaal Recht		Prof. dr. Koen Hermans
· HIVA - Onderzoeksinstituut voor Arbeid en Samenleving		Prof. dr. Johan Put
		Prof. dr. Tine Van Regenmortel
UGent		
· Vakgroep Bestuurskunde en Publiek Management		Prof. dr. Bram Verschuere
· Vakgroep Volksgezondheid en Eerstelijnszorg		Prof. dr. Katrien Vanthomme
· Vakgroep Volksgezondheid en Eerstelijnszorg		Prof. dr. Jeroen Trybou
Vrije Universiteit Brussel		
· Faculteit Psychologie en Educatiewetenschappen		Prof. dr. Johan Vanderfaeillie
Universiteit Antwerpen		
· Faculteit Sociale Wetenschappen		Prof. dr. Koen Decancq
<i>Partnerinstellingen</i>		
Universiteit Hasselt		
· Faculteit Geneeskunde en Levenswetenschappen		Prof. dr. Wim Pinxten
Thomas More Hogeschool		
· Unit Mens & Welzijn		Dr. Peter De Graef
Arteveldehogeschool		
· Sociaal Werk		
Hogeschool Gent		
· Onderzoekscentrum 360° Zorg en Welzijn		
UC Leuven-Limburg		
· Research & Expertise - Expertisecentrum Health Innovation		
Karel De Grote hogeschool		
· Onderzoek en maatschappelijke dienstverlening		
Hogeschool PXL		
· Expertisecentrum PXL Social Work Research		
Odisee co-hogeschool		
· Studiegebied Sociaal-Agogisch Werk		
KeFor		
· Kenniscentrum Forensisch Psychiatrische Zorg		
<i>Coördinatie en secretariaat</i>		
Coördinator SWVG		Dr. Kathleen De Cuyper
Administratieve ondersteuning		Manuela Schröder
		Lut Van Hoof

Rapport 64

April 2026

De ontwikkeling van een leidraad voor effectevaluatie in het Vlaamse preventieve gezondheidsbeleid

Onderzoekers: dr. Eva Rens, dr. Elien Van Hout

Promotor: dr. Evelien Coppens

LUCAS KU Leuven

Samenvatting

Situering

Dit onderzoeksproject vertrekt vanuit de vaststelling dat **effectevaluaties binnen de Vlaamse preventieve gezondheidszorg slechts in beperkte mate plaatsvinden**. In de praktijk ligt de nadruk vooral op procesindicatoren, zoals bereik en tevredenheid, terwijl effectmetingen vaak gebaseerd zijn op onderzoeksmethoden met beperkte bewijskracht. Het vinden van een evenwicht tussen wetenschappelijke validiteit en praktische haalbaarheid vormt een terugkerende uitdaging. Hierdoor is het moeilijk om bestaande interventies te beoordelen, beleidsmatig te verantwoorden of gericht bij te sturen.

Doelstelling onderzoekproject

In dit onderzoeksproject werd gewerkt aan de ontwikkeling van een leidraad voor effectevaluatie in de Vlaamse preventiesector. De leidraad heeft als ambitie om kwaliteitsvolle effectevaluatie van interventies structureel te verankeren binnen het preventieve gezondheidsbeleid. Ze richt zich op universele en selectieve gezondheidsbevorderende en primaire preventieve interventies die de gezondheid van de bevolking versterken door in te spelen op gezondheidsdeterminanten. De leidraad is generiek uitgewerkt en daardoor toepasbaar op de brede diversiteit aan interventies binnen partnerorganisaties en organisaties met terreinwerking.

De leidraad biedt houvast voor zowel beleidsmakers als organisaties in de praktijk. Beleidsmakers krijgen ondersteuning bij het creëren van realistische en passende randvoorwaarden voor evaluatie binnen de preventiesector, terwijl preventieorganisaties worden begeleid om de effectiviteit van hun interventies op een onderbouwde en haalbare manier te onderzoeken.

Onderzoeksmethode

Het onderzoeksproject werd uitgevoerd door LUCAS KU Leuven, in opdracht van het Departement Zorg binnen het Steunpunt Welzijn, Volksgezondheid en Gezin. Methodologisch steunt het op een combinatie van literatuuronderzoek en consultatie van stakeholders uit de preventiesector, beleid en onderzoek.

De literatuurstudie was zowel breed als doelgericht en omvatte wetenschappelijke publicaties en grijze literatuur. Ze had tot doel inzicht te verwerven in verschillende typen evaluatie binnen interventie-onderzoek, in de randvoorwaarden voor kwaliteitsvolle evaluatie van gezondheidsbevordering en in bruikbare evaluatiekaders en methodologische benaderingen. Daarnaast werd aandacht besteed aan prioritering van evaluaties, samenwerking en rolverdeling tussen betrokken actoren, en de aard van passend en bruikbaar bewijs voor preventie, rekening houdend met het contextuele en complexe karakter van gezondheidsbevorderende interventies. Ook benaderingen voor beleidsimpactevaluatie werden meegenomen.

Verder werden twee werkgroepen opgericht: één met actoren uit de preventiesector (o.a. partner-organisaties en organisaties met terreinwerking) en één met actoren uit onderzoek en beleid (o.a. academici en Sciensano). Beide groepen kwamen twee maal samen. De eerste bijeenkomst focuste op randvoorwaarden en hefboomen voor kwaliteitsvolle effectevaluatie, samenwerking en rolverdeling, evaluatiekaders en onderzoeksmethoden. De tweede bijeenkomst richtte zich op de vraag wanneer effectevaluatie aangewezen is, hoe effectevaluatie kan worden gefinancierd, hoe resultaten beter benut kunnen worden en hoe beleidsimpact zichtbaar kan worden gemaakt.

Inhoud van het rapport

Dit rapport, dat de leidraad ondersteunt, heeft een logische opbouw waarbij elk hoofdstuk voortbouwt op de inzichten uit de voorgaande delen. Samen bieden ze een volledig en samenhangend beeld van evaluatie binnen gezondheidsbevordering:

- **Hoofdstuk 1** situeert het onderzoeksproject, formuleert de probleemstelling en geeft een korte toelichting bij het Vlaamse preventieve gezondheidsbeleid.
- **Hoofdstuk 2** beschrijft de gebruikte methodologie en maakt duidelijk hoe de resultaten en aanbevelingen tot stand kwamen.
- **Hoofdstuk 3** gaat dieper in op de fundamentele concepten van interventies en evaluatie binnen gezondheidsbevordering.
- **Hoofdstuk 4** bespreekt de randvoorwaarden die nodig zijn voor kwaliteitsvolle evaluatie.
- **Hoofdstuk 5** belicht het belang van evaluatiekaders en geeft een overzicht van relevante en bruikbare kaders voor interventie-evaluatie.
- **Hoofdstuk 6** behandelt de uitvoering van effectevaluaties in de praktijk, inclusief de keuze van onderzoeksdesigns en indicatoren.
- **Hoofdstuk 7** verduidelijkt wat passend bewijs is voor preventie.
- **Hoofdstuk 8** richt zich op de impactevaluatie van het preventieve gezondheidsbeleid op systeemniveau.
- **Hoofdstuk 9** bespreekt samenwerking en rolverdeling binnen evaluatieprocessen.

- **Hoofdstuk 10** beschrijft manieren om te bepalen welke interventies prioriteit verdienen bij evaluatie.
- **Hoofdstuk 11** bespreekt hoe evaluatiebevindingen optimaal gebruikt kunnen worden voor beleidsontwikkeling en praktijkverbetering.
- **Hoofdstuk 12** geeft een syntheses van de twee bijeenkomsten met de werkgroepen, waarin inzichten uit de literatuur werden afgetoetst aan de praktijk.
- **Hoofdstuk 13** presenteert de volledige leidraad die gebaseerd is op dit onderzoeksrapport.

Leidraad

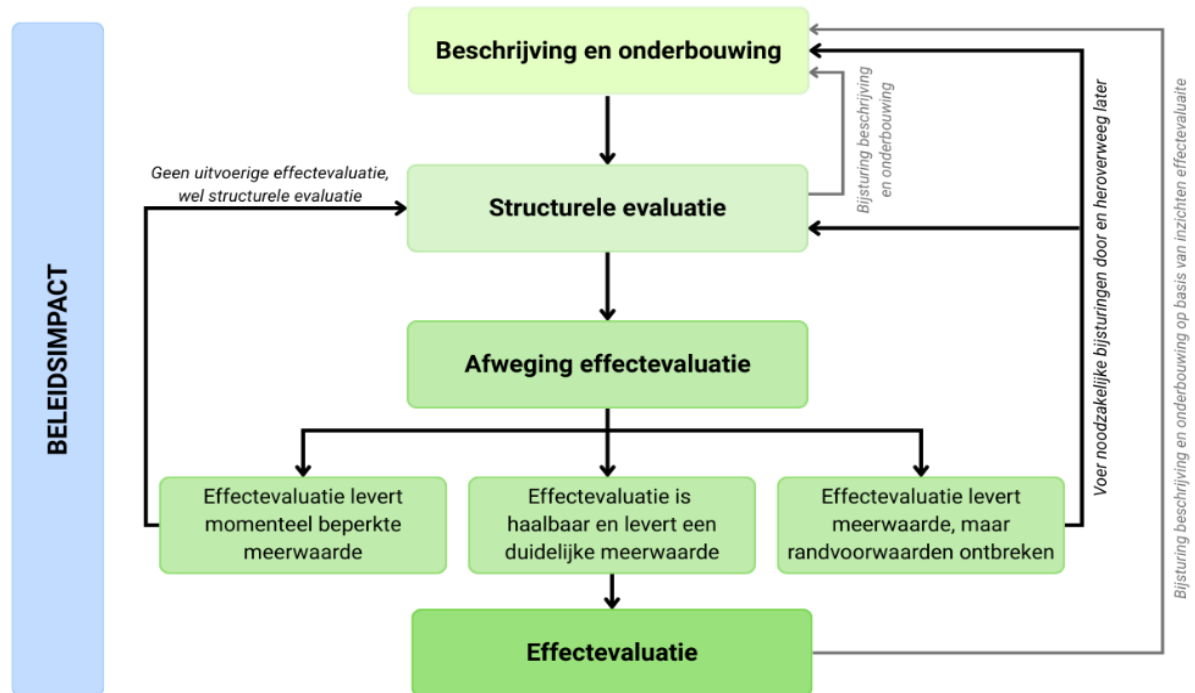
De leidraad bundelt **internationale literatuur** en **input uit twee werkgroepen** met praktijk, beleid en onderzoek tot een **realistisch en haalbaar kader** voor organisaties en beleidsmakers. Ze helpt onderbouwde keuzes maken over **wanneer, waarom en hoe** te evalueren, met aandacht voor **proportionaliteit, kwaliteit** en **leren**.

Effectevaluatie staat in de leidraad **niet op zichzelf**, maar wordt geïntegreerd met **proces- en beleids-impactevaluatie**. Eerst wordt gekeken of een interventie **uitvoerbaar** is en **zoals bedoeld** wordt uitgevoerd; daarna of en in welke mate effecten **aan de interventie** kunnen worden toegeschreven. Het kader is ontwikkeld voor toepassing op **gezondheidsbevorderende en primaire preventie-interventies** (universeel of selectief) in publieke settings.

De leidraad berust op **vijf kernprincipes**:

- **Participatie**: actieve betrokkenheid van doelgroep, intermediairs en stakeholders, met aandacht voor machtsverhoudingen.
- **Equity**: expliciete aandacht voor gezondheidsongelijkheid, inclusief het bereiken en analyseren van kwetsbare groepen.
- **Transparantie**: inzichtelijke en controleerbare verslaggeving van proces, keuzes en beperkingen.
- **Gebruik van inzichten**: evaluatie als basis voor leren, verbeteren en beleidsontwikkeling.
- **Systeemdenken**: interventies plaatsen binnen het bredere complexe systeem van actoren, context en middelen.

De leidraad stelt een iteratieve aanpak voor, opgebouwd uit vier fasen die logisch op elkaar voortbouwen en ruimte laten voor voortdurende bijsturing, zoals weergegeven in hierna volgende figuur.



Fase 1: Beschrijving en onderbouwing van de interventie

In deze eerste fase wordt de basis gelegd voor evaluatie door (1) de interventie en haar onderbouwing helder en narratief te beschrijven, (2) een Theory of Change uit te werken en (3) een Logic Model op te stellen. Het doel is om bij alle betrokkenen een gedeeld en consistent begrip te creëren van wat de interventie omvat, waarom ze nodig is, hoe ze verondersteld wordt te werken en in welke context ze plaatsvindt. Deze fase vormt het startpunt voor verdere evaluatiestappen en wordt gaandeweg verfijnd op basis van nieuwe inzichten.

Fase 2: Structurele evaluatie en bijsturing

In deze fase wordt de uitvoering van de interventie systematisch opgevolgd. Procesevaluaties brengen de kwaliteit, uitvoerbaarheid en aanvaardbaarheid van de interventie in kaart en ondersteunen gerichte bijsturing. Veranderingen in uitkomsten kunnen hierbij pragmatische worden onderzocht, zonder dat uitspraken worden gedaan over causale effecten.

Fase 3: Beoordelen of een robuuste effectevaluatie aan de orde is

Deze fase bestaat uit het afwegen of een meer robuuste effectevaluatie haalbaar is en een duidelijke meerwaarde biedt, gezien de noodzakelijke investering in tijd, middelen en expertise.

Fase 4: Effectevaluatie

Wanneer in fase 3 wordt beslist dat een effectevaluatie aangewezen is, wordt in deze fase een methodologisch onderbouwd onderzoeksopzet uitgewerkt om te onderzoeken of en in welke mate de interventie effecten veroorzaakt. Het evaluatieplan beschrijft minimaal de volgende elementen:

- De nood aan een effectevaluatie
- Noodzakelijke partnerschappen en wijze van samenwerking
- De evaluatiedoelstellingen en onderzoeksvragen
- Het eigenlijke onderzoeksopzet

- De wijze van dataverzameling en -verwerking
- Ethische en juridische aspecten die van kracht zijn
- De wijze waarop de resultaten gebruikt en gedissemineerd zullen worden
- De planning en nodige middelen voor de effectevaluatie

Aanbevelingen

Het rapport formuleert tot slot **twalf beleidsaanbevelingen** die benadrukken dat sterke effectevaluatie niet alleen afhankelijk is van methodologie, maar ook van **realistische beleidsverwachtingen, structurele ondersteuning en een sectorbrede leer- en evaluatiecultuur**:

- 1) **Richt een onafhankelijk expertiseforum op:** Creëer een onafhankelijk forum waar beleid, praktijk en onderzoek samenkomen, en waar een groep experts zich engageert om (plannen voor) effectevaluaties in de preventiesector methodologisch te ondersteunen, beoordelen en prioriteren.
- 2) **Faciliteer structurele samenwerking tussen de preventiesector en onderzoek:** Versterk de samenwerking door netwerking en matching te organiseren op basis van inhoudelijke en methodologische noden.
- 3) **Stimuleer kennisdeling en ontsluiting:** Zorg voor een actieve aanpak om evaluatiekennis te delen, onder meer via publicatie van evaluaties, lerende netwerken en een duidelijk gemandateerde coördinatie van kennisdeling.
- 4) **Ontwikkel een gedeelde onderzoeksagenda:** Werk samen met de preventiesector en onderzoekers aan een agenda die kennishiaten benoemt en beleidsrelevante thema's en interventies prioriteert. Veranker deze prioriteiten in de plannen en beheersovereenkomsten van organisaties.
- 5) **Voorzie realistische basisfinanciering:** De overheid moet voldoende basisfinanciering voorzien, en organisaties moeten deze middelen doelgericht inzetten voor degelijk evaluatieonderzoek.
- 6) **Voorzie voldoende tijd en realistische verwachtingen:** Erken dat effecten van preventieve interventies vaak pas op middellange of lange termijn zichtbaar worden, en beschouw tussentijdse veranderingen in determinanten als betekenisvol.
- 7) **Gebruik gezondheidsdoelstellingen als ankerpunt voor beleidsimpact:** Beleidsimpactevaluatie richt zich op veranderingen op systeem- en populatieniveau, die het resultaat zijn van een complex systeem van actoren, acties en andere maatschappelijke ontwikkelingen. De gezondheidsdoelstellingen dienen als referentiepunt voor beleidsimpact, maar deze mogen niet zomaar beschouwd worden als het directe effect van een bepaalde interventie. Interventie-evaluaties kunnen wel helpen bij het plausibel verklaren van waargenomen veranderingen op populatieniveau.
- 8) **Ontwikkel een Theory of Change op sectorniveau:** Een sectorale of thematische Theory of Change kan verduidelijken hoe beleid via diverse interventies en hefbomen bijdraagt aan de gezondheidsdoelstellingen. In tegenstelling tot een interventiegerichte Theory of Change richt deze Theory of Change zich meer op systeemmechanismen, structurele randvoorwaarden en duurzame impact.
- 9) **Werk toe naar collectieve impactmeting:** Ontwikkel, per thema, gedeelde doelen en kern-indicatoren om de gezamenlijke bijdrage van actoren aan maatschappelijke doelen zichtbaar te maken. Dit vereist intensieve afstemming en coördinatie, waarbij pilots op thema- of sectorniveau in een afgebakende context een eerste stap kunnen zijn.
- 10) **Voorzie een groeipad voor onderbouwing van interventies:** Hanteer een realistische visie op effectiviteitsbewijs, waarbij evidentie geleidelijk kan worden opgebouwd en bestaande en overdraagbare evidentie ten volle wordt benut. In vroege fasen volstaan haalbare evaluaties met pre-experimentele designs om vroege veranderingen te detecteren en aannames uit de Theory of Change te toetsen. Zodra mogelijk wordt (quasi-)experimenteel onderzoek met tegenfeitelijke vergelijking aanbevolen.

- 11) **Ontwikkel een beoordelingskader voor de onderbouwing van preventieve interventies:** Werk naar een gelaagd kader met verschillende niveaus van onderbouwing (bv. goed beschreven, theoretisch onderbouwd, aanwijzingen voor effectiviteit, effectief), met duidelijkheid over welk bewijs per niveau passend is. Dit laat gerichte ondersteuning van interventies toe en verduidelijkt welke vervolgstappen nodig zijn.
- 12) **Versterk effectieve en veelbelovende interventies:** Ontwikkel een beleidssysteem dat veelbelovende interventies ondersteunt en effectieve interventies verduurzaamt. Laat tegelijk ruimte voor innovatieve methodieken en durf interventies stopzetten bij schadelijke effecten of het herhaaldelijk uitblijven van effecten.

Tot slot

De leidraad fungeert als richtinggevend kader en stip op de horizon, en vormt daarmee een startpunt voor verandering en verdere professionalisering van evaluatie in de preventiesector. Afstemming met de preventiesector blijft noodzakelijk om structurele verankering van de leidraad mogelijk te maken. Een langetermijnvisie, gericht op leren, bijsturen en gezamenlijke expertiseopbouw, wordt gezien als essentieel om evaluatiepraktijken te versterken en zo bij te dragen aan duurzame gezondheidswinst.

Inhoud

Hoofdstuk 1	Situering en probleemstelling	15
1	Probleemstelling en doelstellingen van het onderzoeksproject	15
2	Vlaamse cijfers inzake preventieve gezondheidszorg en gezondheidsdeterminanten in een notendop	16
3	Het Vlaamse preventieve gezondheidsbeleid	17
3.1	Partnerorganisaties	18
3.2	Organisaties met terreinwerking en Gezondheidsmakers	19
3.3	Intermediairs	20
Hoofdstuk 2	Methode	21
1	Literatuurstudie	21
2	Consultatie van stakeholders uit de preventiesector, beleid en onderzoek	21
3	Opbouw van het rapport	22
Hoofdstuk 3	Fundamentele concepten van interventie en evaluatie in gezondheidsbevordering	23
1	Gezondheidsbevordering en preventie	23
1.1	Klassieke indeling van preventie	23
1.2	Evolutie naar bredere gezondheidsbevordering	24
1.3	Focus leidraad	25
2	Interventie	25
2.1	Enkelvoudige versus complexe interventies	26
2.2	Multiple health behaviour change interventies	26
2.3	Typologieën van gezondheidsbevorderende interventies	27
3	Evaluatie	28
3.1	Types van evaluatie	28
3.2	Evaluatiebenaderingen	31
4	Effectiviteit	31
5	Implementatie	33
6	Impact	34
Hoofdstuk 4	Randvoorwaarden voor de evaluatie van gezondheidsbevordering	35
1	Evaluatie van gezondheidsbevordering: 'een verplichte last'?	35
2	Governance	37
3	Evaluatiecapaciteit	38
3.1	Individuele evaluatiecapaciteit: vereiste expertise van evaluatoren	38
3.2	Organisatorische evaluatiecapaciteit	38
3.3	Evaluatiecapaciteit op macroniveau: Lessen uit het evaluatiebeleid in Canada	40
4	Datakwaliteit - en management	40

5	Privacy en ethische aspecten	41
6	Principes voor systeemdenken doorheen het ganse evaluatieproces	43
7	Participatie en co-creatie	43
8	Aandacht voor ongelijkheden	44
9	Samenvattend overzicht van noodzakelijke randvoorwaarden voor evaluatie	44
Hoofdstuk 5 Evaluatiekaders voor gezondheidsbevorderende interventies		47
1	Stappen en principes in programmaevaluatie (<i>CDC Program Evaluation Framework 2024</i>)	47
2	Theory of Change en Logic Model als strategisch en operationeel fundament voor interventieplanning- en evaluatie	49
2.1	Theory of change	49
2.2	Logic Model	51
3	RE-AIM voor proces- en effectevaluatie	53
3.1	Practical, Robust Implementation and Sustainability Model (PRISM)	55
4	MRC-framework	57
4.1	Procesevaluatie volgens het MRC-framework	58
5	Realist Evaluation	61
6	Contribution Analysis	63
7	Evaluatiekaders voor implementatieonderzoek	65
7.1	Consolidated Framework for Implementation Research	65
7.2	Proctor's Implementation Outcomes	67
7.3	Implementation Research Logic Model	67
8	Uitgebreide kaders voor het ontwikkelen, plannen en evalueren van interventies	70
8.1	PRECEDE-PROCEED Model	70
8.2	Intervention Mapping	71
Hoofdstuk 6 Uitvoering van effectevaluatie van interventies		73
1	Van piloottest tot impactevaluatie bij gezondheidsbevorderende interventies	73
1.1	Van werkzaamheid naar effectiviteit in preventieonderzoek	73
1.2	Kleine studies, belangrijke gevolgen: misstappen bij het overgaan tot meer grootschalige of complexe evaluatiestudies op basis van resultaten van piloottesten	76
1.3	Wanneer is effectevaluatie van een implementatie nodig?	78
2	Kwantitatieve onderzoeksdesigns voor effectevaluatie	81
2.1	Pre-experimentele designs	82
2.2	Quasi-experimentele designs	82
2.3	Experimentele gerandomiseerde designs (RCT's)	86
2.4	Samenvattend overzicht van de verschillende onderzoeksdesigns	90
2.5	Beslissingsmatrix WHO-Europe voor de selectie van een geschikt onderzoeksdesign	91
3	Kwalitatieve en gemengde methoden	93

4	Uitkomsten en indicatoren van interventies	94
4.1	Selectie van uitkomsten en uitkomstmaten	95
4.2	Types indicatoren	96
Hoofdstuk 7	Passend bewijs voor preventie	97
1	Evidentiestandaarden in de preventiewetenschap	97
2	Richtlijn passend bewijs door Nederlandse Adviescommissie Richtlijn Passend Bewijs Preventie	98
2.1	Overtuigend bewijs van effectiviteit	99
2.2	Toepasbaarheid en contextgevoeligheid	100
2.3	Langetermijneffecten en intermediaire uitkomsten	100
2.4	Andere aanbevelingen	100
3	ADAPT-richtlijn: aanpassen en evalueren van evidence-based interventies in nieuwe contexten	101
4	Aanbevelingen van de onderzoekers inzake passend bewijs voor de Vlaamse preventiesector	104
Hoofdstuk 8	Impactevaluatie van het preventieve gezondheidsbeleid	105
1	Beleidsimpactevaluatie vraagt om een diepgaande analyse	105
2	Randvoorwaarden voor beleidsimpactevaluatie	106
3	Theory of change op sectorniveau	106
4	Indicatoren voor beleidsimpactevaluatie	107
4.1	Indicatoren voor beleidsimpactevaluatie in Vlaanderen	108
5	Rekening houden met beleidsimpact bij de evaluatie van interventies	108
6	Zicht op collectieve impact	109
6.1	Gedeelde uitkomstenkaders en meetsystemen voor organisaties of programma's	110
Hoofdstuk 9	Samenwerking en rollen bij de evaluatie van interventies	113
1	Samenwerkingsmodellen tussen onderzoeks- en beleidsteams (handboek Impact Evaluation in Practice)	113
2	Het samenwerkingsmodel tussen preventieorganisaties en evaluatieonderzoekers volgens het Europees Preventiehandboek (2019)	114
Hoofdstuk 10	Prioritering van evaluaties	117
1	Criteria voor het prioriteren van evaluaties	117
2	Methode: versnelde James Lind Alliance (JLA) aanpak	119
3	Methode: Systematic Screening and Assessment (SSA)	120
4	Methode: strategisch evaluatieportfolio	121
Hoofdstuk 11	Gebruik van evaluatiebevindingen	123
1	Toepassing van onderzoek in beleidsvorming	123

2	De beoordeling en erkenning van effectieve interventies door <i>Samenwerkingsverband Erkenning van Interventies</i> in Nederland	124
2.1	Toelatingscriteria	124
2.2	Erkenningsniveaus	125
3	Programme Budgeting and Marginal Analysis (PBMA)	128
Hoofdstuk 12 Werkgroepen met stakeholders: synthese		131
1	Bijeenkomst 1: noden, randvoorwaarden en rolverdeling	131
1.1	Werkgroep met actoren uit de preventiesector	131
1.2	Werkgroep actoren uit onderzoek en beleid	135
2	Bijeenkomst 2	138
2.1	Werkgroep met actoren uit de preventiesector	139
2.2	Werkgroep met actoren uit onderzoek en beleid	142
Hoofdstuk 13 Leidraad voor effectevaluatie in de Vlaamse preventiesector		147
1	Waarom deze leidraad?	147
2	Evaluatie als geïntegreerd geheel: effect, proces en impact	147
3	Op welke interventies is de leidraad van toepassing?	149
4	Hoe kwam deze leidraad tot stand?	149
5	Kernprincipes van iedere evaluatie in de preventiesector	150
6	Een weloverwogen en transparante inzet van werkmiddelen voor kwaliteitsvolle evaluatie	150
6.1	Scenario 1: Interne inzet (intern VTE inschakelen)	151
6.2	Scenario 2: Inschakelen van externe expertise (bv. onderzoekers uit onderzoeksinstituten)	152
6.3	Scenario 3: Samenwerking tussen interne en externe evaluatoren	153
7	Een evaluatieproces in vier fasen	153
7.1	Fase 1: Beschrijving en onderbouwing van de interventie	155
7.2	Fase 2: Structurele evaluatie en bijsturing	159
7.3	Fase 3: Beoordeel of grondige effectevaluatie zinvol en haalbaar is	162
7.4	Fase 4: Effectevaluatie volgens een duidelijk evaluatieplan	165
8	Aanbevelingen voor het beleid	175
8.1	Versterk de wisselwerking tussen beleid, praktijk en onderzoek	175
8.2	Kwaliteitsvolle effectevaluaties vragen voldoende middelen en tijd	176
8.3	Evalueer de beleidsimpact op systeem- en populatieniveau	177
8.4	Evolueren naar beleid dat effectieve interventies identificeert en versterkt	178
9	Slotwoord	180
10	Bijlagen in leidraad	181
11	Referenties leidraad	199
Bijlagen onderzoeksrapport		201
Referenties		219

Figuren

Figuur 1	Continuüm van effectiviteit volgens the National Center for Injury Prevention and Control.	32
Figuur 2	Zes stappen van programmaevaluatie volgens het CDC Program Evaluation Framework 2024.	48
Figuur 3	Checklist voor rapportering van Theory of Change (ToC) (Nederlandse vertaling van Breuer et al., 2016).	50
Figuur 4	Het Logic Model	51
Figuur 5	Voorbeeld van een Logic Model. Overgenomen van Lawlor et al. (2019).	52
Figuur 6	Het Practical, Robust Implementation and Sustainability Model (PRISM)	55
Figuur 7	Vier contextdomeinen van PRISM.	55
Figuur 8	Aanpassing van PRISM/RE-AIM met een focus op gezondheidsongelijkheden en co-creatie.	56
Figuur 9	Het gereviseerde MRC-framework. Overgenomen van Skivington et al. (2021).	58
Figuur 10	Drie componenten van procesevaluatie volgens het MRC-framework.	59
Figuur 11	Realist Evaluation cyclus.	61
Figuur 12	Verschillende stappen van Contribution Analysis volgens Mayne (2008).	64
Figuur 13	Het Consolidated Framework for Implementation Research (CFIR).	66
Figuur 14	Het Implementation Research Logic Model (IRLM)	69
Figuur 15	Het PRECEDE-PROCEED model.	71
Figuur 16	Stappen Intervention Mapping.	72
Figuur 17	Stappen in een evaluability assessment volgens Leviton et al. (2010).	80
Figuur 18	WHO Beslissingsmatrix voor beleidsmakers.	92
Figuur 19	Redenen om kwalitatieve methoden in verschillende fasen van een interventiestudie te integreren.	93
Figuur 20	Vier hoofdtypen van mixed methods designs volgens Creswell en Plano-Clark (2011).	94
Figuur 21	Het onderscheid tussen efficacy, effectiveness en scale-up in preventieonderzoek volgens Gottfredson et al. (2015)	98
Figuur 22	ADAPT-procesmodel voor het aanpassen van interventies aan nieuwe contexten.	102
Figuur 23	Factoren die bepalen of een sector klaar is voor gedeelde impactmeting.	111
Figuur 24	Zes stappen van de SSA-methode, toegepast op het hospital-based youth violence prevention project (126)	120

Figuur 25	De studiecriteriën voor de verschillende gradaties in bewijskracht. Overgenomen uit Kies erkend, krijg erkenning. Criteria 2024 – 2027 (130)	127
Figuur 26	Verloop en iteratief karakter van de vier fasen in het evaluatieproces.	154
Figuur 28	Procesevaluatie volgens het MRC-framework (70).	160

Tabellen

Tabel 1	Indeling van gezondheidsbevordering en preventie, geïnspireerd op (11).	25
Tabel 2	Overzicht van de verschillende types van evaluatie toegelicht vanuit het preventielexicon	30
Tabel 3	Overzicht van belangrijke randvoorwaarden voor evaluatie.	45
Tabel 4	RE-AIM-dimensies.	53
Tabel 5	Negen domeinen die van belang zijn bij de overgang van werkzaamheidsonderzoek (efficacy) naar effectiviteitsonderzoek (effectiveness) volgens Marchand et al. (2011).	75
Tabel 6	Overzicht van verschillende kwantitatieve onderzoeksdesigns.	90
Tabel 7	Types indicatoren	96
Tabel 8	Drie modellen van samenwerking tussen onderzoek- en beleidsteams (gebaseerd op impactevaluatie in het ontwikkelingswerk)	114
Tabel 9	Samenwerkingsmodel uit het Europees Preventiehandboek.	115
Tabel 10	Domeinen van prioriteringscriteria voor de implementatie en evaluatie van volksgezondheidsprogramma's.	118
Tabel 11	De criteria van de erkenningsniveaus Goed beschreven, Goed onderbouwd, en Effectief (130).	125
Tabel 12	Evaluatieproces in vier fasen.	155
Tabel 13	Overzicht van types kwantitatieve onderzoeksdesigns.	168
Tabel 14	Verwachte expertise inzake evaluatie van de Canadese overheid voor junior, intermediate en senior evaluatiespecialisten	210
Tabel 15	Matrix van Bourgois en Cousins inzake organisatorische evaluatiecapaciteit.	212
Tabel 16	De originele <i>Evaluation Practice Analysis Survey</i> (EPAS).	213

Hoofdstuk 1

Situering en probleemstelling

Effectevaluatie vormt een essentieel onderdeel van een kwalitatief preventief gezondheidsbeleid, maar in Vlaanderen blijkt dit in de praktijk minder vanzelfsprekend. Dit hoofdstuk schetst de context waarbinnen het onderzoeksproject is ontstaan en licht toe waarom een duidelijke leidraad voor effectevaluatie noodzakelijk is. We brengen eerst de huidige knelpunten in kaart, gevolgd door de doelstellingen van het project. Daarnaast wordt een beknopt overzicht gegeven van relevante cijfers over gezondheid en preventie in Vlaanderen en van het beleidskader waarbinnen partnerorganisaties, terreinorganisaties en intermediairs werken. Dit biedt het noodzakelijke vertrekpunt om de verdere uitwerking van de leidraad in de volgende hoofdstukken te kunnen begrijpen en duiden.

1 Probleemstelling en doelstellingen van het onderzoeksproject

Een sterk preventief gezondheidsbeleid steunt op interventies waarvan niet alleen de uitvoering, maar ook de effectiviteit goed onderbouwd is. In de Vlaamse preventiesector blijkt het echter moeilijk om effectevaluatie consequent en kwaliteitsvol vorm te geven. Ondanks formele verankering in beheersovereenkomsten en subsidieaanvragen, blijft in de praktijk vooral aandacht gaan naar procesindicatoren, terwijl systematische effectevaluaties minder vaak plaatsvinden of gebaseerd zijn op methoden met beperkte bewijskracht. Daardoor ontbreekt een solide basis om onderbouwde beleidsbeslissingen te nemen en interventies gericht bij te sturen.

Tegen deze achtergrond werd dit onderzoeksproject opgezet. Het beoogt een generieke en toepasbare leidraad te ontwikkelen die organisaties en beleidsmakers ondersteunt bij het plannen, uitvoeren en benutten van effectevaluaties. De leidraad heeft als finale doelstelling om kwaliteitsvolle effectevaluatie structureel te verankeren binnen het preventieve gezondheidsbeleid.

Inhoudelijk wordt de leidraad generiek uitgewerkt, zodat deze toepasbaar is op de uiteenlopende interventies van partnerorganisaties en organisaties met terreinwerking binnen de preventiesector.

Concreet beoogt de leidraad:

- Het aanreiken van criteria om te bepalen welke interventies prioritair in aanmerking komen voor evaluatie.
- Het aanreiken van evaluatiemethoden die wetenschappelijk onderbouwd en haalbaar zijn in de praktijk.
- Het aanbieden van ondersteunende tools die organisaties helpen bij het ontwikkelen, plannen en uitvoeren van effectevaluaties.
- Het formuleren van aanbevelingen over de rol- en taakverdeling tussen organisaties in de preventiesector, onderzoekspartners en beleid bij de evaluatie van gezondheidsbevorderende interventies.

2 Vlaamse cijfers inzake preventieve gezondheidszorg en gezondheidsdeterminanten in een notendop

De gezondheidstoestand van de Vlaamse bevolking wordt mee beïnvloed door leefstijl, gezondheidsdeterminanten en de mate waarin mensen deelnemen aan preventieve zorgprogramma's. Cijfers tonen vooruitgang op sommige vlakken, zoals rookgedrag en deelname aan kankerscreening, maar ook hardnekkige uitdagingen en duidelijke ongelijkheden in gezondheidsdeterminanten en preventief zorggebruik:

- **Roken:** In 2023 rookte 16% van de Vlamingen van 15 jaar en ouder (waarvan 11% dagelijks), een sterke daling in vergelijking met 2001 (27% rokers, 23% dagelijkse rokers). Vlaanderen kende de grootste daling van alle gewesten¹. Tegelijk gebruikt 5.5% van de Vlamingen een e-sigaret (bv. vapes), met een opvallende stijging vooral bij jongeren².
- **Alcohol:** In 2023 dronk 52% van de Vlamingen vanaf 15 jaar minstens wekelijks alcohol, en 15% kon worden beschouwd als overmatige drinker (>10 consumpties/week). Deze cijfers zijn vrij stabiel sinds 2013, liggen hoger bij mannen en hoogopgeleiden, zonder significante verschillen tussen de gewesten³.
- **Lichaamsbeweging:** In 2023 voldeed slechts 40% van de Vlamingen van 15 jaar en ouder aan de aanbevolen 150 minuten lichaamsbeweging per week, wat nauwelijks verschilt van 2018 (37%). Daarnaast zit of rust 27% langdurig (>8 uur/dag), vooral hogeropgeleiden, wat een bijkomend risico vormt voor de gezondheid⁴.
- **Overgewicht:** Bijna de helft (49%) van de Vlaamse volwassenen had overgewicht (BMI ≥ 25) in 2023, een stijging ten opzichte van 43% in 2001. Het aandeel neemt sterk toe met de leeftijd (26% bij 18-24-jarigen t.o.v. 62% bij 55-64-jarigen), en ligt lager bij hogeschoolden. Het percentage mannen met overgewicht ligt hoger maar blijft stabiel, terwijl dat bij vrouwen gestaag is gestegen⁵.
- **Gezondheidsvaardigheden:** In Vlaanderen beschikte in 2022 81% van de volwassenen over voldoende gezondheidsvaardigheden, terwijl 15% een problematisch en 3% een ontoereikend niveau heeft. Vooral ouderen en lager opgeleiden scoren zwakker⁶.
- **Kankerscreening:** De participatie aan de Vlaamse bevolkingsonderzoeken naar baarmoederhals- en dikkedarmkanker haalt net de doelstelling (>65% en >60%). De deelname aan borstkankerscreening ligt met 65% nog onder de beoogde 75%. Jongere leeftijdsgroepen blijven in alle screeningsprogramma's ondervertegenwoordigd⁷.
- **Overkoepelende gezondheidsindicator:** Voor de gezondheidsdoelstelling 'De Vlaming Leeft Gezonder' wordt een samengestelde indicator gebruikt die het aantal leefstijlgerelateerde

¹ Cijfers roken: https://www.sciensano.be/sites/default/files/ta_report_2023_nl.pdf

² Cijfers vapes: https://www.sciensano.be/sites/default/files/ec_report_2023_nl.pdf

³ Cijfers alcohol: https://www.sciensano.be/sites/default/files/al_report_his2023_nl.pdf

⁴ Cijfers lichaamsbeweging: https://www.sciensano.be/sites/default/files/pa_report_nl_his2023.pdf

⁵ Cijfers overgewicht: https://www.sciensano.be/sites/default/files/ns_report_nl_his2023.pdf

⁶ Cijfers gezondheidsvaardigheden: https://www.sciensano.be/sites/default/files/preventiebarometer_gezondheidsvaardigheden.pdf

⁷ Cijfers kankerscreening: https://dikkedarmkanker.bevolkingsonderzoek.be/sites/default/files/2024-12/CVK24-027%20Verkorte%20Jaarfiche_DDK-BK-BHK%282024%29-DEF%28Hires%29.pdf

gezondheidsrisico's in kaart brengt op basis van de thema's roken, alcohol, gewicht, voeding en beweging. De score varieert van nul (geen risico) tot vijf (alle risico's aanwezig). In de gezondheidsenquête van Sciensano (2018) had 12,7% van de Vlamingen geen enkel risico; slechts 0,2% had alle vijf de risico's. De grootste groepen vertoonden één (32,4%) of twee risico's (32,2%). Door methodologische aanpassingen in enkele indicatoren tussen 2013 en 2018 zijn vergelijkingen met eerdere metingen niet mogelijk⁸.

3 Het Vlaamse preventieve gezondheidsbeleid

Het **preventiedecreet** van 2003⁹ vormt het algemene kader voor het Vlaams preventief gezondheidsbeleid. Om dit beleid en de bijbehorende budgetten een duurzame focus te geven, formuleert de Vlaamse overheid **gezondheidsdoelstellingen**. Dit zijn meetbare doelstellingen die gericht zijn op gezondheidswinst op bevolkingsniveau en het verkleinen van gezondheidsongelijkheden. Anno 2025 zijn er vijf overkoepelende Vlaamse gezondheidsdoelstellingen, elk verder uitgewerkt in strategische en operationele (sub)doelstellingen:

- **Gezondheidsdoelstelling 'De Vlaming leeft gezonder in 2025':** Deze gezondheidsdoelstelling richt zich op meerdere leefstijlfactoren tegelijk: gezonde voeding, beweging, roken en vaperen, middelengebruik (alcohol, drugs, gamen, gokken, psychoactieve medicatie), geestelijke gezondheidsbevordering, mondgezondheid, en valpreventie. In het strategisch plan werden diverse proces- en gezondheidsindicatoren opgenomen, maar zonder concrete streefcijfers.
- **Gezondheidsdoelstelling 'Suïcidepreventie':** De gezondheidsdoelstelling voor het derde Vlaamse Actieplan Suïcidepreventie (2022-2030) luidt als volgt: *'het aantal zelfdodingen in 2030 moet met 10% gedaald zijn t.o.v. het referentiejaar 2020.'*
- **Gezondheidsdoelstelling 'Bevolkingsonderzoeken naar kanker':** In juni 2025 werd een nieuwe gezondheidsdoelstelling voor de bevolkingsonderzoeken naar kanker geformuleerd: *'Van 2025 tot 2035 realiseren we maximale gezondheidswinst voor de bevolking in Vlaanderen door een passend, kwaliteitsvol en wetenschappelijk onderbouwd screeningsaanbod op maat van de doelgroep voor borst-, baarmoederhals- en dikkedarmkanker en volgen we nieuwe ontwikkelingen nauwgezet op.'*
- **Gezondheidsdoelstelling 'Vaccinatie':** De hoofddoelstelling van de gezondheidsdoelstelling vaccinatie 2024-2030 is: *'Tegen 2030 is de bevolking in Vlaanderen nog beter beschermd door vaccinatie. Er is een verbetering van de vaccinatiegraad voor welbepaalde vaccinaties en doelgroepen, er is behoud van de vaccinatiegraad waar die voldoende is, door middel van een kwaliteitsvol en laagdrempelig vaccinatiebeleid, waarbinnen ook de toevoeging van nieuwe vaccins aan het programma wordt overwogen.'*
- **Gezondheidsdoelstelling 'Milieugezondheidszorg':** De gezondheidsdoelstelling milieugezondheidszorg, goedgekeurd in maart 2024, richt zich op leefomgevingskwaliteit, het klimaat en zeer

⁸ Cijfers tussentijdse evaluatie van het strategisch plan 'De Vlaming Leeft Gezonderd in 2025' (gepubliceerd 27/10/2022): https://www.zorg-en-gezondheid.be/sites/default/files/2022-11/Tussentijdsrapport_Gezonderlevenin2025_definitief.pdf

⁹ https://etaamb.openjustice.be/nl/decreet-van-21-november-2003_n2004035090.html

zorgwekkende stoffen. Ze luidt: *‘Tegen 2030 is er een gunstige tendens van de milieugerelateerde ziektelast van de Vlaming.’*

Volgens het preventiedecreet moet het preventieve gezondheidsbeleid rekening houden met de diverse maatschappelijke ‘facetten’ die invloed hebben op gezondheid. Daarbij wordt een **facettenbeleid** gehanteerd: een intersectorale aanpak van preventieve gezondheid, waarbij de uitwerking van gezondheidspromotie en preventie van ziekte niet enkel binnen het beleidsdomein Welzijn, Volksgezondheid en Gezin gebeurt, maar ook in samenwerking verloopt met andere domeinen zoals onderwijs en werk. Dit sluit nauw aan bij het concept ‘Health in All Policies’.

Voor de structurele uitvoering van het beleid doet de overheid onder meer beroep op partnerorganisaties, organisaties met terreinwerking en Gezondheidsmakers. Partnerorganisaties en organisaties met terreinwerking worden hiervoor geselecteerd via een oproep en een beheersovereenkomst. In wat volgt, worden hun rollen verder toegelicht.

3.1 Partnerorganisaties

Partnerorganisaties zijn organisaties met expertise in één of meerdere domeinen van de preventieve gezondheidszorg. Ze ondersteunen Gezondheidsmakers, organisaties met terreinwerking, individuele zorgaanbieders en anderen die instaan voor het veldwerk. Ze verstrekken onder meer informatie en advies, ontwikkelen methodieken, interventies en/of materialen, en bieden ondersteuning bij de implementatie en het gebruik daarvan. Vlaanderen telt 12 partnerorganisaties die actief zijn binnen het preventief gezondheidsbeleid:

- **Diabetes Liga:** Een organisatie die zich richt op de preventie van diabetes via advies, informatie en het ontwikkelen van methodieken en materialen.
- **Eetexpert:** Een onafhankelijk kenniscentrum dat ondersteuning biedt bij preventie, vroegdetectie en behandeling van eet- en gewichtsproblemen, met aandacht voor bio- en psychosociale factoren.
- **Expertisecentrum Val- en fractuurpreventie Vlaanderen (EVV):** Een samenwerkingsverband dat zich inzet voor val- en fractuurpreventie bij ouderen, zowel thuiswonend als in woonzorginstellingen.
- **Sensoa:** De Vlaamse partnerorganisatie voor seksuele gezondheid, actief via educatie, zorg, begeleiding en beleidsmatige ondersteuning.
- **Vlaams expertisecentrum Alcohol en andere Drugs (VAD):** Expertisecentrum voor de preventie van gezondheidsschade door middelengebruik, gamen en gokken, met o.a. De Druglijn als publieksdienst.
- **Vlaams Expertisecentrum Suïcidepreventie (VLESP):** Coördineert het Vlaamse suïcidepreventiebeleid en beheert het platform Zelfmoord1813, in samenwerking met UGent en het CPZ.
- **Vlaams Instituut Gezond Leven:** Partnerorganisatie voor gezondheidsbevordering en ziektepreventie, met focus op voeding, beweging, tabakspreventie, geestelijke gezondheid en settinggericht werken.
- **Vlaams Instituut Mondgezondheid – Gezonde Mond:** Samenwerkingsverband dat zich inzet voor preventieve mondzorg bij risicogroepen, met partners uit de tandheelkundige sector en het hoger onderwijs.

- **Vlaamse Wetenschappelijke Vereniging voor Jeugdgezondheidszorg (VWVJ):** Ondersteunt de CLB's in hun opdracht rond preventieve gezondheidszorg bij kinderen en jongeren.
- **Instituut Tropische Geneeskunde (ITG):** Een onderzoeksinstituut die zich toelegt op onderzoek, onderwijs en dienstverlening rond tropische ziekten en gezondheidszorg in ontwikkelingslanden.
- **Milieugezondheidszorg:** Een samenwerking tussen het Provinciaal Instituut voor Hygiëne, de Vlaamse Instelling voor Technologisch Onderzoek (VITO) en Gezond Leven die zich richt op preventieve milieugezondheidszorg
- **Water Bestemd voor Menselijke Aanwending:** Een samenwerking tussen VITO en Buildwise die wetenschappelijke ondersteuning biedt voor veilig en gezond drinkwaterbeleid.

3.2 Organisaties met terreinwerking en Gezondheidsmakers

Organisaties met terreinwerking zijn, naast individuele zorgaanbieders en andere actoren, verantwoordelijk voor het veldwerk binnen de preventieve gezondheidszorg. Hiervoor doen zij een beroep op de expertise van de partnerorganisaties. Ze vertalen de strategieën en methodieken naar de praktijk door deze zelf te implementeren en/of door regionale en lokale netwerken te coachen of te motiveren bij de uitvoering ervan. Lokaal werken zij samen met Gezondheidsmakers.

Er zijn 11 soorten organisaties met terreinwerking actief binnen het preventief gezondheidsbeleid:

- **Centra voor Leerlingenbegeleiding (CLB):** Bieden begeleiding aan leerlingen, ouders en scholen rond leren, studeren, schoolloopbaan, preventieve gezondheid en psychosociaal welzijn.
- **Centrum ter Preventie van Zelfdoding (CPZ):** Biedt laagdrempelige telezorg via de Zelfmoordlijn en ASPHA, gericht op mensen met suïcidale gedachten en hun omgeving.
- **CGG – Suïcidepreventiewerking:** Preventiewerkers binnen de CGG die vormingen, coaching en beleidsadvies aanbieden aan intermediairs rond suïcidepreventie, verspreid over alle Vlaamse provincies.
- **CGG – Middelenpreventiewerking:** Preventiewerkers binnen de CGG richten zich op intermediairs rond middelengebruik (alcohol, drugs, gamen, gokken) via vorming en ondersteuning.
- **Free Clinic:** Voert het spuitenruilprogramma uit in Vlaanderen, gericht op schadebeperking bij druggebruikers.
- **Vlaamse Vereniging voor Respiratoire Gezondheidszorg en Tuberculosebestrijding (VRGT):** Actief in rookstopbegeleiding en tuberculosebestrijding, met expertise in epidemiologische opvolging en preventie.
- **Centrum voor Kankeropsporing:** Coördineert en ondersteunt de Vlaamse bevolkingsonderzoeken naar kanker, met vijf regionale afdelingen.
- **Fysisch-technische controleorganisaties:** Erkende organisaties die instaan voor de kwaliteitscontrole van mammografietoestellen binnen het bevolkingsonderzoek naar borstkanker.
- **Sciensano:** Ondersteunt het Vlaamse infectieziektebeleid via laboratoriumonderzoek, bronopsporing en wetenschappelijke advisering, o.a. bij uitbraken.
- **Landsbonden van de ziekenfondsen:** Ondersteunen de ziekenfondsen bij hun preventieve gezondheidsopdrachten, via een beheersovereenkomst met de Vlaamse overheid.

- **Pasop:** Werkt rond seksuele gezondheid bij sekswerkers, met focus op HIV/SOA-preventie, seksuele ontwikkeling en welzijn.

De Gezondheidsmakers vzw (de vroegere Logo's) spelen een centrale rol in de uitwerking en implementatie van het preventieve gezondheidsbeleid op het lokale of regionale niveau. Samen met de Brusselse werking "Gezond in Brussel" zetten ze hun expertise in om preventie lokaal verankerd, effectief en toegankelijk te maken. De Gezondheidsmakers ondersteunen intermediairs en actoren zoals gemeenten, scholen, werkgevers en zorgsettings door advies, methodieken, materialen en vorming te bieden rond verschillende gezondheidsthema's.

3.3 Intermediairs

Intermediairs zijn sleutelfiguren zoals werkgevers, leerkrachten, jeugdwerkers en gezondheidscoaches. Ze hebben een invloed op doelgroepen in hun leef-, leer-, werk- en vrijetijdsomgeving. Ze spelen een cruciale rol in het toepassen van preventieve methodieken. In projectmatig werken wordt niet alleen rekening gehouden met de doelgroep, maar ook met de competenties, motivatie en omgeving van intermediairs, die versterkt kunnen worden via gerichte implementatiestrategieën.

Hoofdstuk 2

Methode

Dit hoofdstuk beschrijft de aanpak die gebruikt werd om de leidraad voor effectevaluatie te ontwikkelen. De methodologie combineert een brede verkenning van nationale en internationale literatuur met de expertise van actoren uit praktijk, beleid en onderzoek. Door deze inzichten systematisch te bundelen, ontstaat een stevig fundament voor een leidraad die zowel wetenschappelijk onderbouwd als praktisch toepasbaar is in de Vlaamse preventiesector.

1 Literatuurstudie

Er werd een gerichte en brede, maar niet-systematische literatuurstudie uitgevoerd naar nationale en internationale publicaties uit zowel de wetenschappelijke literatuur (zoals peer-reviewed artikels en reviews) als de grijze literatuur (zoals beleidsdocumenten en rapporten). Het doel van deze literatuurstudie was inzicht te verwerven in de verschillende types evaluatie binnen interventie-onderzoek, in de randvoorwaarden die een kwaliteitsvolle evaluatie ondersteunen, en in bestaande evaluatiekaders en methodologische inzichten die richting geven aan de opzet en uitvoering van evaluaties van gezondheidsbevorderende interventies. Daarbij werd ook aandacht besteed aan vraagstukken rond de prioritering van evaluaties, samenwerking en rolverdeling tussen betrokken actoren bij de evaluatie van interventies, en de aard van passend en bruikbaar bewijs voor gezondheidsbevorderende interventies, rekening houdend met hun contextgebonden en complexe karakter. Daarnaast werd aandacht besteed aan benaderingen voor beleidsimpactevaluatie. Doorheen de literatuurstudie werd bovendien expliciet gezocht naar inspirerende en goed gedocumenteerde voorbeelden van evaluatiepraktijken en beleidsmatige evaluatieaanpakken die relevant en toepasbaar zijn binnen de Vlaamse context.

2 Consultatie van stakeholders uit de preventiesector, beleid en onderzoek

In het kader van dit project werden twee werkgroepen opgericht. Beide werkgroepen kwamen twee keer een halve dag samen, in september en november 2025.

Werkgroep 1: Stakeholders uit de preventiesector

Deze groep bestond voornamelijk uit vertegenwoordigers van partnerorganisaties en organisaties met terreinwerking:

- Gezondheidsmakers
- Vlaams instituut Gezond Leven
- Sensoa
- Gezonde Mond
- Vlaams expertisecentrum Alcohol en andere Drugs
- Intermutualistische vertegenwoordiging
- Diabetes Liga
- Vlaams Expertisecentrum Suïcidepreventie
- Centrum ter Preventie van Zelfdoding
- Centrum voor Kankeropsporing

- Vlaamse Wetenschappelijke Vereniging voor Jeugdgezondheidszorg (VWVJ)
- Expertisecentrum Val- en fractuurpreventie Vlaanderen
- Eetexpert

Werkgroep 2: Actoren in onderzoek en beleid

Deze groep bestond uit experts, onderzoekers en beleidsmedewerkers van onder meer:

- Vlaams Instituut Gezond Leven
- Onderzoeksinstituten (universiteiten, hogescholen, Sciensano)
- Departement Zorg

Inhoud van de twee bijeenkomsten

Tijdens de eerste bijeenkomst werd afgesteld over:

- de randvoorwaarden en hefboomen voor kwalitatieve en haalbare effect- en impactevaluatie;
- de rolverdeling en samenwerkingsmogelijkheden tussen organisaties en onderzoekers;
- het gebruik van evaluatiekaders;
- en de keuze van passende onderzoeksmethoden.

De tweede bijeenkomst focuste op:

- wanneer effectevaluatie aan de orde is en hoe deze best kan worden gefinancierd;
- hoe evaluatieresultaten beter benut kunnen worden;
- hoe meer inzicht kan worden verworven in de beleidsimpact;
- en de verkenning van de behoefte aan een centraal aanspreekpunt en adviesfunctie voor evaluatie binnen de preventiesector.

3 Opbouw van het rapport

Hoofdstukken 3 tot en met 11 bieden een overzicht van de literatuurstudie en behandelen achtereenvolgens:

- de fundamentele concepten van interventies en evaluatie binnen gezondheidsbevordering;
- de randvoorwaarden voor een kwaliteitsvolle evaluatie;
- relevante en bruikbare evaluatiekaders voor interventie-evaluatie;
- de uitvoering van effectevaluaties in de praktijk, inclusief de keuze van onderzoeksdesigns en indicatoren;
- wat geldt als passend bewijs voor preventie;
- de impactevaluatie van het preventieve gezondheidsbeleid op systeemniveau;
- de samenwerking en rolverdeling binnen evaluatieprocessen;
- manieren om te bepalen welke interventies prioriteit krijgen voor evaluatie;
- hoe evaluatiebevindingen optimaal kunnen worden gebruikt voor beleidsontwikkeling en praktijkverbetering.

Hoofdstuk 12 geeft een synthese van de twee bijeenkomsten met de werkgroepen, waarin inzichten uit de literatuur worden afgetoetst aan de praktijk.

Hoofdstuk 13 bevat tot slot de volledige leidraad, gebaseerd op de bevindingen uit dit onderzoeksrapport.

Hoofdstuk 3

Fundamentele concepten van interventie en evaluatie in gezondheidsbevordering

Effectieve gezondheidsbevordering vraagt om goed doordachte interventies, maar ook om een grondige en systematische evaluatie ervan. Dit is essentieel om beleidsmakers te voorzien van betrouwbare en relevante inzichten waarmee zij onderbouwde beslissingen kunnen nemen voor een sterk preventief gezondheidsbeleid. In deze context is het noodzakelijk om vooraf duidelijkheid te scheppen over kernbegrippen zoals *preventie*, *interventie*, *effectiviteit*, *impact*, *evaluatie* en de *mate van bewijskracht*.

1 Gezondheidsbevordering en preventie

Preventie omvat een reeks strategieën die aanvullend zijn op een behandeling en gericht zijn op het bevorderen en behouden van een goede gezondheid. Preventie kan zich richten op de hele bevolking of op specifieke groepen. Het doel is om gezondheidsproblemen te voorkomen, hun ernst te beperken, het ontstaan ervan uit te stellen of het aantal mensen dat ermee te maken krijgt te verminderen (1).

1.1 Klassieke indeling van preventie

Binnen de volksgezondheid wordt preventie traditioneel ingedeeld in drie niveaus, zoals oorspronkelijk beschreven door Caplan (2):

- 1) **Primaire preventie** richt zich op het voorkomen dat gezondheidsproblemen ontstaan door risico-factoren weg te nemen of te beperken. Voorbeelden zijn campagnes rond gezonde voeding of rookverboden in openbare ruimtes.
- 2) **Secundaire preventie** is gericht op het vroegtijdig opsporen en aanpakken van gezondheidsproblemen, bijvoorbeeld via screening op hoge bloeddruk of vroegtijdige interventies bij overgewicht.
- 3) **Tertiaire preventie** is gericht op het beperken van blijvende beperkingen en het bevorderen van herstel bij mensen met een bestaande aandoening, zoals revalidatieprogramma's na een beroerte.

Hoewel deze indeling vaak wordt gebruikt, zijn de grenzen tussen de niveaus in de praktijk soms moeilijk scherp te trekken, zeker bij complexe gezondheidsproblemen die meerdere determinanten en fasen omvatten (3).

Naast deze klassieke indeling ontwikkelde Gordon (4) een aanvullende classificatie op basis van de doelgroep van de interventie:

- 1) **Universele preventie** is gericht op de gehele bevolking, met als doel het algemene risico te verlagen. Bijvoorbeeld algemene voorlichtingscampagnes over lichaamsbeweging.
- 2) **Selectieve preventie** richt zich op groepen met een verhoogd risico, zoals bijvoorbeeld mensen die in een gebied wonen met slechte luchtkwaliteit.
- 3) **Geïndiceerde preventie** is gericht op mensen met een verhoogde kwetsbaarheid, maar zonder duidelijke symptomen, zoals bijvoorbeeld begeleiding van mensen met beginnende symptomen voor diabetes.

In deze benadering valt wat Caplan “tertiaire preventie” noemt onder de noemer behandeling.

1.2 Evolutie naar bredere gezondheidsbevordering

De laatste decennia groeit de aandacht voor preventie op bevolkingsniveau, waarbij naast zorg ook gezondheidsgedrag, sociale omstandigheden en de fysieke leefomgeving centraal staan. Gezondheid wordt zo gezien als het resultaat van een samenspel tussen persoonlijke keuzes en de omgeving (1). Deze benadering sluit aan bij het bredere gezondheidsconcept van de WHO waarin gezondheid wordt opgevat als “a state of complete physical, mental and social well-being” (5). Het impliceert dat **preventie ook moet inspelen op omgevingsfactoren**, zoals huisvesting, onderwijs, arbeidsomstandigheden en sociale ongelijkheid (6, 7). Vanuit deze bredere insteek sluit preventie aan bij **gezondheidsbevordering**, dat zich richt op het versterken van de voorwaarden die mensen en gemeenschappen nodig hebben om hun gezondheid te behouden of te verbeteren. De aanpak is breed en overstijgt individuele gedragsverandering; het gaat ook om het verbeteren van sociale, economische en fysieke omstandigheden (8, 9).

Hoewel **gezondheidsbevordering en preventie vaak in één adem worden genoemd – en ze elkaar zeker versterken – leggen ze elk een ander accent**. Gezondheidsbevordering richt zich op het versterken van mensen en hun omgeving om gezond te leven. Het gaat daarbij niet alleen over gedrag, maar ook over de omstandigheden waarin mensen opgroeien, wonen en werken. Denk aan het creëren van gezonde school- of werkomgevingen of het toegankelijk maken van gezond voeding. *Preventie* focust zoals eerder aangegeven meer op het voorkomen van specifieke ziektes of het beperken van de impact ervan. Het verschil tussen beide benaderingen zit vooral in de invalshoek: gezondheidsbevordering versterkt beschermende factoren en kijkt breed naar gezondheid en welzijn, terwijl preventie eerder inzet op het verminderen van risico’s en het voorkomen en beperken van ziektelast.

De invulling van gezondheidsbevordering wordt vaak gekoppeld aan vijf centrale actieterreinen, zoals omschreven in het Ottawa Charter van de Wereldgezondheidsorganisatie (10):

- 1) Het versterken van persoonlijke vaardigheden
- 2) Het creëren van omgevingen die gezond gedrag ondersteunen
- 3) Het ontwikkelen van gezond beleid
- 4) Het stimuleren van participatie en gemeenschapsacties
- 5) Het heroriënteren van de gezondheidszorg met meer aandacht voor preventie en promotie

Om preventieve strategieën vollediger te beschrijven, kan het nuttig zijn om verschillende typologieën te combineren. Arango et al. (11) stelden een geïntegreerde indeling voor die de klassieke niveaus van Caplan en de doelgroepgerichte indeling van Gordon koppelt aan gezondheidsbevordering. Hoewel deze indeling van Arango oorspronkelijk ontwikkeld werd voor strategieën in de geestelijke gezondheidsbevordering en de preventie van psychische problemen, is ze ook bruikbaar voor algemene gezondheidsbevordering en preventie. Tabel 1 toont hoe dit model kan worden vertaald naar een algemene gezondheidscontext.

Tabel 1 Indeling van gezondheidsbevordering en preventie, geïnspireerd op (11).

Type	Doelgroep	Doel	Voorbeeld
Gezondheidsbevordering	Algemene bevolking	Algemene gezondheid, vitaliteit en veerkracht stimuleren	Campagnes en schoolprogramma's rond gezonde voeding, bewegen en slapen.
Primaire universele preventie	Algemene bevolking, ongeacht risicofactoren	Het ontstaan van ziekte of risico beperken door gezonde keuzes te stimuleren	Anti-rookcampagnes
Primaire selectieve preventie	Subgroepen met verhoogd risico (bv. genetisch, sociaaleconomisch)	Risicofactoren in kwetsbare groepen beperken	Leefstijlprogramma's voor gezinnen met obesitas in de familie
Primaire geïndiceerde preventie	Personen met lichte of beginnende symptomen of risicokenmerken	Verergering of chroniciteit van beginnende gezondheidsklachten voorkomen	Beweegprogramma's bij beginnende rugklachten
Secundaire preventie	Personen met gediagnosticeerde aandoening in vroeg stadium	Snelle diagnostiek, behandeling en begeleiding om verergering te vermijden	Rookstopcampagnes bij COPD-diagnose
Tertiaire preventie	Personen met uitgesproken, langdurige gezondheidsproblemen	Complicaties beperken, kwaliteit van leven verbeteren, zelfredzaamheid vergroten	Revalidatieprogramma's na een hartinfarct

1.3 Focus leidraad

De leidraad richt zich op **gezondheidsbevordering en primaire preventie**, zowel op **universeel niveau** (gericht op de gehele bevolking) als op **selectief niveau** (gericht op specifieke groepen met een verhoogde kwetsbaarheid). Een universele actie kan bijvoorbeeld een campagne zijn die het belang van gezond eten en voldoende bewegen onder de gehele bevolking benadrukt. Een selectieve actie daarentegen kan zich richten op mensen met overgewicht die hierdoor een hoger risico lopen op diabetes type 2, bijvoorbeeld door hen speciale voedingsadviezen en bewegingsprogramma's aan te bieden. Geïndiceerde acties (bv. begeleiding bij beginnende diabetes) vallen niet binnen deze leidraad.

2 Interventie

Een **interventie** is een doelgerichte en geplande ingreep die tot doel heeft een situatie te veranderen. In de context van gezondheidsbevordering en preventie verwijst een interventie naar een doelgerichte, geplande en systematische actie die wordt uitgevoerd binnen een sociaal systeem, met als doel het veranderen van kennis, houding of gedrag van een individu, organisatie of populatie (12). Waar klinische

interventies zich richten op het individu, vinden gezondheidsbevorderende interventies vaak plaats binnen scholen, werkplekken, buurten of gemeenschappen (13).

In de leidraad en het rapport worden de termen ‘interventie’, ‘actie’ en ‘programma’ vaak gebruikt. In de Vlaamse context wordt daarnaast vaak de term **preventiemethodiek** gehanteerd. Een preventiemethodiek is een samenhangend en doordacht geheel van vaste werkwijzen, materialen en diensten, gericht op het bereiken van een specifieke doelstelling. Dit gebeurt via één of meerdere strategieën, zoals educatie, omgevingsinterventies, afspraken en regels, of het aanbieden van preventieve zorg of begeleiding.

2.1 Enkelvoudige versus complexe interventies

Binnen gezondheidsbevordering wordt vaak een onderscheid gemaakt tussen enkelvoudige en complexe interventies. **Enkelvoudige interventies** richten zich op één specifiek gedragsdoel of een beperkte set van determinanten. Ze zijn doorgaans kortdurend, eenvoudig uit te voeren en relatief gemakkelijk te evalueren. Voorbeelden zijn een rookstopcursus of een campagne die mensen aanzet tot meer dagelijkse beweging.

Complexe interventies daarentegen bestaan uit meerdere, onderling samenhangende componenten, worden vaak op verschillende niveaus tegelijk uitgevoerd en vergen meer tijd en middelen. Ze worden meestal ingezet bij gezondheidsproblemen die multifactorieel bepaald zijn en daardoor niet met enkelvoudige maatregelen kunnen worden aangepakt (14). Volgens de Medical Research Council (MRC) is een interventie complex wanneer zij één of meer van de volgende kenmerken bevat (15):

- Bestaat uit meerdere onderling samenhangende onderdelen
- Richt zich op groepen of organisaties (bovenop of in plaats van enkel het individu)
- Beoogt verschillende uitkomsten
- Kan worden aangepast via feedback en leerprocessen
- De effectiviteit wordt mede bepaald door het gedrag van uitvoerders en deelnemers

Kortom: hoe moeilijker het is om precies te bepalen welke onderdelen de interventie doen werken en hoe deze met elkaar samenhangen, hoe complexer de interventie. Een voorbeeld van een complexe interventie is een geïntegreerd obesitasprogramma bestaande uit veranderingen in schoolmaaltijden, bewegingsprogramma's, ouderbetrokkenheid en beleidsmaatregelen om marketing van ongezonde voeding te beperken. Bij dit soort van complexe interventies is aandacht nodig voor zowel effectmetingen als voor implementatie, context, impact en duurzaamheid.

2.2 Multiple health behaviour change interventies

Een recente evolutie binnen gezondheidsbevordering betreft interventies die zich gelijktijdig richten op meerdere gezondheidsgerelateerde gedragingen, bekend als *multiple health behaviour change interventions* (MHBCI). Deze interventies vertrekken vanuit de vaststelling dat risicogedragingen vaak in clusters voorkomen, zoals roken, onvoldoende lichaamsbeweging, ongezonde voeding en problematisch alcoholgebruik (16). Door deze gedragingen niet afzonderlijk, maar in samenhang aan te pakken, kunnen interventies potentieel effectiever en efficiënter zijn (17).

Een belangrijk voordeel van MHBCI is dat gedragsveranderingen elkaar onderling kunnen versterken. Zo kan iemand die stopt met roken verbeteringen merken in fysieke conditie, wat het aannemen van een actievere leefstijl vergemakkelijkt. Tegelijk biedt een geïntegreerde aanpak organisatorische voordelen doordat meerdere doelstellingen binnen één traject kunnen worden verwezenlijkt. Toch brengen dergelijke interventies ook uitdagingen met zich mee. Het risico bestaat dat deelnemers worden overbelast door de veelheid aan gedragsdoelen, en het evalueren van de effectiviteit van dergelijke interventies is methodologisch veeleisender omdat veranderingen op meerdere domeinen tegelijk optreden en minder voorspelbaar zijn (17).

Voorbeelden van MHBCI zijn leefstijlprogramma's die voedings- en bewegingscomponenten combineren, digitale tools die stressmanagement koppelen aan slaapverbetering en alcoholreductie, en gemeenschapsinterventies die zowel focussen op gezonde eetomgeving, bewegingsvriendelijke infrastructuur als tabaksontmoediging.

2.3 Typologieën van gezondheidsbevorderende interventies

Om het brede spectrum van publieke gezondheid beter te kunnen duiden, ontwikkelde het gezondheidsdepartement van Montréal een typologie van interventies die vier grote categorieën van gezondheidsbevorderende interventies onderscheidt (18).

Directe acties gericht op de bevolking

Deze categorie omvat interventies die direct op de bevolking gericht zijn, zoals vaccinatieprogramma's, gezondheidsvoorlichting op school of risicobeoordelingen bij infectieziekten. Hoewel deze acties voor burgers het meest zichtbaar zijn, maken zij slechts een beperkt aandeel uit van het totaal aan publieke gezondheidsactiviteiten.

Belangenbehartiging (advocacy)

Een tweede categorie betreft belangenbehartiging of advocacy, waarbij publieke gezondheidsactoren trachten invloed uit te oefenen op besluitvorming en op organisaties die de randvoorwaarden voor gezondheid vormgeven. Voorbeelden zijn lobbywerk rond rookvrije publieke ruimtes of beleidsadviezen over luchtkwaliteit. Advocacy is vooral nodig wanneer er sprake is van vernieuwende of politiek gevoelige onderwerpen.

Ondersteuning van partners

Een groot deel van het werk van publieke gezondheid richt zich op het versterken van andere organisaties die direct invloed hebben op gezondheid, zoals scholen, lokale besturen of welzijnsorganisaties (m.a.w. intermediairs). Dit kan bijvoorbeeld door het aanbieden van trainingen aan leerkrachten, het ontwikkelen van richtlijnen voor kinderopvang, het organiseren van kennisdelings-evenementen of het geven van advies aan gemeenten over verkeersveiligheid.

Samenwerking

In sommige situaties wordt er echt samen met andere organisaties gewerkt aan een project of programma. Denk hierbij aan gezamenlijke preventieprojecten met ziekenhuizen of het opzetten van partnerschappen met buurthuizen rond gezonde voeding. Hierbij is er sprake van gedeelde verantwoordelijkheid en vaak ook van formele afspraken.

De toepassing van deze typologie op het actieplan van Montréal laat zien dat circa 80% van de publieke gezondheidsinterventies indirect verloopt via partners, waarbij ondersteuning de grootste plaats inneemt, gevolgd door advocacy en samenwerking. Slechts 20% van de interventies is rechtstreeks gericht op de bevolking.

3 Evaluatie

Evaluatie vormt een essentieel onderdeel van het verantwoord ontwikkelen, uitvoeren, bijsturen en onderhouden van gezondheidsbevorderende interventies. Evalueren betekent op een betrouwbare en verantwoorde manier informatie verzamelen over een interventie, methodiek, materiaal of project, met als doel hierover een gefundeerd oordeel te vellen. Een goede evaluatie kijkt verder dan het al dan niet behalen van de doelstellingen. Het biedt inzicht in de redenen waarom bepaalde onderdelen succesvol zijn (of niet), helpt bij het bijsturen van de interventie en levert waardevolle lessen op voor toekomstige projecten (9).

Volgens de WHO-Europe gids *"Evaluating the impact of interventions addressing health behaviour"* (19) zijn er vijf kernredenen om de effecten en impact van interventies te evalueren (19):

- **Vaststellen en vergelijken van (in)effectiviteit:** evaluatie helpt om te bepalen welke interventies werken en welke niet.
- **Kostenbesparing:** door ineffectieve of overbodige interventies te vermijden, kunnen middelen efficiënter worden ingezet.
- **Voorkomen van ongewenste effecten:** evaluatie voorkomt dat interventies onbedoeld negatieve gevolgen hebben, zoals het vergroten van gezondheidsongelijkheden.
- **Aanpassing aan de context:** door evaluatie kunnen interventies beter afgestemd worden op de doelgroep en omgeving.
- **Onderbouwing van beleidskeuzes:** beleidsmakers kunnen gefundeerde beslissingen nemen op basis van betrouwbare evaluatiegegevens.

3.1 Types van evaluatie

De evaluatieliteratuur maakt onderscheid tussen diverse evaluatietypes die elkaar aanvullen en samen een volledig beeld van effectiviteit en impact bieden. Elk type evaluatie heeft zijn eigen doel en specifieke focus, en beantwoordt andere evaluatievragen. Samen zorgen ze voor een diepgaand inzicht in zowel het verloop als het resultaat van een project. Er wordt hierbij een onderscheid gemaakt tussen formatieve en summatieve evaluatie (20, 21).

- **Formatieve evaluatie** vindt plaats tijdens de uitvoering van een interventie en heeft als doel tussentijdse inzichten te genereren die bijsturing mogelijk maken. Procesevaluatie en productevaluatie vallen onder deze categorie.
- **Summatieve evaluatie** vindt plaats na afloop van een interventie en onderzoekt in welke mate de doelstellingen zijn behaald en of de interventie impact heeft gehad. Effectevaluatie, impactevaluatie en economische evaluatie zijn voorbeelden van summatieve evaluatie.
- In wat volgt worden de verschillende types van evaluatie toegelicht vanuit het preventielexicon (9). Tabel 2 toont hiervan een overzicht.

Effectevaluatie

Effectevaluatie of uitkomstevaluatie onderzoekt in hoeverre een interventie heeft geleid tot de gewenste veranderingen bij de doelgroep. Dit kan gaan over kennis, attitude, vaardigheden of gedrag. Het draait hierbij om het meten van concrete resultaten ten opzichte van vooraf opgestelde doelstellingen.

Productevaluatie

Een productevaluatie kijkt naar de kwaliteit en bruikbaarheid van materialen of middelen die in het kader van een project worden ingezet. De tevredenheid van gebruikers over zaken zoals de inhoud, vormgeving en praktische toepasbaarheid wordt geëvalueerd. Het doel is om te achterhalen of de producten geschikt zijn voor de beoogde doelgroep en voldoen aan hun verwachtingen.

Procesevaluatie

Een procesevaluatie richt zich op het traject of proces dat is doorlopen om veranderingen te bereiken. Het onderzoekt hoe de interventie in de praktijk is uitgevoerd (i.e., implementatie) en helpt te begrijpen waarom het project al dan niet effectief is geweest. Dit gebeurt door verschillende onderdelen van het uitvoeringsproces te analyseren, zoals het bereik van de doelgroep, de bereidheid van betrokken personen of organisaties om de methodiek toe te passen, en de kwaliteit van de uitvoering. Daarbij kijkt procesevaluatie naar de contextfactoren en mechanismen die de implementatie en/of effecten beïnvloeden.

Impactevaluatie

Een impactevaluatie kijkt naar veranderingen op langere termijn die aan de interventie kunnen worden toegeschreven, zoals verbeterde gezondheidsuitkomsten of gedragsaanpassingen, die zich pas na verloop van tijd manifesteren.

Economische evaluatie

Bij een economische evaluatie draait het om de afweging tussen de kosten van een interventie en de opgeleverde gezondheidswinst of maatschappelijke meerwaarde. Dit helpt bepalen of de investering verantwoord is ten opzichte van de effecten

Tabel 2 Overzicht van de verschillende types van evaluatie toegelicht vanuit het preventielexicon

Evaluatietype	Doel	Focus	Vragen	Voorbeeld
Procesevaluatie	In kaart brengen hoe een interventie in de praktijk is verlopen	Implementatie, bereik, kwaliteit, en context	Is de interventie uitgevoerd zoals gepland? Wie heeft deelgenomen? Welke factoren beïnvloedden de uitvoering?	Werd de voorlichtingscampagne effectief verspreid onder de doelgroep?
Productevaluatie	Beeoordelen van het materiaal of de middelen die in de interventie zijn gebruikt	Waardering van inhoud, vormgeving, toepasbaarheid	Is het materiaal duidelijk en bruikbaar? Vinden deelnemers het aantrekkelijk en relevant?	Wat vinden jongeren van de lay-out en inhoud van een online zelftest over alcoholgebruik?
Effectevaluatie of uitkomstevaluatie	Nagaan of de gewenste veranderingen zijn bereikt	Resultaten bij deelnemers of omgeving	Is het gedrag of de kennis van de doelgroep veranderd? Zijn de vooropgestelde doelen gehaald?	Zorgde een voedingsworkshop ervoor dat deelnemers vaker gezonde snacks kiezen?
Impactevaluatie	Nagaan of er structurele of blijvende effecten zijn opgetreden	Langetermijnveranderingen in gedrag of gezondheid	Leidt de interventie tot duurzame gedragsverandering? Zijn er blijvende verbeteringen zichtbaar?	Heeft een langdurige campagne geleid tot een daling in obesitascijfers bij jongeren?
Economische evaluatie	Inzicht krijgen in de verhouding tussen kosten en baten	Kostenefficiëntie, gebruik van middelen	Is de gezondheidswinst het geld waard? Hoe verhoudt de investering zich tot het resultaat?	Wat zijn de kosten per extra gevaccineerde persoon bij een huis-aan-huis vaccinatiecampagne?

3.2 Evaluatiebenaderingen

Wanneer we spreken over evaluatiebenaderingen, doelen we op de onderliggende uitgangspunten en principes die richting geven aan hoe een evaluatie wordt opgezet en uitgevoerd. In tegenstelling tot evaluatietypes, die zich richten op wat er geëvalueerd wordt, gaan evaluatiebenaderingen over de **visie, de rol van betrokkenen en het doel van de evaluatie**. Ze bepalen bijvoorbeeld wie inspraak krijgt, welke waarden leidend zijn, en hoe resultaten worden gebruikt.

In het domein van gezondheidsbevordering, waar thema's als participatie, gelijkheid en maatschappelijke relevantie centraal staan, is het belangrijk dat evaluaties deze waarden ook daadwerkelijk weerspiegelen. Twee benaderingen die hier goed bij aansluiten zijn participatieve evaluatie en gebruiksgerichte evaluatie (Utilisation-Focused Evaluation). Dergelijke benaderingen bieden geen vaste methodiek, maar wel een kader dat helpt om evaluaties te ontwerpen. Beide methoden streven naar evaluaties die niet enkel wetenschappelijk waardevol zijn, maar ook maatschappelijk relevant en praktisch bruikbaar.

Participatieve evaluatie vertrekt vanuit het principe dat de mensen die betrokken zijn bij of geraakt worden door een project – zoals deelnemers, intermediairs of lokale organisaties – ook een actieve rol moeten spelen in de evaluatie ervan. Samenwerking staat hierbij centraal. Participatie betekent in dit geval niet enkel feedback geven, maar ook actief meedenken over wat geëvalueerd wordt, hoe dat gebeurt, en hoe de resultaten gebruikt worden (22).

Gebruiksgerichte evaluatie legt de nadruk op het maximaliseren van de bruikbaarheid van de evaluatie door deze af te stemmen op de behoeften van de beoogde gebruikers van de resultaten – bijvoorbeeld beleidsmakers, programmacoördinatoren of lokale partners. Het evaluatieproces wordt in overleg met hen vormgegeven, zodat de opbrengsten van de evaluatie echt gebruikt worden om beleid, programma's of praktijken te verbeteren (23).

Een belangrijk uitgangspunt binnen beide benaderingen is dat **evaluatie niet waardenvrij of neutraal** is. De keuzes over wat en hoe er wordt gemeten wordt beïnvloed door belangen, machtsverhoudingen en morele overwegingen (13). In gezondheidsbevordering, waar gewerkt wordt met diverse en soms kwetsbare populaties, is het van belang om daar bewust en kritisch mee om te gaan.

Ethische principes zoals respect voor personen, het nastreven van welzijn en rechtvaardigheid zijn essentieel, zeker wanneer gewerkt wordt met kwetsbare groepen (24). Evaluaties vragen aldus transparantie, aandacht voor machtsongelijkheid en een streven naar inclusieve processen.

4 Effectiviteit

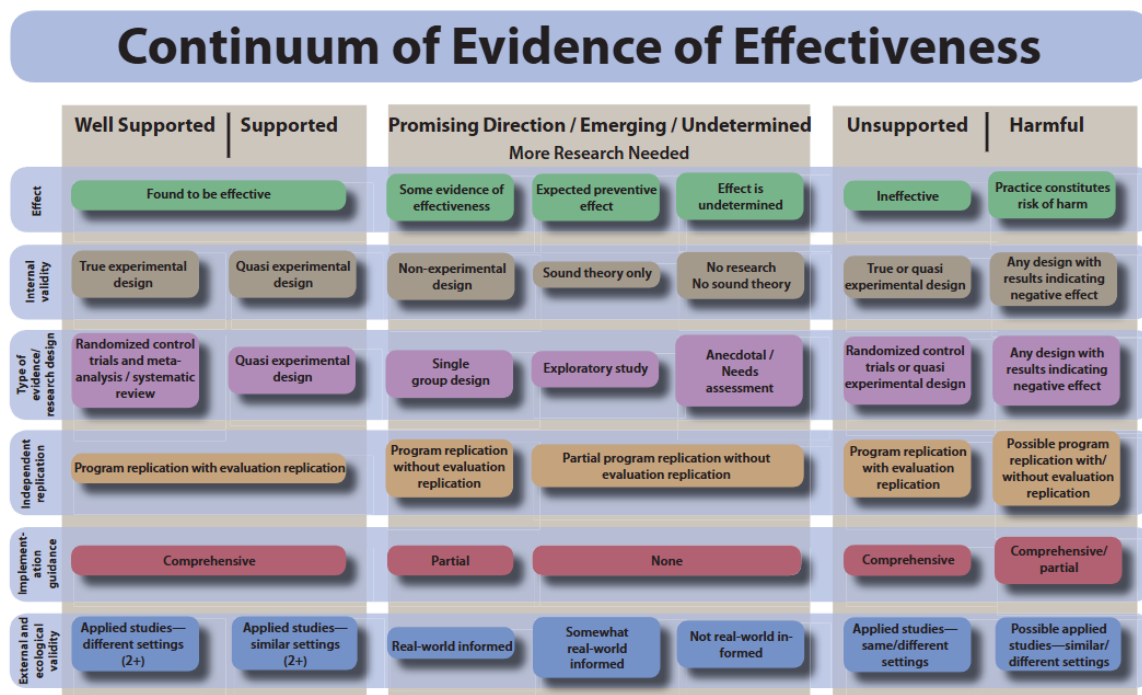
Effectiviteit beschrijft **in welke mate een interventie de doelstellingen behaalt of naar verwachting zal behalen** (25). In de context van gezondheidsbevordering verwijst dit naar de mate waarin een interventie daadwerkelijk leidt tot verbeteringen in gezondheidsgedrag, gezondheidsuitkomsten of determinanten van gezondheid bij de doelgroep waarvoor het bedoeld is.

Effectiviteitsonderzoek heeft tot doel om de resultaten van een interventie op een systematische en betrouwbare manier in kaart te brengen. Dit gebeurt doorgaans aan de hand van vooraf vastgelegde indicatoren die concreet en meetbaar zijn, zoals veranderingen in kennis, houding, gedrag of gezondheidstoestand. In combinatie met procesevaluatie gaat het hierbij niet enkel over het vaststellen van

een effect, maar ook om de vraag voor wie, onder welke omstandigheden en via welke mechanismen de interventie werkt.

Effectiviteit bevindt zich op een continuüm dat varieert van schadelijk tot evidence-based effectief. Het continuüm van het National Center for Injury Prevention and Control uit de VS (26) onderscheidt verschillende dimensies die het niveau van effectiviteit bepalen (zie Figuur 1). Een interventie is effectief wanneer meerdere goed uitgevoerde (quasi-)experimentele studies consistent gunstige effecten hebben aangetoond. Wanneer er aanwijzingen zijn voor negatieve effecten, ongeacht de onderzoeksopzet, wordt een interventie als schadelijk beschouwd. Tussen deze uitersten liggen verschillende gradaties. Sommige interventies tonen enige evidentie van effectiviteit: ze hebben een goede theorie en tonen positieve resultaten in één of enkele evaluaties (bv. pre-post studie in één groep). Andere strategieën hebben een 'verwachte preventieve werking' op basis van theoretische onderbouwing en exploratieve evaluaties. Een interventie wordt pas beschouwd als 'niet effectief' wanneer meerdere (quasi-)experimentele studies aantonen dat er geen gunstige effecten zijn. Het concept van 'best available evidence' erkent bovendien dat evidentie voortdurend evolueert. Zwak of onzeker bewijs, zoals bij interventies die als 'onbepaald' of 'veelbelovend' worden beschouwd, kan na verloop van tijd overgaan tot bewijs van (in)effectiviteit.

Figuur 1 Continuüm van effectiviteit volgens the National Center for Injury Prevention and Control. Overgenomen van: Puddy, Richard W. and Wilkins, Natalie "Understanding Evidence Part 1: Best Available Research Evidence. A Guide to the Continuum of Evidence of Effectiveness", 2011



5 Implementatie

Waar effectiviteitsonderzoek onderzoekt of een interventie werkt onder gecontroleerde omstandigheden, bekijkt implementatieonderzoek *hoe* een interventie wordt ingevoerd, geaccepteerd en duurzaam ingebed in de praktijk (27). Implementatie kan worden gedefinieerd als het proces waarbij een interventie doelgericht wordt ingevoerd en geïntegreerd in de dagelijkse praktijk en het gedrag van zowel gebruikers als intermediairs, met het oog op duurzame en grootschalige toepassing die leidt tot verbeterde gezondheidsuitkomsten.

Implementatie wordt beïnvloed door verschillende factoren. Enerzijds spelen top-down elementen zoals beleid, richtlijnen en financieringsmechanismen een rol, terwijl anderzijds bottom-up factoren zoals de percepties, ervaringen en bereidheid van intermediairs en gebruikers van belang zijn. Daarnaast beïnvloeden contextuele kenmerken van organisaties en/of zorgsystemen de mate waarin implementatie slaagt. Cruciaal zijn bovendien de eigenschappen van de interventie zelf, waaronder acceptatie, fidelity (trouw aan de oorspronkelijke opzet), haalbaarheid, schaalbaarheid en duurzaamheid (20, 28).

Traditioneel worden effectiviteit en implementatie in opeenvolgende fasen onderzocht: eerst wordt de effectiviteit vastgesteld, vervolgens wordt onderzocht hoe zij succesvol kan worden ingevoerd. Deze gefaseerde aanpak kan leiden tot vertraging in het toepassen van bewezen interventies en biedt weinig inzicht in de wisselwerking tussen effectiviteit en implementatie in verschillende contexten (29).

Om deze kloof te overbruggen, werden **effectiveness-implementation hybrid designs** ontwikkeld (30). Dit type studie hanteert vanaf het begin een dubbele focus: het gelijktijdig beoordelen van de effectiviteit van een interventie en de implementatie daarvan. Er zijn drie hoofdtypen van deze hybride designs, elk met een eigen focus en toepassingsgebied.

- **Hybrid Type 1** richt zich **primair op het testen van de effectiviteit van een interventie**, maar verzamelt parallel gegevens over de implementatie, zoals haalbaarheid en acceptatie. Dit kan bijvoorbeeld een studie naar een beweegprogramma voor ouderen zijn, waarbij niet alleen wordt gekeken naar fysieke gezondheidseffecten, maar ook naar praktische barrières voor blijvende deelname aan het dagelijks leven. Door relatief eenvoudige aanpassingen aan een standaard effectiviteitsstudie kan belangrijke informatie over uitvoerbaarheid en opschalingsmogelijkheden worden verkregen.
- **Hybrid Type 2** designs **testen interventies en implementatiestrategieën simultaan**, wat leidt tot een realistischer beeld van effectiviteit en snellere vertaling van evidentie naar de praktijk. Denk hierbij aan een voedingsinterventie voor diabetespreventie, gecombineerd met training van diëtisten en huisartsen om het programma effectief uit te voeren. Deze aanpak is geschikt wanneer er al sterke bewijsbasis is voor zowel de interventie zelf als de implementatiestrategie.
- **Hybrid Type 3** legt de **nadruk op het testen van implementatiestrategieën**, terwijl effecten secundair worden gemonitord. Dit is relevant wanneer de effectiviteit van de interventie al goed is vastgesteld, maar de uitdaging ligt in de opschaling en brede invoering. Een voorbeeld is de implementatie van een vaccinatieprogramma waarbij verschillende implementatiestrategieën, zoals schoolprogramma's versus huisartsbenadering, worden getest. Het blijft hierbij belangrijk om de uitkomsten te monitoren om zo de impact van implementatie in de praktijk te begrijpen.

Bernet et al. (20) onderstrepen dat hybrid designs essentieel zijn om zowel effectiviteit als implementatie in realistische settings te begrijpen. Hoewel hybrid designs veel voordelen bieden, zoals een snellere kennisbenutting en een realistischer beeld van interventies in praktijkomgevingen, wijst de

literatuur ook op belangrijke aandachtspunten. Het combineren van effectiviteits- en implementatie-onderzoek vraagt een complexe studieopzet met voldoende aandacht voor beide onderdelen. Dit kan methodologische uitdagingen met zich meebrengen en vereist zorgvuldige interpretatie van de resultaten.

6 Impact

Impact verwijst naar de bredere en meer duurzame veranderingen die het gevolg zijn van een interventie. Het begrip omvat zowel directe als indirecte, geplande als ongeplande, en positieve of negatieve effecten die zich voordoen op korte, middellange of lange termijn (31). Waar effectiviteit doorgaans kijkt naar het behalen van de vooraf vastgestelde doelstellingen, gaat impact verder door ook te beoordelen of de interventie structurele veranderingen teweegbrengt in gezondheid, gedrag, of maatschappelijke omstandigheden (9).

Impact wordt vaak onderscheiden van effectiviteit doordat het zich richt op de lange termijn en de mate waarin resultaten een blijvende bijdrage leveren aan de gezondheid van individuen en gemeenschappen (25). In de context van gezondheidsbevordering gaat dit bijvoorbeeld over een blijvende afname van rookgedrag in de samenleving of een duurzame daling van obesitasprevalentie. Daarnaast kan impact ook betrekking hebben op maatschappelijke effecten, zoals het verkleinen van gezondheidsverschillen, het verbeteren van sociale cohesie, of het stimuleren van een gezondere fysieke leefomgeving. Dit bredere perspectief maakt impact tot een multidimensionaal concept (32).

Een belangrijk aspect is dat impact niet altijd positief hoeft te zijn. Interventies kunnen ook ongewenste neveneffecten hebben. Zo kan een gezondheidsbevorderende campagne die sterk inzet op persoonlijke verantwoordelijkheid leiden tot stigmatisering van mensen die er niet in slagen hun gedrag te veranderen (33). Ook kan een interventie die vooral de reeds goed geïnformeerde en sociaal sterke groepen bereikt, bijdragen aan het vergroten van gezondheidsongelijkheden (34). Een kritische beoordeling van impact vraagt daarom om aandacht voor zowel de gewenste als de ongewenste, beoogde en onverwachte effecten.

Binnen een meer klassiek perspectief wordt impact vaak sterk verbonden aan causaliteit en bewijsvoering. Een relatief nieuwe benadering, contribution analysis, legt hier een ander accent (35). Deze aanpak erkent dat impact zelden door één enkele oorzaak ontstaat. De kernvraag is niet langer of een interventie de impact heeft veroorzaakt, maar of het aannemelijk is dat het daaraan een betekenisvolle bijdrage heeft geleverd (zie Hoofdstuk 5.6 Contribution Analysis).

Hoofdstuk 4

Randvoorwaarden voor de evaluatie van gezondheidsbevordering

De kwaliteit en bruikbaarheid van evaluaties in de gezondheidsbevordering worden in belangrijke mate bepaald door de voorwaarden waarbinnen zij tot stand komen. Ook wanneer interventies methodologisch degelijk worden opgezet, kunnen gebrekkige randvoorwaarden, zoals onvoldoende capaciteit, onduidelijke governance, beperkte datakwaliteit of onvoldoende aandacht voor ethiek en gelijkheid, de waarde en interpretatie van evaluatieresultaten ondermijnen. Dit hoofdstuk bespreekt daarom de kernvoorwaarden die nodig zijn om tot robuuste, relevante en rechtvaardige evaluaties te komen. Daarbij wordt aandacht besteed aan governance en samenwerking, evaluatiecapaciteit op verschillende niveaus, datakwaliteit en -beheer, ethische en privacygerelateerde aspecten, en het belang van systeemdenken, participatie en aandacht voor ongelijkheden.

1 Evaluatie van gezondheidsbevordering: 'een verplichte last'?

Hoewel er in de literatuur brede consensus bestaat over het belang van evaluatie als integraal onderdeel van programmaontwikkeling binnen de gezondheidsbevordering, blijft evaluatie in de praktijk vaak beperkt tot een administratieve verplichting om te voldoen aan subsidievoorwaarden. Hierdoor wordt vooral gebruik gemaakt van eenvoudig te verzamelen gegevens, zoals deelnemersaantallen of tevredenheidsscores, terwijl meer diepgaande evaluaties die inzicht geven in effectiviteit, mechanismen, uitvoering en context relatief weinig worden uitgevoerd. Australische onderzoekers van het *Centre for Health Promotion Research* beschrijven drie belangrijke groepen van drempels die effectieve evaluaties kunnen belemmeren (36):

1. Organisatorische factoren:

- Evaluatie wordt gezien als een bijkomstige verplichting in plaats van een integraal onderdeel van programmaplanning, waardoor het doel om het programma te verbeteren op basis van de evaluatie verloren gaat.
- De middelen voor gezondheidsbevordering zijn doorgaans beperkt, waardoor de meerderheid van de beschikbare middelen naar de interventie zelf gaat ten koste van de evaluatie.
- Financiers onderschatten vaak de tijd en kosten die nodig zijn voor goede evaluatie, waardoor de middelen en beschikbare tijd in projecten voor evaluaties beperkt zijn. Mede hierdoor worden uitvoerders soms gedwongen om voor uitkomsten op korte termijn of operationele uitkomsten te kiezen, en kan men binnen de looptijd van het project de impact van een programma onvoldoende aantonen.
- Evaluatie wordt beperkt gewaardeerd en beloond (o.a. gebrek aan incentives voor goede evaluatie).

2. Capaciteitsfactoren:

- Gebrek aan kennis en vaardigheden bij actoren in gezondheidsbevordering om gedegen evaluaties uit te voeren en de resultaten te interpreteren (o.a. aanvragen van ethisch comité, opstellen van vragenlijsten, uitvoeren van analyses).

- Politieke druk en de verwachting om snel te handelen bemoeilijken diepgaande evaluatie en zorgen ervoor dat er vaak pragmatische keuzes worden gemaakt (bv. weinig inspanningen leveren om moeilijk te bereiken groepen te betrekken).

3. **Translationele factoren (vertaling van evaluatieresultaten naar de praktijk):**

- Er is weinig aandacht voor het delen van geleerde lessen en om kennis om te zetten in verbeterde praktijkvoering. Ook ontbreekt het vaak aan vaardigheden om evaluatiebevindingen te publiceren.
- Soms is er terughoudendheid om te evalueren uit vrees voor ongunstige resultaten die kunnen leiden tot stopzetting van het programma. Daarom moet er mogelijkheid zijn om het programma te verbeteren bij tegenvallende resultaten.

De auteurs stellen dat evaluatie niet slechts een afsluitende activiteit moet zijn, maar een cyclisch proces dat al in de ontwerpfasen van een project wordt ingebed en doorlopend inzicht biedt in de effectiviteit van een programma. Ze benadrukken de nood aan een multi-levelstrategie en gedeelde verantwoordelijkheid van beleidsmakers, financiers, organisaties, onderzoekers en praktijkwerkers om de drempels aan te pakken en beter te kunnen evalueren. Enkele concrete hefbomen zijn:

- Goede beoordeling van wanneer en waarom een evaluatie nodig is (*evaluability assessment*), om over-evaluatie te voorkomen.
- Stimuleren van directe uitwisseling en samenwerkingen tussen onderzoekers, praktijkwerkers, financiers en beleid.
- **Ondersteuning** door onderzoekers, beleidsmakers en financiers in het **ontwikkelen van meetbare en goede evaluatieplannen** die doelgericht inzicht geven in de effectiviteit en mechanismen van interventies, zodat actoren in de gezondheidsbevordering de resultaten beter kunnen benutten.
- **Leiderschap dat evaluatie waardeert** en dit beschouwt als een onmisbaar instrument voor de kwaliteitsverbetering van interventies.
- **Financiering met budgetten die specifiek zijn bestemd voor evaluatie.** Sommige subsidies verplichten dat 10-15% van het programmabudget aan evaluatie wordt besteed, inclusief kennisverspreiding en vertaling van de bevindingen naar praktijkverbetering.
- **Gerichte training en (gezamenlijke) leertrajecten** voor organisaties voor het versterken van kennis en vaardigheden inzake evaluatie.
- **Flexibele evaluatiebenaderingen die aansluiten bij de praktijk:** financiers en beleidsmakers dienen te erkennen dat niet alle programma's via rigide, experimentele methoden geëvalueerd kunnen worden. Alternatieve evaluatiemethoden zoals participatief actieonderzoek of benaderingen met storytellingen of kunst kunnen eventueel waardevol zijn, in het bijzonder bij kwetsbare groepen.

2 Governance

In de evaluatie van gezondheidsbevordering verwijst governance naar de manier waarop verantwoordelijkheden, samenwerking en besluitvorming zijn georganiseerd. Beleidsmakers en financiers bepalen doorgaans het evaluatiekader en de randvoorwaarden, terwijl uitvoerders cruciale contextinformatie aanleveren en externe onderzoekers de methodologische kwaliteit en onafhankelijkheid waarborgen. Hoofdstuk 9 gaat dieper in op de samenwerking en rolverdeling tussen deze actoren.

Het TAPIC-kader van de WHO verdeelt governance in gezondheidsbeleid in vijf kerndomeinen: Transparantie, Verantwoording (Accountability), Participatie, Integriteit en Beleidscapaciteit (37). Dit biedt handvatten om besluitvorming en implementatie te verbeteren. Elk van deze domeinen kan versterkt of verstoord worden door onder- of overaanwezigheid. Zo kan te veel verantwoording leiden tot overbodige bureaucratie en kan participatie soms zo complex worden ingericht dat enkel goed gefinancierde lobbygroepen effectief kunnen deelnemen. Voor elk domein worden concrete instrumenten voorgesteld om deze te versterken:

Domein	Betekenis	Voorbeelden van instrumenten / uitwerkingen
Transparantie	Helderheid over besluitvorming, argumentatie en verantwoordelijke actoren.	Inspecties en toezichtcommissies, openbare rapportage, opvraagbare informatie, open vergaderingen, duidelijke communicatie.
Accountability (verantwoording)	Actoren leggen uit en verantwoorden hun keuzes en handelingen.	Rapportageverplichtingen, verplichte evaluaties, verslaggeving aan financiers, competitieve financieringsmechanismen.
Participatie	Betrokkenen krijgen de mogelijkheid om inspraak te hebben en gehoord te worden.	Vertegenwoordigers van specifieke groepen, stakeholderconsultaties, adviescommissies, verkiezingen, samenwerking met belangenverenigingen.
Integriteit	Heldere taakomschrijvingen, regels en procedures die eerlijk en zorgvuldig handelen ondersteunen.	Duidelijke mandaten, interne en externe audits, transparante wervingsprocedures, documentbeheer.
Beleidscapaciteit	Beschikbaarheid van middelen, expertise en systemen om beleid te ontwikkelen, uit te voeren en bij te sturen.	Inzet van experts, opleiding en training, strategische analyses, evenwichtige keuze tussen uitbesteding en interne uitvoering.

3 Evaluatiecapaciteit

Evaluatiecapaciteit verwijst naar het vermogen van individuen, organisaties en systemen om interventies binnen de gezondheidsbevordering op een systematische en kwaliteitsvolle manier te evalueren (38). In de literatuur wordt daarbij zowel gewezen op de beschikbaarheid van middelen en expertise als op beleidsstructuren en evaluatiecultuur (39). Evaluation capacity building omvat alle inspanningen om kennis, vaardigheden, motivatie en organisatorische randvoorwaarden te versterken (40).

Onderzoek toont dat effectieve capaciteitsopbouw steeds meerdere niveaus combineert. Op individueel niveau werkt vooral een mix van training, participatie in evaluatieprocessen en begeleiding. Op organisatieniveau zijn ondersteuning van leidinggevendenden, ingebedde evaluatieprocessen, een leer-cultuur en voldoende middelen bepalend. Eenmalige initiatieven volstaan niet: duurzame capaciteits-versterking vraagt tijd, herhaling en een ondersteunende omgeving waarin samenwerking en netwerken centraal staan (40).

3.1 Individuele evaluatiecapaciteit: vereiste expertise van evaluatoren

Ondanks het belang van individuele evaluatiecapaciteit bestaat er weinig consensus over welke vaardigheden en kennis evaluatoren precies nodig hebben. Ignacio et al (41) identificeerden 30 kerncompetenties, gegroepeerd in beschrijvende en analytische vaardigheden. Beschrijvende competenties omvatten onder meer projectmanagement, dataverzameling en -analyse, methodologische keuzes en communicatieve vaardigheden. Analytische competenties richten zich op het formuleren van evaluatievragen, het ontwikkelen van programmatheorieën, het interpreteren van bevindingen en het integreren van verschillende perspectieven.

Deze competenties illustreren dat evaluatie meer vereist dan technische vaardigheden alleen. Evaluatoren moeten ook de context begrijpen, met stakeholders samenwerken en evaluatiemodellen kunnen ontwikkelen die lokale doelen en realiteiten weerspiegelen. De Canadese federale overheid¹⁰ legt dergelijke competenties expliciet vast per functieniveau (zie Bijlage 2), wat systematische kwaliteitsbewaking ondersteunt.

3.2 Organisatorische evaluatiecapaciteit

3.2.1 Raamwerk van Bourgeois en Cousins (2013)

Organisaties spelen een cruciale rol in het mogelijk maken van kwaliteitsvolle evaluaties. Het raamwerk van Bourgeois en Cousins (42) beschrijft organisatorische evaluatiecapaciteit aan de hand van volgende zes dimensies:

- **Menselijke middelen:** Dit gaat over het personeel dat instaat voor evaluatie: hun aantal, vaardigheden, loopbaanontwikkeling en leiderschap. Zonder voldoende en goed opgeleide mensen, kan evaluatie niet op een kwaliteitsvolle manier gebeuren.

¹⁰ <https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/evaluation-government-canada/evaluation-competencies.html>

- **Organisatorische middelen:** Hier draait het om de steunstructuren: budget, toegang tot gegevens, ICT-systemen, en een duidelijke governance. Een organisatie kan pas duurzaam evalueren als evaluatie structureel ingebed is in de middelen en processen.
- **Evaluatieplanning & uitvoering:** Dit gaat over hoe evaluaties gepland en uitgevoerd worden. Beschikt de organisatie over een meerjarenplan? Worden risico's en prioriteiten in kaart gebracht? Hoe wordt informatie gedeeld? Hoe sterk zijn de netwerken met interne en externe partners?
- **Evaluatiekennis:** Evaluatiekennis betekent dat medewerkers begrijpen wat evaluatie is, hoe het werkt en hoe resultaten geïnterpreteerd kunnen worden. Het gaat dus om kennis en kunde, evenals om de vaardigheid om bevindingen toe te passen in de dagelijkse werking.
- **Besluitvorming:** De kernvraag hier is: worden evaluaties ook echt gebruikt? Evaluaties hebben pas meerwaarde als ze een rol spelen in strategische keuzes, budgetverdeling en programma-vormgeving. Dit vraagt ook steun en betrokkenheid van leidinggevenden.
- **Leer- en verbeterkansen:** Evaluaties moeten leiden tot een beter beleid, effectievere programma's en een organisatie die leert uit ervaring. Deze dimensie gaat na of er een leercultuur bestaat, waarin systematisch reflectie en verbetering centraal staan.

Het raamwerk is helpend voor organisaties om hun sterktes en tekortkomingen op vlak van evaluatiecapaciteit in kaart te brengen en vervolgens systematisch in te zetten op capaciteitsopbouw (42). Bij het raamwerk hoort ook een beoordelingsmatrix die voor elke dimensie vier niveaus van evaluatiecapaciteit onderscheidt (laag, ontwikkelend, gemiddeld, voorbeeldig) (zie Bijlage 3).

3.2.2 Actieonderzoek uit Ontario (Canada) om organisatorische evaluatiecapaciteit te verbeteren

Een actieonderzoek uitgevoerd in tien public health units in Ontario (Canada) toonde dat succesvolle strategieën om evaluatiecapaciteit in de organisatie te versterken steunt op vier mechanismen: leiderschap dat evaluatie actief ondersteunt, een organisatorische omgeving die systematisch gebruikmaakt van resultaten, versterking van individuele kennis en vaardigheden, en een helder, participatief ontwikkeld evaluatiekader (Bourgeois et al., 2018).

Daarnaast formuleerde het onderzoek vijf aanbevelingen voor een succesvolle versterking van evaluatiecapaciteit:

- Behoeftanalyse is cruciaal, bij voorkeur aan de hand van een zelfevaluatie-instrument helpt bij het kiezen van passende strategieën.
- Participatie van medewerkers en managers vergroot het draagvlak en de effectiviteit van strategieën.
- Strategieën die meerdere elementen of mechanismen combineren zijn krachtiger.
- Leiders op alle niveaus spelen een sleutelrol in het verankeren van evaluatiepraktijken.
- Wees realistisch: capaciteitsopbouw kost tijd, middelen en doorzetting. Kleine stappen en langetermijndenken zijn nodig.

3.2.3 Instrument om evaluatiecapaciteit te meten

De vragenlijst 'Evaluation Practice Analysis Survey' (EPAS) is een gevalideerd instrument dat ontwikkeld werd om evaluatiecapaciteit op individueel, organisatorisch en systeemniveau in kaart te brengen (43). Het instrument omvat 25 schalen met in totaal 125 items, en is voorlopig enkel in het Engels beschikbaar (zie Bijlage 4).

Op basis van de EPAS-vragenlijst ontwikkelden (38)) een conceptueel kader dat aantoont dat vooral organisatorische factoren, zoals leiderschap, evaluatiecultuur en systemen, bepalend zijn voor evaluatiecapaciteit. Individuele kennis en motivatie spelen ook wel een rol, maar hun impact wordt sterk bepaald door de ondersteuning en verwachtingen binnen de organisatie. Systeemfactoren, zoals beleidsprioriteiten en financieringsmodellen, vormen daarbij de structurele randvoorwaarden waarbinnen organisaties opereren.

Kortom organisaties zijn de spil in het versterken of belemmeren van evaluatiecapaciteit. Medewerkers kunnen gemotiveerd en bekwaam zijn, maar hun effect is beperkt als organisaties geen middelen, tijd of steunende structuren beschikbaar stellen. Tegelijkertijd bepaalt het bredere beleid de voorwaarden waaronder organisaties opereren.

3.3 Evaluatiecapaciteit op macroniveau: Lessen uit het evaluatiebeleid in Canada

Het Canadese evaluatiebeleid laat zien hoe beleid evaluatiecapaciteit op systeemniveau kan versterken. Met het *Policy on Evaluation* (2009) en later het flexibelere *Policy on Results* (2016) evolueerde Canada van rigide verplichtingen – zoals het vijfjaarlijks evalueren van alle programma's – naar een benadering die departementen meer ruimte geeft om relevante en haalbare evaluaties uit te voeren. Deze verschuiving illustreert het belang van realistische eisen die aansluiten bij de uitvoeringscapaciteit van organisaties (44).

Daarnaast toont de Canadese ervaring dat de kwaliteit en bruikbaarheid van evaluaties niet vanzelfsprekend leiden tot beter strategisch gebruik. Evaluaties worden vaak ingezet voor verantwoording en financieringsaanvragen, maar minder voor strategische besluitvorming, prioritering en budgetverdeling.

4 Datakwaliteit - en management

Betrouwbare data zijn essentieel voor het evalueren van interventies en het nemen van beleidsbeslissingen. Voor een betrouwbare impactevaluatie van gezondheidsbevorderende interventies zijn zowel kwaliteitsbewaking als datagovernance belangrijk.

Kwaliteitsbewaking zorgt ervoor dat de interventie wordt uitgevoerd zoals bedoeld en dat de verzamelde gegevens valide, betrouwbaar en reproduceerbaar zijn. Het focus op drie dimensies (45):

- **Kwaliteit van de data zelf:** dit verwijst naar de waarden en structuren in databases of tabellen. Het gaat om de inhoud en vorm van de gegevens zelf. Bv. volledigheid, nauwkeurigheid, tijdigheid en bruikbaarheid.
- **Kwaliteit van datagebruik:** dit verwijst naar de mate waarin gegevens worden gebruikt door eindgebruikers, zoals beleidsmakers, onderzoekers of preventieactoren. Bv. gebruikerstevredenheid, gebruiksintentie, disseminatie van data.

- **Kwaliteit van dataverzameling:** dit omvat alle stappen voorafgaand aan het opslaan van data, zoals het genereren, verzamelen, beschrijven en onderhouden van gegevens. Bv. data flow, metadata-management, feedbackmechanismen.

Datakwaliteit is een belangrijk aspect van **datagovernance**. Dit laatste verwijst naar het geheel aan afspraken, procedures en verantwoordelijkheden rond bescherming, beheer, toegankelijkheid en hergebruik van data. Een goed ingericht datagovernanceproces waarborgt dat gegevens bruikbaar zijn voor analyse en toekomstige beleidsvorming, en voorkomt dat gegevens versnipperd raken of inconsistent worden gedefinieerd.

De '**FAIR Guiding Principles**' bieden richtlijnen om de vindbaarheid (Findability), toegankelijkheid (Accessibility), interoperabiliteit (Interoperability) en herbruikbaarheid (Reuse) van onderzoeksdata te verbeteren (<https://www.go-fair.org/fair-principles/>) (46). Belangrijke elementen voor FAIR-data zijn o.a. rijke metadata en duidelijke documentatie, het gebruik van gestandaardiseerde bestandsformaten en het toekennen van 'persistente identificatoren' (zoals doi's) aan data-objecten.

FAIR ondersteunt de reproduceerbaarheid en validatie van resultaten, zodat andere onderzoekers of beleidsmakers analyses kunnen repliceren of verifiëren. Belangrijk om te benadrukken is dat FAIR-data niet hetzelfde is als open data. FAIR richt zich op het optimaal beschikbaar, begrijpelijk en herbruikbaar maken van data, binnen de grenzen van wat juridisch en ethisch verantwoord is. Het FAIR-principe 'Accessible' betekent dus niet dat iedereen vrij toegang heeft, maar dat duidelijk is hoe en onder welke voorwaarden toegang mogelijk is.

Een **Data Management Plan** (DMP) is cruciaal: hierin wordt vooraf vastgelegd hoe data worden verzameld, wie verantwoordelijk is voor beheer, hoe metadata worden aangemaakt, wie toegang krijgt, en onder welke voorwaarden hergebruik mogelijk is. Daarnaast is het belangrijk gebruik te maken van betrouwbare data-repositories. Door deze principes te integreren in het onderzoeksproces wordt de kwaliteit, bruikbaarheid en impact van gezondheidsbevorderende interventies aanzienlijk versterkt.

5 Privacy en ethische aspecten

De ethische dimensies van interventieonderzoek binnen de volksgezondheid blijven vaak onderbelicht, ondanks de groeiende aandacht voor ethiek in onderzoek en praktijk. **Ethische reflexiviteit** omvat het anticiperen op ethische kwesties, het reflecteren op ethische momenten tijdens het onderzoek, en het overwegen van de sociale en politieke implicaties van het onderzoek (47). Ethische uitdagingen in onderzoek naar gezondheidsbevordering zijn:

- **Geïnformeerde toestemming**, wat betekent dat deelnemers helder en volledig geïnformeerd moeten worden over het doel, de inhoud, de mogelijke risico's en de gevolgen van hun deelname aan het onderzoek. Hoewel dit een fundamenteel principe, is volledige informatieverstrekking vaak moeilijk bij complexe real-life interventies.
- **Categorisering van doelgroepen** kan leiden tot stigmatisering. Het definiëren van een probleem impliceert wie het probleem heeft en wie moet veranderen, wat ethische implicaties heeft. Interventies kunnen onbedoeld bijdragen aan **stereotypering en stigmatisering**, vooral bij kwetsbare groepen (bv. mensen met overgewicht).

- **Analyse en disseminatie** van onderzoeksresultaten brengt ethische verantwoordelijkheden met zich mee. Het niet rapporteren van nulresultaten of negatieve effecten kan de integriteit van het onderzoek schaden.

Daarnaast spelen **ethische commissies** een belangrijke rol in het beoordelen van onderzoeksvoorstellen. Ze zorgen ervoor dat studies voldoen aan richtlijnen en nationale wetgeving, waaronder de GDPR/AVG (Algemene Verordening Gegevensbescherming). Deze regelgeving vereist dat persoonsgegevens zorgvuldig worden behandeld, geanonimiseerd waar nodig, en enkel worden gebruikt voor het beoogde onderzoeksdoel.

Een ander ethisch aandachtspunt betreft het gebruik van **controlegroepen** in interventieonderzoek. Hoewel het gebruik van controlegroepen het onderzoeksdesign versterken, roept het ethische vragen op wanneer deelnemers in de controlegroep een mogelijk nuttige interventie wordt onthouden. Dit kan deels worden ondervangen door de interventie op een later moment alsnog aan de controlegroep aan te bieden (deze designs worden besproken in Hoofdstuk 6).

Box 1: Ethische code voor en door de Vlaamse sector preventieve gezondheid

De ethische code voor en door de Vlaamse sector preventieve gezondheid werd in 2019 ontwikkeld en vormt een gezamenlijk moreel kompas voor organisaties die werken aan gezondheidsbevordering en preventie¹¹. De code bevat twaalf ethische principes en werd onderschreven door 35 organisaties uit de sector. Ze vraagt dat organisaties systematisch reflecteren over de ethische implicaties van hun interventies: zijn ze niet stigmatiserend of discriminerend, waarborgen ze vrijwillige deelname, en zijn ze toegankelijk voor alle doelgroepen?

De principes zijn gegroepeerd in drie categorieën:

- 1) **Deontologisch handelen:** Dit omvat transparantie naar alle betrokkenen, publieke verantwoording en het naleven van relevante wetgeving, waaronder de GDPR.
- 2) **Sociaal en rechtvaardig handelen:** Organisaties moeten gemeenschapsgericht werken, gelijke toegang garanderen, negatieve stereotypering vermijden en respect tonen voor autonomie en onafhankelijkheid. Ook duurzaamheidsprincipes in productie en aankoop maken deel uit van deze categorie.
- 3) **Kernprincipes van gezondheidsbevordering:** Deze verwijzen naar het nastreven van gezondheidswinst, het verkleinen van gezondheidsongelijkheid, en het versterken van gezondheidsvaardigheden en omgevingsfactoren.

Het Vlaams Instituut Gezond Leven biedt ondersteuning bij de toepassing van de code. De begeleiding verloopt stapsgewijs:

- 1) Een praktisch werkinstrument vertaalt de ethische criteria naar concrete toetsingsvragen voor intern gebruik.
- 2) Organisaties kunnen informeel 1-op-1 advies krijgen van medewerkers uit de sector over ethische dilemma's rond interventies of partnerschappen.
- 3) Indien nodig kan een ethische adviescommissie worden samengeroepen. Deze bestaat uit interne en eventueel externe experts die een niet-bindend advies formuleren.

¹¹ Ethische code voor de sector gezondheidsbevordering en ziektepreventie in Vlaanderen
<https://www.gezondleven.be/files/Gezond-Leven-ethische-code.pdf>

6 Principes voor systeemdenken doorheen het ganse evaluatieproces

Systeemdenken is een benadering die complexe situaties benadert door aandacht te besteden aan samenhang, dynamiek en perspectieven. Evaluaties vinden vaak plaats in complexe sociale systemen waarin interventies, context en actoren elkaar beïnvloeden. Traditionele evaluatiemethoden schieten tekort in het begrijpen van deze complexiteit.

De *Systems in Evaluation Topical Interest Group (SETIG)* van de American Evaluation Association biedt richtlijnen voor het toepassen van systeemdenken in evaluatiepraktijken (48). Het doel is om evaluatoren en stakeholders te ondersteunen bij het integreren van kernconcepten uit systeemdenken in alle fasen van evaluatie. Ze stellen vier kernprincipes voor die elkaar wederzijds beïnvloeden en samen een systemische benadering vormen:

- **Interrelaties:** richt zich op het identificeren en begrijpen van de relaties tussen elementen binnen en buiten het systeem. Deze relaties zijn niet statisch, maar dynamisch en beïnvloeden zowel het evaluatieobject als de evaluatie zelf. Evaluatoren moeten interrelaties in kaart brengen en transparant zijn over keuzes. Dit principe helpt om verbanden en verandertheorieën beter te begrijpen.
- **Perspectieven:** benadrukt het belang van diverse perspectieven. Elke stakeholder ziet een situatie op een bepaalde manier, wat invloed heeft op hoe grenzen worden getrokken en hoe resultaten worden geïnterpreteerd. Evaluaties moeten actief op zoek gaan naar verschillende perspectieven en rekening houden met machtsverhoudingen.
- **Grenzen:** gaat over het expliciet maken van keuzes over wat binnen en buiten de evaluatie valt. Omdat niet alles geëvalueerd kan worden, moeten evaluatoren transparant zijn over deze keuzes en hun consequenties. Grenzen beïnvloeden wie betrokken is, welke data wordt verzameld en hoe conclusies worden getrokken. Het principe vraagt om reflectie op ethische implicaties, omdat grenzen bepalen wie of wat wordt uitgesloten of gemarginaliseerd.
- **Dynamieken:** richt zich op patronen van verandering binnen systemen en hun invloed op de evaluatie. Complexe systemen zijn dynamisch, met interacties die vaak leiden tot onvoorspelbare uitkomsten. Evaluatieplannen moeten flexibel zijn en inspelen op nieuwe ontwikkelingen. Dit principe stimuleert een houding van continu leren en aanpassen, zodat evaluaties relevant blijven in een veranderende context.

7 Participatie en co-creatie

Participatieve benaderingen en co-creatieve methoden zijn steeds belangrijker geworden in onderzoek naar gezondheidsbevordering. Ze maken het mogelijk om complexe gezondheidsvraagstukken beter te begrijpen door de perspectieven en ervaringen van diverse betrokkenen (bv. praktijkwerkers, beleidsmedewerkers, burgers, kwetsbare doelgroepen) te integreren in het onderzoeksproces. Stakeholders kunnen problemen anders definiëren, oorzaken realistischer duiden en haalbare oplossingen aanwijzen. Hun deelname versterkt eveneens het draagvlak van beslissingen. Co-creatie wordt dus gezien als een manier om onderzoek te democratiseren en interventies beter af te stemmen op de context en behoeften van populaties en (vaak kwetsbare) groepen. Andere voordelen zijn o.a. empowerment en eigenaarschap van deelnemers, verhoogde innovatie en meer impact van interventies (49). Participatie maakt ook dat evaluaties oog hebben voor de wisselwerking tussen interventie en context, en dat de perspectieven van lokale stakeholders kunnen leiden tot bijsturing van de interventies (50). Op basis

van een systematische review kwamen vier kernprocessen naar voren die samen co-creatie van nieuwe kennis vormgeven (51):

- **Co-ideatie:** het gezamenlijk genereren van ideeën en het identificeren van problemen en noden.
- **Co-design:** het gezamenlijk ontwerpen van interventies of beleid, inclusief de onderzoeksmethoden.
- **Co-implementatie:** het gezamenlijk uitvoeren van de ontworpen interventie of beleidsmaatregel.
- **Co-evaluatie:** het ontwerpen of aanpassen van dataverzamelingssystemen, het gezamenlijk verzamelen, analyseren en interpreteren van de resultaten over effectiviteit en impact van de interventie.

8 Aandacht voor ongelijkheden

Een evaluatie met een expliciete focus op equity of rechtvaardigheid heeft tot doel om de verdeling van effecten tussen verschillende sociale groepen zichtbaar te maken en ongelijkheden te verkleinen. Gezondheidsbevorderende interventies zijn niet altijd even effectief voor alle bevolkingsgroepen. Soms werken ze beter bij hoger opgeleiden of mensen met meer middelen, wat kan leiden tot het vergroten van gezondheidsongelijkheden (34). Waar traditionele evaluaties vooral de gemiddelde effecten benadrukken, laat een equity-gedreven benadering zien dat deze gemiddelden kunnen verschillen tussen groepen, waardoor ongelijkheid in stand blijft of zelfs toeneemt (52). Participatie, inclusie en (cultuur)sensitiviteit zijn hierin essentieel.

Daarom werden evaluatiemodellen ontwikkeld met een focus op equity (53). Deze modellen leveren extra inspanningen om kwetsbare groepen te betrekken en helpen om te evalueren welke onderdelen van een interventie invloed hebben op wie meedoet en welke resultaten worden bereikt. Daarbij wordt specifiek gekeken naar verschillen tussen groepen, bijvoorbeeld op basis van sociaaleconomische positie, etniciteit, leeftijd of gender. Equity-gedreven evaluaties zijn vaak beter in staat om drempels inzake toegankelijkheid te signaleren, om interventies te vertalen naar de lokale context en om duurzame veranderingen op meerdere niveaus te stimuleren.

In de praktijk betekent dit dat evaluatoren vanaf het begin moeten nadenken over equity: welke groepen dreigen buitengesloten te raken, welke barrières spelen een rol in toegang en acceptatie, en hoe kunnen deze worden gemeten (54). Dit houdt ook in dat evaluatieplannen aangeven voor welke subgroepen aparte analyses worden gedaan, zowel voor effect- als voor procesindicatoren. Zo wordt duidelijk voor wie de interventie werkt en voor wie niet, waardoor onderzoekers en beleidsmakers beter kunnen inspelen op verschillen en ongelijkheden kunnen verkleinen.

9 Samenvattend overzicht van noodzakelijke randvoorwaarden voor evaluatie

Randvoorwaarden zijn de **noodzakelijke factoren** die aanwezig moeten zijn om kwaliteitsvolle evaluaties binnen de gezondheidsbevordering mogelijk te maken. Het is mede de taak van het Departement Zorg om deze voorwaarden te faciliteren. Daarnaast is er nood aan samenwerking op meerdere niveaus tussen beleid, onderzoekers, financiers en uitvoerders.

Over het algemeen dient evaluatie in gezondheidsbevordering te **evolueren van een verplichting naar een lerend en rechtvaardig proces**, dat bijdraagt aan **continue verbetering**. Evaluatie dient ingebed te worden in het ontwerp en de uitvoering van programma's, zodat leren en bijsturen structureel

verankerd raken. Evalueren **mag geen doel op zich worden**: het moet steeds doelgericht gebeuren en enkel wanneer het daadwerkelijk een meerwaarde biedt voor het versterken van kennis, beleid en praktijk. Tabel 3 vat de randvoorwaarden die aan bod komen in dit hoofdstuk samen.

Tabel 3 Overzicht van belangrijke randvoorwaarden voor evaluatie.

Randvoorwaarde	Toelichting / concretisering
Sterke governance en samenwerking tussen actoren	• Beleidsmakers en financiers bepalen het evaluatiekader en voorzien toereikende middelen. • Uitvoerders leveren contextkennis en data. • Externe onderzoekers bewaken methodologische kwaliteit en onafhankelijkheid.
Evaluatiecapaciteit op individueel niveau	• Evaluatoren beschikken over beschrijvende competenties (opzetten en uitvoeren van evaluaties, dataverzameling en -analyse) en analytische competenties (beoordelen van kwaliteit, waarde en impact van interventies).
Evaluatiecapaciteit op organisatorisch niveau	• Organisaties voeren behoeftanalyses uit om na te gaan of versterking van hun evaluatiecapaciteit nodig is. • Leiderschap is cruciaal voor het verankeren van evaluatiepraktijken. Investeren in een combinatie van strategieën is het meest effectief om evaluatiecapaciteit te versterken. • Capaciteitsopbouw vergt tijd, middelen en langetermijnvisie.
Evaluatiecapaciteit op macroniveau	• Beleidsprioriteiten en financieringsmodellen bepalen de ruimte voor evaluatie. • Centrale expertise verankert kennis binnen het beleid. • Duurzame capaciteitsopbouw vraagt stabiele financiering en continuïteit in het beleid.
Betrouwbare en kwaliteitsvolle data en datagovernance	• Heldere afspraken over beheer, beveiliging, kwaliteit, toegankelijkheid en gebruik van data zijn noodzakelijk voor betrouwbare evaluaties.
Ethische kaders en wettelijke vereisten rond privacy	• Evaluaties voldoen aan ethische principes en wetgeving, waaronder geïnformeerde toestemming en GDPR.
Aandacht voor systeemdenken	• Interventies staan nooit op zichzelf, maar maken deel uit van een complex systeem waarin context, actoren en middelen elkaar beïnvloeden. • Plaats bevindingen steeds in een groter geheel door rekening te houden met het beleid, organisatieprocessen en maatschappelijke trends. • Houd rekening met interrelaties tussen elementen binnen en buiten het systeem die elkaar wederzijds beïnvloeden. • Betrek verschillende perspectieven (relevante betrokkenen). • Baken bewust de grenzen van de evaluatie af. • Schenk aandacht aan dynamieken en onvoorspelbaarheid door voldoende flexibiliteit in te bouwen in evaluatieplannen.
Participatieve methoden en co-design van evaluaties	• Betrokkenheid van burgers, doelgroepen en intermediairs bij ontwerp en uitvoering verhoogt relevantie, kwaliteit en draagvlak van evaluaties.
Aandacht voor gezondheidsongelijkheden (equity)	Inclusie van onderbereikte groepen en analyse van verschillen tussen bevolkingsgroepen zijn essentieel om ongelijkheden zichtbaar te maken en te verkleinen.

Hoofdstuk 5

Evaluatiekaders voor gezondheidsbevorderende interventies

Het evalueren van gezondheidsbevorderende interventies vraagt om meer dan het meten van effecten alleen. Preventieve programma's worden immers ontwikkeld en uitgevoerd binnen complexe sociale, organisatorische en beleidscontexten, waardoor inzicht nodig is in zowel resultaten als de processen die tot deze resultaten leiden. Evaluatiekaders bieden hierbij houvast: ze structureren de wijze waarop interventies begrepen, onderzocht en verbeterd kunnen worden, en helpen om aannames expliciet te maken. In dit hoofdstuk geven we een overzicht van vaak gebruikte en relevante evaluatiekaders binnen de gezondheidsbevordering. We bespreken zowel kaders die focussen op programma- en theorie-gebaseerde evaluatie, als modellen die inzicht geven in implementatieprocessen en context. Samen ondersteunen ze onderzoekers en praktijkwerkers bij het ontwikkelen van robuuste evaluaties en het versterken van de kwaliteit, effectiviteit en duurzaamheid van interventies.

1 Stappen en principes in programmaevaluatie (*CDC Program Evaluation Framework 2024*)

In 2024 introduceerde de Amerikaanse *Centers for Disease Control and Prevention* (CDC) een vernieuwd en praktisch evaluatiekader voor uiteenlopende programma's binnen de publieke gezondheidszorg (55). Het CDC Program Evaluation Framework is ontworpen als een flexibel hulpmiddel dat de essentiële elementen van programmaevaluatie helder samenvat en structureert¹².

Het kader bestaat uit zes stappen met verschillende onderdelen (zie Figuur 2). Doorheen alle stappen worden drie kernprincipes geïntegreerd: (1) samenwerking met belanghebbenden, (2) aandacht voor gezondheidsongelijkheid en inclusie, en (3) het systematisch gebruik van inzichten voor programma-verbetering.

¹² CDC Program Evaluation Framework <https://www.cdc.gov/evaluation/php/evaluation-framework/index.html>

Figuur 2 Zes stappen van programmaevaluatie volgens het CDC Program Evaluation Framework 2024.



2 Theory of Change en Logic Model als strategisch en operationeel fundament voor interventieplanning- en evaluatie

Voordat een interventie geëvalueerd kan worden, is het belangrijk om helder te krijgen hoe deze geacht wordt verandering teweeg te brengen. Methoden zoals de Theory of Change en het Logic Model helpen om die redenering expliciet te maken. Een Theory of Change is een overkoepelende theorie over hoe en waarom veranderingen optreden, terwijl een Logic Model een meer praktische weergave is van de stappen, activiteiten en verwachte resultaten van een interventie. Beide kunnen elkaar aanvullen en zijn nuttig voor het plannen, uitvoeren en evalueren van interventies.

2.1 Theory of change

Een verandertheorie of Theory of Change is een model dat beschrijft hoe een programma of interventie leidt tot een reeks veranderingen die uiteindelijk het gewenste eindresultaat opleveren. Het model maakt oorzaak-gevolgrelaties expliciet en beschrijft hoe en waarom bepaalde acties tot veranderingen zouden moeten leiden. (56).

Een belangrijk kenmerk van Theory of Change is dat het een keten van resultaten toont, waarbij kortetermijnveranderingen leiden tot middellangetermijnresultaten en uiteindelijk tot langetermijndoelen. Dit model maakt ook expliciet welke aannames ten grondslag liggen aan het programma, zoals de veronderstelling dat deelnemers nieuwe kennis zullen toepassen of beleidsmakers openstaan voor verandering. Door deze aannames zichtbaar te maken, kunnen ontwikkelaars beter reflecteren op hun aanpak en mogelijke risico's herkennen (57, 58).

Theory of Change baseert zich vaak op wetenschappelijk onderzoek en eerdere ervaringen om de verbanden tussen acties en uitkomsten te onderbouwen. Een veelgebruikte methode is 'backward mapping': beginnen bij het langetermijndoel en terugwerken om te bepalen welke veranderingen daarvoor nodig zijn. Dit helpt om prioriteiten te stellen en te focussen op cruciale activiteiten (58).

Het is belangrijk te benadrukken dat een model dat logisch lijkt opgebouwd, niet automatisch correct of realistisch is. Veel programma's vertrekken bijvoorbeeld van de veronderstelling dat gedragsverandering kan worden bereikt door enkel informatie te geven over gezondheidsrisico's, terwijl dergelijke informatie zelden voldoende is om daadwerkelijk gedrag te veranderen. Ook kunnen er lacunes ontstaan in de keten van verwachte uitkomsten, zoals bij campagnes die onmiddellijke gezondheidsverbetering verwachten zonder eerst de noodzakelijke tussenliggende gedragsveranderingen te adresseren. Daarnaast kan een interventie zich op het verkeerde probleem richten, bijvoorbeeld door gedrag te willen veranderen terwijl de echte belemmering ligt in aanbod, toegankelijkheid of beschikbaarheid. Een grondige situatieanalyse en het verzamelen van baselinegegevens zijn daarom essentieel om aannames te toetsen en realistische verwachtingen te formuleren (58).

Publieke gezondheidsinterventies zijn vaak complex en sterk afhankelijk van hun context, waaronder sociale, culturele, economische en beleidsmatige factoren. Traditionele evaluatiemethoden, zoals RCT's houden vaak onvoldoende rekening met deze contextuele factoren, waardoor het moeilijk wordt om te begrijpen hoe en waarom een interventie werkt. Een Theory of Change is in dit opzicht nuttig, omdat het expliciet maakt hoe contextfactoren, aannames en tussenliggende stappen samenhangen met de gewenste uitkomsten (59).

Hoewel Theory of Change vaak participatief wordt ontwikkeld, varieert de mate van betrokkenheid van stakeholders sterk. In de praktijk blijft de Theory of Change soms oppervlakkig of weinig expliciet, en wordt ze tijdens de implementatie niet systematisch gebruikt, terwijl ze net waardevol kan zijn voor monitoring en bijsturing. Omdat Theory of Change methode-neutraal is, kunnen diverse evaluatiemethoden worden toegepast, wat flexibiliteit biedt maar soms ook leidt tot een minder optimale aansluiting tussen programmatheorie en evaluatieopzet.

Ten slotte onderscheidt Theory of Change zich van andere denkkaders doordat het de causale paden expliciet maakt en breed inzetbaar is, terwijl logische modellen vaak rigider en lineair van aard zijn. Het is daarom aanbevolen om een Theory of Change grondig en transparant te rapporteren, inclusief het ontwikkelproces, de betrokkenheid van stakeholders, een duidelijke diagram en de manier waarop de Theory of Change is gebruikt bij de ontwikkeling, uitvoering en evaluatie van de interventie. Checklists, zoals die van Breuer et al. 2016, kunnen hierbij richtinggevend zijn (zie Figuur 3) (59).

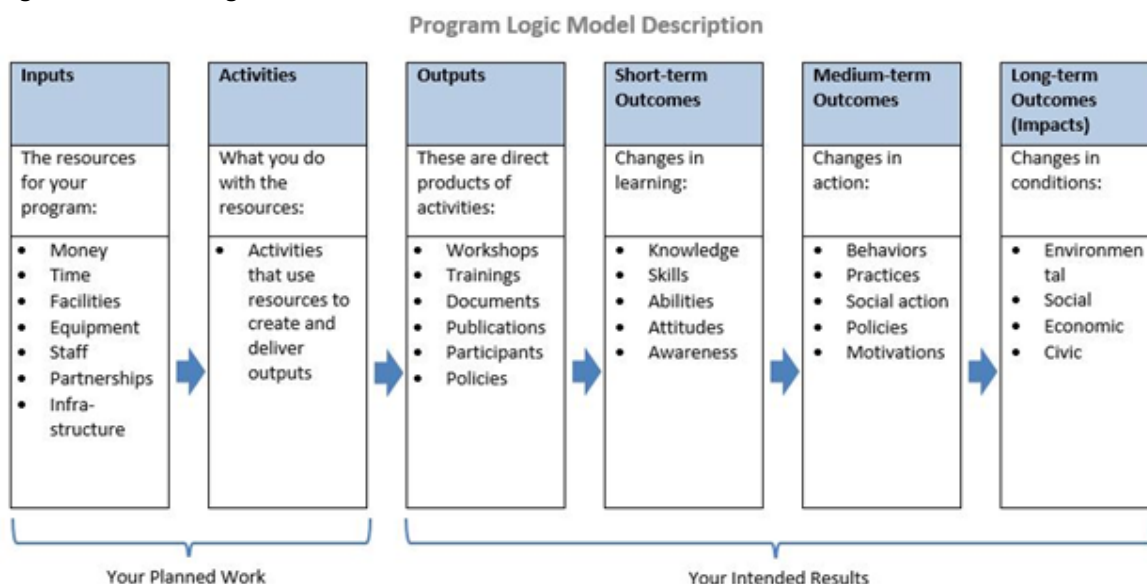
Figuur 3 Checklist voor rapportering van Theory of Change (ToC) (Nederlandse vertaling van Breuer et al., 2016).

1. Is de ToC-benadering gedefinieerd? a. Wordt er door de auteurs een definitie van ToC gegeven? b. Leggen de auteurs hun redenen uit voor het gebruik van een ToC-benadering?
2. Is het ontwikkelingsproces van de ToC beschreven? a. Worden de gebruikte methoden om de ToC te ontwikkelen beschreven, zoals bijeenkomsten en interviews met belanghebbenden, documentanalyse, observatie van het programma, bestaande conceptuele kaders of gepubliceerde onderzoeken? b. Indien belanghebbenden betrokken waren: is dan duidelijk hoeveel belanghebbenden deelnamen, welke rol ze hadden in relatie tot de interventie, hoe ze werden geraadpleegd (bv. aantal interviews, focusgroepen, ToC-workshops) en hoe participatief deze consultaties waren? c. Wordt beschreven hoe gegevens werden samengebracht tot één ToC (inclusief hoe men omging met meningsverschillen tussen belanghebbenden)? d. Wordt beschreven in welke mate belanghebbenden de resulterende ToC konden valideren en er mede-eigenaarschap over hadden?
3. Wordt de resulterende ToC (of een samenvatting ervan) weergegeven in diagramvorm en bevat deze de volgende elementen: a. Het langetermijnresultaat of de beoogde impact van de interventie. b. De verwachte korte- en middellangetermijnresultaten en het veranderingsproces. c. De interventiecomponenten in verschillende fasen of stappen van de resultatenketen. d. De context waarbinnen de interventie plaatsvindt. e. Aannames over hoe verandering tot stand zou moeten komen. f. Aanvullende ToC-elementen zoals indicatoren, ondersteunend onderzoeksbewijs, begunstigden, actoren in de context, invloedssfeer en tijdlijnen waar relevant.
4. Is het proces van interventieontwikkeling o.b.v. de ToC beschreven? a. Worden de methoden beschreven om interventies te verfijnen o.b.v. de ToC? (bv. verdere workshops met belanghebbenden, interviews, systematische literatuurreviews)
5. Wordt beschreven hoe de ToC is gebruikt voor de ontwikkeling en uitvoering van de evaluatie? a. Zijn de evaluatievragen afgeleid van de ToC? b. Is de rol van de ToC bij het ontwerp, de planning en de uitvoering van de evaluatie helder omschreven? c. Wordt aangegeven in welke mate de belangrijkste elementen uit de ToC werden gemeten in de evaluatie (impact, korte- en middellangetermijnresultaten, veranderingsproces, context, aannames en interventiecomponenten)? d. Wordt beschreven of en hoe procesindicatoren zijn gebruikt om de kwaliteit van de interventie te verbeteren? e. Is de rol van de ToC bij de analyse van de evaluatieresultaten duidelijk omschreven? f. Wordt de rol van de ToC bij de interpretatie van de evaluatieresultaten beschreven (inclusief uitsplitsing van de programmatheorie, onverwachte uitkomsten, verduidelijking van causaliteit, met aandacht voor richting en sterkte van causale relaties)?

2.2 Logic Model

Een Logic Model is een gestructureerd hulpmiddel om de stappen, activiteiten en verwachte resultaten van een programma te visualiseren. Het ondersteunt zowel planning, communicatie als evaluatie doordat het inzicht geeft in de samenhang tussen de ingezette middelen, de uitgevoerde activiteiten en de beoogde resultaten (60). Figuur 4 toont hoe een Logic Model eruit kan zien.

Figuur 4 Het Logic Model

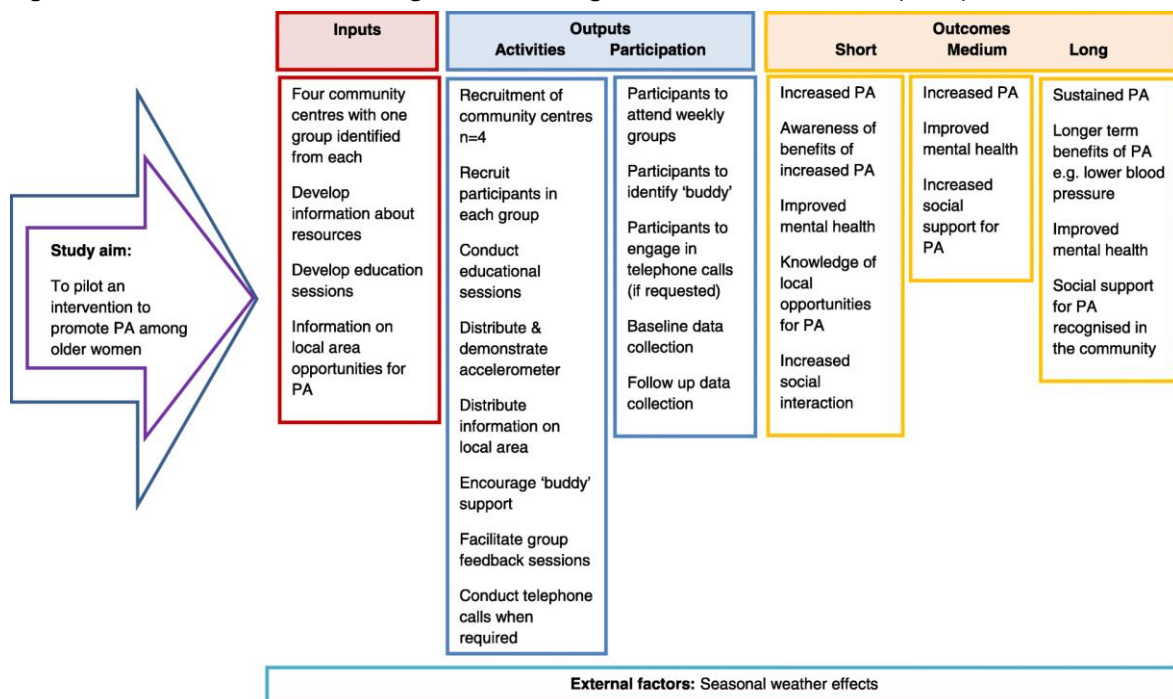


- **Inputs** zijn de middelen die een programma nodig heeft om te kunnen functioneren, zoals financiële middelen, menskracht, materiaal, apparatuur of tijd. Ze vormen de basis waarop activiteiten worden opgebouwd.
- **Activiteiten** zijn de concrete acties, diensten of interventies die het programma aanbiedt. Voorbeelden zijn het opzetten van samenwerkingen, het informeren van medewerkers, het organiseren van workshops of het verspreiden van educatief materiaal. Vaak wordt onderscheid gemaakt tussen 'vroegtijdige' en 'latere' activiteiten, omdat sommige acties de basis leggen voor andere.
- **Outputs** zijn de direct meetbare resultaten van deze activiteiten, zoals het aantal bijeenkomsten, het aantal deelnemers aan een cursus, of het aantal verspreide brochures. Outputs geven een eerste indicatie van wat het programma heeft gerealiseerd.
- **Kortetermijnresultaten** verwijzen naar de directe effecten van het programma op bijvoorbeeld kennis, houding of bewustwording bij de doelgroep. Deze uitkomsten worden vaak kort na de activiteiten zichtbaar.
- **Middellangetermijnresultaten** omvatten veranderingen in gedrag, sociale normen of beleid die voortvloeien uit de kortetermijnresultaten. Dit zijn vaak de centrale veranderingen waarop het programma mikt.
- **Langetermijnresultaten of impacts** zijn de uiteindelijke effecten waar het programma naar streeft, zoals verbeterde gezondheid, welzijn of maatschappelijke omstandigheden. Deze resultaten worden doorgaans pas na langere tijd zichtbaar.

Door middelen, activiteiten, resultaten en context helder in kaart te brengen, helpt een Logic Model betrokkenen om samen te begrijpen hoe een programma werkt en wat het beoogt te bereiken. Het is ook een krachtig communicatiemiddel richting stakeholders, financiers en deelnemers (61). Voor de praktische uitwerking van een Logic Model bestaan verschillende online hulpmiddelen. Een voorbeeld hiervan is Dylomo, een interactieve tool die gebruikers begeleidt bij het systematisch in kaart brengen van inputs, activiteiten en resultaten (Dylomo.com, 2016).

Figuur 5 toont een voorbeeld van een Logic Model van een interventie om fysieke activiteit te bevorderen bij oudere vrouwen in sociaaleconomisch kwetsbare buurten, uitgevoerd binnen vrouwengroepen in vier gemeenschapscentra (62).

Figuur 5 Voorbeeld van een Logic Model. Overgenomen van Lawlor et al. (2019).



3 RE-AIM voor proces- en effectevaluatie

Het RE-AIM-kader werd ontwikkeld om de impact van gezondheidsbevorderende interventies systematisch te evalueren aan de hand van vijf dimensies (63, 64). Het richt zich zowel op individuele als organisatorische aspecten van implementatie. Tabel 4 geeft een overzicht van de vijf dimensies, hun definities en aanbevelingen voor rapportage.

Tabel 4 RE-AIM-dimensies.

Dimensie	Definitie	Aanbevelingen voor rapportage
Bereik (individueel niveau)	Aantal, proportie en representativiteit van individuen die deelnemen aan de interventie, inclusief redenen voor deelname of niet-deelname.	Rapporteer percentage deelnemers op basis van een correcte noemer en vergelijk kenmerken van deelnemers met niet-deelnemers. Gebruik exclusiecriteria en kwalitatieve methoden om werving en bereik te duiden.
Effectiviteit (individueel niveau)	Impact van de interventie op relevante uitkomsten, inclusief mogelijke negatieve effecten, kwaliteit van leven, economische uitkomsten en variaties tussen subgroepen.	Rapporteer primaire uitkomst, bredere uitkomsten, uitval op korte termijn en verschillen tussen patiëntkenmerken. Gebruik indien mogelijk ook kwalitatieve methoden.
Adoptie (organisatie en uitvoerders)	Aantal, proportie en representativiteit van organisaties en uitvoerders die bereid zijn een programma te implementeren, inclusief redenen voor deelname en niet-deelname.	Rapporteer het percentage benaderde organisaties en uitvoerders dat deelneemt, en vergelijk kenmerken van deelnemers met niet-deelnemers.
Uitvoering (organisatie en uitvoerders)	Mate van getrouwheid aan interventiecomponenten, consistentie van uitvoering, benodigde tijd en kosten, en aangebrachte aanpassingen aan de interventie.	Rapporteer consistentie van uitvoering, aanpassingen, kosten (tijd, geld), en redenen voor aanpassingen.
Behoud (organisatie en individueel)	Mate waarin een programma wordt verankerd in de organisatie en de langetermijneffecten bij individuen na afloop van het programma.	Organisatieniveau: rapportage over voortzetting, aanpassingen en afstemming met organisatiedoelen. Individueel niveau: uitkomsten op lange termijn (aanbevolen na ± 2 jaar), uitval, en heterogeniteit van effecten.

Uit een systematische review blijkt dat RE-AIM vandaag breed wordt toegepast in uiteenlopende gezondheidsdomeinen, zowel voor evaluatie, planning en voortgangsmeting als in rapportering van resultaten en literatuurreviews (65). Wel blijkt dat in veel studies niet alle dimensies worden meegenomen. De auteurs stellen dat niet elke studie alle vijf RE-AIM dimensies hoeft te behandelen,

maar dat het wel belangrijk is om expliciet te maken welke dimensies zijn opgenomen en waarom. Idealiter wordt RE-AIM zo volledig mogelijk toegepast, aangezien de dimensies onderling samenhangen: een hoge participatiegraad (bereik) kan bijvoorbeeld onbedoelde gevolgen hebben voor de effectiviteit of voor de kwaliteit van uitvoering.

De website <https://re-aim.org/> biedt uitgebreide informatie over het kader, inclusief hulpmiddelen zoals checklists, een interactieve planning tool, een vragenlijst om de verschillende dimensies te scoren en voorbeelden uit de literatuur die illustreren hoe RE-AIM effectief kan worden toegepast in diverse settings.

In Vlaanderen wordt het RE-AIM-kader toegelicht in de leidraad 'Evalueren en impact meten binnen gezondheidsbevordering en ziektepreventie' van het Vlaams Instituut Gezond Leven. Daarin wordt de methodiek *10.000 stappen* als voorbeeld gebruikt¹³. Gezond Leven ontwikkelde bovendien een Nederlandstalige RE-AIM-checklist, met per dimensie relevante evaluatievragen, meetmethoden, beïnvloedende factoren en strategieën ter verbetering¹⁴. Dit biedt een waardevol hulpmiddel voor praktijktoepassing.

Hoewel RE-AIM erg bruikbaar is voor het plannen en evalueren van implementaties, is het minder geschikt om factoren te identificeren die implementatie beïnvloeden of om te verklaren hoe en waarom effecten ontstaan. Voor deze onderzoeksvragen biedt het PRISM-model soelaas (Feldstein & Glasgow, 2008). PRISM is een uitbreiding van RE-AIM, waarin contextuele determinanten een meer centrale plaats krijgen.

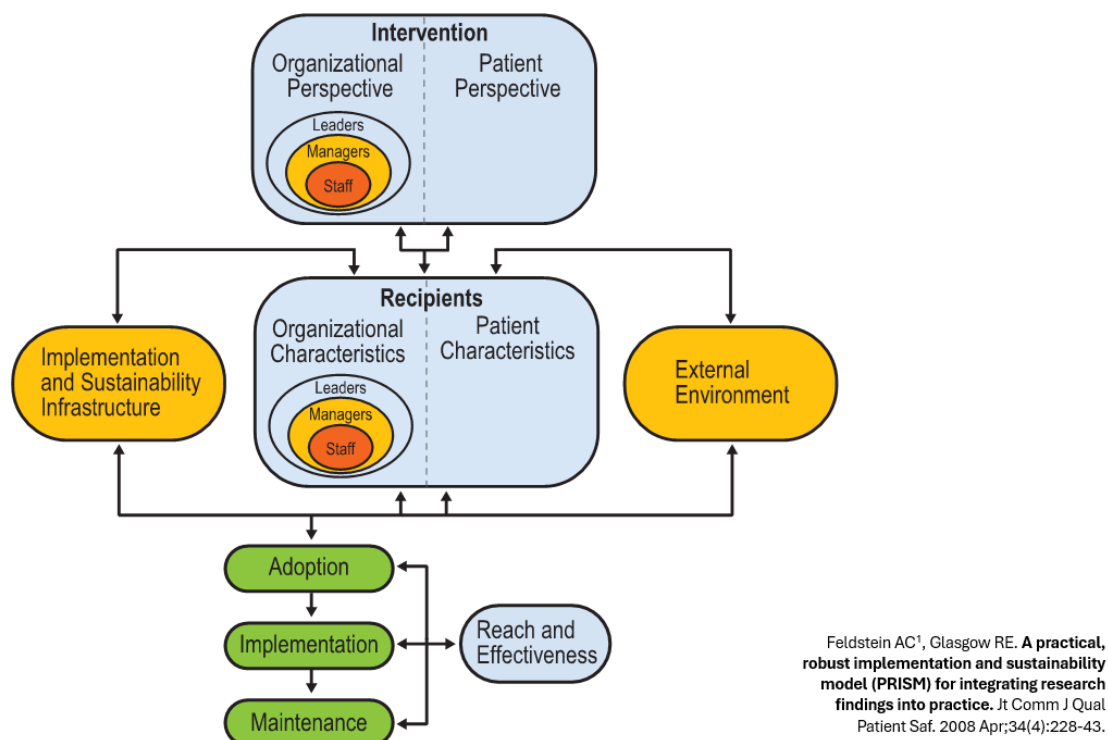
¹³ <https://www.gezondleven.be/files/gezondheidsbevordering/GezondLeven-leidraad.pdf>

¹⁴ https://www.gezondleven.be/files/gezondheidsbevordering/RE-AIM_Gezond-Leven.pdf

3.1 Practical, Robust Implementation and Sustainability Model (PRISM)

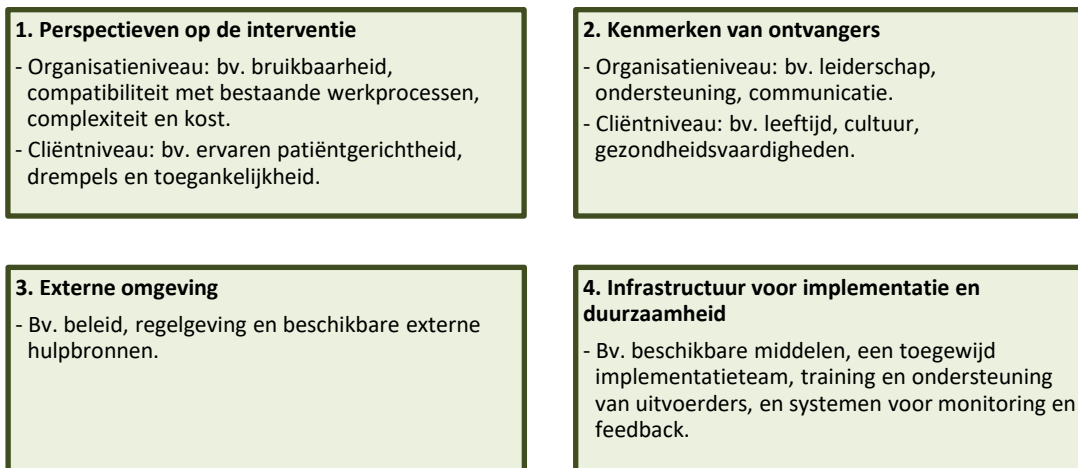
Het **Practical, Robust Implementation and Sustainability Model (PRISM)** is een contextgebaseerd kader voor implementatieonderzoek dat kan worden beschouwd als een uitgebreidere versie van RE-AIM (66) (Figuur 6). PRISM integreert de vijf RE-AIM-dimensies en voegt daar contextuele factoren aan toe, waardoor het meer geschikt is voor complexe interventies dan het klassieke RE-AIM-kader. Concreet helpt PRISM bij het identificeren van factoren die het succes en de duurzaamheid van interventies beïnvloeden.

Figuur 6 Het Practical, Robust Implementation and Sustainability Model (PRISM)



PRISM onderscheidt daarbij vier contextdomeinen (Figuur 7):

Figuur 7 Vier contextdomeinen van PRISM.

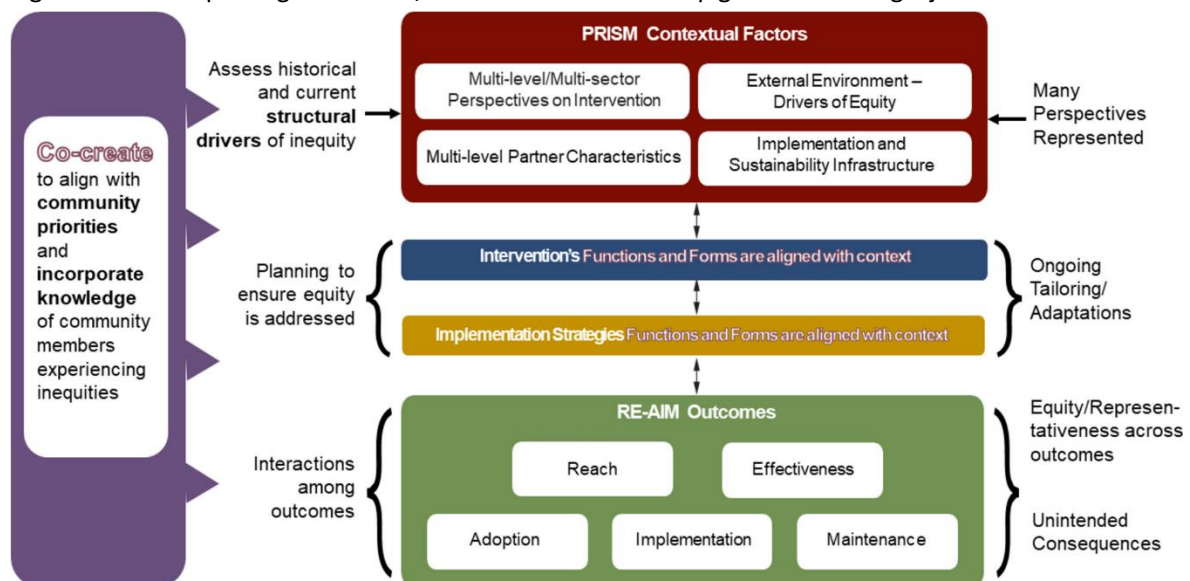


Door deze domeinen te koppelen aan RE-AIM-uitkomsten, kunnen onderzoekers en praktijk-professionals beter inspelen op de context. Het “Guidebook to the Pragmatic and Iterative Use of PRISM and RE-AIM for Planning, Implementation, and Sustainment” (67) en de interactieve [iPRISM webtool](#) bieden praktische handvatten en diverse ondersteunende tools om PRISM toe te passen.

Een systematische review naar het gebruik van PRISM sinds de introductie in 2008 toont aan dat het model vooral wordt ingezet voor de evaluatie van interventies (68). De toepassing varieert echter sterk, en contextdomeinen worden zelden expliciet gekoppeld aan RE-AIM-uitkomsten. De auteurs pleiten dan ook voor een meer geïntegreerde toepassing van PRISM, zodat de samenhang tussen contextfactoren en implementatieresultaten beter zichtbaar wordt.

Recent stelden Jolles et al. (2024) een geïntegreerde aanpak voor om complexe gezondheidsinterventies beter af te stemmen op lokale contexten, met expliciete aandacht voor gezondheidsgelijkheid en participatie (69). Ze introduceerden daarbij het onderscheid tussen kernfuncties (core functions) en vormen (forms). Kernfuncties verwijzen naar het doel of de essentie van een interventie, terwijl vormen de concrete, contextspecifieke manieren zijn waarop deze functies worden uitgevoerd. Door deze differentiatie kunnen interventies flexibel worden aangepast aan lokale omstandigheden zonder hun kern te verliezen. Co-creatie, waarbij lokale partners en doelgroepen actief betrokken worden in alle fasen van het proces, versterkt dit principe. De benadering wordt weergegeven in Figuur 8.

Figuur 8 Aanpassing van PRISM/RE-AIM met een focus op gezondheidsongelijkheden en co-creatie.



4 MRC-framework

Het in 2021 gereviseerde MRC-framework (*Medical Research Council guidance*) is een richtlijn voor het ontwikkelen en evalueren van complexe interventies in de gezondheidszorg en in het gezondheids- en sociaal beleid (15). Complexe interventies bestaan uit meerdere componenten, beïnvloeden verschillende niveaus (individu, gemeenschap, systeem) en interageren sterk met de context waarin ze worden uitgevoerd.

Het framework erkent dat interventies zich moeten kunnen aanpassen aan verschillende contexten, terwijl de kernfuncties behouden blijven. Standaardisatie richt zich dus niet op de vorm, maar op de onderliggende processen en doelen. Complexe interventies worden beschouwd als gebeurtenissen binnen complexe adaptieve systemen (CAS), waarin dynamiek en onvoorspelbaarheid centraal staan.

Waar eerdere versies van het MRC-framework hoofdzakelijk focusten op de vraag of een interventie effectief is, schenkt de gereviseerde versie meer aandacht aan **hoe, waarom en wanneer een interventie werkt**. De nieuwe versie introduceert vier gelijkwaardige en overlappende perspectieven op evaluatie:

- **Efficacy:** werkt de interventie onder ideale of experimentele omstandigheden?
- **Effectiviteit:** werkt de interventie in de praktijk ('de echte wereld')?
- **Theoriegebaseerd:** wat werkt onder welke omstandigheden en hoe?
- **Systeem:** hoe beïnvloeden de interventie en het bredere systeem elkaar?

Deze perspectieven kunnen gecombineerd worden. Zo kan effectiviteitsonderzoek worden versterkt door theoriegebaseerde methoden die mechanismen en context verduidelijken.

Het framework beschrijft een iteratief proces met vier fasen (zie Figuur 9), waarbij onderzoekers kunnen instappen in elke fase, afhankelijk van bestaande kennis en onzekerheden:

- **Ontwikkeling/identificatie van een interventie:** Interventies kunnen nieuw worden ontwikkeld of bestaande interventies kunnen worden aangepast aan een andere context of populatie. Identificatie van een interventie houdt in dat een reeds uitgevoerde of geplande interventie (in de praktijk of beleid) geselecteerd wordt en de opties voor evaluatie worden onderzocht.
- **Haalbaarheid:** Een haalbaarheidsstudie beoordeelt vooraf bepaalde criteria invloed hebben op de evaluatie (bv. dataverzameling, uitkomsten) of de interventie zelf (bv. optimale inhoud en uitvoering, acceptatie, kosteneffectiviteit en capaciteit van uitvoerders). Een voorafgaande *evaluability assessment* kan bepalen of en hoe een interventie zinvol geëvalueerd kan worden. Het eindresultaat is een aanbeveling of de interventie en de evaluatie haalbaar zijn, of het kan worden uitgevoerd tegen een redelijke kost en met welke methode.
- **Evaluatie:** In deze fase kunnen o.a. de effectiviteit, werkingsmechanismen, interactie met de context, procesfactoren en systeemveranderingen worden onderzocht. Het is belangrijk om in samenspraak met stakeholders belangrijke uitkomstmaten te kiezen en te bepalen wat 'voldoende bewijs' is. Voor complexe interventies volstaat een louter kwantitatieve experimentele aanpak meestal niet; kwalitatieve of mixed-methods designs zijn vaak nodig om vragen te beantwoorden over mechanismen en context. Procesevaluatie speelt een rol bij het beantwoorden van vragen over implementatietrouw en kwaliteit, werkingsmechanismen en

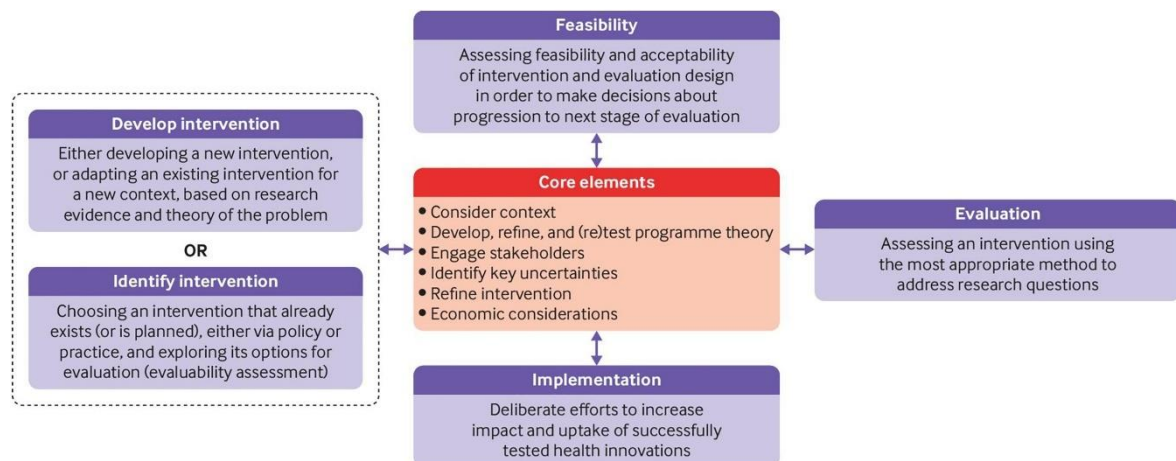
context. Het helpt te verklaren waarom een interventie al dan niet werkt, en draagt bij aan verdere ontwikkeling van de programmatheorie.

- **Implementatie:** Vroege aandacht voor implementatie vergroot de kans op brede en duurzame toepassing. Belangrijke aandachtspunten zijn de gekozen implementatiestrategieën en de contextfactoren die effecten kunnen beïnvloeden. Stakeholderbetrokkenheid helpt om mogelijke uitdagingen vooraf te identificeren.

In alle fases spelen zes transversale elementen een centrale rol. Zij bepalen mee of kan worden overgegaan naar een volgende fase, of herhaling, terugkeer of stopzetting aangewezen is:

- **Interactie met de context:** Hoe past de interventie in de context
- **Programmatheorie:** Beschrijft hoe en onder welke voorwaarden een interventie naar verwachting werkt; moet in alle fasen worden getest, verfijnd en gebruikt om onzekerheden en onderzoeksvragen te bepalen.
- **Betrekken van stakeholders:** Hoe worden verschillende perspectieven meegenomen, van zowel betrokkenen als belanghebbenden?
- **Onzekerheden:** Welke kennislücken blijven bestaan o.b.v. bestaande kennis, programmatheorie en stakeholderinput, en hoe sturen zij de onderzoeksvragen?
- **Verbetering:** Hoe kan de interventie verder worden geoptimaliseerd?
- **Kosten-baten:** Rechtvaardigen de verwachte of aangetoonde effecten de kosten?

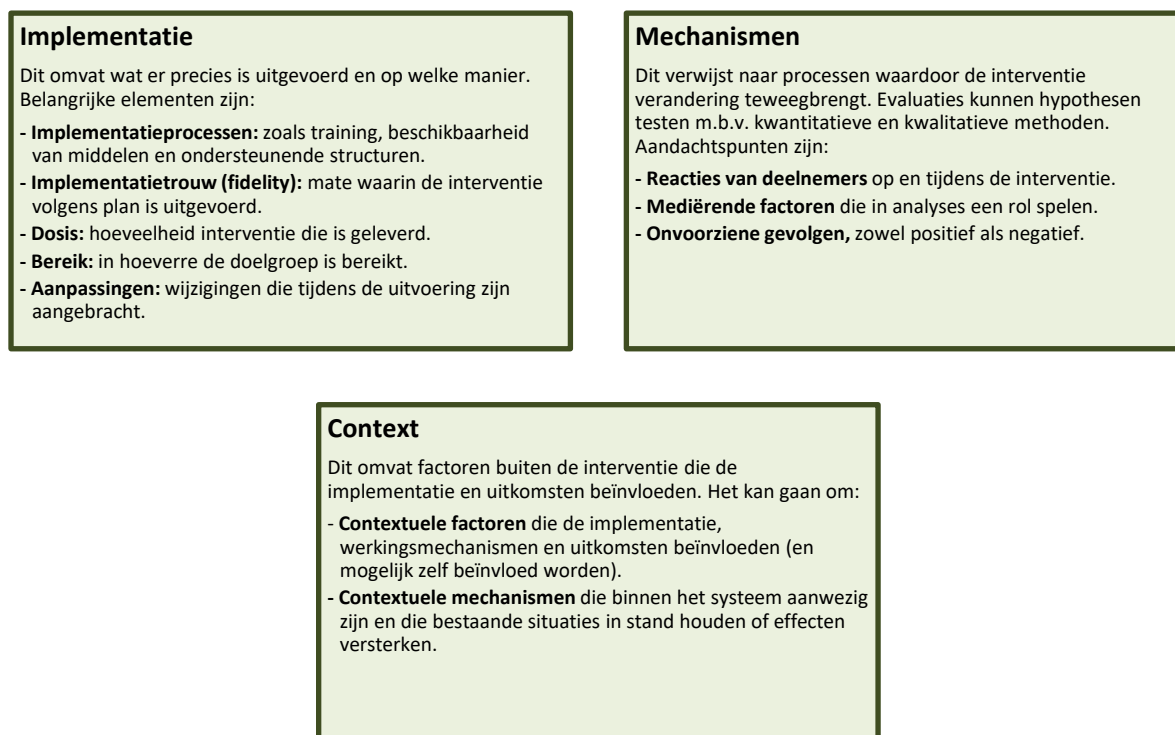
Figuur 9 Het gereviseerde MRC-framework. Overgenomen van Skivington et al. (2021).



4.1 Procesevaluatie volgens het MRC-framework

Procesevaluatie is essentieel voor het begrijpen van complexe interventies. Het helpt te verklaren waarom een interventie wel of niet werkt, maakt implementatieproblemen zichtbaar en ondersteunt de beoordeling of een interventie geschikt is voor toepassing in andere contexten. De gids voor procesevaluatie van het oorspronkelijke MRC-framework richt zich hierbij op drie componenten (Figuur 10) (70):

Figuur 10 Drie componenten van procesevaluatie volgens het MRC-framework.



Procesevaluaties kunnen niet alle vragen rond een complexe interventies beantwoorden. Het is efficiënter om een beperkt aantal kernvragen diepgaander te onderzoeken dan vele vragen slechts oppervlakkig. Om die kernvragen te bepalen, kunnen evaluatoren nagaan **welke aannames in de programmatheorie het minst goed onderbouwd zijn**. Dit kan gebeuren a.d.h.v. literatuuronderzoek en overleg met stakeholders uit beleid en praktijk.

De focus van procesevaluatie kan verschillen per fase van de interventie. Tijdens een pilootfase ligt de nadruk vaak op haalbaarheid, terwijl in de effectiviteitsfase vooral wordt gekeken naar werkingsmechanismen, uitvoering en de invloed van context.

De analyse van kwantitatieve procesdata start doorgaans met beschrijvende statistieken en indicatoren voor implementatie. Vervolgens worden deze gegevens gekoppeld aan uitkomstmaten van de interventie. Dit maakt het mogelijk te onderzoeken of variaties in implementatie, bijvoorbeeld verschillen tussen locaties in interventietrouw of dosis, samenhangen met variaties in effectiviteit. Dit gebeurt vaak via *on-treatment analyses*, waarin de relatie tussen kwaliteit van uitvoering en uiteindelijke effecten wordt onderzocht.

De koppeling van procesgegevens aan uitkomsten biedt ook ruimte om hypothesen te testen die voortkomen uit kwalitatieve bevindingen. Als interviews bijvoorbeeld suggereren dat sociale steun een belangrijk mechanisme is, kunnen kwantitatieve variabelen over sociale steun worden gekoppeld aan uitkomsten om te verifiëren of dit mechanisme inderdaad een rol speelt.

Procesevaluatie volgens het MRC-framework **maakt doorgaans deel uit van dezelfde studie die ook de effectiviteit onderzoekt**. Sommige evaluatoren kiezen ervoor om de teams voor proces- en uitkomstonderzoek te scheiden, terwijl anderen deze teams combineren. Gescheiden teams vereisten

sterke communicatie om overlap of inconsistenties te vermijden. Bij gecombineerde teams hebben baat bij transparantie over hoe de integratie van rollen de uitvoering en interpretatie van de evaluatie kan beïnvloeden.

Box 2: Procesevaluatie in de praktijk: het MRC-framework toegepast op een Welsh 'bewegen op verwijzing' programma

Het *National Exercise Referral Scheme* (NERS) in Wales – vergelijkbaar met het Vlaamse "Bewegen op Verwijzing" – werd door (29) gebruikt als illustratie van het MRC-kader voor procesevaluatie. Dit programma heeft als doel de fysieke activiteit van volwassenen te bevorderen via verwijzingen van huisartsen en andere zorgprofessionals naar lokale sportcentra voor gesuperviseerde groepsactiviteiten.

De procesevaluatie van NERS maakte gebruik van kwantitatieve en kwalitatieve methoden om inzicht te krijgen in (1) de uitvoering van de interventie, (2) de mechanismen die tot verandering leidden en (3) de rol van contextfactoren.

Kwantitatieve gegevens:

- Gestructureerde observaties van audio-opnames van consultaties tussen patiënten en zorgverleners om implementatietrouw te meten (bv. mate waarin gemotiveerde gespreksvoering werd toegepast) en de dosis van de interventie te bepalen (bv. duur consultaties).
- Monitoringdata om het bereik van het programma te onderzoeken, met aandacht voor sociaal-demografische patronen in deelname.
- Psychologische variabelen zoals motivatie en zelfvertrouwen van deelnemers om mogelijke mechanismen van gedragsverandering te toetsen.

Kwalitatieve gegevens:

- Interviews met patiënten, sportprofessionals, programmacoördinatoren en zorgverleners om zicht te krijgen op uitdagingen tijdens de implementatie in verschillende lokale contexten en hoe het programma in de praktijk werd beleefd.

Belangrijkste bevindingen

De evaluatie toonde dat de kerncomponenten van NERS goed werden uitgevoerd, maar dat andere elementen, zoals motiverende gespreksvoering, minder consequent aanwezig waren. Toch bleken deelnemers vooral gemotiveerd door sociale steun binnen de groepen – een mechanisme dat niet expliciet was voorzien. Het koppelen van proces- en uitkomstdata liet zien dat verbeteringen in fysieke activiteit vooral werden verklaard door toegenomen motivatie, gevoed door zowel formele als informele programma-elementen.

Conclusie

Dit voorbeeld illustreert dat een aanvullende procesevaluatie essentieel is om te begrijpen waarom een interventie werkt. Zonder procesevaluatie had men foutief kunnen concluderen dat gemotiveerde gespreksvoering doorslaggevend is voor het effect, terwijl sociale steun een veel belangrijkere rol bleek te spelen. Het toont aan dat interventies niet altijd werken zoals oorspronkelijk gedacht en dat een goed begrip van uitvoering, mechanismen en context noodzakelijk is om interventieresultaten correct te interpreteren.

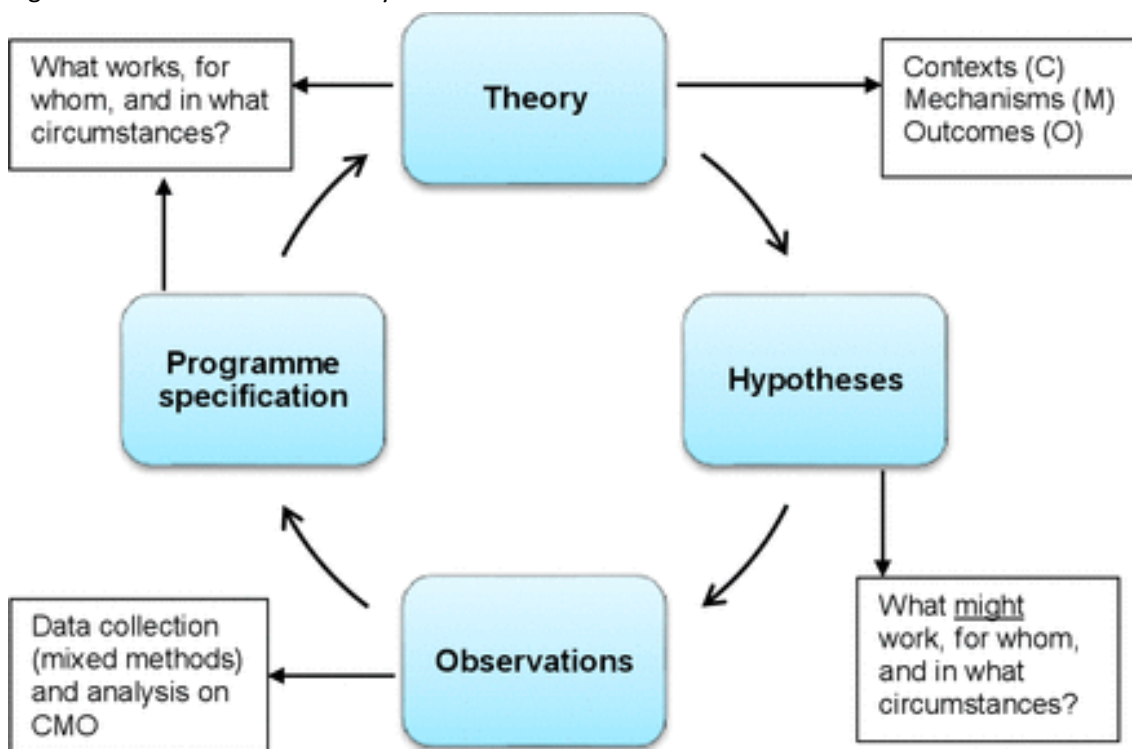
5 Realist Evaluation

De centrale vraag in Realist Evaluation is niet zozeer of een interventie werkt, **maar vooral wat werkt, voor wie, onder welke omstandigheden en waarom**. Deze benadering is een specifieke vorm van theoriegebaseerd evaluatieonderzoek, waarbij de focus ligt op de interactie tussen **Context (C)**, **Mechanismen (M)** en **Uitkomsten (O)** – samengevat als **C + M = O** (71-73). Figuur 11 toont de cyclus van Realist Evaluation.

De drie kernbegrippen:

- **Context** verwijst naar de bestaande omstandigheden waarin de interventie wordt geïmplementeerd: kenmerken van de setting, sociale relaties, normen, cultuur, verwachtingen en beschikbare hulpbronnen.
- **Mechanismen** zijn de onderliggende processen die in werking treden bij betrokkenen (bv. cognitieve of emotionele reacties, ervaren steun), en verklaren hoe de interventie tot bepaalde uitkomsten leidt.
- **Uitkomsten** zijn de effecten van de interventie op korte en (middel)lange termijn.

Figuur 11 Realist Evaluation cyclus.



Interventies leiden tot succesvolle uitkomsten wanneer ze de juiste hulpbronnen en kansen (mechanismen) aanbieden aan de juiste personen of groepen en onder geschikte sociale en culturele omstandigheden (context). Een programmatheorie beschrijft de onderliggende aannames over hoe en waarom een interventie geacht wordt te werken. In de beginfase wordt een initiële programmatheorie opgesteld, gebaseerd op bestaande kennis, ervaring en aannames van de interventie-ontwikkelaars. Deze theorie stuurt de dataverzameling, die gericht is op het testen van specifieke CMO-hypothesen.

Realist Evaluation is methodologisch flexibel doordat er geen specifieke onderzoeksmethoden worden voorgeschreven. De keuze van methoden hangt af van wat nodig is om de programmatheorie te toetsen. Vaak wordt gewerkt met doelgerichte casusstudies, waarbij kwantitatieve data vooral worden gebruikt om de context en uitkomsten te beschrijven, en kwalitatieve data inzicht bieden in de onderliggende mechanismen. Vergelijkingen gebeuren meestal tussen doelgroepen (bv. leeftijdsgroepen) in plaats van met controlegroepen.

De analyse focust op het identificeren van CMO-configuraties: welke mechanismen werden geactiveerd, in welke context, en met welke uitkomsten? Op basis hiervan wordt de programmatheorie verfijnd. Idealiter vindt men een balans tussen specificiteit (voor een bepaalde doelgroep of setting) en abstractie (toepasbaarheid in andere contexten).

Realist Evaluation gaat ervan uit dat er geen kant-en-klaar oplossing bestaat: wat in de ene groep of regio werkt, werkt niet noodzakelijk overal of voor iedereen. Dit maakt de aanpak bijzonder geschikt voor het begrijpen van complexe interventies binnen gezondheidsbevordering (74). Daarnaast is het waardevol voor interventies die moeten worden aangepast aan specifieke contexten, voor veelbelovende interventies die lijken te werken maar waarvan het werkingsmechanisme nog onvoldoende begrepen is, of voor interventies die wisselende effecten laten zien. Kwaliteitsstandaarden voor Realist Evaluations zijn beschikbaar via het Rameses-II project¹⁵ (75).

Hoewel Realist Evaluation veel flexibiliteit biedt, brengt de aanpak ook uitdagingen met zich mee:

- Het opstellen van een goede programmatheorie is complex en tijdsintensief.
- De analyse van meerdere CMO-configuraties kan arbeidsintensief zijn.
- De relatie tussen context en uitkomst is niet altijd gemakkelijk empirisch te onderbouwen.
- De terminologie en methodologie kunnen voor niet-experts moeilijk toegankelijk zijn.

Box 3: Evaluatieprotocol voor gezondheidsbevordering in scholen met een Realist Evaluation aanpak en gemengde methoden

Pommier et al. (76) beschrijven een evaluatieprotocol van een grootschalig gezondheidsbevorderend schoolprogramma in Franse basisscholen. Het voorbeeld illustreert hoe realistic evaluation en mixed methods kunnen worden ingezet om complexe schoolinterventies te begrijpen. De onderzoekers kiezen voor Realist Evaluation omdat deze benadering bijzonder goed aansluit bij de complexiteit en contextafhankelijkheid van gezondheidsbevordering in scholen. Bovendien past Realist Evaluation bij kernwaarden van gezondheidsbevordering, zoals participatie, empowerment en samenwerking.

De onderzoekers ontwikkelden een programmatheorie waarin interventiestrategieën, mechanismen, contextfactoren en uitkomsten systematisch worden beschreven. Het evaluatieontwerp volgt een **mixed-methods design**. De **kwantitatieve component** vormt het zwaartepunt van de evaluatie en omvatten vier verschillende vragenlijsten voor leerlingen, leerkrachten, ouders en schoolteams, aangevuld met contextformulieren over schoolkenmerken en implementatieaspecten. De **kwalitatieve component** bestaat uit open vragen en interviews om mechanismen, percepties en implementatieprocessen te verhelderen.

De gegevens worden verzameld aan de start van het project en het einde van elk schooljaar, gedurende drie opeenvolgende jaren. Tussentijdse en jaarlijkse rapportages worden teruggekoppeld aan de deel-

¹⁵ Zie https://ramesesproject.org/media/RE_Quality_Standards_for_evaluators_and_peer_reviewers.pdf

nemende scholen, wat zowel de betrokkenheid als het leerproces ondersteunt. De evaluatie richt zich op **drie uitkomstniveaus**: (1) schoolniveau (i.e., de ontwikkeling van een gezondheidsbevorderende schoolaanpak); (2) leerkrachtenniveau (i.e., veranderingen in praktijken en het gedrag van leerkrachten); en (3) leerlingenniveau (i.e., effecten op het welzijn van kinderen op school). Deze uitkomsten worden telkens geïnterpreteerd in relatie tot mechanismen en contextfactoren, die van school tot school kunnen verschillen. De kwantitatieve en kwalitatieve resultaten worden geïntegreerd om veranderingen in scholen en klaspraktijken te beschrijven, onderliggende mechanismen te identificeren en het theoretisch model stapsgewijs te verfijnen.

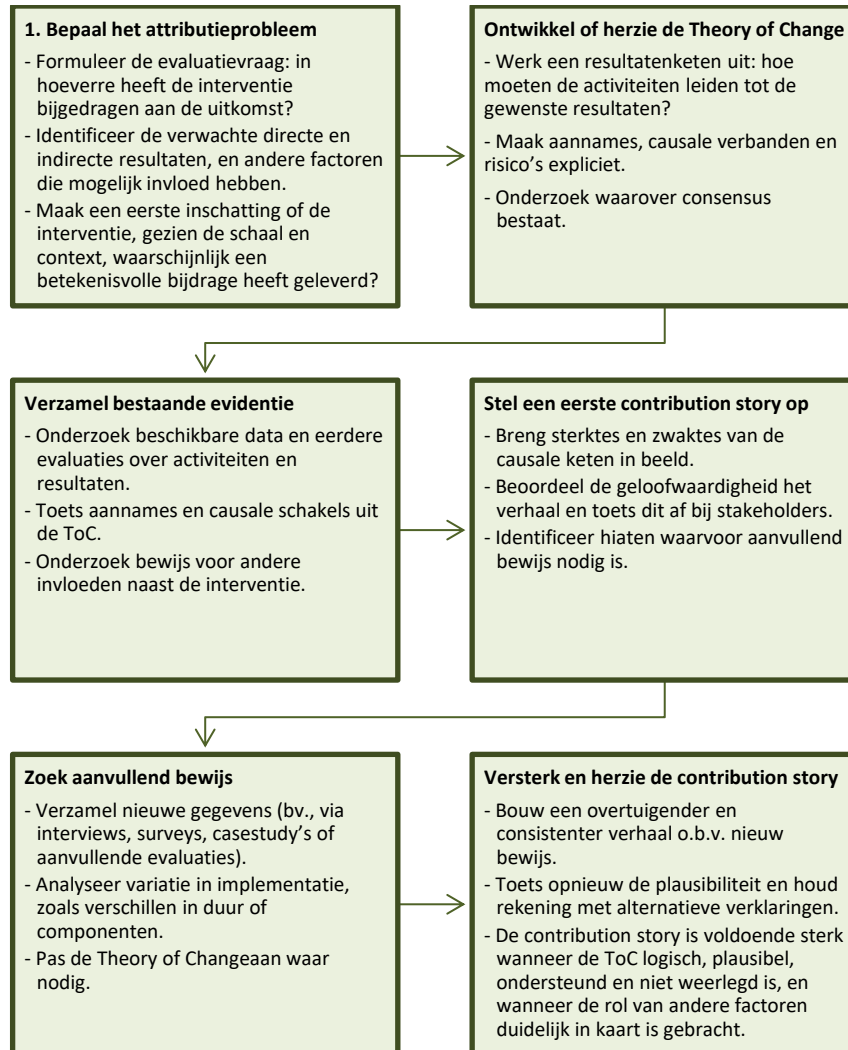
Dit voorbeeld toont hoe Realist Evaluation gecombineerd met mixed-methods onderzoek een rijk en contextgevoelig inzicht kan bieden in de werking van complexe schoolinterventies.

6 Contribution Analysis

Contribution Analysis is een evaluatiebenadering die onderzoekt in welke mate een interventie heeft bijgedragen aan bepaalde resultaten. In plaats van te streven naar het aantonen van absolute causaliteit, richt Contribution Analysis zich op het onderbouwen van de plausibele bijdrage van een interventie. Dit maakt de methode bijzonder geschikt voor complexe beleids- en gezondheidsinterventies, waarbij klassieke kwantitatieve designs vaak tekortschieten.

De aanpak vertrekt vanuit een Theory of Change waarin wordt geanalyseerd welke factoren een rol hebben gespeeld en welke verklaringen mogelijk zijn. Het resultaat is een 'contribution story': een onderbouwd narratief dat uitlegt hoe, waarom en voor wie resultaten tot stand zijn gekomen en in welke mate de interventie hieraan heeft bijgedragen. Contribution Analysis biedt geen absolute zekerheid, maar wel een realistisch en gefundeerd beeld door theorie, empirische data en context te combineren. Ze is dan ook niet geschikt om de mate waarin een interventie een resultaat veroorzaakt te kwantificeren. De bedenker van de methode, John Mayne, formuleerde een reeks iteratieve stappen voor Contribution Analysis (Figuur 12). Tijdens het proces kan op elk moment worden teruggekeerd naar eerdere stappen om theorie, bewijs of het verhaal te versterken (35):

Figuur 12 Verschillende stappen van Contribution Analysis volgens Mayne (2008).



Een causal-pathway-benadering binnen Contribution Analysis draait om het begrijpen hoe, waarom en onder welke omstandigheden verandering plaatsvindt (77). Het gaat om de bredere keten van causale verbanden, waarin meerdere actoren, gebeurtenissen en factoren samen leiden tot uiteenlopende uitkomsten en impacts. Het toepassen van causale inferentietechnieken zoals *process tracing* kan de contribution story verder versterken. Process tracing is een kwalitatieve onderzoeksmethode die de tussenliggende stappen en mechanismen in een oorzakelijk verband aannemelijk maken.

Ondanks het brede gebruik van Contribution Analysis in internationale ontwikkelingsprogramma's, zijn er slechts weinig peer-reviewed studies over evaluaties van gezondheidsinterventies met Contribution Analysis beschikbaar (78). Toch blijkt uit onderzoek dat het toepasbaar is op diverse gezondheidsinterventies, zoals gezondheidsbevordering en beleidsevaluaties.

7 Evaluatiekaders voor implementatieonderzoek

Evaluatiekaders die zich uitsluitend richten op implementatie zijn ontwikkeld om systematisch inzicht te krijgen in hoe een interventie in de praktijk wordt uitgevoerd, zonder daarbij de effectiviteit van de interventie zelf te beoordelen. Deze kaders leggen de nadruk op factoren zoals haalbaarheid en acceptatie, en maken zichtbaar welke drempels en randvoorwaarden een rol spelen bij succesvolle integratie in de praktijk. Ze zijn bijzonder waardevol wanneer het primaire doel is te begrijpen of en hoe een interventie ingebed raakt in de praktijk.

7.1 Consolidated Framework for Implementation Research

Het Consolidated Framework for Implementation Research (CFIR) biedt onderzoekers en professionals een gestructureerde manier om determinanten van implementatieprocessen te identificeren en analyseren (79). Het kader helpt bij het in kaart brengen van contextuele factoren die de implementatie van innovaties (zoals interventies) beïnvloeden, ondersteunt het opsporen van drempels en bevorderende factoren, en draagt bij aan het selecteren van passende implementatiestrategieën. Binnen CFIR worden constructen gedefinieerd als specifieke, meetbare elementen die invloed uitoefenen op implementatie. Deze zijn gegroepeerd in vijf hoofddomeinen: interventiekenmerken, externe setting, interne setting, individuele kenmerken en proces (zie Figuur 13)¹⁶.

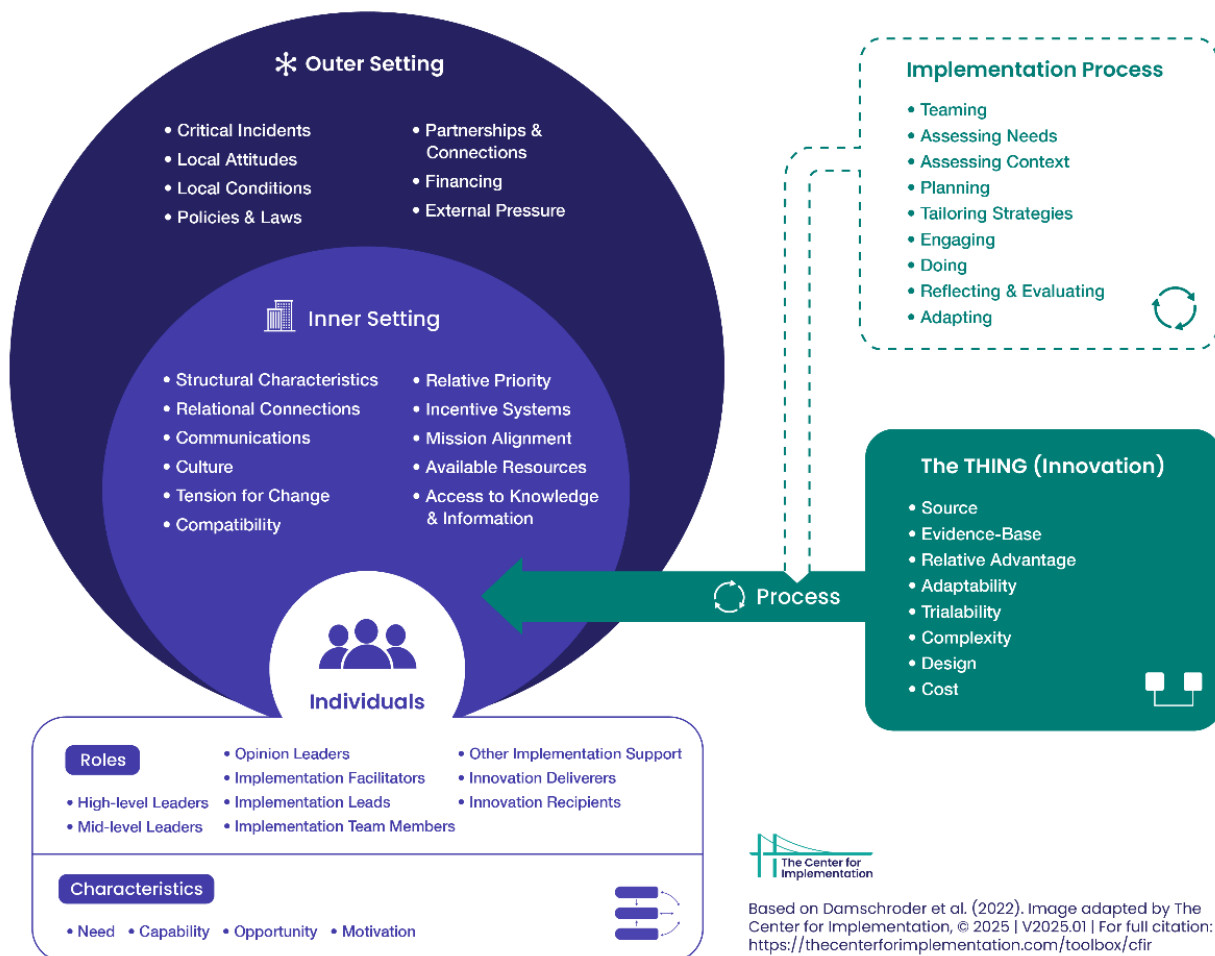
Om het gebruik van CFIR in de praktijk te faciliteren, publiceerde het CFIR-team in 2025 een praktische gids met diverse tools en templates die het toepassen van CFIR in onderzoek en praktijk in vijf stappen ondersteunen (80):

1. **Studieontwerp:** formuleren van onderzoeksvragen en kiezen van relevante implementatie-uitkomsten.
2. **Dataverzameling:** selecteren van geschikte methoden en instrumenten.
3. **Data-analyse:** structureren en interpreteren van data per construct.
4. **Interpretatie:** koppelen van uitkomsten aan relevante determinanten.
5. **Disseminatie:** verspreiden van bevindingen en aanbevelingen.

¹⁶ Overzicht van alle constructen: <https://cfirguide.org/wp-content/uploads/2024/03/CFIR-Updated-List-of-Constructs-2024-03-26.pdf>

Figuur 13 Het Consolidated Framework for Implementation Research (CFIR).

Consolidated Framework for Implementation Research (CFIR)



De auteurs benadrukken dat CFIR geen ontwikkelkader is voor interventies en geen instrument voor procesevaluatie. Het richt zich uitsluitend op het identificeren van determinanten die de implementatie beïnvloeden. Uitkomstmeting valt buiten het oorspronkelijk kader, maar kan worden aangevuld via het CFIR Outcomes Addendum (81), dat een onderscheid maakt tussen:

- Verwachte uitkomsten (prospectief): adoptability, implementability, sustainability
- Feitelijke uitkomsten (retrospectief): adoptie, implementatie, sustainment

De CFIR kan bovendien gekoppeld worden aan **ERIC (Expert Recommendations for Implementing Change)**: een set van strategieën om implementatieproblemen aan te pakken. Dit helpt bij de stap van diagnose naar actie: vindt men een probleem in een CFIR-construct, dan zijn er aanbevolen ERIC-strategieën om dat probleem aan te pakken¹⁷ (82).

¹⁷ Meer informatie: <https://cfirguide.org/choosing-strategies/>

7.2 Proctor's Implementation Outcomes

Proctor en collega's merkten op dat onderzoek naar interventies vaak sterk gericht is op effectiviteit en gezondheidsuitkomsten, terwijl er weinig aandacht is voor de mate waarin interventies daadwerkelijk en succesvol worden geïmplementeerd in de praktijk (83). Om dit hiaat te dichten, ontwikkelden ze een raamwerk dat uitkomsten op drie niveaus onderscheidt: cliënt- of patiëtniveau (bv. tevredenheid, symptomen), dienstenniveau (bv. effectiviteit, equity) én implementatieniveau. Voor het implementatieniveau ontwikkelden ze een **raamwerk** bestaande uit acht implementatie-uitkomsten die samen een volledig beeld geven van de kwaliteit van implementatie:

1. **Acceptability** (acceptatie): ervaren geschiktheid van de interventie door doelgroep, uitvoerders en andere betrokkenen.
2. **Adoption/Uptake** (toepassing): mate waarin de interventie daadwerkelijk wordt gebruikt.
3. **Appropriateness** (geschiktheid): ervaren relevantie van de interventie voor een specifieke context, doelgroep of probleem.
4. **Feasibility** (haalbaarheid): praktische uitvoerbaarheid van de interventie binnen een organisatie of setting.
5. **Fidelity** (getrouwheid): mate waarin de interventie wordt uitgevoerd zoals bedoeld.
6. **Cost** (kosten): financiële implicaties van implementatie.
7. **Penetration/Reach** (bereik): mate waarin de interventie geïntegreerd is in de organisatie of doelgroep.
8. **Sustainability** (duurzaamheid): mate waarin de interventie op lange termijn behouden blijft.

Het gebruik van deze uitkomsten maakt zichtbaar **waarom een interventie werkt, onder welke omstandigheden en in welke mate ze succesvol geïmplementeerd is**. Het raamwerk helpt bij het systematisch evalueren van implementatieprocessen, het identificeren van barrières en facilitators, en het verbeteren van effectiviteit en duurzaamheid. Cruciaal is het onderscheid tussen falen door een slecht ontwerp versus falen door gebrekkige implementatie.

Een scoping review toonde aan dat vooral acceptatie, haalbaarheid, geschiktheid, toepassing en fideliteit vaak worden bestudeerd, terwijl bereik, duurzaamheid en kosten minder aandacht krijgen. Hoewel het aantal studies groeit, blijft rigoureu empirisch onderzoek naar de onderlinge relaties tussen implementatie-uitkomsten beperkt (84). Proctor en collega's pleiten daarom voor meer theoretisch onderbouwde modellen en objectieve meetinstrumenten.

7.3 Implementation Research Logic Model

Het **Implementation Research Logic Model (IRLM)** werd ontwikkeld om de complexiteit van implementatieonderzoek beter te structureren en te rapporteren (85). Veel bestaande raamwerken behandelen slechts onderdelen van implementatieonderzoek (zoals determinanten of uitkomsten), terwijl een model dat deze elementen samenbrengt vaak ontbreekt. Het IRLM vult deze leemte op door de volgende componenten logisch met elkaar te verbinden: determinanten (barrières en facilitatoren), implementatiestrategieën, mechanismen achter strategieën en uitkomsten (zie Figuur 14).

De determinantendomeinen zijn rechtstreeks overgenomen uit CFIR. Voor de uitkomstenkolom kan het raamwerk van Proctor worden gebruikt, met een onderscheid tussen: implementatie-uitkomsten (bv. acceptatie, adoptie, haalbaarheid, kosten, enz.), dienstuitkomsten en cliënt- of patiëntuitkomsten.

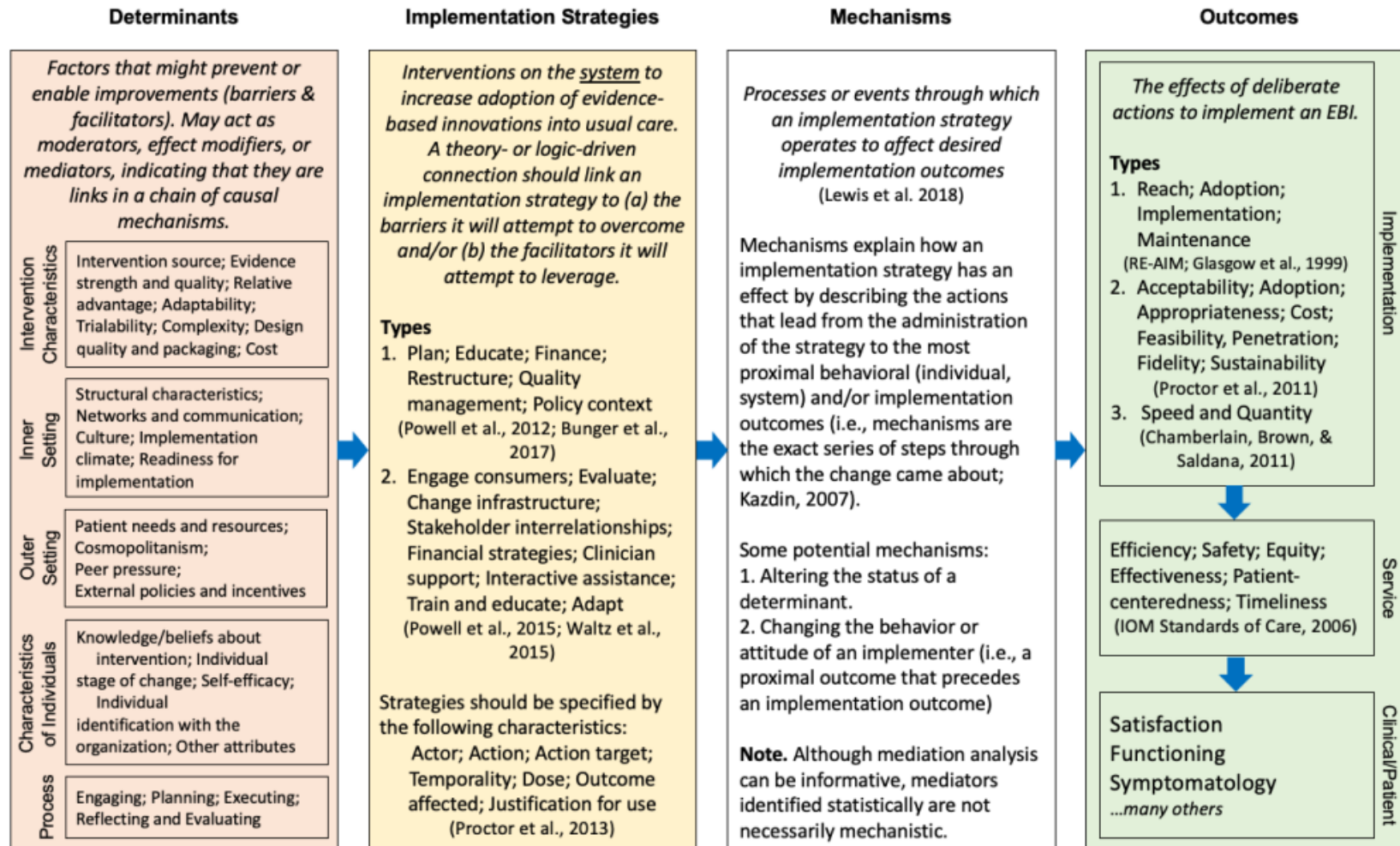
Het uitgangspunt van het IRLM is dat een interventie pas effect kan hebben wanneer ze correct wordt uitgevoerd. Implementatie-uitkomsten vormen daarom voorwaarden voor latere veranderingen in uitkomsten op het niveau van individuen en diensten. Het onderscheid tussen **implementatie-effectiviteit** en **interventie-effectiviteit** is cruciaal om te kunnen bepalen of falen het gevolg is van een gebrekkige uitvoering (implementatiefalen) of van een interventie die in een bepaalde context niet werkt (interventiefalen).

De auteurs formuleren drie leidende principes voor het gebruik van het IRLM:

- **Streven naar volledigheid:** alle relevante determinanten, strategieën en uitkomsten opnemen.
- **Duidelijk maken van conceptuele relaties:** bijvoorbeeld via kleuren of notaties, zodat veronderstelde causale verbanden en hypothesen zichtbaar worden.
- **Specificeren van ontwerpvariabelen:** zoals primaire uitkomsten en de strategieën of variabelen die gemanipuleerd worden.

Ter ondersteuning stellen de auteurs verschillende sjablonen en voorbeelden beschikbaar die onderzoekers houvast bieden bij het toepassen van het IRLM (85).

Figuur 14 Het Implementation Research Logic Model (IRLM)



Smith, J.D., Li, D., & Rafferty, M.R. (2020)

8 Uitgebreide kaders voor het ontwikkelen, plannen en evalueren van interventies

Hoewel dit hoofdstuk focust op evaluatiekaders, bespreken we ook twee meer omvattende kaders die naast evaluatie ook de ontwikkeling en planning van interventies structureren.

8.1 PRECEDE-PROCEED Model

Het PRECEDE-PROCEED model is een uitgebreid raamwerk voor het plannen, implementeren en evalueren van gezondheidsbevorderende interventies (86). Het model hanteert een ecologische en participatieve benadering en integreert zowel individuele als omgevingsfactoren. Er wordt een onderscheid gemaakt tussen de PRECEDE- en de PROCEED fase, die samen acht stappen omvatten, zoals weergegeven in Figuur 15 (87):

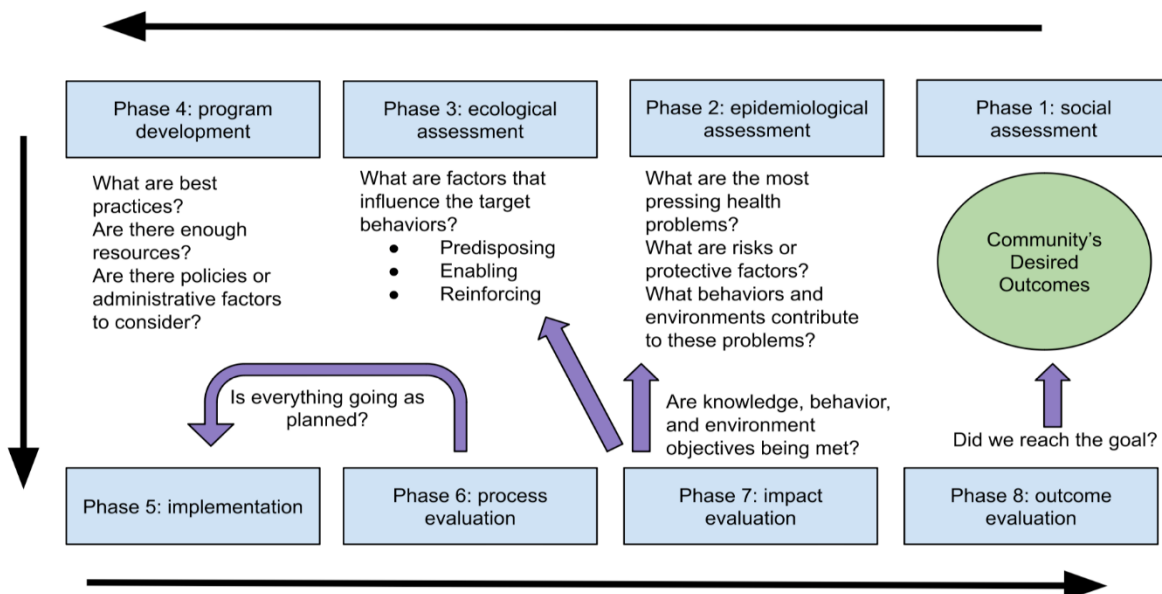
PRECEDE staat voor *Predisposing, Reinforcing, and Enabling Constructs in Educational Diagnosis and Evaluation* en richt zich op:

1. **Sociale analyse:** Identificeren van sociale problemen, noden en prioriteiten van de doelgroep en vastleggen van de ultieme gewenste uitkomst.
2. **Epidemiologische analyse:** In kaart brengen van gezondheidsproblemen en prioritaire gedrags- en omgevingsdeterminanten die moeten worden aangepakt.
3. **Educatieve en ecologische analyse:** In kaart brengen van predisponerende, versterkende en faciliterende factoren die invloed hebben op de gedrags- en omgevingsdeterminanten.
4. **Programma- en beleidsanalyse:** Inventariseren van goede praktijken voor het ontwerp van de interventie en analyseren van middelen, capaciteit, beleid en organisatorische factoren die implementatie kunnen beïnvloeden.

PROCEED staat voor *Policy, Regulatory, and Organizational Constructs in Educational and Environmental Development* en richt zich op:

5. **Implementatie:** Uitvoeren van de interventie.
6. **Procesevaluatie:** Nagaan of de interventie wordt uitgevoerd zoals gepland.
7. **Impactevaluatie (korte termijn):** Meten van de effect op gedrags- en omgevingsdeterminanten uit stap 2.
8. **Uitkomstevaluatie (lange termijn):** Evalueren van veranderingen in gezondheidstoestand (ultiem gewenst resultaat uit stap 1).

Figuur 15 Het PRECEDE-PROCEED model.



Het model biedt een gestructureerd, maar flexibel proces. Een systematische review toont aan dat PRECEDE-PROCEED vooral wordt toegepast bij het ontwerp en de evaluatie van educatieve interventies, en dat het gebruik ervan leidt tot significante verbeteringen in kennis en gedragsverandering (88).

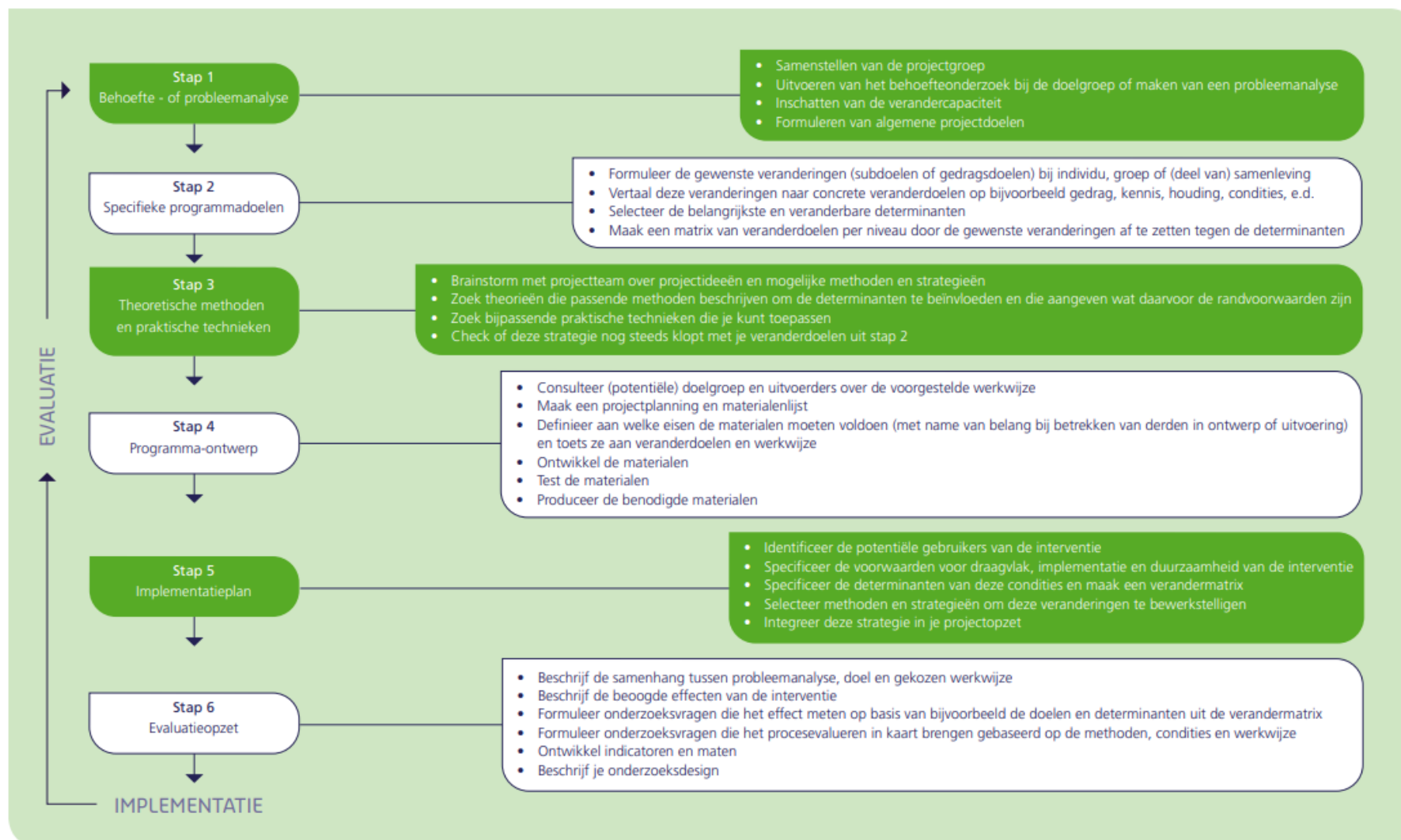
Tegelijkertijd brengt het model enkele uitdagingen met zich mee. De uitgebreide analyse- en planningsfasen vereisen aanzienlijke capaciteit, samenwerking en tijd, wat in preventieve settings met beperkte middelen soms moeilijk haalbaar is. De PRECEDE-fasen vragen bovendien om robuuste sociale, gedragsmatige en beleidsmatige data (89).

8.2 Intervention Mapping

Het Intervention Mapping-protocol is een systematische, gedetailleerde aanpak voor het ontwikkelen, uitvoeren en evalueren van gezondheidsbevorderende interventies (90). Het protocol integreert theoretische inzichten, participatie van stakeholders en een ecologische benadering.

Intervention Mapping gebruikt sociaal-cognitieve theorieën om gedrag te verklaren en kiest op basis daarvan bewezen methoden voor gedragsverandering. De veranderingsmethoden worden vertaald naar concrete toepassingen die passen bij de doelgroep en de context. Intervention Mapping erkent dat individuen functioneren binnen een bredere sociale en organisatorische omgeving. Daarom richt een interventie zich ook op intermediairs en omgevingsactoren (bv. collega's en leidinggevendenden bij interventies in de werksetting) (91). Figuur 16 toont de verschillende stappen in Intervention Mapping. Het proces bestaat uit zes stappen en is iteratief waarbij nieuwe inzichten kunnen leiden tot herziening van eerdere stappen.

Figuur 16 Stappen Intervention Mapping. Overgenomen uit: de Groot, Nada, and Hanneke Mateman. *Zicht op effect: Een overzicht van instrumenten om zelf het effect van je aanpak te meten*. Movisie, kennis en aanpak van sociale vraagstukken, 2014.



Hoofdstuk 6

Uitvoering van effectevaluatie van interventies

Dit hoofdstuk gaat in op de operationele uitvoering van effectevaluaties van gezondheidsbevorderende interventies. Waar eerdere hoofdstukken focusten op evaluatieplanning en theoretische kaders, wordt hier concreet gemaakt welke gegevens nodig zijn, wanneer deze moeten worden verzameld en op welke manier dit het best kan gebeuren. Effectevaluatie is geen éénmalige stap, maar een proces dat idealiter de volledige levenscyclus van een interventie bestrijkt: van vroege pilootstudies en haalbaarheidsonderzoek tot grootschalige effectiviteitsstudies en duurzame implementatie. In dit hoofdstuk worden de belangrijkste onderzoeksdesigns, dataverzamelingsstrategieën en methodologische overwegingen besproken die nodig zijn om evaluatievragen betrouwbaar te beantwoorden. Hiermee vormt het hoofdstuk een praktisch hulpmiddel voor onderzoekers en beleidsmakers die robuust en onderbouwd willen vaststellen of, hoe en waarom interventies werken in de praktijk.

1 Van piloottest tot impactevaluatie bij gezondheidsbevorderende interventies

De evaluatie van gezondheidsbevorderende interventies doorloopt doorgaans meerdere fasen die aansluiten bij de ontwikkelings- en implementatiecyclus van een interventie (92). In de vroege **ontwikkelingsfase** wordt nagegaan of de inhoud, structuur en uitvoering van de interventie daadwerkelijk bijdragen aan het beoogde effect. Dit gebeurt doorgaans via pilootstudies en haalbaarheids- en werkzaamheidsonderzoek (*efficacy trials*), waarbij in een gecontroleerde en kleinschalige setting wordt onderzocht of de interventie werkt onder ideale omstandigheden. Deze fase biedt vooral inzicht in interne validiteit en laat toe om gerichte verbeteringen door te voeren vóór bredere uitrol.

Wanneer de interventie voldoende is ontwikkeld, volgt **implementatie op grotere schaal** en wordt de effectiviteit geëvalueerd in realistische omstandigheden. Effectiviteitsonderzoek (*effectiveness trials*) onderzoekt of de interventie werkt bij diverse doelgroepen en in uiteenlopende settings, en geeft zicht op uitkomsten op korte en langere termijn.

Sommige onderzoekers vinden dit lineaire model in het preventieonderzoek te rigide. Zij pleiten ervoor om effectiviteit reeds gedeeltelijk te onderzoeken tijdens de efficacy-fase, bijvoorbeeld via hybride designs waarin zowel interne validiteit (controle over variabelen) als externe validiteit (realistische setting) wordt meegenomen (93).

Na afronding van de interventie wordt de totaalwaarde van de interventie beoordeeld. Hierbij wordt gekeken naar duurzaamheid, acceptatie door de doelgroep en eventuele onverwachte effecten. Deze inzichten ondersteunen verdere optimalisering en innovatie van gezondheidsbevorderende interventies.

1.1 Van werkzaamheid naar effectiviteit in preventieonderzoek

Het artikel “Moving from efficacy to effectiveness trials in prevention research” bespreekt cruciale overwegingen bij de overgang van gecontroleerde werkzaamheidsstudies naar effectiviteitsonderzoek in realistische settings (93). De auteurs benadrukken dat preventieonderzoek specifieke uitdagingen kent: interventies vinden plaats in complexe, minder gecontroleerde omgevingen en bereiken diverse

doelgroepen. Daardoor is het niet vanzelfsprekend dat positief bevonden interventies onder gecontroleerde omstandigheden ook werkzaam blijken in de praktijk.

Op basis van literatuur en eigen onderzoekservaring identificeren de auteurs negen domeinen die bepalend zijn voor een succesvolle overgang van efficacy- naar effectiveness-onderzoek. Een overzicht hiervan wordt gegeven in Tabel 5. De auteurs gaan niet uitgebreid in op disseminatie, maar beschouwen deze fase als een logisch vervolg op het proces, waarin bewezen effectieve interventies duurzaam worden geïmplementeerd.

Tabel 5 Negen domeinen die van belang zijn bij de overgang van werkzaamheidsonderzoek (efficacy) naar effectiviteitsonderzoek (effectiveness) volgens Marchand et al. (2011).

	Efficacy / werkzaamheid	Effectiveness / effectiviteit
Onderzoeksprogramma	Gecontroleerde trials die interne validiteit garanderen, maar beperkte generaliseerbaarheid hebben.	Hybride of pragmatische trials die zowel experimentele controle als realistische variatie omvatten en daarmee meer generaliseerbaar zijn.
Design en conceptualisering van interventie en evaluatie	RCT's zijn het meest voorkomende design om de interventie als belangrijkste variabele te isoleren, vaak met weinig aandacht voor generaliseerbaarheid.	Designs bevatten meer variatie in deelnemers, aanbieders en settings, maar zijn niet altijd gerandomiseerd. Implementatievermogen en betrokkenheid en motivatie van betrokkenen moeten worden meegenomen.
Selectie en kenmerken van deelnemers	Kleine, homogene steekproef qua geografie, risicofactoren of andere variabelen. Onderzoekers zijn verantwoordelijk voor rekrutering.	Grotere, heterogene steekproef om verschillen tussen groepen en contexten te kunnen analyseren in functie van generaliseerbaarheid. Praktijkpartners rekruteren vaak deelnemers.
Verstrekken van de interventie	Interventies uitgevoerd door goed getrainde onderzoekers/assistenten met voldoende tijd en motivatie voor voorbereiding en uitvoering.	Uitvoering door getrainde leken of professionals in het veld, die conflicterende opvattingen, minder tijd, supervisie of motivatie kunnen hebben en soms weerstand ervaren.
Context van interventie-uitvoering	Beperkt aantal gecontroleerde locatie(s), zoals een onderzoeksruimte of specifieke klas.	Meerdere, diverse en realistische locaties met variatie in omstandigheden.
Metingen en methodologie	RCT als gouden standaard; lange vragenlijsten; nauwgezette opvolging door onderzoekers om uitval en procedurele afwijkingen te minimaliseren.	Alternatieve designs zijn mogelijk; kortere metingen zijn nodig; onderzoekers moeten voorbereid zijn op hogere uitval en onregelmatigheden door meerdere uitvoerders.
Selectie van uitkomsten	Focus op preventieve effecten na afloop en follow-up.	Naast preventieve effecten ook haalbaarheid door uitvoerders, modererende effecten, getrouwheid aan implementatie en andere implementatie-uitkomsten.
Kosten	Kosten worden zelden meegenomen in dit stadium. Interventie wordt doorgaans volledig gefinancierd met subsidies.	Aanbevolen om de kosten voor uitvoering en potentiële besparingen te onderzoeken. Interventies kunnen deels of volledig zelf gefinancierd zijn.
Duurzaamheid in de praktijk	Wordt meestal niet meegenomen in dit stadium.	Cruciaal: haalbaarheid in gemeenschappen, acceptatie door stakeholders en langdurige implementatie.

Box 4: De overgang van werkzaamheid (efficacy) naar effectiviteit geïllustreerd met onderzoek naar eetstoornispreventie

De overgang van efficacy- naar effectiveness-onderzoek wordt door Marchand (93) geïllustreerd aan de hand van twee studies naar eetstoornispreventie.

Het eerste voorbeeld betreft het *Body Project*, dat een klassiek, lineair traject volgde. In een reeks werkzaamheidsstudies konden studenten (eerst universitair, later middelbaar) zichzelf inschrijven voor een interventie gericht op *body positivity*. De interventie werd uitgevoerd door getrainde onderzoekers (o.a. doctoraatsstudenten) in een gecontroleerde labosetting. Na positieve resultaten in deze trials werd een effectiviteitsstudie opgezet in scholen. Hierbij lag de verantwoordelijkheid voor zowel de rekrutering als de uitvoering bij het schoolpersoneel, ondersteund door onderzoekers via materiaal, workshops en supervisie. Ook in deze praktijksetting bleken de effecten stand te houden: het Body Project werkte niet alleen onder ideale omstandigheden, maar ook in realistische situaties.

Het tweede voorbeeld richtte zich op studentenverenigingen (sororities) en volgde een meer participatieve en hybride aanpak. In een eerste kleine pilot werden leden van een lokale studentenvereniging willekeurig toegewezen aan de interventie, een actieve controlegroep of een wachtlijst. Deze opzet combineerde elementen van gecontroleerde efficacy-studies met kenmerken van effectiveness-onderzoek, aangezien werving en uitvoering grotendeels door de studenten zelf werden verzorgd en er geen financiële compensatie of externe financiering beschikbaar was.

De positieve resultaten leidden ertoe dat de studentenvereniging besloot om het programma verplicht te maken voor alle nieuwe leden, waarbij studenten zelf konden kiezen of zij deelnamen aan het onderzoek. Door gebrek aan middelen voor professionele aanbieders, werden studenten gerekruteerd en opgeleid als peer-aanbieders. Dit vereiste meer training, maar maakte de interventie wel duurzaam en breed inzetbaar. Het draagvlak groeide uiteindelijk zo sterk dat een grote nationale studentenvereniging voor vrouwen het programma overnam.

1.2 Kleine studies, belangrijke gevolgen: misstappen bij het overgaan tot meer grootschalige of complexe evaluatiestudies op basis van resultaten van piloottesten

Het artikel "*Small studies, big decisions: the role of pilot/feasibility studies in incremental science and premature scale-up of behavioral interventions*" pleit voor een meer strategische benadering van piloot- en haalbaarheidsstudies in gedragsinterventies (94). De auteurs benadrukken dat beleidsmakers vooral zouden moeten investeren in **kwalitatief sterke, realistische pilots in de beginfase van een interventie**.

Nieuwe interventies starten doorgaans met kleinschalige piloot- of haalbaarheidsstudies. Een veelvoorkomende misvatting is dat een interventie die goed werkt in een kleine gecontroleerde studie automatisch ook effectief zal zijn in een grootschalige evaluatie. Kleine studies kunnen echter tot **misleidend optimistische conclusies** leiden, zowel wat betreft haalbaarheid als effectiviteit. Hierdoor wordt vaak te snel opgeschaald, waardoor interventies in latere, duurdere evaluatiefases falen.

Een goed uitgevoerde pilotstudie helpt onderzoekers beter voorbereid te zijn op de uitdagingen van opschaling, maar dit vergt tijd, middelen en een zorgvuldig ontwerp. Het artikel bespreekt veelvoorkomende fouten in pilootstudies, hoe deze het bewijs vertekenen en welke richtlijnen gevolgd moeten worden om te beslissen of grootschaliger onderzoek verantwoord is.

1.2.1 Steekproefgrootte

Pilootstudies werken vaak met kleine steekproeven (de literatuur vermeldt een mediaan van ongeveer 60 deelnemers), omdat rekrutering en uitvoering arbeids- en kostenintensief zijn. In deze fase is statistische power minder doorslaggevend: zodra een studie voldoende power heeft om effecten aan te tonen, is het strikt genomen geen pilotstudie meer.

Belangrijker dan power is inzicht in factoren die bij een latere grootschalige uitrol verschillen in effectiviteit of uitvoerbaarheid kunnen veroorzaken. Onderzoekers moeten daarom vooraf nadenken over de interactie tussen interventie, doelgroep en setting en over de vraag **hoeveel variatie de pilot al moet kunnen detecteren om toekomstige opschaling te kunnen onderbouwen**. De benodigde steekproef moet dus afgestemd zijn op de variatie in contexten waarin de interventie bij opschaling zal plaatsvinden.

Er bestaat geen vaste vuistregel om de ideale steekproefgrootte te bepalen; een duidelijke onderbouwing door de onderzoekers blijft noodzakelijk. Voor interventies met homogene doelgroepen in uniforme settings kunnen kleinere pilots volstaan, terwijl complexe interventies in diverse contexten vragen om een grotere en meer gelaagde pilootopzet.

1.2.2 Opschaling en aanpassing

Een veelvoorkomende denkfout is dat een interventie die effectief blijkt in een pilot bij opschaling automatisch hetzelfde effect zal hebben. In werkelijkheid ondergaat een interventie vaak wijzigingen wanneer zij verschuift van een kleine, gecontroleerde setting naar een grotere en realistischere omgeving. Denk aan aanpassingen in inhoud of componenten, verschillen in intensiteit of duur en andere uitvoerders of doelgroepen. Wanneer deze aanpassingen substantieel zijn, zijn resultaten uit de oorspronkelijke pilot mogelijk niet langer representatief. Toch worden dergelijke interventies vaak opgeschaald zonder herpilotering, wat de kans op mislukking aanzienlijk vergroot. Opschaling zonder extra pilot is enkel verantwoord wanneer aanpassingen minimaal zijn.

Veelvoorkomende (denk)fouten in piloot- en haalbaarheidsstudies die de interpretatie van haalbaarheid en effectiviteit vertekenen, zijn onder andere:

- **Neglecting to repilot:** Grote interventieaanpassingen worden opgeschaald zonder herpilotering.
- **Delivery agent bias:** In de pilot gebeurt de uitvoering van de interventie vaak door de ontwikkelaar zelf in plaats van door de bedoelde uitvoerders (bv. leerkrachten), wat onrealistisch positieve resultaten oplevert.
- **Implementation support bias:** Pilots krijgen vaak intensieve begeleiding en feedback, terwijl dergelijke ondersteuning in grootschalige implementaties niet haalbaar is.
- **Target audience bias:** Pilots rekruteren gemotiveerde, responsieve deelnemers, waardoor haalbaarheid en effectiviteit overschat worden.
- **Intervention duration bias:** Pilots zijn vaak korter dan de uiteindelijke interventie, waardoor novelty-effecten of hoge initiële motivatie de resultaten vertekenen.
- **Conflating feasibility with impact:** Sommige pilootstudies focussen enkel op procesindicatoren (opkomst, tevredenheid en acceptatie), zonder effectmaten te evalueren. Hierdoor is er onvoldoende basis om opschaling te rechtvaardigen.

De kern van het probleem is dat pilootresultaten soms worden geïnterpreteerd als bewijs voor effectiviteit, waarna cruciale pilot-specifieke omstandigheden wegvallen bij de grootschalige trial. Dit leidt vaak tot teleurstellende resultaten. Pilootstudies zijn waardevol, maar mogen niet de enige basis zijn voor beslissingen over opschaling.

1.2.3 Financiering van kleine studies

Hoewel piloot- en haalbaarheidsstudies essentieel zijn voor de verantwoorde opschaling van evaluaties, worden deze studies beperkt gefinancierd. Grote fondsen richten zich eerder op effectstudies, terwijl de middelen voor pilootstudies beperkt zijn. Dit leidt ertoe dat onderzoekers vanwege beperkte middelen suboptimale keuzes moeten maken in het ontwerp van hun pilot/haalbaarheidsstudie (bv. met eigen personeel als uitvoerder, te korte duur, onvoldoende diversiteit in deelnemers), welke kunnen leiden tot een te optimistisch beeld van de haalbaarheid en effectiviteit van de interventie. Dit creëert een paradoxale situatie: grootschalige evaluaties zijn afhankelijk van robuuste en realistische data uit pilootstudies, maar de middelen om die data verantwoord te genereren zijn ontoereikend.

Beleidsmakers en financiers moeten het strategische belang van pilootstudies erkennen en structurele middelen voorzien voor deze fase in de ontwikkelcyclus. Daarnaast zouden financieringssystemen best kwaliteitscriteria of richtlijnen bevatten die de genoemde (denk)fouten voorkomen. Zonder kwalitatieve pilootstudies bestaat het risico dat grootschalige implementaties vaker falen, wat op lange termijn zorgt voor een inefficiënt gebruik van middelen.

1.3 Wanneer is effectevaluatie van een implementatie nodig?

1.3.1 Evaluability assessment: wanneer is effectevaluatie zinvol?

Een *evaluability assessment* is een verkennend vooronderzoek dat beoordeelt of en hoe een interventie op een zinvolle manier kan worden geëvalueerd. Het doel is te bepalen of een interventie voldoende ontwikkeld, duidelijk beschreven en stabiel geïmplementeerd is om een grondige effectevaluatie te rechtvaardigen, en welke vorm van evaluatie dan het meest geschikt is (55).

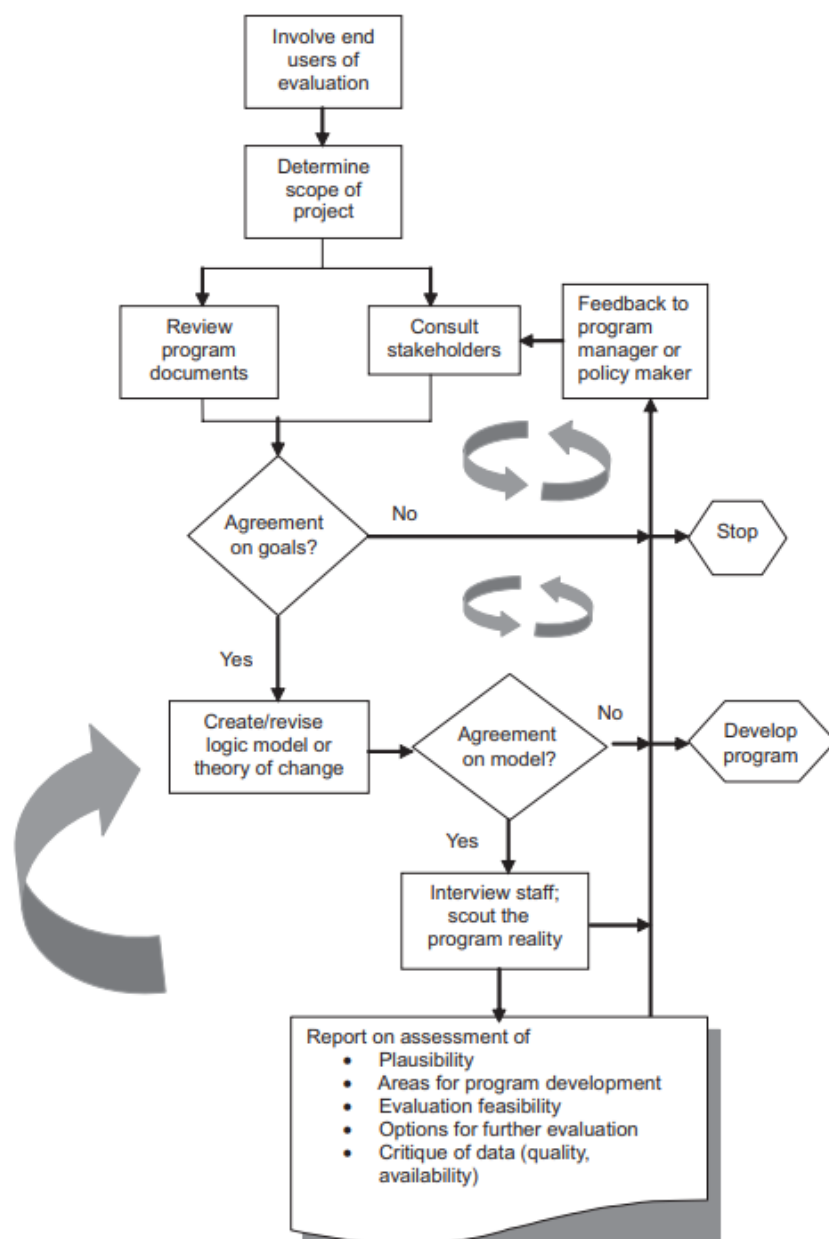
Een evaluability assessment onderzoekt onder meer:

- **Duidelijkheid van het programma:** Zijn doelstellingen, doelgroep, activiteiten en verwachte effecten helder omschreven?
- **Beschikbaarheid van een Theory of Change of Logic Model:** Is het mechanisme achter de interventie expliciet gemaakt?
- **Implementatiekwaliteit:** Wordt de interventie voldoende uitgevoerd zoals bedoeld?
- **Beschikbaarheid en kwaliteit van data:** Zijn er al bruikbare gegevens beschikbaar of is aanvullende dataverzameling haalbaar?
- **Stakeholderbetrokkenheid en draagvlak:** Zijn betrokken partijen bereid én in staat om mee te werken aan dataverzameling en de resultaten te gebruiken?
- Een evaluability assessment levert op zich geen bewijs van effectiviteit, maar helpt om de interventie indien nodig te verfijnen, haalbare en relevante evaluatievragen te formuleren, een passend evaluatiedesign te kiezen en middelen doelgericht in te zetten.

De uitgebreide stappen in een evaluability assessment worden beschreven door Leviton et al. (95). Een samenvatting hiervan wordt weergegeven in Figuur 17. Het eindoordeel van een evaluability assessment bevat volgens het artikel vijf kernonderdelen:

- **Beoordeling van de plausibiliteit van de interventie:** In hoeverre is het realistisch dat het programma de gewenste resultaten kan behalen?
- **Aanbevelingen ter verbetering:** Wat zijn suggesties om onderdelen van het programma (bv., activiteiten, timing) te optimaliseren.
- **Beoordeling van de haalbaarheid van een volledige evaluatie:** In welke mate is een formele effectevaluatie uitvoerbaar, gezien context, middelen en beschikbare data?
- **Analyse van datakwaliteit en beschikbaarheid:** Welke bestaande gegevens kunnen worden gebruikt en welke bijkomende dataverzameling is nodig?
- Voorstellen voor relevante evaluatievragen en geschikte onderzoeksdesigns.

Figuur 17 Stappen in een evaluability assessment volgens Leviton et al. (2010).



Box 5: Beoordeling van de evalueerbaarheid van complexe volksgezondheidsinterventies: vijf reflectievragen voor onderzoekers, financiers en beleidsmakers (96).

Ogilvie et al. (96) formuleren vijf richtinggevende vragen die kunnen helpen bij het bepalen of een complexe gezondheidsbevorderende interventie evalueerbaar is. De vragen vormen geen checklist, maar een instrument voor gezamenlijke reflectie, kennisdeling en beleidsvertaling tijdens verschillende fasen van een interventie.

1) In welke fase van ontwikkeling bevindt de interventie zich?

Is de interventie nog in concept- of ontwikkelingsfase, of al in implementatie?

Vroege effectevaluatie kan misleidend zijn wanneer interventies nog volop worden aangepast. In die fase zijn procesvragen (bv. drempels, doelgroepbereik) relevanter dan effectmetingen.

2) Hoe kan een effectevaluatie beleidsbeslissingen beïnvloeden?

Is een effectevaluatie beleidsrelevant op dit moment? Soms is een meer genuanceerde analyse nodig: een inschatting van wat beleidsmakers waarschijnlijk zouden doen of aanbevelen op basis van bestaande kennis (policy prior), en een inschatting van wat ze zouden beslissen als nieuwe onderzoeksresultaten beschikbaar zijn (policy post). Onderzoekers moeten tegelijk vermijden dat onderzoek te instrumenteel wordt ingezet en dus enkel als waardevol wordt beschouwd als het direct een beleidsvraag beantwoordt.

3) Wat zijn de te verwachten grootte en verdeling van de impact?

Een interventie met een groot effect voor een kleine groep heeft mogelijk beperkte impact op populatieniveau, terwijl een interventie met een klein effect maar een groot bereik wél substantieel kan bijdragen. Toch kunnen ook nieuwe kleinschalige interventies de moeite waard zijn om te evalueren, zelfs als de verwachte populatiewinst beperkt is. Dat geldt vooral als er aanwijzingen zijn voor ongewenste effecten, als de interventie bijdraagt aan andere beleidsdomeinen, wanneer er impact op ongelijkheid mogelijk is, of als opschaling kansrijk is.

4) Hoe zal de effectevaluatie waarde toevoegen aan bestaande wetenschappelijke evidentie?

Niet elke interventie hoeft uitgebreid onderzocht te worden, zeker niet als de effectiviteit al voldoende bewezen is. Evaluatie kan wel nuttig zijn wanneer het onduidelijk is of de interventie effect heeft op andere groepen, op andere uitkomsten of op ongelijkheid, of wanneer er informatie ontbreekt over het onderliggende mechanismen of de duurzaamheid van de interventie.

5) Is het uitvoerbaar om de interventie te evalueren binnen de beschikbare tijd?

Tot slot zijn er praktische overwegingen die de uitvoerbaarheid van een effectevaluatie bepalen, zoals: de benodigde tijd voor dataverzameling en follow-up, stabiliteit van de interventie (blijft ze lang genoeg bestaan om effecten te kunnen meten), mogelijkheid om de uitkomsten van de interventie te isoleren van andere factoren, en veranderingen in context of inhoud van de interventie over de tijd.

2 Kwantitatieve onderzoeksdesigns voor effectevaluatie

Het kiezen van een geschikt onderzoeksdesign hangt af van de evaluatievraag en van verschillende praktische overwegingen. Het ontwerp moet steeds een balans vinden tussen wetenschappelijke bewijskracht en uitvoerbaarheid in de praktijk. In dit onderdeel worden de belangrijkste kwantitatieve onderzoeksdesigns besproken die kunnen worden ingezet voor effectevaluatie. De beschrijvingen zijn gebaseerd op volgende bronnen:

- WHO-Europe gids voor de impactevaluatie van gezondheidsbevorderende interventies (19)
- Passend bewijs voor preventie door de Adviescommissie richtlijn passend bewijs voor preventie in Nederland (97)
- Handreiking onderzoeksmethoden leefstijlinterventies in de zorg door Coalitie Leefstijl in de Zorg in Nederland: mogelijkheden voor het evalueren van effectiviteit (98)
- Europees Preventiehandboek Middelen - Hoofdstuk 4: Monitoren en evalueren (92)

2.1 Pre-experimentele designs

Pre-experimentele designs worden gebruikt wanneer geen controlegroep of randomisatie wordt toegepast. Ze zijn relatief eenvoudig en goedkoop uit te voeren, maar bieden beperkte mogelijkheden om causale conclusies te trekken. Ze zijn vooral geschikt als eerste verkenning of pilot.

2.1.1 Enkelvoudige posttest studie in één groep

Bij dit design wordt slechts één meting bij één groep uitgevoerd na afloop van de interventie (bijvoorbeeld via een vragenlijst bij deelnemers). Dit kan informatie opleveren over ervaringen of tevredenheid, maar laat geen uitspraken toe over effectiviteit omdat veranderingen niet kunnen worden vastgesteld en alternatieve verklaringen niet kunnen worden uitgesloten. Het is daarom enkel geschikt als eerste pilottering.

2.1.2 Pretest-posttest studie in één groep

Dit studiedesign is veelgebruikt, omdat het relatief snel, goedkoop en gemakkelijk te organiseren is. Het verzamelt dezelfde gegevens voor en na de interventie bij dezelfde deelnemers. Dit maakt het mogelijk om veranderingen binnen de groep vast te stellen, maar zonder controlegroep kan het effect niet eenduidig aan de interventie worden toegeschreven. Externe factoren of natuurlijke trends kunnen een rol spelen.

2.2 Quasi-experimentele designs

Quasi-experimentele designs proberen de impact van een interventie te meten door de interventiegroep te vergelijken met een hypothetische situatie zonder interventie, een zogenaamde **counterfactual**. Vaak is dit een aparte controlegroep.

In tegenstelling tot experimenteel onderzoek, maken quasi-experimentele designs echter geen gebruik van randomisatie (willekeurig toewijzing aan een interventie- of controlegroep). Hierdoor spreken we van designs met **niet-equivalente groepen**, wat betekent dat er mogelijk al verschillen bestonden tussen de groepen vóór de interventie plaatsvond. Om toch een inschatting te kunnen maken van het causale effect van een interventie, wordt gebruik gemaakt van statistische technieken zoals *matching* (Box 6).

De counterfactual hoeft niet noodzakelijk een aparte controlegroep te zijn. In sommige designs wordt de interventiegroep vergeleken met zichzelf over de tijd. Aan de hand van herhaalde metingen vóór en na de interventie kan men trends analyseren en inschatten of er een verandering is opgetreden die kan worden toegeschreven aan de interventie.

Box 6: Keuze en matching van een controlegroep

Er zijn verschillende opties voor het selecteren van een controleconditie bij gedragsmatige interventies (99). Meestal worden **inactieve of passieve controlegroepen** gebruikt, waarbij de interventiegroep wordt vergeleken met een groep zonder interventie. Soms betekent dit dat ze helemaal geen interventie of de standaardpraktijkvoering ontvangen. Bij een *wachtlijstcontrole* ontvangen de deelnemers de interventie pas na afloop van de studieperiode. Een risico van inactieve controle is dat placebo- en verwachtingseffecten meespelen.

Daarnaast worden soms **actieve controlegroepen** gebruikt, vergelijkbaar met placebo's in farmacologisch onderzoek (i.e., lijkt qua presentatie op de echte interventie). Dit zijn meestal condities waarbij

de controlegroep een alternatieve activiteit krijgt met een gelijkaardige mate van aandacht en contact, maar zonder de werkzame componenten van de interventie.

Om de werkzame componenten van gedragsinterventies te achterhalen worden **component controls** gebruikt, zoals *additive controls* (toevoegen van interventie-onderdelen) en *dismantling controls* (weglaten van interventie-onderdelen).

Voor quasi-experimentele methoden met controlegroep is het belangrijk dat de interventiegroep en de controlegroep gelijkaardig zijn. Matchingtechnieken creëren gebalanceerde groepen door de deelnemers te koppelen op basis van relevante covariaten of kenmerken, zoals leeftijd, geslacht of sociaaleconomische status (100). Een veelgebruikte vorm is **Propensity Score Matching (PSM)**. Een *propensity score* staat voor de kans dat een individu de behandeling krijgt, gegeven de covariaten, berekend aan de hand van een statistisch model. Op basis van deze score worden deelnemers van beide groepen gematcht. Dit vermindert bias en maakt effectschattingen betrouwbaarder wanneer randomisatie niet haalbaar is. Matching is echter complex, en de resultaten zijn gevoelig voor de gekozen matchingmethode.

2.2.1 Gecontroleerde enkelvoudige posttest studie (static group comparison design)

Deelnemers worden zonder randomisatie toegewezen aan een interventiegroep of controlegroep, en er enkel na de afloop van de interventie vindt een vergelijkende meting plaats¹⁸. Dit design is eenvoudig en minder belastend voor deelnemers, maar zonder pretest is het niet mogelijk om individuele veranderingen te meten. Een goede matching van de groepen kan het design versterken.

2.2.2 Gecontroleerde pretest-posttest studie

Een pretest-posttest studie is een quasi-experimenteel onderzoeksdesign waarin deelnemers zonder randomisatie door de onderzoeker worden toegewezen aan een interventie- of controlegroep. Er wordt een nulmeting uitgevoerd voorafgaand aan de interventie en er vinden één of meer nametingen plaats. Op deze manier is het mogelijk om de effectiviteit van een interventie relatief eenvoudig te evalueren.

Omdat er geen randomisatie is, kunnen systematische verschillen bestaan tussen de interventie- en controlegroep, die een invloed hebben op de uitkomst. Matching o.b.v. baselinedata of demografische gegevens kan de validiteit verhogen, maar het is niet mogelijk om rekening te houden met onderliggende trends die al vóór de interventie bestonden.

Aanvullende strategieën om de interne en externe validiteit van deze designs te vergroten zijn: het vergelijken van de interventie met meerdere controlegroepen, vergelijking in een kritisch tijdsvenster (waarin de interventie verwacht wordt het meeste effect te hebben), of het vergelijken van subgroepen die een lage of hoge blootstelling aan de interventie kregen (i.e., dosis-responsanalyse) (Handley et al., (101).

¹⁸ De enkelvoudige posttest studie met controlegroep (static group comparison design) wordt soms tot de pre-experimentele designs gerekend vanwege de beperkingen, en soms tot de quasi-experimentele designs vanwege het gebruik van een controlegroep.,

2.2.3 Difference-in-differences

Difference-in-Differences (DiD) is een veelgebruikte analysemethode voor pretest-posttest studies om causale effecten van beleidsmaatregelen of interventies in te schatten wanneer randomisatie niet mogelijk of wenselijk is. De methode vereist minimaal twee groepen en twee meetmomenten (pre en post). Het basisprincipe is dat de verandering in de uitkomstvariabele in de interventiegroep wordt vergeleken met die in de controlegroep.

Een cruciale voorwaarde is de **parallele-trendaanname**: in afwezigheid van de interventie zouden beide groepen een gelijkaardige trend hebben gevolgd. Wanneer deze assumptie plausibel is, levert DiD een robuuste schatting op van het gemiddelde effect van de interventie. In tegenstelling tot een eenvoudige vergelijking van verschijscores, waarin veranderingen vaak zonder expliciete interactieterm worden geanalyseerd, modelleert DiD expliciet het interactie-effect tussen tijd en groep. Hierdoor worden zowel tijdsinvariante groepsverschillen als algemene tijdseffecten systematisch gecontroleerd in de schatting van het interventie-effect.

Hoewel DiD vaak wordt toegepast op bestaande groepen die door natuurlijke of externe omstandigheden verschillend worden blootgesteld aan een interventie, kan DiD ook worden toegepast bij interventies met toewijzing. Daarnaast kan de methode worden uitgebreid naar meerdere interventiegroepen, meerdere tijdsperiodes of interventies die op verschillende momenten worden ingevoerd.

2.2.4 Regressie-discontinuïteit

Een *regression discontinuity design* (RDD) maakt gebruik van een drempelwaarde op een continue variabele (bv. leeftijd, testscore, BMI, inkomen) om deelnemers in te delen in een interventiegroep en een controlegroep. Het idee achter RDD is dat deelnemers net boven en net onder de drempelwaarde sterk op elkaar lijken, waardoor verschillen in uitkomst na de interventie met grote waarschijnlijkheid kunnen worden toegeschreven aan de interventie. Dit maakt RDD bijzonder geschikt voor situaties waarin beleid of programma's worden toegekend op basis van een score of criterium. Enkel de uitkomst na de interventie wordt vergeleken bij mensen onder en boven de drempel.

Er bestaan twee vormen van RDD. Bij *sharp RDD* krijgt iedereen boven de drempel de interventie en iedereen eronder niet. Bij *fuzzy RDD* neemt de kans op interventie toe boven de drempel, maar is dit niet volledig bepaald. Hoewel RDD een hoge interne validiteit heeft, is de generaliseerbaarheid beperkt tot personen rond de drempelwaarde. Het kiezen van een geschikte bandbreedte is essentieel: een te brede bandbreedte vergroot het risico op heterogeniteit, terwijl een te smalle bandbreedte kan leiden tot te weinig data. RDD combineert de voordelen van observationeel onderzoek met de validiteit van experimentele designs en is een krachtige methode voor causale inferentie wanneer randomisatie niet mogelijk is, mits bepaalde voorwaarden vervuld zijn (zoals de afwezigheid van andere veranderingen rond de drempelwaarde) (102).

2.2.5 Interrupted time series

Een *interrupted time series* (ITS) is een quasi-experimenteel onderzoeksdesign waarbij in één groep op meerdere tijdstippen metingen plaatsvinden voor en na de implementatie van een interventie. De trends in de uitkomsten na de interventie worden vergeleken met de trends vóór de interventie, zonder controlegroep. Om verschillen tussen beide perioden te corrigeren, kunnen controlevariabelen worden toegevoegd.

In de praktijk wordt aanbevolen om minimaal drie meetmomenten voor en drie na de interventie te voorzien, maar een groter aantal verhoogt de betrouwbaarheid, vooral wanneer de effecten klein zijn of andere factoren (bv. seizoensinvloeden) een rol spelen. ITS is methodologisch sterker dan een enkelvoudige pre-postmeting en is bijzonder geschikt wanneer gegevens routinematig worden verzameld zoals in gezondheidszorgregistraties. Het ontbreken van een controlegroep blijft echter een beperking, omdat externe ontwikkelingen en placebo-effecten niet kunnen worden uitgesloten.

2.2.6 Comparative time series

Een comparative time series bouwt voort op een ITS-opzet, maar voegt een controlegroep toe waarin dezelfde herhaalde metingen plaatsvinden. Hierdoor kunnen alternatieve verklaringen, zoals seizoensinvloeden of veranderingen in zorgkwaliteit, beter worden uitgesloten. Meestal is er geen randomisatie¹⁹. In tegenstelling tot Difference-in-Differences vereist dit design niet dat de trends voor de interventie gelijklopen.

2.2.7 Synthetische controle

Voor beleidsinterventies op hoger niveau (bv. regionale maatregelen) wordt soms een synthetische controlemethode gebruikt. In plaats van één enkele controlegroep te selecteren, wordt een synthetische vergelijkingseenheid geconstrueerd uit een gewogen combinatie van meerdere controle-eenheden (bv. regio's, groepen) die vóór de interventie sterk lijken op de interventiegroep. Deze methode is vooral nuttig bij grootschalige interventies, zoals bij de evaluatie van een gezondheidszorgbeleid.

Een voordeel van de synthetische controlemethode is dat het geen aanname vereist van parallelle trends tussen de interventie-eenheid en de controle-eenheden, zoals wel het geval is bij de difference-in-difference methode. Wel is de methode afhankelijk van uitgebreide en kwalitatief goede data. Wanneer de interventie-eenheid te sterk afwijkt van de controle-eenheden, bijvoorbeeld door unieke kenmerken of beleid, kan het moeilijk zijn om een geschikte synthetische controle te construeren (103).

2.2.8 Natuurlijk experiment

Natuurlijke experimenten ontstaan wanneer gebeurtenissen of beleidsveranderingen plaatsvinden buiten de controle van onderzoekers, zoals bij invoering, aanpassing of afschaffing van maatregelen, diensten, of zorgsystemen. De evaluatie maakt gebruik van data die voortkomen uit deze verandering, waardoor het mogelijk is om effecten van grootschalige of ethisch moeilijk te randomiseren interventies te onderzoeken. Er wordt dus een interventie-effect geëvalueerd zonder willekeurige toewijzing door onderzoekers. Toch zijn er bepaalde toewijzingsprocessen in het beleid die de exposure kunnen bepalen, zoals de interventie enkel uitrollen in bepaalde regio's of bepaalde voorwaarden stellen om recht te hebben op de interventie.

Natuurlijke experimenten hebben als voordelen dat ze interventies onder realistische omstandigheden kunnen bestuderen, lange termijneffecten kunnen meten en ook onbedoelde gevolgen zichtbaar kunnen maken. Het is wel belangrijk dat de groepen vergelijkbaar zijn en dat er voldoende kwalitatieve gegevens beschikbaar zijn vóór en na de verandering.

¹⁹ Indien er wel randomisatie is, is het een experimenteel design.

2.3 Experimentele gerandomiseerde designs (RCT's)

2.3.1 Traditionele en pragmatische (cluster-)RCT

Een gerandomiseerde gecontroleerde studie of *randomized controlled trial* (RCT) is een experimenteel onderzoeksdesign waarbij deelnemers volledig willekeurig worden toegewezen aan een interventiegroep of controlegroep. Door deze zuivere randomisatie heeft elke deelnemer een gelijke kans om in elke conditie terecht te komen. Dit minimaliseert confounding en selectiebias, waardoor verschillen in uitkomsten met grotere zekerheid kunnen worden toegeschreven aan de interventie. Idealiter worden RCT's uitgevoerd in (dubbel)blinde omstandigheden, waarbij de deelnemers en/of onderzoekers niet weten wie in welke groep zit, maar dit is bij gezondheidsbevorderende interventies vaak moeilijk te realiseren.

Wanneer randomisatie op individueel niveau niet mogelijk is, bijvoorbeeld omdat een interventie wordt aangeboden op niveau van een school, organisatie of regio, kan cluster-randomisatie worden toegepast. Hierbij worden grotere eenheden zoals klassen, team of huisartsenpraktijken willekeurig toegewezen aan een interventie- of controlegroep. Het effect wordt vervolgens op individueel niveau gemeten, met correctie voor clustering omdat deelnemers binnen een cluster meer op elkaar lijken.

Cluster-RCT's zijn organisatorisch vaak makkelijker uitvoerbaar, maar vereisen doorgaans grotere steekproeven en zijn daarmee duurder dan een klassieke RCT. Daarnaast kunnen verschillen tussen clusters die niet met de interventie samenhangen de resultaten beïnvloeden.

Binnen klinisch onderzoek worden RCT's beschouwd als de gouden standaard voor het aantonen van causale effecten dankzij hun hoge interne validiteit. Voor onderzoek naar gezondheidsbevordering in real-life settings zijn traditionele RCT's echter niet altijd geschikt. Gedragsinterventies zijn vaak complex, contextafhankelijk en bestaan uit meerdere samenhangende componenten. Traditionele RCT's trachten deze context te neutraliseren om interne validiteit te maximaliseren, maar dit gaat ten koste van de generaliseerbaarheid naar de praktijk. Bovendien vergen RCT's veel tijd, middelen en expertise, wat de haalbaarheid kan beperken.

Om de kloof tussen interne validiteit en toepasbaarheid te overbruggen, zijn pragmatische RCT's ontwikkeld (104). Pragmatische RCT's evalueren interventies in realistische omstandigheden: met flexibelere uitvoering, minder strikte inclusiecriteria en meer ruimte voor contextuele variatie. Dankzij randomisatie bieden ze sterker bewijs voor causaliteit dan quasi-experimentele designs, terwijl ze beter aansluiten bij de complexiteit van preventieve interventies in de praktijk.

Box 7: "Test, learn, adapt": negen stappen voor het uitvoeren van een RCT voor beleidsinterventies

De publicatie "Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials" van het Britse Behavioural Insights Team²⁰ pleit voor het gebruik van RCT's in beleidsvorming (105). Het rapport beschrijft een negenstappenplan dat RCT's verankert in beleidsontwikkeling. De methode omvat drie fasen – Test, Learn en Adapt – en bestaat uit de volgende negen stappen:

²⁰ Het Behavioral Insights Team (BIT) is een organisatie die gedragsinterventies uitvoert en evalueert inzake overheidsbeleid en publieke dienstverlening. Het werd in 2010 opgericht binnen het Cabinet Office van de Britse overheid. Sinds 2014 is het een sociale onderneming, onafhankelijk van de overheid.

Ontwerpen en uitvoeren van de interventie (Test)**1) Identificeer twee of meer beleidsinterventies (bv. oude en nieuwe beleidsmaatregel)**

Bepaal welke beleidsopties of acties met elkaar vergeleken worden. Dit kan een vergelijking zijn tussen bestaande en nieuwe beleidsmaatregelen of verschillende varianten van een nieuwe aanpak. Test enkel opties die in de praktijk uitvoerbaar en schaalbaar zijn.

2) Bepaal de uitkomst die beïnvloed moet worden

Definieer vooraf duidelijk de doelstelling en selecteer een concrete, meetbare en relevante uitkomstmaat. Surrogaatuitkomsten (i.e., indirecte uitkomsten die makkelijker meetbaar of beschikbaar zijn) zijn zinvol wanneer ze betrouwbare voorspellers zijn van het uiteindelijke beleidsdoel.

3) Kies de randomisatie-eenheid

Dit kan een individu zijn, maar ook een organisatie of geografisch gebied, afhankelijk van de aard van de interventie.

4) Bepaal de benodigde steekproefgrootte

De steekproefgrootte moet voldoende groot zijn om een significant robuust effect te kunnen meten. Gebruik poweranalyse om de minimale steekproefgrootte te bepalen op basis van het verwachte effect, de variabiliteit in de populatie en de gewenste mate van significantie of statistische zekerheid. Wanneer men een klein verschil verwacht tussen de interventie- en controlegroep, is een grotere steekproef nodig om dat verschil aan te tonen.

5) Gebruik een robuuste randomisatiemethode

Toewijzing aan de interventie- en controlegroep moet volledig willekeurig en niet beïnvloedbaar zijn. 'Allocation concealment' voorkomt dat de onderzoekers verantwoordelijk voor de toewijzing weten in welke groep deelnemers zullen terechtkomen.

6) Implementeer de interventie zoals bedoeld

Voer de beleidsmaatregel of interventie uit volgens plan en representatief voor hoe deze in werkelijkheid zou worden uitgevoerd.

Analyse en interpretatie (Learn)**7) Meet en analyseer de resultaten**

Meet de geselecteerde uitkomst op verschillende tijdstippen bij de interventie- en controlegroep. De timing van de meting is afhankelijk van wanneer de interventie verwacht wordt effect te hebben. Vergelijk vervolgens de uitkomsten van de verschillende groepen met gepaste statistische methoden.

Verbeter het beleid (Adapt)**8) Pas het beleid aan op basis van de bevindingen**

Wanneer uit de trial blijkt dat een maatregel effectief is, kan deze worden opgeschaald. Ook niet-significante of negatieve resultaten leveren waardevolle inzichten op.

9) Keer terug naar stap 1

Beleidsvorming is geen eindpunt, maar een continu proces van leren en bijsturen. Door nieuwe varianten te testen, andere doelgroepen te onderzoeken of aanvullende inzichten in te bouwen, kan beleid systematisch worden verbeterd.

2.3.2 Cross-over design

De meest gebruikelijke vorm van een cross-overstudie is het zogeheten AB/BA-design. In deze opzet doorlopen alle deelnemers (individueel of op clusterniveau) zowel de interventieconditie als de controleconditie, maar in een verschillende volgorde. De ene groep ontvangt eerst de interventie en daarna de controleconditie, terwijl de andere groep start met de controleconditie en vervolgens de interventie krijgt. Bij een cross-over RCT worden deelnemers willekeurig toegewezen aan een van beide volgordecondities. Wanneer deze toewijzing niet willekeurig gebeurt, spreekt men van een quasi-experimenteel cross-over design.

Voor en na elke fase (interventie en controle) wordt dezelfde uitkomstmaat gemeten bij elke deelnemer of cluster. Omdat elke deelnemer beide condities doorloopt, kunnen de uitkomsten binnen dezelfde persoon of hetzelfde cluster worden vergeleken: hoe scoort iemand na de interventie vergeleken met hoe diezelfde persoon scoort na de controleconditie? Dit heeft als voordeel dat de statistische power vergroot en het benodigde aantal deelnemers vermindert.

Een essentieel aandachtspunt is het risico op een carry-overeffect, waarbij het effect van de interventie uit de eerste fase gedeeltelijk blijft doorwerken in de tweede fase. Om dit te voorkomen kan een uitwasperiode nodig zijn, zodat de effecten van de eerste fase voldoende zijn uitgewerkt voor de start van de volgende. Hierdoor is een cross-overdesign minder geschikt voor leefstijlinterventies die gericht zijn op langdurige gedragsverandering, omdat nieuw gedrag niet eenvoudig kan worden “uitgewassen” binnen de studieperiode. Het design kan wel nuttig zijn voor interventies waarbij effecten kortdurend en omkeerbaar zijn, zoals bij de evaluatie van omgevingsfactoren (bv. de beschikbaarheid van gezonde voeding in een bedrijfsrestaurant).

2.3.3 Stepped-wedge design

Een stepped-wedge design is een onderzoeksopzet waarbij een interventie sequentieel wordt uitgerold in verschillende clusters (bv. groepen, teams, organisaties) over meerdere tijdsperiodes. Aan het begin bevinden alle clusters zich in de controleconditie. Vervolgens schakelt telkens één cluster in een vooraf bepaald tijdsblok over naar de interventie, totdat uiteindelijk alle clusters de interventie hebben ontvangen. Bij elke meetronde worden alle deelnemers gemeten, zowel clusters die al aan de interventie zijn begonnen als clusters die nog in de controlefase zitten.

Zodra een cluster met de interventie start, blijft deze actief tot het einde van de studie. Hierdoor genereren vroeg overstappende clusters meer interventiedata, terwijl laat overstappende clusters meer controlegegevens opleveren. Dit gefaseerde ontwerp spreidt de implementatielast en biedt organisatorische voordelen, vooral wanneer een gelijktijdige uitrol praktisch niet haalbaar is.

Een bijkomend voordeel is dat een stepped-wedge design doorgaans meer statistische power oplevert dan een klassieke cluster-RCT. Omdat elke cluster zowel controle- als interventiewaarnemingen levert, zijn er minder deelnemers nodig om betrouwbare effecten aan te tonen. Dit design wordt echter alleen als een cluster-RCT beschouwd wanneer het tijdstip van overstap willekeurig wordt toegewezen; gebeurt dit niet, dan is het een quasi-experimenteel stepped-wedge design.

Box 8: Debat over RCT's voor leefstijlinterventies en gezondheidsbevordering

RCT's worden traditioneel gezien als de gouden standaard voor effectiviteitsonderzoek vanwege hun sterke interne validiteit. Voor leefstijlinterventies en gezondheidsbevordering is echter ook externe validiteit (i.e., toepasbaarheid in diverse contexten) cruciaal. Daarom zijn onderzoekopzetten nodig die een betere balans bieden tussen interne en externe validiteit, zoals quasi-experimentele designs en pragmatische RCT's (101).

Een WHO-werkgroep concludeerde al in 1998 dat RCT's voor veel vormen van gezondheidsbevordering **ongeschikt, onnodig duur en beperkt generaliseerbaar** zijn, omdat wat werkt in gecontroleerde omstandigheden vaak niet werkt in real-life situaties (13). Onderdeel van deze kritiek is de aanname van o.a. beleidsmakers dat “geen bewijs van effectiviteit” gelijkstaat aan “bewijs van geen effectiviteit”.

Als tegenreactie benadrukten verschillende onderzoekers dat gezondheidsbevordering **net wél** baat heeft bij methodologische sterkte. Zij stellen dat RCT's ook in realistische omstandigheden kunnen worden uitgevoerd, en dat veel praktische bezwaren oplosbaar zijn, bijvoorbeeld via **cluster-randomisatie** (106). Bovendien is randomisatie op zichzelf niet duur; de kosten liggen vooral in steekproefgrootte en dataverzameling, wat bij alle betrouwbare evaluatiemethoden meespeelt.

Leefstijlinterventies verschillen echter wezenlijk van medische behandelingen: ze vragen actieve betrokkenheid van deelnemers en beïnvloeden uiteenlopende processen zoals motivatie, attitude en kennis (98). Daardoor is het identificeren van geschikte uitkomstmaten complexer.

Door de grote variatie in complexiteit en context van leefstijlinterventies bestaat er **geen universeel beste onderzoeksdesign**. De klassieke hiërarchie van bewijsvoering, waarin RCT's bovenaan staan, is daarom niet altijd toepasbaar. Hoewel RCT's waardevol blijven om causaliteit aan te tonen, is het voor onderzoek naar gezondheidsbevordering essentieel dat ze worden uitgevoerd in **realistische settings**, zodat zowel interne als externe validiteit gewaarborgd blijft.

2.4 Samenvattend overzicht van de verschillende onderzoeksdesigns

Een overzicht van de verschillende besproken onderzoeksdesigns wordt weergegeven in Tabel 6.

Tabel 6 Overzicht van verschillende kwantitatieve onderzoeksdesigns.

Design	Controlegroep?	Randomisatie?	Beschrijving
RCT	Ja	Ja	Willekeurige toewijzing van deelnemers aan een interventie- of controlegroep.
Cluster-RCT	Ja	Ja	Willekeurige toewijzing van groepen (bv. scholen, regio's) aan interventie- of controlegroep.
Cross-over design	Ja	Van volgorde: ja (experimenteel), nee (quasi)	Alle deelnemers ontvangen zowel de interventie als de controleconditie, maar in een verschillende volgorde.
Stepped-wedge design	Ja	Van tijdstip: ja (experimenteel), nee (quasi)	Alle clusters starten in de controleconditie en ontvangen de interventie op verschillende tijdstippen.
Natural experiment	Ja	Nee	Interventie wordt niet willekeurig toegewezen; natuurlijke omstandigheden bepalen de groepen.
Synthetic control	Ja	Nee	De controlegroep wordt samengesteld uit een gewogen combinatie van meerdere groepen. Het gemiddelde van de controlegroep wordt vergeleken met de interventiegroep.
Regressie discontinuïteit	Ja	Nee	Groepen worden bepaald door een drempelwaarde op een continue variabele (bv. leeftijd). De uitkomst na de interventie wordt vergeleken bij deelnemers net onder/boven de drempel.
Interrupted time series	Nee	Nee	Meerdere metingen vóór en na de interventie binnen één groep om trends en interventie-effecten te analyseren.
Gecontroleerde pretest-posttest (difference-in-difference analyse)	Ja	Nee	Vergelijking van veranderingen tussen interventie- en controlegroep o.b.v. voor- en nametingen.
Gecontroleerde enkelvoudige posttest studie (static group comparison design)	Ja	Nee	Alleen vergelijking van nameting tussen interventie- en controlegroep
Pretest-posttest in één groep	Nee	Nee	Vergelijking van voor- en nameting binnen één interventiegroep
Enkelvoudige posttest in één groep	Nee	Nee	Alleen nameting binnen een interventiegroep.

2.5 Beslissingsmatrix WHO-Europe voor de selectie van een geschikt onderzoeksdesign

De WHO-Europe gids *“Evaluating the impact of interventions addressing health behaviour: Considerations and tools for policy-makers”* (2024) introduceert een beslissingsmatrix die beleidsmakers en onderzoekers ondersteunt bij het kiezen van het meeste geschikte onderzoeksdesign voor het evalueren van een gezondheidsinterventie op beleidsniveau (bv. vergelijken van regio’s waar een beleidsmaatregel al dan niet wordt ingevoerd). De matrix helpt om een weloverwogen keuze te maken, rekening houdend met:

- **Fase van implementatie:** wanneer de interventie al volledig is uitgerold, zijn experimentele designs niet langer mogelijk.
- **Mogelijkheid tot randomisatie:** als randomisatie niet haalbaar is of wanneer vooraf vaststaat in welke regio’s of groepen de interventie wordt uitgerold, zijn experimentele designs niet meer mogelijk.
- **Beschikbaarheid van een controlegroep** (en bijbehorende data): als er geen controlegroep is (bv. de maatregel wordt breed geïmplementeerd) of er zijn geen data beschikbaar van afzonderlijke groepen of regio’s, zijn enkel pre-poststudies in één groep en interrupted time series (ITS) mogelijk.
- **Beschikbaarheid van baselinemetingen:** klassieke RCT’s, cluster-RCT’s en natuurlijke experimenten vereisen geen baselinemeting.
- **Reversibiliteit van de interventie:** wanneer het niet mogelijk is om een interventie ongedaan te maken (bv. bij educatieve programma’s), zijn cross-over designs en ITS niet uitvoerbaar.

Het doel van de matrix is om binnen de gegeven omstandigheden het sterkst mogelijke onderzoeksdesign te selecteren, in termen van zowel haalbaarheid als wetenschappelijke bewijskracht (causaliteit).

De beslissingsmatrix bestaat uit een stappenplan met zes ja/nee-vragen. Bij elke stap worden bepaalde designs uitgesloten, zodat men geleidelijk toewerkt naar het meest geschikte design voor de specifieke situatie (Figuur 18).

Figuur 18 WHO Beslissingsmatrix voor beleidsmakers.

Decision Matrix for policy-makers Evaluating the impact of interventions addressing health behaviour Excluded ✖				Impact evaluation						
				Randomized controlled trial			Natural experiment			
				Crossover	Stepped-wedge	Randomized controlled trial	Natural experiment	Difference-in-differences	Pre-post	Interrupted time series
Step	Exclusion criteria	Answer: YES or NO								
1	The intervention has already been implemented	If NO, then go to the next step	If YES, then the following designs are excluded →	✖	✖	✖				
2	Group comparison is possible (you may answer NO if, for example, data is only available for entire population)	If YES, then go to the next step	If NO, then the following designs are excluded →	✖	✖	✖	✖	✖		
3	Randomization is possible (you may answer NO if, for example, a policy intervention is already planned in a fixed region)	If YES, then go to the next step	If NO, then the following designs are excluded →	✖	✖	✖				
4	Outcome can be measured before implementation (you may answer NO if, for example, no data is available from before implementation)	If YES, then go to the next step	If NO, then the following designs are excluded →	✖	✖			✖	✖	✖
5	Intervention can be undone (you may answer NO if, for example, the intervention will not or cannot be interrupted)	If YES, then go to the next step	If NO, then the following designs are excluded →	✖						✖
If, after the previous five steps, multiple designs are available, the decision can be based on the following step 6 on the strength of the design.				Crossover	Stepped-wedge	Randomized controlled trial	Natural experiment	Difference-in-differences	Pre-post	Interrupted time series
6	Weak	Medium	Strong							

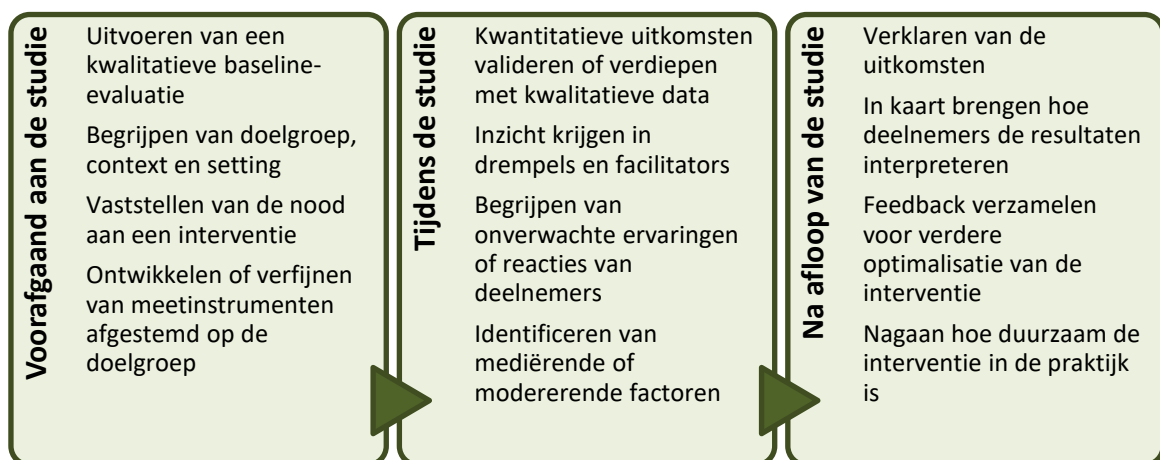
3 Kwalitatieve en gemengde methoden

Hoewel effectiviteitsonderzoek sterk leunt op kwantitatieve onderzoeksmethoden, kunnen kwalitatieve methoden een belangrijke aanvullende rol spelen om dieper inzicht te krijgen in waarom interventies al dan niet werken. Fetter en Molina-Azorin (107) betogen zelfs dat interventiestudies in de onderwijs-, gezondheids- en sociale wetenschappen idealiter standaard een **mixed-methodsbenadering** zouden moeten hanteren. Kwantitatieve data tonen aan of een interventie effect heeft, maar geven vaak weinig informatie over de onderliggende mechanismen, hoewel mediatie- en moderatieanalyses dit inzicht deels kunnen vergroten.

Het grote voordeel van een mixed-methodsbenadering is dat men niet alleen kan nagaan of een interventie werkt, maar ook hoe, voor wie en waarom. Kwalitatieve data maken het mogelijk om onverwachte barrières, contextuele factoren en ervaringen van deelnemers bloot te leggen die de uitkomsten beïnvloeden. Hierdoor kan men verklaren waarom een interventie in de ene situatie wel werkt en in de andere niet. Bovendien vergroot deze aanpak de aansluiting van de interventie bij de noden en verwachtingen van deelnemers, stakeholders en beleidsmakers, wat het draagvlak en de maatschappelijke meerwaarde van het onderzoek vergroot.

Fetter en Molina-Azorin (107) bespreken eveneens redenen om kwalitatieve methoden in verschillende fasen van een interventiestudie te integreren (Figuur 19):

Figuur 19 Redenen om kwalitatieve methoden in verschillende fasen van een interventiestudie te integreren.

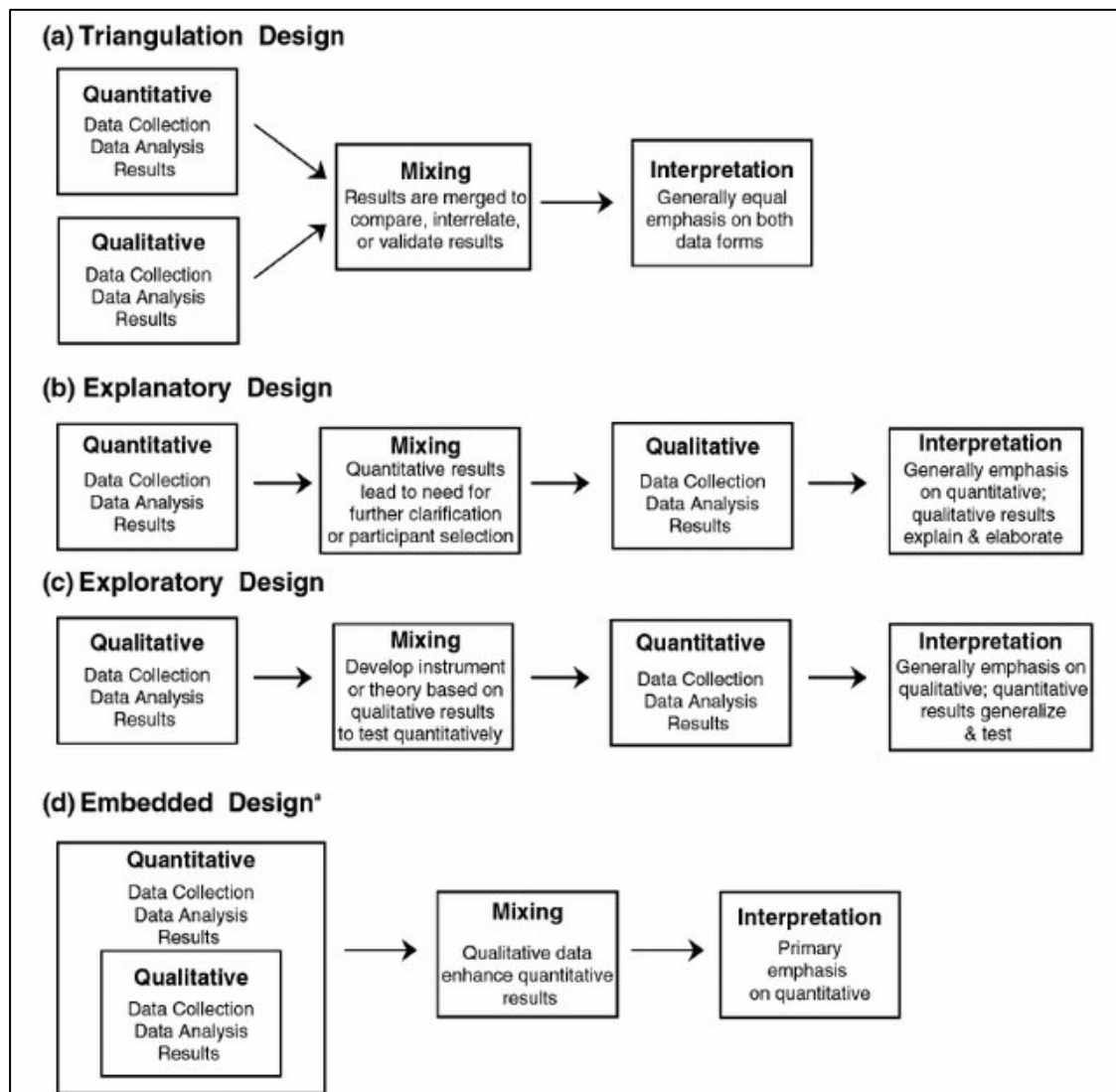


De manier waarop kwalitatieve en kwantitatieve methoden worden gecombineerd, kan sterk variëren. Creswell en Plano-Clark (108) onderscheiden vier hoofdtypen van mixed-methodsbenaderingen (zie Figuur 20):

- **Triangulatie design:** beide datatypen worden evenwaardig verzameld om een completer beeld te krijgen van hetzelfde fenomeen.
- **Verklarend design:** kwantitatieve data worden eerst verzameld, waarna kwalitatieve data helpen om deze resultaten te verduidelijken of te verklaren.
- **Exploratief design:** kwalitatieve data worden eerst verzameld om de ontwikkeling van kwantitatieve instrumenten of analyses te ondersteunen.

- **Ingebed design:** één type data (meestal kwalitatief) heeft een ondersteunende, secundaire rol binnen een studie die primair kwantitatief is of omgekeerd.

Figuur 20 Vier hoofdtypen van mixed methods designs volgens Creswell en Plano-Clark (2011).



4 Uitkomsten en indicatoren van interventies

Een goede effectevaluatie koppelt beoogde uitkomsten aan meetbare indicatoren, gebaseerd op zorgvuldig gekozen en onderbouwde meetinstrumenten (109).

Uitkomsten verwijzen naar de veranderingen die een interventie wil realiseren. Ze kunnen betrekking hebben op de korte, middellange en lange termijn en moeten duidelijk beschrijven wat verandert, bij wie, en in welke richting. De volgende stap is om voor elke uitkomst een passende indicator te kiezen.

Indicatoren zijn de concrete gegevens waarmee wordt gemeten of een uitkomst wordt bereikt. Ze bevatten geen veranderwoorden, maar verwijzen naar objectief meetbare grootheden zoals percentages, aantallen, scores of niveaus.

Voorbeeld: een daling van het aantal rokers is een uitkomst; het percentage deelnemers dat gestopt is met roken is een indicator.

4.1 Selectie van uitkomsten en uitkomstmaten

Interventie-uitkomsten kunnen zich situeren op verschillende niveaus: bij de einddoelgroep, bij de intermediairs (bv. professionals) of bij de organisatie. Complexe interventies kunnen uitkomsten op meerdere niveaus beogen, afhankelijk van de onderzoekseenheden.

Effectevaluaties maken doorgaans een onderscheid tussen primaire en secundaire uitkomsten. **Primaire uitkomsten** zijn de kernveranderingen waarop de interventie gericht is. **Secundaire uitkomsten** zijn gerelateerde of ondersteunende effecten. Afhankelijk van de interventie en programmatheorie (bv. Theory of Change of Logic Model) kunnen uiteenlopende domeinen van primaire en secundaire uitkomsten worden overwogen, zoals:

- **Gezondheidsuitkomsten:** verbetering van mentale of fysieke gezondheid
- **Gedragsverandering:** toename in frequentie, duur of intensiteit van gewenst gedrag
- **Intentie of motivatie:** bereidheid tot gedragsverandering
- **Attitudes en overtuigingen:** wijziging in houding tegenover het doelgedrag
- **Zelfeffectiviteit:** vertrouwen dat men het doelgedrag kan veranderen
- **Kennis of bewustzijn:** toename van kennis of bewustzijn over het gezondheidsthema
- **Zorggebruik:** veranderingen in het gebruik van (preventieve) zorg
- **Contextverandering:** bv. verbeterde werkcultuur of sociale cohesie

Gezondheidsbevorderende interventies kunnen inhoudelijk sterk verschillen, maar hebben vaak vergelijkbare doelstellingen rond gedragsverandering en beïnvloeden gelijkaardige determinanten zoals motivatie, kennis en zelfeffectiviteit.

Hoewel het nodig kan zijn om nieuwe gegevens te verzamelen om de evaluatievragen te beantwoorden, is het toch aangewezen om eerst na te gaan of bestaande registraties al relevante informatie bevatten die (gedeeltelijk) kunnen voorzien in de informatiebehoefte. Vooral bij beleidsinterventies met brede implementatie is dit nuttig.

Wanneer zelf data worden verzameld, is het cruciaal om goede uitkomstmaten te selecteren. Waar mogelijk en haalbaar is het aangewezen om bestaande gevalideerde vragenlijsten te gebruiken. Deze verhogen de betrouwbaarheid, maken vergelijkingen tussen doelgroepen en interventies mogelijk en bieden zo een onderbouwde basis voor beleidskeuzes.

Toch is het gebruik van gevalideerde instrumenten niet altijd haalbaar in de praktijk. Niet elke uitkomst heeft gevalideerde instrumenten beschikbaar in het Nederlands of in begrijpelijke taal. Sommige vragenlijsten zijn ook lang en complex. In dat geval is het een optie om te kiezen voor verkorte of aangepaste versies van bestaande instrumenten, of voor zorgvuldig opgestelde eigen vragen. Het is belangrijk om deze keuzes goed te motiveren. Voor het opstellen van vragenlijsten verwijzen we naar “Eerste hulp bij vragenlijsten – Vlaams Instituut Gezond Leven (2018)”²¹.

21 https://www.gezondleven.be/files/gezondheidsbevordering/Gezond-Leven-2018_-Eerste-hulp-bij-...-vragenlijsten.pdf

4.2 Types indicatoren

In de gezondheidsdoelstelling De Vlaming Leeft Gezonder wordt een onderscheid gemaakt tussen gezondheidsindicatoren en procesindicatoren. Voor effectevaluaties is een driedeling nuttig, zoals beschreven door de WHO (110): procesindicatoren, uitkomstindicatoren en impactindicatoren. Deze worden toegelicht in Tabel 7.

Het is belangrijk om het aantal indicatoren beperkt en relevant te houden, zodat dataverzameling haalbaar blijft en deelnemers niet onnodig belast worden. De programmatheorie en de evaluatievragen bieden richting bij de selectie van geschikte indicatoren.

Tabel 7 Types indicatoren

Type indicator	Wat wordt gemeten	Doel	Voorbeelden
Procesindicatoren	Uitvoering en implementatie van de interventie, inclusief factoren die deze beïnvloeden	Nagaan of de interventie uitvoerbaar is en volgens plan wordt uitgevoerd	- Aantal bereikte deelnemers - Mate van uitvoering volgens protocol - Tevredenheid, acceptatie, haalbaarheid
Uitkomstindicatoren	Directe effecten op uitkomsten op korte tot middellange termijn bij deelnemers	Vaststellen of de interventie het gewenste effect oplevert	- Toename in kennis over gezonde voeding - Afname van alcoholgebruik - Verbetering van welzijn
Impactindicatoren	Langetermijneffecten op gezondheid en welzijn van de doelgroep	Nagaan of het beleid bijdraagt aan duurzame gezondheidsverbetering	- Aantal rokers - Kwaliteit van leven

Hoofdstuk 7

Passend bewijs voor preventie

Preventieve interventies worden in uiteenlopende contexten ontwikkeld en toegepast, en vereisen daarom een doordachte benadering van welke vormen van bewijs geschikt zijn om hun effectiviteit en toepasbaarheid te beoordelen. Passend bewijs verwijst niet alleen naar hoogwaardig onderzoek uit de literatuur, maar ook naar een realistische en proportionele aanpak voor het genereren van nieuw bewijs via evaluaties. In dit hoofdstuk wordt toegelicht hoe binnen de preventiewetenschap verschillende niveaus van evidentie worden onderscheiden, welke standaarden richting geven aan kwaliteitsvol onderzoek, en hoe beleidsmakers en onderzoekers kunnen bepalen welk evaluatiedesign het meest geschikt is in een specifieke situatie. Daarnaast wordt ingegaan op hoe bestaande interventies verantwoord kunnen worden aangepast aan nieuwe contexten. Samen bieden deze inzichten een fundament om preventieve maatregelen in Vlaanderen op een onderbouwde, haalbare en contextgevoelige manier te beoordelen en te versterken.

1 Evidentiestandaarden in de preventiewetenschap

In 2005 publiceerde de Society for Prevention Research een eerste set standaarden voor het beoordelen van de effectiviteit van preventieve interventies. Tien jaar later werden de standaarden herzien en werd een onderscheid gemaakt tussen **efficacy (werkzaamheid)**, **effectiveness (effectiviteit in de praktijk)** en **scale-up (opschaling)** (111).

Deze standaarden zijn bedoeld als **groeikader**: interventies hoeven niet onmiddellijk aan alle criteria te voldoen, maar de standaarden zijn richtinggevend voor beleidsmakers en professionals bij het bepalen welke interventies voldoende onderbouwd zijn om grootschalig toe te passen.

Preventieonderzoek werd aanvankelijk opgevat als een lineair proces startend met fundamenteel onderzoek, gevolgd interventieontwikkeling, pilootstudies, efficacy-onderzoek in gecontroleerde omstandigheden, effectiveness-onderzoek in realistische settings en uiteindelijk grootschalige implementatie. Dit lineaire model bleek echter te beperkt. De herziening van de standaarden hanteert daarom een flexibeler en meer dynamisch model, waarbij onderzoek naar scale-up en contextuele factoren al in de vroege ontwikkelfasen wordt geïntegreerd.

Dit vernieuwde kader biedt ook ruimte om interventies te onderzoeken die reeds in de praktijk worden toegepast, en stimuleert feedbackloops tussen de verschillende onderzoeksfasen. De volgorde van onderzoek wordt daarbij minder strikt gehanteerd: de focus ligt op het beantwoorden van de meest relevante onderzoeksvragen op het juiste moment.

In Figuur 21 worden belangrijke kenmerken van efficacy, effectiveness en scale-up samengevat. De criteria zijn cumulatief: een interventie die effectief is, voldoet ook aan de efficacy-standaarden; interventies klaar voor scale-up voldoen aan alle voorgaande criteria. Een volledige en gedetailleerde beschrijving van de standaarden is te vinden in *Standards of Evidence for Efficacy, Effectiveness, and Scale-up Research in Prevention Science: Next Generation* (111).

Figuur 21 Het onderscheid tussen efficacy, effectiveness en scale-up in preventieonderzoek volgens Gottfredson et al. (2015)

Efficacy: werkzaamheid onder gecontroleerde omstandigheden	Effectiveness: werkzaamheid in de praktijk	Scale-up: opschaling en brede uitrol
- Interventies worden volledig en nauwkeurig beschreven zodat replicatie mogelijk is (inhoud, doelgroep, setting, verwachte effecten). - Theoretische modellen van causale mechanismen worden geformuleerd en getest. - Implementatie (kwaliteit, hoeveelheid, acceptatie door doelgroep) wordt gemeten en gerapporteerd. - Betrouwbare en valide meetinstrumenten worden gebruikt. - Langetermijneffecten worden onderzocht (minimaal 6 maanden follow-up). - Sterke onderzoeksdesigns zijn vereist voor causale claims: altijd een controlegroep zonder interventie; randomisatie heeft de voorkeur. - Alle uitkomsten worden gerapporteerd. - Er zijn geen ernstige negatieve of onbedoelde gevolgen.	- Er is voldaan aan alle noodzakelijke criteria van efficacy - De interventie wordt uitgevoerd en getest in realistische omstandigheden (de echte wereld; zoals deze in de praktijk zou worden uitgerold). - Er is een handleiding en/of training beschikbaar voor implementatie. - De interventie wordt getest in diverse populaties en settings; generaliseerbaarheid wordt gerapporteerd. - Er worden subgroepanalyses uitgevoerd (bv. effectiviteit per leeftijdsgroep of SES). - De implementatiekwaliteit wordt onderzocht en vergeleken met die in de efficacy-studies. - De kosteneffectiviteit wordt gerapporteerd. - Uitvoering door onafhankelijke onderzoekers wordt aanbevolen.	- Er is voldaan aan alle noodzakelijke criteria van efficacy en effectiveness. - Er is duidelijke informatie over de kosten en er zijn tools die kostenramingen mogelijk maken voor implementatoren. - Er zijn materialen beschikbaar die optimale uitvoering specificeren. - Er is training en technische ondersteuning beschikbaar voor de implementatie van de kerncomponenten van de interventie. - Er is een systeem voor monitoring, feedback en documentatie van aanpassingen aan de interventie, met aandacht voor implementietrouwheid. - Er is een systeem voor planning en monitoring van de rekrutering van deelnemers. - Scale-up wordt geëvalueerd om te verzekeren dat de beoogde directe effecten ook op populatieniveau worden bereikt.

2 Richtlijn passend bewijs door Nederlandse Adviescommissie Richtlijn Passend Bewijs Preventie

In Nederland werd de Adviescommissie Richtlijn Passend Bewijs Preventie ingesteld door de Staatssecretaris van Volksgezondheid, Welzijn en Sport. In 2025 publiceerde de Adviescommissie een vernieuwende richtlijn waarin wordt omschreven wat geldt als passend bewijs bij de beoordeling van preventieve maatregelen. Deze richtlijn biedt een helder, samenhangend en praktisch toepasbaar kader voor onderzoekers en beleidsmakers die preventieve interventies of beleidsmaatregelen willen beoordelen op hun effectiviteit en toepasbaarheid.

De Nederlandse richtlijn vormt een sterk en actueel referentiekader voor de onderbouwing van het preventieve beleid. Voor Vlaanderen biedt deze richtlijn een waardevolle inspiratiebron voor het ontwikkelen van een consistent en onderbouwd beoordelingskader voor preventieve maatregelen.

2.1 Overtuigend bewijs van effectiviteit

De richtlijn vertrekt vanuit de eis dat **passend bewijs steeds inzicht moet bieden in de ‘tegenfeitelijke situatie’**: wat zou er gebeurd zijn indien de maatregel niet was ingevoerd? Deze vereiste geldt voor alle vormen van preventie, ook voor complexe of meervoudige maatregelen. De commissie verwijst hierbij naar de *Maryland Scientific Methods Scale* (zie Box 9) en definieert dat passend bewijs overeenstemt met minimaal niveau 3 op deze schaal (112, 113). Dit betekent dat veranderingen in uitkomsten voor en na een interventie moeten worden vergeleken tussen groepen die de maatregel wel of niet hebben ondergaan, zonder dat er noodzakelijk sprake is van een gerandomiseerd experiment.

De richtlijn benadrukt daarmee dat passend bewijs niet beperkt is tot RCT's. Hoewel RCT's de hoogste bewijskracht voor causaliteit bieden, is het uitvoeren ervan voor veel preventiemaatregelen niet haalbaar en niet nodig. Daarom geeft de richtlijn richting aan het gebruik van verschillende vormen van quasi-experimentele designs en natuurlijke experimenten. Deze kunnen eveneens overtuigend bewijs bieden mits ze voldoende nauwkeurig de tegenfeitelijke situatie benaderen.

Box 9: De Maryland Scientific Methods Scale

De Maryland Scientific Methods Scale is een internationale beoordelingsschaal om de sterkte van bewijs over de effectiviteit van beleid of interventies in te schatten (112, 113). De schaal bestaat uit **vijf niveaus**, waarbij elk hoger niveau een sterkere methodologische basis biedt om causaliteit te onderbouwen.

Niveau	Kenmerken van het onderzoeksdesign
1	Cross-sectionele analyse (vergelijkt de interventiegroep met zichzelf vóór en na de interventie) of voor- en na vergelijking tussen groepen (met één meting per moment) zonder controlevariabelen of matchingtechnieken
2	Cross-sectionele analyse (vergelijkt de interventiegroep met zichzelf vóór en na de interventie) of voor-en na vergelijking (met één meting per moment) met passende controlevariabelen of matchingtechnieken . Bv. interrupted time series, pre-posttest studie.
3	Analyse maakt gebruik van een geloofwaardige controlegroep en meerdere metingen vóór en na de interventie (met trendanalyse). Bv. difference in differences.
4	Analyse maakt gebruik van (bijna) willekeurige toewijzing ontstaan door bestaande gebeurtenissen of regelgeving . Bv. regressie discontinuïteit, natuurlijke experimenten, indeling op basis van instrumentele variabelen (die wel van invloed zijn op de toewijzing aan de interventie- of controlegroep, maar die geen direct effect hebben op gezondheidsuitkomsten).
5	Gerandomiseerde gecontroleerde experimenten (RCT).

2.2 Toepasbaarheid en contextgevoeligheid

Preventieve maatregelen worden vaak onderzocht in andere landen, settings of bevolkingsgroepen. Deze resultaten kunnen niet zonder meer worden overgenomen. De richtlijn benadrukt dat evaluatoren altijd moeten nagaan **in hoeverre de context van de oorspronkelijke studie vergelijkbaar is met de context waarin de interventie geïmplementeerd wordt of zou worden**.

De **context** van een interventie omvat factoren die invloed hebben op de uitvoering en de effectiviteit ervan, zoals populatiekenmerken, sociaaleconomische en geografische omstandigheden, epidemiologische en culturele factoren, ethische, juridische en politieke omstandigheden.

Bijvoorbeeld bij een interventie rond overgewicht kunnen contextfactoren zijn: beschikbaarheid van sportfaciliteiten, culturele eetgewoonten en inkomens- en opleidingsniveau van de doelgroep.

Beoordelen of de onderzoeksresultaten toepasbaarheid zijn in een andere context, vereist inzicht in de werkzame elementen van een interventie: de mechanismen of componenten die noodzakelijk zijn voor effectiviteit (bv. minimale intensiteit, uitvoerder, gedragsveranderingstechnieken). Hiervoor is een sterke theoretische onderbouwing cruciaal, zoals een interventielogica of Theory of Change. Methoden zoals Realist Evaluation en procesevaluaties (met kwalitatieve én kwantitatieve data) helpen deze elementen te identificeren.

2.3 Langetermijneffecten en intermediaire uitkomsten

Veel preventieve maatregelen hebben effecten die zich pas op langere termijn manifesteren. De richtlijn stelt dat het ontbreken van langetermijngegevens geen reden mag zijn om preventieve interventies af te wijzen. In die gevallen kunnen intermediaire uitkomstmaten of modelanalyses worden gebruikt als alternatief om aannemelijke langetermijneffecten te schatten.

2.4 Andere aanbevelingen

Daarnaast bevat de richtlijn beleidsaanbevelingen rond onderzoeksfinanciering en data-infrastructuur. Omdat robuust effectonderzoek vaak hogere kosten met zich meebrengt, raadt de commissie aan om onderzoeksbudgetten af te stemmen op designs met controlegroepen of vergelijkende opzetten. Een goed uitgebouwde data-infrastructuur en de toegang tot longitudinale en populatiegegevens, vergroot de mogelijkheden voor quasi-experimenteel onderzoek en voor het schatten van langetermijneffecten.

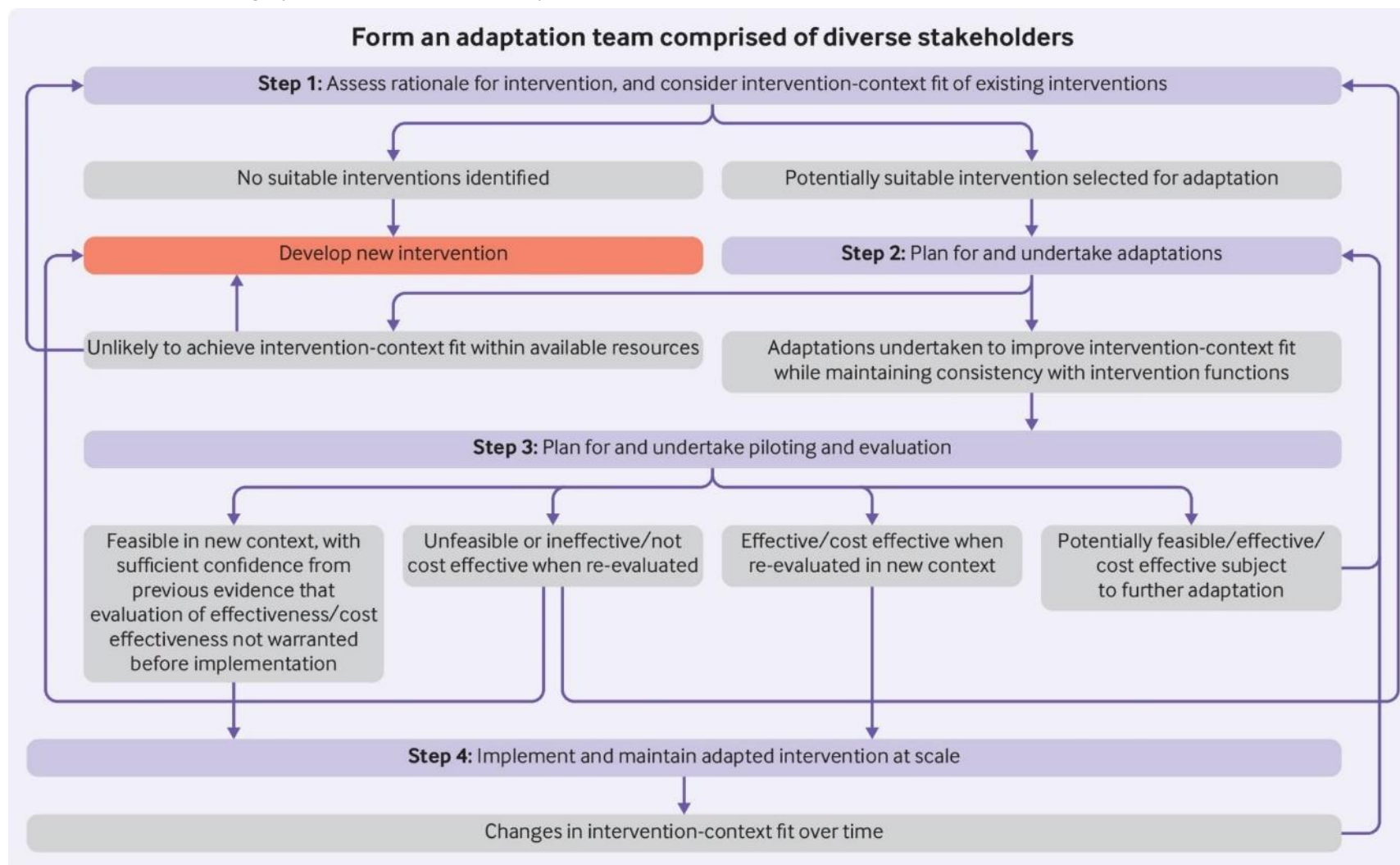
Voor de toepassing van de richtlijn geldt het “pas toe of leg uit” principe: de richtlijn is het uitgangspunt, maar mag worden verlaten wanneer dit goed gemotiveerd is. Tegelijk benadrukt de richtlijn dat innovatie niet belemmerd mag worden: onderzoeken die nog niet volledig aan alle criteria voldoen, maar wel richting passend bewijs evolueren, moeten mogelijk blijven. Zo blijft er ruimte bestaan om nieuwe preventieve aanpakken te verkennen en onderbouwing gradueel te versterken.

3 **ADAPT-richtlijn: aanpassen en evalueren van evidence-based interventies in nieuwe contexten**

Het implementeren van interventies waarvoor reeds sterke evidentie bestaat, is vaak efficiënter dan voor elke context een volledig nieuwe interventie te ontwikkelen. Toch is effectiviteit niet altijd zomaar overdraagbaar: de effectiviteit en uitvoerbaarheid van interventies hangen sterk af van contextuele factoren. De **ADAPT-richtlijn** biedt daarom een gestructureerd en verantwoord kader om evidence-informed interventies systematisch en iteratief aan te passen aan nieuwe contexten (114). De richtlijn is ontwikkeld op basis van literatuurreviews, interviews met onderzoekers en beleidsmakers, en een Delphi-consensusstudie.

Het procesmodel van ADAPT bestaat uit vier stappen (Figuur 22), waarbij **stakeholderbetrokkenheid** een centraal principe is: doelgroep, uitvoerders en beleidsmakers worden van bij de start en gedurende het hele proces betrokken om een goede aansluiting tussen interventie en context te waarborgen.

Figuur 22 ADAPT-procesmodel voor het aanpassen van interventies aan nieuwe contexten. Paarse vakken = stappen van de stapsgewijze ADAPT-richtlijn; Grijsze vakken = mogelijke uitkomsten van elke stap (114).



- **Probleemanalyse en context-fit**

In de eerste stap wordt het probleem in de nieuwe populatie helder gedefinieerd. Er worden geschikte interventies geïdentificeerd op basis van de sterkte van het bewijs, theoretische onderbouwing en overeenkomsten tussen de oorspronkelijke en nieuwe context. Hierbij wordt ook rekening gehouden met praktische haalbaarheid én met intellectuele eigendomsrechten.

De mate waarin een interventie moet worden aangepast, hangt af van de verschillen tussen contexten. Belangrijke aandachtspunten zijn onder meer beschikbare middelen, uitvoerbaarheid binnen bestaande rollen, acceptatie door doelgroep en uitvoerders, aansluiting bij bestaande systemen en het onderscheid met de huidige praktijk. Kwalitatief onderzoek kan helpen deze context- en implementatiefactoren beter te begrijpen.

- **Plannen en uitvoeren van adaptaties**

Het onderzoeksteam identificeert drempels en hefboomen voor implementatie en stelt een lijst op met mogelijke aanpassingen. Belangrijk is dat de kernmechanismen van de interventie behouden blijven. Aanpassingen kunnen variëren van kleine wijzigingen in materialen tot substantiële herzieningen van onderdelen van het programma. Soms vraagt dit ook een wijziging in wie de interventie uitvoert. Na het selecteren van de nodige aanpassingen worden de programmatheorie, handleidingen en protocollen geactualiseerd.

- **Pilotering en evaluatie**

In deze fase wordt bepaald in **welke mate her-evaluatie nodig is**. Dit hangt af van drie overwegingen:

- **Kwaliteit en overdraagbaarheid van bestaand bewijs:** bij grote onzekerheid is een robuuste effectiviteitsstudie nodig; bij beperkte onzekerheid is evaluatie niet altijd nodig of kan evaluatie geïntegreerd worden in de implementatie.
- **Waarde van nieuwe informatie:** het team moet afwegen of de kosten en vertraging van een uitgebreide evaluatie opwegen tegen de voordelen voor beleid en praktijk.
- **Beschikbare middelen:** bij beperkte budgetten is het een optie om te kiezen voor pragmatische evaluaties, zoals ongecontroleerde effectmetingen of kwalitatieve studies naar haalbaarheid en onbedoelde effecten.

Feedback uit deze fase kan aanleiding geven tot verdere adaptaties of zelfs tot stopzetting van implementatie. Tijdens de uitvoering moeten zowel geplande als responsieve aanpassingen systematisch worden gedocumenteerd om te garanderen dat de kernmechanismen behouden blijven.

- **Implementatie en opschaling**

Voor duurzame implementatie zijn sterke partnerschappen, capaciteitsopbouw en monitoring essentieel. De ADAPT-richtlijn biedt hiervoor een set van **checklist-** en **rapportage-items** die transparantie garanderen, waaronder: een beschrijving van de oorspronkelijke en nieuwe context, details van het adaptatieproces, betrokken stakeholders, en de onderbouwing van evaluatiekeuzes.

4 Aanbevelingen van de onderzoekers inzake passend bewijs voor de Vlaamse preventiesector

Binnen de Vlaamse preventiesector zijn realistische en toekomstgerichte verwachtingen nodig over welk bewijs haalbaar en passend is. Interventies kunnen niet in één stap als effectief worden beschouwd. Eerst moet de evidentie voor interventies geleidelijk worden opgebouwd, waarna in de toekomst een formeel beoordelingskader kan worden uitgewerkt.

- **In de vroege fasen staat haalbaarheid centraal: implementatie-inzichten primeren**

Evaluaties dienen methodologisch degelijk te zijn, maar tegelijk uitvoerbaar binnen de beschikbare tijd en middelen. In vroege ontwikkelings- en implementatiefasen ligt de nadruk op uitvoerbaarheid, implementatiekwaliteit en succes- en faalfactoren.

In deze fase bieden pre-experimentele en pragmatische designs waardevolle eerste inzichten in effectiviteit, helpen ze aannames uit de Theory of Change te toetsen en ondersteunen ze bij het identificeren van noodzakelijke bijsturingen. Deze bevindingen vormen een belangrijke basis voor latere, robuustere effectstudies.

- **Op langere termijn is het belangrijk om effectevaluatie niet uit het oog te verliezen: causale evidentie vraagt immers een doordachte aanpak**

Zodra dit haalbaar is, verdient het aanbeveling om over te stappen op onderzoekopzetten die sterkere causale uitspraken toelaten. Een interventie moet wel voldoende rijp zijn voor een robuuste effectevaluatie. Randomisatie is daarbij geen verplichting, maar een (quasi-)experimenteel design dat de tegenfeitelijke situatie benadert, wordt sterk aanbevolen wanneer de omstandigheden dit toelaten.

- **Effectiviteit bouwt op volgens een groeipad**

Een interventie is niet van de ene dag op de andere bewezen effectief. Bewijs voor effectiviteit wordt geleidelijk opgebouwd. Tijdens dit groeipad is ruimte nodig om interventies bij te sturen en verder te ontwikkelen. Daarbij blijft aandacht nodig voor de vragen *wat werkt, voor wie en onder welke omstandigheden*.

- **Internationale evidentie als vertrekpunt**

Omdat Vlaanderen een relatief kleine preventiecontext vormt, is het aangewezen om bestaande internationale evidentie te benutten. Interventies die elders hun effectiviteit hebben aangetoond, kunnen worden aangepast aan de Vlaamse context met behulp van kaders zoals ADAPT.

Na contextuele aanpassing is herevaluatie in Vlaanderen noodzakelijk om haalbaarheid, acceptatie en effectiviteit te bevestigen. De mate van herevaluatie hangt af van de omvang van de aanpassingen en de vergelijkbaarheid tussen contexten.

- **Naar een Vlaams beoordelingskader voor onderbouwde interventies**

Een volledig beoordelingskader, zoals dat van Nederland, is op dit moment nog niet haalbaar. Eerst zijn stappen nodig om de algemene evaluatiepraktijk te versterken en meer ervaring op te bouwen met effectonderzoek. Op langere termijn kan worden gewerkt aan een breed gedragen, realistisch Vlaams kader voor de onderbouwing en beoordeling van preventieve interventies.

Hoofdstuk 8

Impactevaluatie van het preventieve gezondheidsbeleid

Waar de eerdere hoofdstukken focusten op de evaluatie van afzonderlijke gezondheidsbevorderende interventies, zoomt dit hoofdstuk uit naar een breder niveau: de impactevaluatie van het preventieve gezondheidsbeleid als geheel. Daarbij verschuift de aandacht van effecten die aan één specifieke interventie kunnen worden toegeschreven naar de vraag hoe een geheel van maatregelen, programma's en beleidsinstrumenten gezamenlijk bijdraagt aan betere gezondheidsuitkomsten op populatie-, sector- of systeemniveau.

Impactevaluatie op beleidsniveau vereist een andere benadering dan interventie-evaluatie. Het gaat niet alleen om het meten van veranderingen, maar ook om het begrijpen van de complexe dynamiek tussen verschillende beleidsacties, de context waarin ze plaatsvinden en de manieren waarop ze elkaar versterken of juist verzwakken. Dit hoofdstuk verkent die bredere blik en bespreekt benaderingen, hulpmiddelen en randvoorwaarden om de impact van preventief gezondheidsbeleid op een onderbouwde manier in kaart te brengen.

1 Beleidsimpactevaluatie vraagt om een diepgaande analyse

Preventief gezondheidsbeleid bestaat doorgaans uit een geheel van acties en maatregelen die samen een **complex systeem** vormen. Beleidsimpactevaluatie richt zich op de **structurele en duurzame effecten** van dit beleid op een hoger niveau, zoals de gezondheid van de bevolking of de organisatie en werking van preventieve diensten binnen verschillende settings.

Impactevaluatie verschilt van effectevaluatie doordat ze verder kijkt dan de impact van één afzonderlijke interventie. Waar causale effecten van interventies vaak onderzocht worden met (quasi-)experimentele designs en kwantitatieve effectgroottes, is **beleidsimpact moeilijker te isoleren**. De effecten van beleid ontstaan niet alleen uit individuele acties, maar vooral uit de **interacties tussen beleidsinstrumenten** (zoals regelgeving, opdrachten, financiering en partnerschappen), **contextuele factoren** en **de gezamenlijke inspanningen van actoren** in het veld.

Belangrijke kernvragen in beleidsimpactevaluatie zijn onder andere:

- Hoe dragen verschillende interventies **samen** bij aan het realiseren van beleidsdoelstellingen?
- Welke **structurele veranderingen** (bv. in gedrag, organisatie, samenwerking of gezondheidstoestand) zijn zichtbaar op populatie- of sectorniveau?
- Welke **contextuele factoren** (bv. politieke, sociale of economische omstandigheden) beïnvloeden de mate waarin beleid zijn doelen bereikt?

Box 10: Whole systems approaches (WSA) voor preventie

De complexiteit van (preventieve) volksgezondheidsproblemen heeft geleid tot een groeiende belangstelling voor *whole systems approaches* (WSA). Deze benaderingen erkennen dat gezondheidsproblemen niet voortkomen uit één oorzaak, maar uit onderling verbonden factoren binnen een breder systeem van beleid, context en actoren.

In plaats van te focussen op één interventie, onderzoekt een WSA hoe meerdere maatregelen en factoren gezamenlijk veranderingen tot stand brengen. De aanpak stimuleert gecoördineerde en intersectorale samenwerking tussen o.a. gezondheidszorg, onderwijs, welzijn en lokale besturen.

Een systematische review toont dat WSA's **vaak positieve effecten** op gezondheidsuitkomsten opleveren (115). Tegelijk signaleert de review dat niet alle interventies volledig systemisch zijn (sommige zijn eerder multicomponent interventies), en dat er nood is aan heldere definities van WSA's. Uit de procesevaluaties kwamen negen sleutelprincipes naar voren die succesvolle implementatie ondersteunen:

- Voldoende middelen (financieel en personeel)
- Sterk leiderschap en betrokkenheid van partners
- Betrokkenheid van de gemeenschap (lokale burgers) om behoeften te identificeren en oplossingen te ontwikkelen
- Tijd en ruimte voor het opbouwen van relaties en vertrouwen
- Goede governance en gedeelde waarden (gezamenlijke visie, transparante besluitvorming)
- Duurzame samenwerking via stakeholdersnetwerken
- Consistente communicatie en gedeelde taal
- Inbedding in een breder beleid
- Lokale evaluatie en monitoring

2 Randvoorwaarden voor beleidsimpactevaluatie

Beleidsimpactevaluatie is complex, kostbaar en langdurig, en vereist daarom specifieke randvoorwaarden om betrouwbaar uitgevoerd te kunnen worden:

- **Heldere beleidsdoelstellingen:** duidelijke prioriteiten en passende indicatoren zijn essentieel.
- **Continuïteit van data:** populatie- en systeemindicatoren moeten langdurig en consistent beschikbaar zijn.
- **Intersectorale samenwerking:** gezondheidsimpact wordt vaak mede bepaald door beleid in domeinen zoals onderwijs, werk en milieu.
- **Gemengde methoden en participatieve benaderingen:** kwalitatieve en kwantitatieve gegevens moeten worden gecombineerd; triangulatie uit diverse bronnen is cruciaal voor een rijk en betrouwbaar inzicht (116).

3 Theory of change op sectorniveau

Hoewel een Theory of Change traditioneel wordt toegepast op het niveau van afzonderlijke interventies of programma's, kan dit instrument ook worden ingezet op hoger systemisch niveau, zoals voor sectorbrede veranderingen of impactmeting op het niveau van organisaties, netwerken of preventiesectoren. Een sectorale Theory of Change brengt meerdere interventies, programma's en beleidsmaatregelen samen in één logisch en samenhangend model, met tussentijdse en einddoelen op systeemniveau. Hierdoor wordt zichtbaar hoe verschillende initiatieven en beleidsacties elkaar beïnvloeden en gezamenlijk bijdragen aan gewenste veranderingen. Dit ondersteunt een meer geïntegreerde aanpak en maakt systematische evaluatie van sectorbrede impact mogelijk.

Het ontwikkelen van een Theory of Change op sectorniveau is echter **complex** dan op interventieniveau. Een sectorale Theory of Change is ruimer, omvat meer actoren en vraagt om een grondige analyse van onderlinge relaties. Het vereist bovendien intensieve stakeholderbetrokkenheid, zodat de diversiteit aan perspectieven, doelen en rollen binnen de sector wordt meegenomen. Volgens New Philanthropy Capital (NPC) is het relatief eenvoudig om de uiteindelijke doelen van een sector te identificeren, maar ligt de werkelijke uitdaging in het ontrafelen van de verschillende paden waarlangs uiteenlopende interventies aan die doelen bijdragen (117).

Een Theory of Change op sectorniveau kan verschillende belangrijke voordelen bieden:

- **Stimuleren van samenwerking:** organisaties stemmen hun aannames, inspanningen en evaluaties beter op elkaar af.
- **Identificeren van hiaten:** het model helpt tekorten in beleid, praktijk of evidentie bloot te leggen.
- **Versterken van gedeelde meetsystemen:** een sectorale Theory of Change faciliteert de ontwikkeling van gemeenschappelijke indicatoren, wat vergelijkbaarheid bevordert en gezamenlijke leerprocessen stimuleert.

Omdat het toepassen van een Theory of Change op sectorniveau nog relatief nieuw is, is de literatuur hierover voorlopig beperkt. Dit illustreert dat sectorale Theory-of-Change-ontwikkeling een opkomende en evoluerende praktijk is, maar wel een met aanzienlijk potentieel voor geïntegreerd beleid en impactevaluatie.

4 Indicatoren voor beleidsimpactevaluatie

Bij beleidsimpactevaluatie worden indicatoren doorgaans geformuleerd op **populatie- of sectorniveau**. Ze geven inzicht in hoe beleid doorwerkt in gezondheid, gedrag en de organisatie van preventie. Mogelijke types van indicatoren zijn:

- **Gezondheidsindicatoren:** bv. prevalentie van depressie, suïcidecijfers, welbevinden
- **Determinanten op populatieniveau:** bv. attitudes, kennis en motivatie rond gezondheidsbevorderend gedrag
- **Systeemindicatoren:** bv. samenwerking tussen actoren, integratie van preventie in verschillende settings
- **Proces- en governance-indicatoren:** bv. budgetverdeling, beleidscontinuïteit, mate van implementatie

Een robuuste impactevaluatie combineert kwantitatieve trendanalyse (zoals gezondheidsmonitoring) met kwalitatieve analyse (bv. interviews, netwerk- of contextanalyses) om zowel patronen als onderliggende mechanismen in kaart te brengen.

Belangrijke databronnen kunnen zijn:

- Representatieve bevragingen en gezondheidsonderzoeken
- Registratiedata (bv. uit settings)
- Beleidsdocumenten en begrotingen
- Interviews of focusgroepen met beleidsmakers en actoren uit de preventiesector
- Evaluatieverslagen van afzonderlijke interventies

4.1 Indicatoren voor beleidsimpactevaluatie in Vlaanderen

In Vlaanderen vormt de opvolging van **populatiebrede gezondheidsdoelstellingen** een belangrijke vorm van impactevaluatie van het preventieve gezondheidsbeleid. Hierbij wordt nagegaan welke effecten beleid heeft op de gezondheidsuitkomsten, gedragsverandering en het bereiken van doelgroepen.

Betrouwbare en representatieve databronnen zijn hiervoor cruciaal. Relevante Vlaamse databronnen zijn onder meer:

- Gezondheidsenquête (Sciensano)
- Preventiebarometer (Sciensano)
- HBSC- studie (Health Behaviour in School-aged Children) (UGent)
- VAD-leerlingenbevragingen
- Preventiepeiling (Gezond Leven), die data verzamelt over het preventieve gezondheidsbeleid in scholen, ondernemingen, lokale besturen, kinderopvang en zorg- en welzijnsinstellingen

Hoewel deze databronnen elk hun eigen focus en doelgroep hebben, vormen ze samen een complementair geheel voor de opvolging van beleidsimpact. Ze bieden inzichten in de doeltreffendheid van acties binnen verschillende thema's, settings en doelgroepen en helpen beleidsmakers zo bepalen waar bijsturing nodig is.

5 Rekening houden met beleidsimpact bij de evaluatie van interventies

Effectevaluaties leveren waardevolle input voor impactevaluaties op beleidsniveau. Ze tonen wat werkt, voor wie en in welke context, en kunnen dienen als casestudy's binnen bredere beleidsanalyses. Door resultaten van meerdere interventies te bundelen analyseren of te vergelijken, ontstaat inzicht in hoe deze samen bijdragen aan de overkoepelende gezondheidsdoelstellingen.

Daarnaast kunnen interventie-evaluaties helpen verklaren waarom impact verschilt tussen regio's, doelgroepen of implementatiestrategieën, bijvoorbeeld wanneer procesevaluaties tonen dat bepaalde groepen minder goed bereikt worden.

Om beleidsimpact en gezondheidsdoelstellingen adequaat mee te nemen, moet de Theory of Change niet alleen focussen op gedrags- of omgevingsverandering binnen de doelgroep, maar ook op het bredere beleidskader waarin de interventie opereert. Hierdoor ontstaat een Theory of Change die het interventieniveau en beleidsniveau verbindt.

Bij het ontwikkelen van een Theory of Change of Logic Model voor een interventie kan beleidsimpact expliciet worden geïntegreerd door duidelijk te maken:

- welke langetermijndoelstellingen de interventie ondersteunt (bv. betere gezondheidsuitkomsten, minder gezondheidsverschillen);
- aan welke Vlaamse gezondheidsdoelstelling(en) de interventie bijdraagt;
- welke databronnen gebruikt kunnen worden om de impact op beleidsniveau te monitoren;
- hoe de interventie kan bijdragen aan impact in andere sectoren (bv. onderwijs, cultuur).

Door deze bredere impact expliciet mee te nemen, wordt duidelijk hoe interventies passen binnen het ruimere preventiebeleid en bijdragen aan intersectorale strategieën.

6 Zicht op collectieve impact

Collective Impact (CI) is een aanpak waarbij diverse actoren samenwerken rond gemeenschappelijke doelen om complexe maatschappelijke problemen aan te pakken (118, 119). Deze problemen hebben doorgaans meerdere oorzaken en worden beïnvloed door wisselende contextfactoren. CI-initiatieven worden gekenmerkt door vijf condities: een gemeenschappelijke agenda, gedeelde meetsystemen, wederzijds versterkende activiteiten, continue communicatie en een coördinerende organisatie.

Het evalueren van CI-initiatieven is bijzonder uitdagend. Ze omvatten meerdere projecten, stakeholders en veranderende strategieën, waardoor het moeilijk is om causale verbanden vast te stellen of effecten toe te schrijven aan afzonderlijke interventies. Externe factoren, zoals beleidswijzigingen of demografische veranderingen, beïnvloeden de uitkomsten en versterken deze complexiteit. Daarnaast is evaluatie vaak afhankelijk van zelfrapportage, kleine steekproeven en onvolledige datasets, wat de betrouwbaarheid van conclusies kan beperken. Organisatorische obstakels spelen eveneens een rol: sommige organisaties ervaren weerstand, beperkte capaciteit of extra administratieve lasten. Ook is financiering van CI-initiatieven vaak kortdurend, wat longitudinale evaluaties bemoeilijkt.

In het artikel *“Evaluating Collective Impact: Five Simple Rules”* bespreekt Cabaj (119) vijf praktische principes om de evaluatie van CI-initiatieven te versterken:

- **Gebruik evaluatie om strategisch leren te stimuleren**

Evaluatie moet niet beperkend werken, maar juist leren en adaptief handelen ondersteunen. Traditionele evaluaties met vaste plannen en vooraf bepaalde doelen, falen vaak in zulke dynamische omgevingen. Een effectieve aanpak bevat real-time feedback en flexibele designs, met aandacht voor wat werkt, waarom en onder welke omstandigheden.

- **Ontwerp meerdere evaluaties voor meerdere gebruikers**

CI-initiatieven kennen veel stakeholders (bv. uitvoerders, beleid) met uiteenlopende informatienoden. Een uniform evaluatiemodel volstaat daarom niet. Cabaj (2014) spreekt over een “patch evaluation design”: een samenstelling van meerdere, soms overlappende evaluatieprocessen die diverse methoden gebruiken en de resultaten op maat communiceren.

- **Gedeelde meting indien nodig, maar niet noodzakelijk gedeelde meting**

Hoewel gedeelde indicatoren nuttig kunnen zijn voor afstemming, mogen ze geen dogma worden. Een te sterke focus op uniforme indicatoren kan innovatie en strategisch denken beperken. Gedeelde meetsystemen vergen bovendien veel tijd en middelen. Evaluatie moet pragmatisch blijven: meet wat nodig is, en begin klein en werkbaar in plaats van groot en perfect.

- **Wees volledig: let ook op onbedoelde effecten**

In complexe systemen hebben interventies vrijwel altijd onverwachte effecten, zowel positief als negatief. Traditionele evaluaties zijn vaak blind voor deze effecten omdat ze te sterk focussen op vooraf vastgestelde doelen. Door systematisch te vragen naar alle effecten van collectief handelen, ontstaat een rijker en realistischer beeld van impact.

- **Focus op bijdrage (contributie), niet op causale toeschrijving (attributie)**

Bij CI-initiatieven is het vrijwel onmogelijk om veranderingen rechtstreeks toe te schrijven aan één interventie, omdat strategieën verschuiven, interventies elkaar beïnvloeden en vergelijkingsgroepen ontbreken. Contribution analysis is daarom geschikter: niet *bewijzen* dat iets uitsluitend door een

interventie komt, maar aantonen *hoe* verschillende inspanningen samen bijdragen aan verandering. Dit levert meer realistische inzichten op en ondersteunt onderbouwde strategische keuzes.

6.1 Gedeelde uitkomstenkaders en meetsystemen voor organisaties of programma's

Een gedeeld meetsysteem (shared measurement system) biedt een gezamenlijke aanpak om gegevens te verzamelen en te rapporteren op basis van gestandaardiseerde indicatoren en methoden. Het doel is om resultaten van verschillende programma's of organisaties te vergelijken en een volledig beeld te krijgen van hun gezamenlijke impact. Het uitgangspunt is dat organisaties die rond soortgelijke doelen werken, niet elk afzonderlijk hun eigen meetsysteem ontwikkelen, maar een gezamenlijk raamwerk creëren voor het meten van voortgang, resultaten en collectieve impact.

Het ontwikkelen van een gedeeld meetsysteem is echter een intensief en ambitieus proces dat langdurige investering en betrokkenheid vraagt (120, 121). Het proces start met het bepalen van wat gemeten moet worden. Daarbij wordt een onderscheid gemaakt tussen:

- **Een gemeenschappelijk uitkomstenkader:** dit legt gezamenlijke eindresultaten vast, maar bepaalt niet hoe deze worden gemeten.
- **Een gedeeld meetsysteem:** dit bevat gezamenlijk meetmethoden, instrumenten en procedures om de resultaten te evalueren.

Gedeelde meetsystemen hebben verschillende voordelen:

- **Betere samenwerking en afstemming:** organisaties werken vanuit één gedeeld begrip van succes. Door gezamenlijk een Theory of Change te ontwikkelen begrijpen organisaties beter hoe hun activiteiten elkaar aanvullen.
- **Meer inzicht in wat werkt:** doordat data vergelijkbaar en betrouwbaar zijn, kan men patronen herkennen, effecten vergelijken en effectieve strategieën identificeren.
- **Efficiëntie:** In plaats van dat elke organisatie afzonderlijk meetinstrumenten ontwikkelt of rapportages opstelt, kunnen meerdere organisaties gebruikmaken van gedeelde tools en indicatoren. Dit vermindert de administratieve lasten en voorkomt dubbel werk.
- **Hogere meetkwaliteit:** gezamenlijke ontwikkeling bevordert transparantie en wederzijdse verantwoording.

Er zijn echter aandachtspunten. Bij organisaties ontstaat mogelijks de bezorgdheid dat gedeelde meet-systemen vooral door financiers zullen worden gebruikt om hun prestaties te vergelijken en beoordelen, wat weerstand kan oproepen. Daarom is een bottom-up aanpak cruciaal, waarbij sectoren zelf betrokken zijn bij de ontwikkeling van het meetsysteem. Bovendien moet rekening worden gehouden met verschillen in doelen, doelgroepen en activiteiten, om eerlijke en zinvolle vergelijkingen te garanderen.

Het document 'The future of shared measurement: A guide to assessing individual sector readiness' (120) onderscheidt drie factoren met bijhorende criteria, drivers en barrières die bepalen of een sector klaar is voor gedeelde impactmeting (Figuur 23).

Figuur 23 Factoren die bepalen of een sector klaar is voor gedeelde impactmeting.

Sector en infrastructuur	Evidentie en meetpraktijken	Momentum voor gedeelde meting
- Criteria: mate van coherentie en consistentie in interventies, infrastructuur voor samenwerking, bestaande samenwerking en inzicht in de noden binnen de sector. - Drivers: gelijkaardige doelen en vergelijkbare uitkomsten, gedeelde definities, onafhankelijke coördinatie en eerdere succesvolle samenwerking. - Barrières: gebrek aan consensus, grote variatie in interventies, weinig samenwerking, concurrentie om middelen, geen goed begrip van noden.	- Criteria: beschikbaar bewijs van effectiviteit, huidige meetmethoden, druk om effectiviteit aan te tonen. - Drivers: gedeeltelijk ontwikkeld bewijs, academische interesse, diverse meetmethoden zonder dominante standaard, sterke vraag van financiers naar vergelijkbare data. - Barrières: beperkte rapportagevereisten, weinig robuuste meetmethoden, lage prioriteit voor impactmeting bij financiers.	- Criteria: bewustzijn van nut en noodzaak, bereidheid tot deelname, betrokken financiers. - Drivers: bewustzijn van de voordelen, beleids- of financieringsveranderingen die adoptie stimuleren, bereidheid van financiers om ontwikkeling van systeem te financieren. - Barrières: lage bekendheid met de voordelen, versnipperde financiering.

Box 11: Voorbeeld: een gedeeld meetsysteem van LEAP voor ontwikkeling in de vroege kindertijd

Het gedeelde meetsysteem van LEAP (Lambeth Early Action Partnership) werd ontwikkeld om monitoring, evaluatie en leren binnen het programma te versterken (121). LEAP maakt deel uit van het nationale *A Better Start*-initiatief en richt zich op het verbeteren van ontwikkeling in de vroege kindertijd en het verminderen van gezondheidsongelijkheden.

De ontwikkeling van het gedeelde meetsysteem was een intensief proces dat grote uitdagingen dienden te overwinnen: een gefragmenteerde Theory of Change, inconsistente dataverzameling en beperkte capaciteit. Diensten betrokken in LEAP werkten rond 200 verschillende uitkomsten, vaak gemeten met eigen niet-gevalideerde meetinstrumenten. Het ontwikkelingsproces omvatte onder andere: een evidence-based review van determinanten inzake ontwikkeling in de vroege kindertijd, system mapping om overlap en synergie te identificeren, en het creëren van een **vernieuwde Theory of Change met een geneste structuur op drie niveaus**:

1. **Programma-niveau:** Een overkoepelende Theory of Change die het algemene veranderingsmodel van LEAP beschrijft. Dit bestaat uit acht onderling verbonden uitkomstdomeinen die samen de uiteindelijke impact van het programma realiseren.
2. **Domein-niveau:** Voor elk van de acht uitkomstdomeinen werden lange-termijnuitkomsten gedefinieerd die verduidelijken welke verbeteringen LEAP nastreeft en hoe diensten, activiteiten en werkwijzen hieraan bijdragen.
3. **Dienstoniveau:** Meer dan 20 diensten werkten hun eigen Theory of Change uit om te verduidelijken hoe hun specifieke activiteiten (zoals groepsessies of huisbezoeken) bijdragen aan de uitkomsten op middellange en lange-termijn binnen hun domein.

Een voorbeeld: de dienst *Breastfeeding Peer Support* richtte zich op de uitkomst “Meer vrouwen geven borstvoeding”, onderdeel van het domein *Versterking van kennis, vaardigheden en gedrag van gezinnen*.

Deze structuur maakt zichtbaar hoe diensten samen bijdragen aan collectieve impact, biedt een logisch kader voor evaluatie en ondersteunt standaardisatie van metingen doordat diensten binnen hetzelfde domein vergelijkbare indicatoren en tools gebruiken.

De implementatie van het meetsysteem vereiste drie belangrijke veranderingen:

- **Capaciteitsopbouw:** uitbreiding van het LEAP-team en het aanstellen van een Data Collection Manager om diensten actief te ondersteunen.
- **Praktische hulpmiddelen:** op maat gemaakte Monitoring, Evaluation and Learning (MEL)-frameworks en een digitale Outcomes Assessment Tool om het registreren van gegevens mogelijk te maken.
- **Data-integratie:** het Data Integration Platform zorgde voor een veilige koppeling van gegevens en een centrale dataset voor analyse en rapportage, wat leidde tot betere datakwaliteit, lagere werkdruk en zichtbare collectieve impact.

Hoofdstuk 9

Samenwerking en rollen bij de evaluatie van interventies

Een kwaliteitsvolle evaluatie van preventieve interventies vraagt niet alleen om de juiste methoden, maar ook om een doordachte samenwerking tussen alle betrokken partijen en om een goede rolverdeling om de wetenschappelijke kwaliteit en de beleidsrelevantie van de evaluatie te waarborgen. Dit hoofdstuk bespreekt twee benaderingen uit de literatuur die handvatten bieden voor het organiseren van zulke samenwerkingen.

1 Samenwerkingsmodellen tussen onderzoeks- en beleidsteams (handboek Impact Evaluation in Practice)

Het handboek Impact Evaluation in Practice, gericht op impactevaluatie in ontwikkelingswerk, bespreekt verschillende modellen om de samenwerking tussen onderzoeksteams en beleidsteams²² te organiseren. De meest geschikte aanpak hangt af van de context, de doelstellingen van de evaluatie en de risico's die men wil vermijden. Te weinig samenwerking kan ertoe leiden dat een evaluatie onvoldoende aansluit bij beleidsvragen of bij de praktijk, terwijl te sterke integratie net kan leiden tot belangenconflicten.

Het handboek onderscheidt drie samenwerkingsmodellen, elk met specifieke voor- en nadelen (Tabel 8). Bij **uitbesteding** besteedt het beleidsteam de evaluatie volledig of gedeeltelijk uit aan een externe partij, zoals een onderzoeksinstituten. Volledige uitbesteding is zelden ideaal: modellen waarbij enige samenwerking behouden blijft, leveren doorgaans betere resultaten op. In een **partnerschapsmodel** werken onderzoeksteam en beleidsteam nauwer samen op basis van wederzijds belang. Deze modellen werken goed wanneer er duidelijke afspraken en gezamenlijke doelen zijn. In een **volledig geïntegreerd model** zijn het onderzoeksteam en het uitvoerend team identiek. Dit is vooral geschikt voor gecontroleerde experimentele settings waarin onderzoekers maximale controle nodig hebben, maar minder bruikbaar voor de onderbouwing van beleid in realistische omstandigheden.

²² Beleidsteam omvat hier zowel de overheid als de uitvoerende organisatie (in de Vlaamse context: de partnerorganisatie of organisatie met terreinwerking)

Tabel 8 Drie modellen van samenwerking tussen onderzoek- en beleidsteams (gebaseerd op impactevaluatie in het ontwikkelingswerk)

Model	Voordelen	Risico's
Uitbesteding	• Verhoogt onafhankelijkheid doordat uitvoering gescheiden is van beleid. • Praktisch wanneer uitvoerders beperkte tijd of evaluatie-expertise hebben.	• Beperkte samenwerking kan leiden tot gebrekkige aansluiting van evaluatie op beleidsvragen. • Minder invloed van onderzoekers op ontwerp of implementatie, zeker wanneer programma's al lopen. • Contractuele afspraken kunnen flexibiliteit beperken.
Partnerschap	• Gezamenlijke inzet bevordert kwaliteit en beleidsrelevantie. • Onderzoekers kunnen cofinanciering aantrekken.	• Onderzoekers kunnen academische elementen toevoegen die lokaal minder nuttig zijn. • Beleidsmakers onderschatten soms het belang van strikte methodologie.
Volledig geïntegreerd model	• Onderzoekers hebben maximale controle over design van de evaluatie en uitvoering van programma's. • Geschikt voor het testen van causale mechanismen en theorieën onder ideale omstandigheden.	• Vooral bruikbaar in gecontroleerde settings: resultaten hebben beperkte externe validiteit in real-life beleidssituaties.

2 Het samenwerkingsmodel tussen preventieorganisaties en evaluatieonderzoekers volgens het Europees Preventiehandboek (2019)

Het Europees Preventiehandboek is de Nederlandstalige vertaling van het European Prevention Curriculum (EUPC), ontwikkeld door het Europees Monitoring Centre for Drugs and Drug Addiction (EMCDDA) (92). Het handboek biedt beleidsmakers, preventiewerkers en andere betrokken professionals een praktische leidraad voor wetenschappelijk onderbouwde preventie van middelengebruik. Een interessant onderdeel van het handboek is het uitgewerkte samenwerkingsmodel voor evaluatie, waarin preventieprofessionals en gespecialiseerde evaluatieonderzoekers in één team samenwerken voor de evaluatie van preventieve interventies, met een duidelijke afbakening van rollen en verantwoordelijkheden.

De externe evaluator is meestal verbonden aan een universiteit of hogeschool en beschikt over de nodige expertise in het evalueren van gezondheidsbevorderende interventies. Regelmatige afstemming tussen alle betrokken partijen vormt een belangrijk onderdeel van dit model. Deze overlegmomenten bewaken de voortgang van het evaluatieproces, signaleren knelpunten tijdig en zorgen ervoor dat de evaluatie goed wordt ingebed in de werking van de interventie. Daardoor is de kans groter dat bevindingen effectief worden benut voor verdere verbetering. Tabel 9 toont de mogelijke taakverdeling, waarbij beide partijen hun eigen expertise inbrengen.

Tabel 9 Samenwerkingsmodel uit het Europees Preventiehandboek.

Evaluatiestap	Preventieteam	Extern evaluatieteam
Stakeholders betrekken	Leidende rol: Kent de stakeholders en bepaalt wie moet worden betrokken.	Ondersteunende rol: Zorgt voor effectieve betrokkenheid.
Interventie omschrijven	Gedeelde rol: Levert informatie aan over de interventie.	Gedeelde rol: Omschrijft de interventie en betreft daarbij het programmeam.
Focus van de evaluatie bepalen	Gedeelde rol: Bepaalt de belangrijkste onderzoeksvragen.	Gedeelde rol: Scherpt de onderzoeksvragen aan en ontwikkelt het onderzoeksdesign.
Gegevens verzamelen	Ondersteunende rol: Faciliteert de toegang tot onderzoeksgegevens en deelnemers voor de gegevensverzameling.	Leidende rol: Voert de gegevensverzameling uit in samenwerking met het preventieteam.
Conclusies formuleren	Gedeelde rol: Helpt bij het interpreteren van de resultaten en het formuleren van aanbevelingen.	Gedeelde rol: Leidt de analyse van de onderzoeksgegevens, werkt daarvoor samen met het preventieteam.
Gebruik en delen van resultaten	Leidende rol: Zorgt ervoor dat de resultaten gebruikt worden voor de voortzetting van de interventie.	Ondersteunende rol: Presenteert de resultaten die het gebruik van het programma ondersteunen.

Hoofdstuk 10

Prioritering van evaluaties

In een context waarin middelen, tijd en evaluatiecapaciteit beperkt zijn, is het essentieel om strategisch te bepalen welke interventies prioritair geëvalueerd moeten worden. Niet elke interventie vraagt dezelfde diepgang of urgentie van evaluatie, en niet elke evaluatie levert een even grote meerwaarde op voor beleid en praktijk. Dit hoofdstuk bespreekt welke criteria richting kunnen geven aan deze prioritering, welke methoden beschikbaar zijn om keuzes systematisch te maken, en hoe deze in de praktijk zijn toegepast.

1 Criteria voor het prioriteren van evaluaties

Beleidsmakers, onderzoekers en uitvoerende instanties moeten strategisch keuzes maken over welke gezondheidsbevorderende interventies ze in eerste instantie evalueren of verder onderbouwen. In de praktijk is het vaak onduidelijk welke criteria daarbij worden gebruikt en hoe systematisch dit gebeurt.

Een recente scoping review bracht 40 studies samen waarin meer dan 300 unieke prioriteringscriteria werden beschreven voor implementatie en evaluatie van volksgezondheidsprogramma's (122). Deze criteria werden geclusterd in 16 domeinen (Tabel 10). Veelgenoemde domeinen zijn onder meer: ziektelast, sociale overwegingen en gezondheidseffecten van interventies.

Een opvallende bevinding van de review is dat de literatuur zich vooral richt op prioritering van implementatie, terwijl nauwelijks beschreven wordt wanneer en waarom interventies prioritair geëvalueerd moeten worden. Slechts één studie stelde expliciete evaluatiecriteria op om te bepalen wanneer programma's geëvalueerd zouden moeten worden (96)²³. Hoewel formele evaluaties van de effectiviteit van deze prioriteringsraamwerken schaars zijn, tonen de ervaringen die voorhanden zijn dat het gebruik van zulke raamwerken wel degelijk leiden tot meer systematische en transparante besluitvorming, en meer draagvlak door betrokkenheid van stakeholders.

Sommige organisaties gebruiken systematische methoden om prioriteiten te bepalen, vooral bij implementaties. Deze methoden worden gewaardeerd om hun transparantie, maar zijn vaak tijdsintensief. Bovendien bestaat het risico dat programma's eerder voorrang krijgen op basis van populariteit dan op impact.

²³ Ogilvie et al. (2011) focust op evalueerbaarheid van interventies en wordt beschreven in Box 5, hoofdstuk 6.

Tabel 10 Domeinen van prioriteringscriteria voor de implementatie en evaluatie van volksgezondheidsprogramma's.

Domein	Criteria
Ziektelast	Aantal mensen dat getroffen wordt door een ziekte, ernst van een ziekte, o.a. zorgnoden, zorgvraag en zorggebruik Exclusief economische ziektelast.
Gelijkheid/Eerlijkheid/Ethiek	Rechtvaardigheid, onevenredige impact van een ziekte op kwetsbare groepen, mate waarin een interventie ongelijkheid vermindert.
Kosteneffectiviteit	Verhouding tussen kosten en baten van een interventie (waarde voor geld).
Budgetimpact	Financiële gevolgen voor de financierende instantie van (gebrek aan) investeringen in een interventie, beschikbaarheid van middelen voor evaluatie.
Kosten van ziekte voor gezondheidssysteem	Kosten voor het gezondheidssysteem om zorg te leveren aan mensen met de ziekte (economische ziektelast voor het systeem).
Kosten van ziekte voor patiënt	Kosten die patiënten maken om zorg te verkrijgen voor een ziekte.
Kosten van ziekte voor maatschappij	Kosten van een ziekte voor de maatschappij, o.a. door verzuim of productiviteitsverlies.
Gezondheids- en welzijnseffecten	Positieve en negatieve effecten van de interventie op gezondheid en welzijn, zoals effectiviteit, vergelijkende voordelen/nadelen, bijwerkingen en veranderingen in kennis of gedrag.
Sociale overwegingen	Sociale aspecten van de ziekte of de interventie, zoals stigma en maatschappelijke aanvaardbaarheid.
Overwegingen van organisaties/aanbieders	Aanvaardbaarheid, uitvoerbaarheid van een interventie voor uitvoerende organisaties, zorgverleners en industrie, o.a. aansluiting bij richtlijnen of organisatiedoelen.
Wettelijk en regelgevend kader	Wet- en regelgeving die van invloed is op de uitvoering van de interventie.
Politieke overwegingen	Aansluiting van de interventie bij beleid of prioriteiten.
Internationale steun	Aansluiting van de interventie bij internationale strategieën.
Uitvoerbaarheid van implementatie	Haalbaarheid van implementatie van de interventie, o.a. beschikbaarheid en capaciteit, bereik van de doelgroep.
Huidige dienstverlening	Bestaande voorzieningen in de gemeenschap, o.a. beschikbaarheid van interventies of alternatieven en hun toegankelijkheid.
Evidentie	Beschikbaarheid en kwaliteit van bestaande evidentie.

2 Methode: versnelde James Lind Alliance (JLA) aanpak

De JLA-aanpak is een internationaal erkend model om onderzoeks- en evaluatieprioriteiten vast te stellen. Wereldwijd zijn er al meer dan 150 Priority Setting Partnerships uitgevoerd met deze aanpak. De methode combineert brede betrokkenheid van stakeholders, consensusvorming en transparante verslaggeving. Een stapsgewijze gids is vrij beschikbaar ²⁴.

Traditionele JLA-trajecten duren 12 tot 18 maanden, maar recent ontwikkelde versnelde varianten, zoals ontwikkeld door Health and Care Research Wales en Social Care Wales, kunnen binnen één jaar worden doorlopen²⁵. Hierbij wordt gebruik gemaakt van gerichte vragenlijsten, bestaande data en compacte workshops met zorgvuldig samengestelde groepen. De aanpak helpt beleidsmakers om snel én participatief tot prioriteiten te komen, op voorwaarde dat doelen vooraf helder worden geformuleerd en neutrale begeleiding wordt voorzien. Box 12 en 13 illustreren hoe verkorte JLA-processen met succes zijn toegepast in respectievelijk de Belgische geestelijke gezondheidszorg (MIND10) en het Engelse sociaal werk.

Box 12: Prioritering van onderzoeksvragen voor het Fund for Mental Health Research in het Belgische MIND10-project

Het Belgische MIND10-project (2025) identificeerde via een participatief traject de belangrijkste onderzoeksprioriteiten binnen de geestelijke gezondheidszorg voor het Fund for Mental Health Research van de Koning Boudewijnstichting. De aanpak was gebaseerd op een verkorte JLA-methodiek en omvatte drie stappen:

- **Inventarisatie van onderzoeksvragen:** Met input van 60 patiënten, familieleden en zorgverleners, aangevuld met literatuur, werd een brede lijst van 769 onderzoeksvragen opgesteld.
- **Selectie van een shortlist:** De vragen werden thematisch gebundeld tot 103 bredere vragen, waarna 35 stakeholders ze beoordeelden op prioriteit. Op basis van prioriteitsscores, herkomst van de vragen, omvang van de vragen en bestaande onderzoekslacunes werd een shortlist van 24 vragen samengesteld.
- **Workshops en finale prioritering:** Tijdens drie interactieve workshops rangschikten stakeholders de shortlist via groepsdiscussies en prioriteitspunten. Dit resulteerde in een top tien van onderzoeksprioriteiten, die de basis vormt voor een oproep naar praktijkgericht onderzoek.

Box 13: Snelle prioritering van evaluaties: case van innovaties in sociaal werk

Het artikel van Cowan et al. (2021) beschrijft een versnelde methode, gebruik makend van de JLA-aanpak, om innovaties in de volwassenenzorg en het sociaal werk te identificeren en te prioriteren voor evaluatie (123). In vier maanden tijd werd een prioriteringsproces doorlopen dat uit vier stappen bestond. Eerst werd via een brede consultatie van stakeholders een longlist van 158 innovaties verzameld en gegroepeerd in thematische categorieën. Vervolgens werden op basis van literatuur en eerdere ervaringen selectiecriteria opgesteld (o.a. relevantie, concreetheid en geschiktheid voor snelle evaluatie) om de lijst te verkleinen. In een derde stap werden vergelijkbare innovaties samengevoegd

²⁴ <https://www.jla.nihr.ac.uk/jla-guidebook>

²⁵ Zie <https://www.jla.nihr.ac.uk/news/reflections-rapid-priority-setting-and-james-lind-alliance-method> voor uitgebreide bespreking en casussen

en initiatieven die al uitgebreid waren geëvalueerd of onvoldoende vernieuwend waren, verwijderd. Dit resulteerde in een shortlist van 20 innovaties. Tot slot werden deze 20 innovaties besproken in een eendaagse workshop met 23 stakeholders (gebruikers, mantelzorgers, professionals, aanbieders, onderzoekers en beleidsmakers). In opeenvolgende rondes, telkens in nieuwe groepjes en begeleid door onafhankelijke facilitators, bespraken deelnemers hun prioriteiten en kwamen ze tot een gezamenlijke top vijf. De vijf prioritaire innovaties om te evalueren vertegenwoordigen een mix van zorgcoördinatie, familiegericht ondersteuning, technologische oplossingen en interventies voor onderbediende groepen. Dit voorbeeld illustreert dat de JLA-aanpak bruikbaar is om in korte tijd en met een brede betrokkenheid van stakeholders een prioriteitenlijst op te stellen, hoewel de tijdsdruk het moeilijk maakte om sommige stakeholdergroepen te betrekken.

3 Methode: Systematic Screening and Assessment (SSA)

De Systematic Screening and Assessment (SSA) methode biedt een relatief snelle en kostenefficiënte manier om interventies te identificeren die voldoende potentieel hebben om verder te evalueren (124, 125). De methode bestaat uit een gefaseerde aanpak, opgebouwd uit zes stappen. In Figuur 24 worden deze stappen toegelicht aan de hand van een concreet voorbeeld.

Figuur 24 Zes stappen van de SSA-methode, toegepast op het hospital-based youth violence prevention project (126)



4 Methode: strategisch evaluatieportfolio

Het strategisch evaluatieportfolio van Spilsbury et al. (127) beschrijft een methode om evaluatiekansen binnen organisaties systematisch te analyseren en te prioriteren, met als doel een portfolio samen te stellen dat maximaal bijdraagt aan beleidsvorming, programmabeheer en verantwoording. De methode vertrekt niet vanuit het uitsluiten van evaluaties, maar vanuit het optimaliseren van de beschikbare middelen.

Het raamwerk onderscheidt twee hoofddimensies: **aantrekkelijkheid** (de potentiële meerwaarde van een evaluatie, bijvoorbeeld voor beleidsverantwoording, leren en programma-verbetering) en **haalbaarheid** (de praktische uitvoerbaarheid, zoals beschikbare tijd, middelen en expertise).

Evaluatiemogelijkheden worden beoordeeld via een scoringsmatrix met drie gewogen componenten:

- **Strategisch belang (20%):** afzonderlijk gewogen omdat het een unieke rol speelt in besluitvorming.
- **Aantrekkelijkheid (50%):** o.a. effectgrootte, spreiding, mate van causaliteit, financieringskansen en potentieel voor organisatorisch leren.
- **Haalbaarheid (30%):** o.a. vereiste technische en evaluatie-expertise, en benodigde middelen en tijd.

De methode werd getest in een milieuorganisatie, waar evaluatieprofessionals onafhankelijk scoringslijsten invulden, verschillen bespraken en tot gezamenlijke scores per evaluatie kwamen. Uit de toepassing bleek dat vooral evaluaties die vroegtijdige feedback bieden op lopende initiatieven en een hoog strategisch belang hebben, de grootste meerwaarde leveren voor operationele verbetering en leren.

Het raamwerk helpt organisaties een gebalanceerd evaluatieportfolio samen te stellen, waarin zowel heel strategische als haalbare evaluaties worden geïdentificeerd. Daarbij benadrukt het dat ook tactische evaluaties, bijvoorbeeld vanuit wettelijke verplichtingen of verantwoordingsvereisten, een blijvende plaats hebben binnen het geheel.

Hoofdstuk 11

Gebruik van evaluatiebevindingen

De waarde van een evaluatie ligt niet enkel in het uitvoeren ervan, maar vooral in de manier waarop de bevindingen daadwerkelijk worden gebruikt. Evaluatie-inzichten kunnen leiden tot beter beleid, effectievere interventies en meer gerichte inzet van middelen. Dit hoofdstuk bespreekt hoe evaluatieresultaten worden toegepast in beleidsvorming, welke structuren en beoordelingskaders bestaan om interventies te erkennen, en welke methoden beleidsmakers kunnen ondersteunen bij beslissingen over investeringen en prioriteiten.

1 Toepassing van onderzoek in beleidsvorming

La Brooy en Kelaher (128) onderzochten hoe onderzoeksbevindingen effectief kunnen worden ingezet in beleidsvorming aan de hand van casussen van gezondheidsbevorderende interventies. Ze gebruikten hiervoor de typologie van Murray (129), oorspronkelijk ontwikkeld voor transportbeleid, om drie modellen van onderzoekstoepassing te onderscheiden:

- **Customer/client-model:** Onderzoekers leveren gegevens aan beleidsmakers, die deze vervolgens zelfstandig gebruiken zonder verdere betrokkenheid van de onderzoekers. Onderzoek vult vooral hiaten in het beleidsproces op. Dit model wordt bekritiseerd, omdat het onvoldoende aansluit bij de dynamiek en complexiteit van beleidsprocessen.
- **Interactief model:** Onderzoekers en beleidsmakers werken samen met regelmatige overlegmomenten en overlappende rollen. Onderzoekers kunnen bijvoorbeeld deelnemen aan adviescommissies. Dit model bevordert wederzijds begrip en zorgt ervoor dat onderzoeksresultaten beter aansluiten bij beleidsvragen.
- **Joint construction-model:** Onderzoekers, beleidsmakers en burgers werken samen in een proces van gezamenlijke kennisopbouw. Bijvoorbeeld via lokale onderzoekscomités waarin alle partijen samen problemen definiëren en oplossingen ontwikkelen.

In enkele casestudies gaven deelnemers aan dat uitwisseling tussen beleidsmakers en onderzoekers een duidelijke meerwaarde biedt. Bevorderende factoren voor een sterke relatie waren onder meer regelmatige vergaderingen, duidelijke rolverdeling, betrokkenheid van partners, en een leidende organisatie die de samenwerking faciliteert. Uitdagingen hadden vooral te maken met het samenstellen van een representatieve groep van belanghebbenden en het managen van verwachtingen.

De auteurs stellen dat de traditionele, lineaire relatie tussen onderzoek en beleid, waarbij beleidsmakers pas laat in het proces onderzoek raadplegen, niet langer volstaat. Hoewel het joint construction-model theoretisch het meest ideaal is, bleek in de praktijk het interactief model het meest werkbaar: het biedt ruimte voor dialoog en samenwerking, maar blijft uitvoerbaar binnen de beperkingen van tijd, middelen en politieke prioriteiten. De auteurs benadrukken dat succesvolle toepassing van onderzoek in beleid vooral afhankelijk is van sterke governance: duidelijke rolafbakening, gedeelde besluitvorming en continue communicatie.

2 De beoordeling en erkenning van effectieve interventies door *Samenwerkingsverband Erkenning van Interventies* in Nederland

In Nederland vormen zes landelijke kennisinstituten samen het Samenwerkingsverband Erkenning van Interventies. Ze hanteren een uniforme werkwijze en criteria voor de erkenning van interventies in zorg, preventie, sport, welzijn en jeugdhulp. Een onafhankelijke erkenningscommissie beoordeelt de interventies op kwaliteit, effectiviteit en uitvoerbaarheid. Het doel is om de zichtbaarheid van goed onderbouwde interventies te vergroten en een lerende praktijk te stimuleren.

Erkende interventies maken duidelijk wat werkt, voor wie en onder welke omstandigheden, en worden ontwikkeld door praktijkorganisaties of onderzoeksinstituten. Gemeenten en zorg- en welzijnsorganisaties gebruiken de databanken met erkende interventies om onderbouwde keuzes te maken. Hoewel het aantal erkende interventies inmiddels boven de 700 ligt, blijft het systeem dynamisch. Er is blijvend ruimte voor de beoordeling van nieuwe of vernieuwde interventies, zodat het aanbod actueel en innovatief blijft.

Voor sommige thema's is het aanbod aan erkende interventies nog beperkt, wat de kenniscentra ertoe aanzet om de ontwikkeling van interventies op die gebieden te stimuleren. Daarnaast worden alle erkende interventies periodiek opnieuw beoordeeld om de erkenning te verlengen. Sommige interventies verliezen hierbij hun erkenning, terwijl andere juist worden opgeschaald of doorontwikkeld.

Het aanvragen van erkenning helpt om een interventie beter te onderbouwen, verscherpen en breder toe te passen. Het Samenwerkingsverband ondersteunt dit proces met een duidelijke erkenningsprocedure. De aanvraag voor erkenning start met een aanmelding via de website van het kenniscentrum, waarna het kenniscentrum beoordeelt of de interventie aan de toelatingscriteria voldoet. Vervolgens beschrijft de interventie-eigenaar de interventie aan de hand van een vast format. Hierbij krijgt hij advies en begeleiding van het kenniscentrum. Een onafhankelijk panel of commissie beoordeelt de interventie tot slot op basis van de erkenningscriteria. Bij toekenning is de erkenning drie tot vijf jaar geldig, met de mogelijkheid om deze te verlengen. In wat volgt, worden de verschillende toelatings- en erkenningscriteria voor de periode 2024-2027 beschreven.

2.1 Toelatingscriteria

De kenniscentra screenen zelf of een interventie voldoet aan de toelatingscriteria voordat deze in aanmerking komt voor erkenning. Daarbij wordt minimaal gekeken naar de volgende aspecten:

- De interventie is planmatig, doelgericht, en meer dan een eenmalige activiteit. Online interventies moeten gekoppeld zijn aan een fysieke praktijk (bv. offline begeleiding).
- De eigenaar is bereid om de interventie actief te beheren en te verspreiden.
- De interventie wordt op het moment van aanmelding in Nederland uitgevoerd, bij voorkeur op meerdere locaties en beschikt over een Nederlandstalige handleiding voor de uitvoerders.
- Er is reeds een procesevaluatie uitgevoerd.

Bij overlap met bestaande erkende interventies wordt in overleg bekeken wat de toegevoegde waarde van de interventie is.

2.2 Erkenningsniveaus

Vanaf het niveau 'Goed Onderbouwd' worden interventies getoetst door een erkenningscommissie (met leden uit onderzoek, praktijk en beleid). Sommige kenniscentra hanteren een extra instapniveau 'Goed beschreven', waarbij interventies worden getoetst door praktijkpanels met relevante inhoudelijke kennis en ervaring.

Indien de erkenningscommissie of het praktijkpanel de interventie heeft getoetst, zijn er drie erkenningsniveaus mogelijk (Tabel 11):

- **Goed beschreven:** duidelijke omschrijving, uitvoerbaar, actueel en ondersteund met procesevaluatie.
- **Goed onderbouwd:** voldoet aan alle bovengenoemde criteria én beschikt over een sterke theoretische onderbouwing.
- **Effectief:** effectonderzoek toont aan dat de interventie de gewenste uitkomsten bereikt.

Tabel 11 De criteria van de erkenningsniveaus Goed beschreven, Goed onderbouwd, en Effectief (130).

Erkenningsniveau	Criteria
Goed beschreven	✓ Overdraagbaar: doel, doelgroep, aanpak, benodigde expertise, kwaliteitsbewaking, context, randvoorwaarden en kosten zijn beschreven. Handleiding voor uitvoerders is beschikbaar. ✓ Actueel: interventie wordt in Nederland uitgevoerd en sluit aan bij de huidige context. ✓ Gebaseerd op praktijkervaringen via procesevaluatie (kwalitatief of kwantitatief): ○ De onderzoeksopzet is beschreven (methode van onderzoek, doelgroep, aantal respondenten). ○ Er is inzicht in de mate waarin de activiteiten zijn uitgevoerd zoals bedoeld (programma integriteit) en er zijn gegevens beschikbaar over: bereik van deelnemers, succes- en faalfactoren, waardering en ervaring van uitvoerders én doelgroep. ○ Het is duidelijk hoe de interventie wordt aangepast aan de hand van de resultaten van het onderzoek (indien van toepassing).
Goed onderbouwd	✓ Voldoet aan alle criteria van <i>Goed beschreven</i>. ✓ Theoretisch onderbouwd met modellen, theorieën of onderzoek: ○ Er is adequaat beschreven welke factoren van invloed zijn op het vraagstuk of probleem. ○ De factoren die met de interventie worden aangepakt zijn benoemd en gekoppeld aan de (sub)doelen van de interventie (doelverantwoording). ○ Er is adequaat verantwoord hoe met de gekozen aanpak daadwerkelijk de gestelde (sub)doelen bij de beschreven doelgroep(en) bereikt kunnen worden. Er is een logische samenhang tussen doelen, doelgroep en aanpak. ○ De werkzame elementen (of technieken of principes) van de aanpak zijn benoemd en verantwoord, gebaseerd op onderzoek of op een veranderingsmodel of interventietheorie. ○ Er worden relevante bronvermeldingen gegeven

Effectief	✓ Voldoet aan alle criteria van <i>Goed onderbouwd</i>. ✓ Effectonderzoek (kwantitatief en/of kwalitatief) is uitgevoerd naar de interventie zoals deze in de praktijk is, en toont aan dat de interventie leidt tot de gewenste effecten bij de doelgroep. ○ De interventie werd uitgevoerd zoals beschreven in de interventiebeschrijving. ○ De uitkomstmaten in de studie zijn relevant, gegeven de (sub)doelen en doelgroep van de interventie. ○ De meting of waarneming is verricht met betrouwbare en valide methoden en/of instrumenten. ○ De respondenten in de studie zijn representatief voor de doelgroep (respons, uitval). ○ De resultaten zijn met een gepaste statistische techniek geanalyseerd. (indien van toepassing) ○ De grootte van de effecten is redelijk gezien de aard, het doel en de doelgroep van de interventie. (indien van toepassing) ○ Eventuele negatieve effecten zijn gerapporteerd. ○ Het onderzoek is zodanig gedocumenteerd (in een rapport of artikel), dat replicatie van de studie mogelijk is. ✓ De bewijskracht van de effectstudie(s) bepaalt op welk niveau van effectiviteit een interventie wordt erkend (zie Figuur 25): ○ Geen aanwijzingen ○ Eerste aanwijzingen (2 studies met lichte bewijskracht of 1 studie met vrij sterke bewijskracht) ○ Goede aanwijzingen (2 studies met vrij sterke bewijskracht of 1 studie met sterke tot zeer sterke bewijskracht) ○ Sterke aanwijzingen voor effectiviteit (2 studies met sterke tot zeer sterke bewijskracht)
------------------	---

Binnen 'effectief' wordt een verdere indeling gemaakt volgens de bewijskracht van de studies (Figuur 25). De basis hiervoor werd gelegd in het boek *Zicht op effectiviteit* (131). Om de bewijskracht te bepalen, maakt het samenwerkingsverband gebruik van een lichtere versie van de internationale GRADE-systematiek²⁶, waarbij niet alleen de kracht van het onderzoeksdesign als beoordelingscriterium telt, maar ook de kwaliteit van de uitvoering en de relevantie van de uitkomsten.

²⁶ www.gradeworkinggroup.org

Figuur 25 De studiecriteriën voor de verschillende gradaties in bewijskracht. Overgenomen uit Kies erkend, krijg erkenning. Criteria 2024 – 2027 (130)

Bewijskracht design	Kenmerken studie	Kern
Zeer sterk	Hier gelden dezelfde criteria als bij het onderliggende niveau, met als verschil:	
	Er is een experimentele onderzoekopzet (d.w.z. er is een aselechte toewijzing van subjecten aan onderzoeksgroepen) <i>of</i> er is een andere opzet die de causale relatie tussen interventie en effect voldoende overtuigend aantoont.	Randomisatie (of alternatieve opzet)
Sterk	Hier gelden dezelfde criteria als bij het onderliggende niveau, met als toevoeging:	
	Er is een follow-up meting (vuistregel: tenminste 6 maanden na afloop van de interventie) <i>of</i> er is een andere opzet die voldoende zicht biedt op de stabiliteit van de resultaten.	Follow-up meting (of alternatieve opzet)
Vrij sterk	Hier gelden dezelfde criteria als bij het onderliggende niveau, met als verschil:	
	Het is een onderzoek met een interventie- en een controlegroep (care as usual) waarbij niet is gerandomiseerd <i>of</i> er is een andere opzet met vergelijkbare bewijskracht, bijvoorbeeld: – herhaalde casestudies met een baseline <i>of</i> – een time series design met een (multiple) baseline <i>of</i> – alternating treatments <i>of</i> – een studie naar de samenhang tussen de mate waarin een interventie is toegepast en de mate waarin bedoelde uitkomsten zijn opgetreden	Vergelijking met een controlegroep (of alternatieve opzet)
Matig	Hier gelden dezelfde criteria als bij het onderliggend niveau, met als toevoeging:	
	De resultaten zijn vergeleken met ander onderzoek naar de effecten van de gebruikelijke situatie, handelwijze of zorg (care as usual) <i>of</i> de effecten van een andere interventie bij een soortgelijke doelgroep.	Vergelijking met gegevens uit ander onderzoek
Licht	De eisen die op dit niveau gelden zijn:	
	Er zijn resultaten beschikbaar voorafgaand aan, bij de start of tijdens de interventie en bij einde van de interventie (zgn. voor- en nameting), zonder controleconditie. Ook gegevens uit bv. observationeel onderzoek of monitoring zijn hiervoor geschikt. Bij onderzoek naar de ervaren effectiviteit volstaan de resultaten van een nameting.	Vergelijking in de tijd (voor/na)

3 Programme Budgeting and Marginal Analysis (PBMA)

Programme budgeting and marginal analysis (PBMA) is een methode die een gestructureerde en wetenschappelijke aanpak biedt om schaarse middelen optimaal te verdelen over verschillende concurrerende prioriteiten. PBMA wordt veel toegepast in de gezondheidszorg (bv. ziekenhuizen of regionale zorgorganisaties), maar ook in bredere publieke en beleidscontexten. Hoewel de methode doorgaans wordt gebruikt om middelen over diensten of programma's te herverdelen, kan ze ook ingezet worden voor het prioriteren van de evaluatie van interventies.

PBMA steunt op twee kernconcepten uit de economie: opportuiniteitskost en marginale analyse (132). Opportuiniteitskost verwijst naar de waarde die verloren gaat door niet in het beste alternatief te investeren. Dit vereist een afweging van kosten en baten tussen verschillende opties. Marginale analyse verwijst naar de winst (of het verlies) dat ontstaat door kleine veranderingen in middelen. Centraal staat de vraag: "Wat gebeurt er als we extra middelen toevoegen of wegnemen bij een programma?" Dit maakt het mogelijk om beslissingen te nemen zonder dat een volledige herziening van het totale budget nodig is.

HealthKnowledge, een leerplatform voor public health professionals in het Verenigd Koninkrijk, biedt een module met tools en praktijkvoorbeelden voor PBMA²⁷.

Het PBMA-proces bestaat uit zeven stappen (132):

- 1) **Bepalen van doel en reikwijdte:** Vastleggen of de prioritering betrekking heeft op de herverdeling binnen één programma of over meerdere programma's, en afbakenen welke onderdelen binnen of buiten scope vallen.
- 2) **Samenstellen van een programmabudget:** Beschrijven hoe de huidige middelen worden verdeeld.
- 3) **Vorming van een adviespanel:** Betrekken van relevante stakeholders die de marginale analyse ondersteunen.
- 4) **Definiëren van besluitvormingscriteria:** Criteria bepalen op basis van de doelen (bv. gezondheidsdoelstellingen) en de input van het adviespanel (bv. maximaliseren van baten, verminderen van gezondheidsongelijkheid).
- 5) **Identificeren van opties:** Opties bepalen voor investeringen of uitbreidingen, vrijmaken van middelen via efficiëntiewinsten en door afbouw of stopzetting van activiteiten.
- 6) **Marginale analyse:** Voor elke optie inschatten welke bijkomende winsten of kosten ontstaan bij verandering in de inzet van middelen.
- 7) **Valideren en herverdelen van middelen:** Beslissingen finaliseren en budgetverschuivingen doorvoeren.

Voor het afwegen van opties in stap 5 kan gebruik worden gemaakt van **Multi Criteria Decision Analysis (MCDA)** (133). MCDA helpt besluitvormers systematisch keuzes te maken wanneer meerdere criteria en belangen spelen. De methode combineert kwalitatieve elementen (stakeholderperspectieven) met

²⁷ HealthKnowledge - Programme Budgeting and Marginal Analysis
<https://www.healthknowledge.org.uk/interactive-learning/pbma>

kwantitatieve elementen (voorkeuren en prestaties van opties uitgedrukt in scores en gewichten), en berekent voor elke optie een totaalscore. De Government Analysis Function (een netwerk van en voor overheidspersoneel in het Verenigd Koninkrijk) biedt een gids voor het gebruik van MCDA²⁸, met bijzondere aandacht voor besluitvorming in de publieke sector.

Box 14: PBMA voor gezondheidsbevordering in Wales

In 2013 voerde Public Health Wales een PBMA-proces uit in opdracht van de minister van Gezondheid en Sociale Zorg. Het doel was om de nationale uitgaven aan gezondheidsbevordering kritisch te evalueren en te verbeteren. Voor dit proces werd een panel samengesteld van twaalf experts uit diverse sectoren, waaronder beleid, lokale overheden en volksgezondheid. Ze beoordeelde 25 gezondheidsbevorderende initiatieven, met een totaalbudget van £15,1 miljoen (134).

Vijf reviewsubgroepen verzamelden volgende informatie over de interventies: effectiviteit, kosten-effectiviteit, impact op gezondheidsongelijkheid, stakeholder feedback, en alternatieve opties. Interventies kregen een kleurcode (rood, oranje of groen) op basis van sterkte van het bewijs binnen elk domein:

- Rood: weinig of geen bewijs dat de interventie gezondheidswinst oplevert.
- Oranje: beperkt of onduidelijk bewijs, of noodzaak tot herziening van het programma.
- Groen: redelijke tot goede evidentie, mits blijvende aandacht voor bereik en kosteneffectiviteit.

De kleurbeoordelingen werden uiteindelijk samengevoegd tot één eindscore per initiatief. Tijdens drie plenaire sessies besprak het panel de resultaten. Via elektronische stemrondes werden criteria vastgelegd voor evaluatie, zoals prioriteit voor de overheid, expert- en stakeholdermeningen, effectiviteit, kosteneffectiviteit en (potentiële) impact op het verkleinen van gezondheidsverschillen. In de derde sessie werd gestemd over afbouw of stopzetting van interventies. Zeven initiatieven kregen de aanbeveling tot volledige desinvestering, drie initiatieven tot gedeeltelijke desinvestering en geen enkel initiatief werd aanbevolen voor extra middelen. Hoewel de gezondheidsdoelen op zich zinvol waren, ontbrak voor veel initiatieven overtuigend bewijs voor effectiviteit, kosteneffectiviteit of impact op ongelijkheid.

Na afloop voerde het panel een marginale analyse uit, waarbij leden een hypothetisch budget van £5 miljoen mochten herverdelen. De voorkeur ging uit naar vroege interventies (bijv. prenatale zorg, jonge kinderen) en de thema's obesitas, geestelijke gezondheid en tabakscontrole.

Ondanks enkele beperkingen concludeerden betrokkenen dat het PBMA-proces succesvol was: het leidde tot een beter begrip van bestaande uitgaven en stimuleerde een cultuur van evidence-based besluitvorming en kosteneffectieve inzet van middelen.

²⁸ Government Function Analysis - An Introductory Guide to Multi-Criteria Decision Analysis (MCDA)
<https://analysisfunction.civilservice.gov.uk/policy-store/an-introductory-guide-to-mcda/>

Hoofdstuk 12

Werkgroepen met stakeholders: synthese

Om een breed en gedragen beeld te vormen van de noden, verwachtingen en mogelijkheden rond evaluatie binnen de Vlaamse preventiesector, werden in het kader van dit project twee werkgroepen georganiseerd: één met actoren uit de preventiesector (i.e., partnerorganisaties, organisaties met terreinwerking en Gezondheidsmakers) en één met vertegenwoordigers uit beleid en onderzoek. Beide groepen kwamen twee keer samen en reflecteerden over haalbare evaluatiepraktijken, randvoorwaarden voor kwaliteitsvol onderzoek, samenwerking en rolverdeling, en de toekomst van effectevaluatie binnen de sector. Dit hoofdstuk vat de belangrijkste inzichten samen en schetst de rode lijnen die door beide groepen gedeeld werden.

1 Bijeenkomst 1: noden, randvoorwaarden en rolverdeling

Tijdens de eerste bijeenkomst werd afgestemd over de randvoorwaarden en hefboomen voor kwalitatieve en haalbare effectevaluatie, het gebruik van evaluatiekaders, onderzoeksdesigns, en de rolverdeling en samenwerkingsmogelijkheden tussen organisaties en onderzoekers.

1.1 Werkgroep met actoren uit de preventiesector

De eerste bijeenkomst met actoren uit de preventiesector vond plaats op donderdag 11 september 2025 van 12u tot 16u30 in het VAC Mechelen. In totaal namen 17 deelnemers deel aan de werkgroep.

1.1.1 Huidige ervaringen en percepties inzake effectevaluatie

Hoewel evaluatie breed wordt erkend als waardevol, ervaren organisaties in de preventiesector dat het in de praktijk vaak **tijdintensief, complex en moeilijk uitvoerbaar** is met de beschikbare middelen. Hierdoor worden voornamelijk procesevaluaties uitgevoerd, omdat deze intern haalbaar zijn en geen gespecialiseerde methodologische expertise vereisen. Robuuste effectevaluaties vormen een grotere uitdaging door **praktische beperkingen**, een gebrek aan data en capaciteit, en ethische drempels. Rollen en verwachtingen zijn bovendien niet altijd duidelijk: door beleidswijzigingen wordt vaker naar effecten gevraagd, ook wanneer deze oorspronkelijk niet waren voorzien, en bestaat onzekerheid over welke evaluatietaken bij partnerorganisaties horen en welke bij onderzoekers. Daarnaast leeft er frustratie over de beleidsfocus op **harde eindpunten**, terwijl tussentijdse veranderingen in kennis, motivatie of omgeving eveneens belangrijke indicatoren zijn. De geringe beschikbaarheid van frequente en representatieve populatiedata bemoeilijkt bovendien het opvolgen van snel evoluerende trends. Tot slot heerst er onduidelijkheid over **wat precies geëvalueerd moet worden** wanneer interventies al wetenschappelijk onderbouwd zijn, en hoe langetermijneffecten kunnen worden opgevolgd met de beperkte middelen die er zijn.

1.1.2 Randvoorwaarden en hefboomen voor kwaliteitsvolle evaluatie

1.1.2.1 *Prioritaire randvoorwaarden*

Kwaliteitsvolle evaluatie vraagt om een aantal structurele randvoorwaarden die vandaag niet altijd aanwezig zijn. Zo benadrukken organisaties dat **structurele en voldoende financiering**, idealiter expliciet ingecalculeerd binnen projecten, essentieel is, al mag dit het implementatiebudget niet onder druk zetten. Daarnaast is er nood aan een **meerjarenperspectief en beleidsmatige stabiliteit**, omdat effectmetingen tijd vergen en moeilijk te realiseren zijn binnen korte financieringscycli. Organisaties vragen **duidelijke verwachtingen** over welk type van evaluatie wanneer nodig is, aangevuld met **ondersteuning bij de keuze van indicatoren** en ruimte voor dialoog met het Departement Zorg hierover. Ook een **betere monitoringinfrastructuur** is cruciaal: betrouwbare registratiedata zijn schaars en digitale tools die over thema's en settings heen bruikbaar zijn ontbreken vaak.

Verder wordt **directe samenwerking met onderzoekers** gezien als een belangrijke hefboom, maar partnerorganisaties willen tegelijk voldoende autonomie behouden en benadrukken de waarde van interne evaluatie-expertise. Ze vragen ook dat de overheid **alternatieve en kwalitatieve evaluatiemethoden erkent**, zodat impact niet uitsluitend wordt beoordeeld via harde effectmetingen. Tot slot wijzen organisaties op het belang van het **systematisch gebruik en delen van evaluatieresultaten**; vandaag blijven inzichten te vaak intern of bij de administratie zonder externe disseminatie. Dit alles leidt tot bredere vragen over proportionaliteit: wanneer is een uitgebreide effectevaluatie nodig, en wanneer volstaat het om te vertrouwen op bestaande (inter)nationale evidentie, in het bijzonder wanneer methodieken voor de Vlaamse context aangepast zijn? De werkgroep benadrukt dat expertise voor specifieke doelgroepen en intermediairs eveneens moeten worden meegenomen in evaluatie-inspanningen.

1.1.2.2 *Bestaande hefboomen voor evaluatie*

Binnen de preventiesector bestaan vandaag al verschillende hefboomen die organisaties ondersteunen bij het uitvoeren van kwaliteitsvolle evaluaties. Zo biedt de **werkgroep evaluatoren van Gezond Leven** een platform waar medewerkers methodologische en praktische uitdagingen kunnen bespreken en expertise kunnen uitwisselen. Instrumenten zoals het **Gedragswiel**, de **leidraad voor evaluatie en impactmeting** en het **RE-AIM-sjabloon** helpen organisaties om hun evaluatie-inspanningen van bij de ontwikkeling van een interventie gestructureerd aan te pakken. Daarnaast leveren grote databronnen, zoals de **Gezondheidsenquête** en de **Preventiebarometer**, waardevolle populatiegegevens, al is verdere afstemming met partnerorganisaties wenselijk. Tot slot zou ook **BELRAI**, wanneer systematisch toegepast in diverse settings, een rijke bron van relevante data kunnen vormen. Samen bieden deze bestaande elementen een basis waarop verdere evaluatie-inspanningen in de sector kunnen voortbouwen.

1.1.3 Evaluatiekaders

1.1.3.1 *Gebruik van verandertheorie (Theory of Change)*

Binnen de preventiesector wordt Theory of Change erkend als een waardevol instrument om de logica achter interventies zichtbaar te maken, maar het gebruik ervan en de expertise hierover verschillen sterk tussen organisaties. Vaak wordt een Theory of Change vooral ingezet tijdens de ontwikkeling van een interventie, terwijl de toepassing ervan in evaluaties minder ingeburgerd is. Preventiemedewerkers

geven aan dat **basiskennis van Theory of Change noodzakelijk** is om interventies beter te kunnen plannen en monitoren, maar dat de methodiek complex is en daarom baat heeft bij **externe ondersteuning**. Sommige organisaties, zoals VAD, het Expertisecentrum Val- en Fractuurpreventie Vlaanderen en Gezonde Mond, gebruiken het **Intervention Mapping-protocol**. In deze aanpak wordt teruggegrepen naar determinanten van gedrag, contextfactoren en theoretische aannames, wat helpt om zowel de verwachte werking van een interventie als de evaluatie hiervan beter te structureren.

1.1.3.2 *Gebruik evaluatiekaders*

Evaluatiekaders worden binnen de preventiesector algemeen erkend als nuttige hulpmiddelen om evaluatie-inspanningen te structureren, maar hun toepassing vraagt tijd, expertise en verdieping die in veel organisaties vandaag beperkt beschikbaar zijn. De keuze voor een kader hangt sterk af van het type interventie, de beschikbare competenties en de fase van het project. Kaders zoals **RE-AIM** worden gebruikt voor een globaal overzicht: dit is prima voor de evaluatie van eenvoudige acties, maar minder geschikt voor complexe interventies. **CFIR (Consolidated Framework for Implementation Research)** biedt meer diepgang maar wordt als complex ervaren, en het meer compacte **MIDI-instrument (Meetinstrument Determinanten van Innovaties)** wordt door sommige organisaties gebruikt als een haalbaar alternatief om determinanten van implementatie in kaart te brengen. Werkgroepleden benadrukken dat **één uniform kader niet haalbaar of wenselijk** is: eenvoudige acties vragen een andere aanpak dan complexe interventies, en een vast kader kan onnodige registratielast veroorzaken. Belangrijk is dat evaluatiekaders voldoende flexibiliteit bieden om, afhankelijk van context en beschikbare evidentie, de focus te leggen op implementatie, effect of een combinatie van beide. Een evaluatiekader kan beiden bevatten, maar de focus hoeft niet op beiden te liggen. Bijvoorbeeld, wanneer er veel internationale evidentie voor een methodiek bestaat, ligt de focus eerder op het evalueren van de implementatie in de huidige context. Wanneer er weinig evidentie voorhanden is, wordt effectiviteit belangrijker.

1.1.4 *Onderzoeksdesigns*

1.1.4.1 *Huidige ervaringen en visie op gecontroleerd en/of gerandomiseerd onderzoek*

Binnen de werkgroep bestaat brede consensus dat gecontroleerd onderzoek waardevol kan zijn, maar alleen wanneer het **haalbaar, ethisch verantwoord en passend bij de interventiecontext** is. In het bijzonder voor grootschalige interventies mag men wel meer gecontroleerd onderzoek verwachten. Veel partnerorganisaties beschikken bovendien niet over de nodige wetenschappelijke expertise, waardoor samenwerking met onderzoeksinstituten noodzakelijk blijft. Zelfs wanneer expertise aanwezig is, botsen organisaties op praktische uitdagingen zoals beperkte middelen, tijdsintensieve procedures en moeilijkheden om voldoende grote steekproeven te werven. Ook **randomisatie** blijkt in realistische preventiesettings vaak niet haalbaar, bijvoorbeeld wanneer scholen of organisaties zelf willen kiezen of ze deelnemen aan een interventie. Daarnaast spelen **ethische bezorgdheden**: het is niet altijd verantwoord om deelnemers een interventie te onthouden, en alternatieve designs zoals stepped-wedge of uitgestelde deelname zijn niet altijd toepasbaar, zeker niet bij langetermijnonderzoek. Tot slot benadrukken deelnemers dat niet elke interventie een streng onderzoeksdesign vereist en dat een goed uitgevoerde quasi-experimentele studie vaak waardevoller is dan een zwakke RCT. Wat vooral nodig is, is een **duidelijk overzicht van mogelijke designs**, zodat organisaties en onderzoekers samen een realistische en passende aanpak kunnen kiezen.

1.1.4.2 *Indicatoren*

Binnen de werkgroep leefde brede consensus dat **vaste thematische effectindicatoren niet haalbaar of wenselijk** zijn. Door de grote variatie in doelgroepen, settings en interventiedoelen moeten organisaties zelf indicatoren kunnen kiezen die aansluiten bij hun specifieke context. Een uniforme lijst zou in de praktijk te weinig toepasbaar zijn en eerder registratielast dan bruikbare inzichten opleveren. Wel werd voorgesteld om belangrijke indicatoren **per doelgroep of setting** naar voren te schuiven, omdat dit geïntegreerde en samenhangende acties stimuleert en samenwerking tussen partners bevordert. Daarnaast vinden deelnemers het belangrijk dat organisaties voldoende ruimte hebben om in **dialog met het Departement Zorg** passende effectindicatoren en meetinstrumenten te selecteren, zodat de gekozen indicatoren zowel beleidsrelevant als uitvoerbaar zijn.

1.1.5 Samenwerking en rolverdeling

1.1.5.1 *Huidige samenwerkingen tussen de preventiesector en onderzoeksinstituten*

De samenwerking tussen preventieorganisaties en onderzoeksinstituten kent vandaag een zeer **uiteenlopend karakter**. Organisaties die ingebed zijn in een universitaire omgeving beschikken duidelijk over meer mogelijkheden: zij kunnen rekenen op onderzoekers, doctoraatsstudenten en masterstudenten om evaluaties uit te voeren, zoals bij Gezonde Mond, VLESP en het Expertisecentrum Val- en Fractuurpreventie Vlaanderen. Voor andere organisaties verloopt samenwerking echter vaker ad hoc en is het sterk afhankelijk van persoonlijke netwerken. Uitbesteding vormt een drempel, gezien de nodige budgetten en overheadkosten, terwijl formele oproepen of Europese projecten voor veel partnerorganisaties moeilijk toegankelijk zijn door tijdsdruk, beperkte capaciteit en gebrek aan informatie. Masterproeven worden regelmatig ingezet, maar zijn vooral geschikt voor kleinschalige evaluaties, omdat de kwaliteit en timing niet altijd gegarandeerd zijn. Hierdoor ontstaat een **ongelijk speelveld** waarin sommige organisaties structureel toegang hebben tot onderzoek en anderen vooral op eigen kracht of minder structurele samenwerkingskansen zijn aangewezen.

1.1.5.2 *Stimuleren van samenwerking door het Departement Zorg*

De werkgroep benadrukte dat het Departement Zorg een belangrijke rol kan spelen in het versterken van samenwerking tussen preventieorganisaties en onderzoeksinstituten. Een vaak genoemd voorstel is de oprichting van een **onafhankelijk evaluatiefonds voor preventie** waarop organisaties samen met onderzoekers kunnen intekenen, zodat kwaliteitsvolle effectevaluaties niet langer volledig op de schouders van partnerorganisaties terechtkomen. Daarnaast kan het Departement samenwerking vergemakkelijken door een **gebundeld budget** te voorzien voor organisaties die geen universitaire inbedding hebben, door **gerichte oproepen** te lanceren waarop onderzoekers kunnen intekenen, en door een **overzicht van onderzoeksgroepen met expertise** per thema beschikbaar te stellen. Ook **co-financieringstrajecten**, waarbij partnerorganisaties en universiteiten gezamenlijk onderzoeksprojecten opzetten, kunnen samenwerking stimuleren. Ten slotte kan het Departement bijdragen door **onderzoeksprioriteiten en een gedeelde onderzoeksagenda** mee te bepalen, in samenwerking met de PO's en eventueel via de netwerkgroepen. Thema-overschrijdend onderzoek en gezamenlijke focus op kwetsbare doelgroepen kunnen hieruit voortvloeien.

1.1.5.3 Oefening rolverdeling

Een oefening met de werkgroepleden rond rolverdeling maakte duidelijk dat een strikte taakafbakening van rollen in de praktijk niet werkbaar is, omdat preventieorganisaties en interventies sterk van elkaar verschillen en personen vaak meerdere rollen combineren. Toch kwamen enkele algemene tendensen naar voren. **Gedeelde verantwoordelijkheid** tussen partnerorganisaties en onderzoekers is wenselijk voor het uitwerken van een Theory of Change, het bepalen van de evaluatiefocus en het interpreteren en verspreiden van de resultaten. **Partnerorganisaties** nemen doorgaans het voortouw in het betrekken van stakeholders en in het benutten van de bevindingen om interventies bij te sturen. **Onderzoekers** hebben een leidende rol in methodologische aspecten, zoals de keuze van het onderzoeksdesign en het uitvoeren van de data-analyse. **Terreinorganisaties** leveren tot slot cruciale praktijkkennis en faciliteren vaak de dataverzameling, onder meer door doelgroepen te bereiken en inzichten te geven in implementatie en context.

1.2 Werkgroep actoren uit onderzoek en beleid

De eerste bijeenkomst met de werkgroep met actoren uit onderzoek en beleid vond plaats op dinsdag 16 september 2025 van 12u tot 16u30 in het VAC Mechelen. In totaal namen 12 deelnemers deel.

1.2.1 Huidige ervaringen en percepties inzake evaluatie

De werkgroep onderzoek en beleid gaf aan dat de huidige evaluatiepraktijk in de preventiesector sterk verschilt en dat **effectevaluatie vaak onderbelicht** blijft. Processen zoals bereik worden doorgaans goed opgevolgd, maar het aantonen van effecten is moeilijk door complexe ketens van betrokken organisaties, afhankelijkheid van lokale implementatie en het ontbreken van robuuste designs. Daarbij worden nog regelmatig interventies ingezet zonder bewezen effectiviteit, of zelfs met aanwijzingen van ineffectiviteit. **Onderzoekers vervullen een belangrijke rol** in dit landschap: ze ondersteunen organisaties met methodologische expertise, ontwikkelen soms samen met hen een Theory of Change en voeren uiteenlopende evaluaties uit, variërend van proces- en effectevaluaties tot participatieve en theoretische evaluaties zoals realist evaluation. Onderzoeksinstellingen krijgen regelmatig de vraag van organisaties binnen en buiten de preventiesector (PO's, zorg- en welzijnsorganisaties, social profit sector, steden en gemeenten...) voor dienstverlening of contractonderzoek inzake evaluatie. De concrete vragen en invullingen van de dienstverlening gaan van methodologische en statistische ondersteuning tot meer uitgebreide evaluaties. Samenwerking kan organisaties ook versterken: door samen te evalueren, kunnen organisaties later zelfstandig aan de slag om interventies met specifieke onderzoeksmethoden te evalueren.

1.2.2 Randvoorwaarden en hefboomen voor kwaliteitsvolle evaluatie

1.2.2.1 Prioritaire randvoorwaarden (prioriteringsoefening)

De werkgroep benadrukte dat kwaliteitsvolle evaluatie alleen mogelijk is wanneer enkele cruciale randvoorwaarden vervuld zijn. Zo moet **voldoende tijd** worden voorzien binnen projectplanning, zowel voor voorbereiding (zoals ethische goedkeuring en rekrutering) als voor langdurige opvolging van effecten. Een grote uitdaging blijft de **financiering**: middelen voor evaluatie zijn schaars, waarbij bestaande methodieken soms onvoldoende worden onderbouwd en het ontwikkelen van nieuwe

interventies de voorkeur krijgt. Een richtlijn voor financiering van evaluatie (bv. 10% van het projectbudget voor evaluatie) kan helpend zijn, maar belangrijker is dat organisaties vooraf goed inschatten welke evaluatie nodig is en welke middelen daarvoor vereist zijn. De vereiste middelen variëren: soms is een grondige evaluatie pas uitgevoerd, soms is een grondige evaluatie juist essentieel. **Interne expertise**, bijvoorbeeld via een wetenschappelijk medewerker, is een belangrijke hefboom, maar samenwerking met externe onderzoekers blijft belangrijk om onafhankelijkheid en neutraliteit te waarborgen. Ook binnen het Departement Zorg is er interne expertise nodig om de evidentie van interventies te kunnen beoordelen en organisaties te kunnen ondersteunen.

Evaluatie vraagt bovendien om **vertrouwen en draagvlak**, aangezien organisaties soms onzeker zijn of evaluatie bedoeld is om uit te leren dan wel voor verantwoording. De werkgroep pleit eveneens voor erkenning van **alternatieve en flexibele evaluatiemethoden**. Meetinstrumenten die opgebouwd zijn vanuit ervaringskennis en dus aansluiten bij wat de doelgroep zelf belangrijk vindt zijn een meerwaarde. Ook de ontwikkeling van **passende indicatoren** en een betere monitoringinfrastructuur zijn noodzakelijke voorwaarden. Ten slotte werd het idee van een **'Steunpunt Evalueren'** naar voren geschoven, dat expertise bundelt en organisaties ondersteunt met advies over onderzoeksdesign en evaluatiekosten, matchmaking tussen organisaties en onderzoekers, administratieve ondersteuning en bruikbare sjablonen.

1.2.2.2 *Bestaande hefbomen voor evaluatie*

Binnen de werkgroep werd benadrukt dat er vandaag al verschillende waardevolle hefbomen bestaan die organisaties ondersteunen bij hun evaluatiepraktijk. Zo kunnen onderzoekers terugvallen op **statistische expertise** uit onderzoekinstellingen, bijvoorbeeld voor advies over geschikte onderzoeksdesigns, analyses of het bepalen van de benodigde steekproefgrootte. Binnen **Gezond Leven** speelt een **evaluatorenoverleg** een belangrijke rol: in elk team is iemand verantwoordelijk voor evaluatie, en via regelmatige bijeenkomsten worden casussen besproken, noden gedeeld en wordt evaluatie als thema binnen de organisatie versterkt. Tot slot werd ook verwezen naar het belang van **lerende netwerken**, zoals die in de social-profitsector al bestaan. Dergelijke netwerken, waarin organisaties van elkaar leren en ervaringen uitwisselen, zouden ook in de preventiesector een belangrijke meerwaarde kunnen bieden om evaluatiepraktijken te versterken.

1.2.3 Evaluatiekaders

1.2.3.1 *Gebruik van verandertheorie (Theory of Change)*

Het opstellen van een Theory of Change wordt door de werkgroep beschouwd als een **waardevolle maar veeleisende** oefening die zorgt voor transparantie en continuïteit. Organisaties hanteren verschillende benaderingen: sommigen reconstrueren eerst de praktijk om onderliggende aannames zichtbaar te maken, terwijl anderen vertrekken vanuit de projectopzet en die vertalen naar een eerste Theory of Change. Omdat deze oefening vaak een helikopterperspectief en explicitering van vanzelfsprekende processen vraagt, is **externe ondersteuning** voor veel organisaties essentieel. Een Theory of Change kan ook opgebouwd worden op basis van een bevraging over determinanten bij de doelgroep. Gezond Leven tracht de Theory of Change inmiddels systematisch toe te passen op methodieken om werking en beoogde impact transparanter te maken. Dit moet niet alleen leiden tot meer helderheid voor aanbieders, maar ook tot betere evaluaties doordat **duidelijker wordt wat gemeten moet worden en waarom**.

1.2.3.2 *Gebruik van evaluatiekaders*

De werkgroep benadrukte dat evaluatiekaders een **waardevol hulpmiddel** zijn om evaluaties systematisch en doordacht aan te pakken, maar dat de **keuze en toepassing ervan sterk afhangen van de context, beschikbare middelen en doelstellingen** van een interventie. Idealiter besteden kaders aandacht aan zowel implementatie als effect, waarbij implementatie meer omvat dan louter uitvoering en ook inzicht vraagt in onderliggende mechanismen. Tegelijk werd benadrukt dat het vertalen van een kader naar concrete methoden en meetinstrumenten vaak moeilijk is en ondersteuning vereist. In de discussie rond consistentie versus flexibiliteit erkenden deelnemers de waarde van een gedeeld kader of een kernset van indicatoren voor beleidsdoeleinden, maar benadrukten ze dat **maatwerk essentieel blijft** voor interventieverbetering. Een haalbare middenweg is **werken met een aantal vaste principes**, zoals het gebruik van een Theory of Change, een passend evaluatiekader en aandacht voor meerdere perspectieven, inclusief ervaringskennis, gecombineerd met voldoende flexibiliteit. Wel werd benadrukt dat bij effectiviteitsonderzoek bepaalde methodologische eisen onvermijdelijk zijn, wat betekent dat voldoende financiering en ondersteuning cruciaal blijven.

1.2.4 Onderzoeksdiseins en effectindicatoren

1.2.4.1 *Huidige ervaringen en visie op gecontroleerd en/of gerandomiseerd onderzoek*

De werkgroep bevestigde dat de keuze van een onderzoeksdisein altijd afhankelijk is van de aard van de interventie, de context en de beschikbare middelen. Hoewel **gecontroleerd of gerandomiseerd onderzoek** in theorie wenselijk is om interne validiteit te versterken, blijft het in praktijkgerichte preventiesettings **vaak moeilijk haalbaar** door beperkte expertise, tijd, budgetten en ethische bezwaren. Voor veel organisaties is randomisatie bijvoorbeeld niet mogelijk omdat intermediairs of settings zelf willen bepalen of ze de interventie ontvangen. Daarbij vraagt gecontroleerd onderzoek goede methodologische ondersteuning, grote steekproeven, goede matching tussen groepen en langdurige opvolging, wat in de preventiesector vaak niet realistisch is. Deelnemers benadrukten dat **alternatieve designs** (zoals sense-making methoden, dagboekstudies of experience sampling) soms beter aansluiten bij de doelgroep en de context. Een gefaseerde aanpak (bv. eerst piloot met controlegroep, daarna bredere uitrol) kan zinvol zijn, zeker bij interventies met een lanceringsmoment (bv. Tournée Minerale). Ook blijft het onderscheid tussen **effect** en **impact** belangrijk: een interventie kan effectief zijn zonder brede maatschappelijke impact. Tot slot werd gewezen op bijkomende aandachtspunten zoals **natuurlijke trends op populatieniveau, betekenisvolle uitval** van deelnemers uit een studie en **selectie-effecten** waarbij bepaalde groepen niet worden bereikt met een interventie. Naast het meten van effecten en monitoren van impact dient er ook aandacht te zijn voor **internationale evidentie en standaarden** (bv. Internationale standaarden drugpreventie), zodat de kwaliteit van het preventief beleid voor een bepaald thema in kaart kan worden gebracht. Tot slot is er ook nood aan **centrale expertise binnen de overheid**, zodat organisaties ondersteund kunnen worden bij het kiezen van haalbare en kwaliteitsvolle designs.

1.2.4.2 *Indicatoren*

De werkgroep benadrukte voldoende **ruimte voor flexibiliteit en maatwerk** mogelijk moet zijn. Effectcijfers kunnen immers misleidend zijn wanneer ze losgekoppeld worden van het onderzoeksdisein of wanneer kleine steekproeven of selectie-effecten de resultaten vertekenen. Deelnemers pleiten daarom voor **meer duidelijkheid vanuit de overheid over welke uitkomsten belangrijk worden gevonden** en welke indicatoren als richtinggevend kunnen dienen. Tegelijk moet worden vermeden dat

organisaties verplicht worden om met vaste indicatoren te werken. Een haalbare tussenweg is het definiëren van **types indicatoren** (zoals gedragsverandering, kennis of motivatie) of het werken met **uniforme meetinstrumenten** waar dat zinvol en uitvoerbaar is. Voor procesindicatoren lijkt het wel mogelijk te streven naar een beperkte set van minimale indicatoren die voor alle projecten relevant zijn, aangevuld met contextspecifieke elementen. Het blijft bovendien een uitdaging om interventie-uitkomsten te vertalen naar de gezondheidsdoelstellingen en inzichten op beleidsniveau, wat interpretatie en duiding vraagt.

1.2.5 Samenwerking en rolverdeling

Binnen de werkgroep werd duidelijk dat samenwerking tussen PO's en onderzoekers vandaag zeer **uiteenlopend verloopt, vaak op ad-hoc basis** en sterk afhankelijk van bestaande contacten. Sommige onderzoeken worden opgezet doordat onderzoekers zelf organisaties benaderen, terwijl anderen gebruikmaken van masterproeven, al blijft dit risicovol door wisselende kwaliteit en timing. De groep wees ook op **structurele hindernissen**: bestaande financieringslijnen binnen de onderzoekswereld sluiten niet altijd aan bij de noden van de overheid, omdat fondsen zoals het FWO zich vooral richten op innovatie. Om samenwerking te versterken ziet de werkgroep vooral potentieel in **gerichte financieringsmechanismen**, zoals een evaluatiefonds waarin partnerorganisaties en onderzoekers gezamenlijk projecten kunnen indienen. Het zou tot slot goed zijn om ziektelast te vertalen naar prioriteiten voor interventies, om een mismatch te vermijden tussen wat beleidsmatig belangrijk is en wat er op het terrein gebeurt.

1.2.5.1 Oefening rond rolverdeling

De oefening rond rolverdeling maakte duidelijk dat een strikte afbakening van taken in de praktijk niet haalbaar is, omdat interventies en organisaties sterk van elkaar verschillen en medewerkers vaak meerdere rollen combineren. Toch kwamen enkele algemene tendensen naar voren. PO's nemen vooral het voortouw in het **betrekken van stakeholders**, het **opstellen van de Theory of Change** en het **benutten van resultaten** om de interventie te verbeteren. Onderzoekers hebben een leidende rol in de **methodologische aspecten**, zoals het bepalen van het onderzoeksdesign en het uitvoeren van de analyses. Terreinorganisaties spelen vaak een belangrijke rol bij de **dataverzameling** en het vertalen van de praktijkcontext naar het onderzoek. De focus van de evaluatie, zoals het bepalen van relevante uitkomstindicatoren, is doorgaans een **gedeelde verantwoordelijkheid** tussen partnerorganisaties en onderzoekers, met een belangrijke rol voor het Departement Zorg om richting te geven.

2 Bijeenkomst 2

De tweede bijeenkomst zoemde in op wanneer een effectevaluatie aan de orde is, hoe de financiering idealiter gebeurt en op welke manier evaluatieresultaten beter benut kunnen worden. Daarnaast focuste de bespreking op hoe men meer inzicht kan verwerven in beleidsimpact en of de oprichting van een centraal aanspreekpunt voor evaluatie binnen de preventiesector een meerwaarde zou zijn.

2.1 Werkgroep met actoren uit de preventiesector

De tweede bijeenkomst met de werkgroep met actoren uit de preventiesector vond plaats op donderdag 6 november 2025 van 12u tot 16u30 in het VAC Mechelen en telde 12 deelnemers.

2.1.1 Nood en meerwaarde van effectevaluatie

2.1.1.1 *Wanneer is effectevaluatie nodig?*

Effectevaluatie dient **doelgericht** te zijn en is afhankelijk van de aard van de interventie. Idealiter wordt al vóór de implementatie nagedacht over meetbaarheid, en worden effecten op meerdere momenten opgevolgd zodat een doorlopende evaluatiecyclus ontstaat. Het **vooraf bepalen van duidelijke en realistische evaluatiedoelen** wordt gezien als een randvoorwaarde. Zonder vooraf vastgelegde doelen heeft een effectevaluatie weinig zin. Voor het formuleren van doelen is soms al verkennend onderzoek nodig, bijvoorbeeld via literatuurstudie of een pilootstudie.

Effecten kunnen gemeten worden op verschillende niveaus: bv. op gezondheidsniveau, gedragsniveau en niveau van intermediairs. Dit hangt af van de aard en doelstellingen van de interventie. Bij een interventie gericht op intermediairs is het in eerste instantie haalbaar en nodig om effecten bij deze groep te meten. Het meten van effecten bij de einddoelgroep is een meerwaarde, maar complexer en niet altijd haalbaar.

De noodzaak van effectevaluatie hangt ook samen met de **kwaliteit en relevantie van bestaand onderzoek**: sterke internationale evidentie kan richting geven, maar wat in het buitenland werkt, werkt niet automatisch in Vlaanderen. Verschillen in beleid, regelgeving, zorgorganisatie, training en beschikbaarheid van intermediairs maken een vertaalslag nodig. Dit geldt zowel voor de haalbaarheid van implementatie als culturele verschillen die de effectiviteit kunnen beïnvloeden. Vele interventies kunnen gedeeltelijk worden overgenomen met aanpassingen aan de lokale context, maar in dat geval blijft een nieuwe evaluatie belangrijk. Een interventie kan eveneens effectief zijn in een **pilootfase**, maar dat betekent niet automatisch dat ze implementeerbaar is op grote schaal. Pilootprojecten worden vaak uitgevoerd in zeer gemotiveerde settings, zoals scholen die al met het thema bezig zijn. Dit maakt resultaten niet altijd generaliseerbaar.

2.1.1.2 *Prioritering van interventies*

De prioriteringsoefening toont dat interventies vooral als belangrijk worden beschouwd wanneer ze **hoog inzetten op kosten-batenverhouding**, inspelen op **hiaten in de preventiesector**, **bewezen effectief zijn in andere contexten** of **veelbelovende resultaten** laten zien in kleinere studies. Ook interventies die **beleidsbeslissingen ondersteunen**, **beperkte bestaande evidentie** hebben, of bijdragen aan het **verminderen van gezondheidsongelijkheden** kregen veel stemmen. Verder worden interventies die aansluiten bij **gezondheidsdoelstellingen**, of zich richten op **problematieken met een hoge ziektelast** worden als prioritair gezien.

Ook **interventies met een kleiner bereik** verdienen een effectevaluatie. Interventies met een groot bereik zijn vaker universele interventies die dezelfde groep mensen bereiken, meestal diegenen die al ontvankelijk zijn voor preventie waardoor de impact soms kleiner is. Selectieve interventies richten zich op risicogroepen. Hoewel hun bereik daardoor kleiner is, kunnen ze een groter effect en rendement hebben en bijdragen aan het verminderen van gezondheidsongelijkheden.

Het uitgangspunt dat evaluatie vooral zinvol is wanneer het kan bijdragen aan beleidsbeslissingen of strategische keuzes', is problematisch als dit te strikt wordt opgevat. Dit kan leiden tot een cirkelredenering: wat geëvalueerd wordt, hangt dan volledig af van waar het beleid op dat moment over wil beslissen. Hierdoor komen thema's die niet in de huidige beleidsscope zitten moeilijker aan bod, zelfs als het werkveld dit waardevol vindt. Effectevaluatie moet vooral bijdragen aan **leren, verbeteren en onderbouwen**, binnen een realistisch kader en met duidelijke doelen.

2.1.2 Gebruik van evaluatieresultaten

De werkgroep benadrukte dat evaluatieresultaten pas echt waardevol zijn wanneer ze actief worden gebruikt voor leren en verbeteren, zowel binnen organisaties als in beleid. Hoewel aanbevelingen vaak zorgvuldig worden geformuleerd, blijft de **opvolging** ervan regelmatig achterwege, waardoor inzichten het risico lopen verloren te gaan of "in de schuif te verdwijnen". Een vroege dialoog met de overheid over hoe resultaten beleid en praktijk kunnen beïnvloeden, wordt gezien als cruciaal.

Ook **ongunstige resultaten** verdienen aandacht: ze bieden belangrijke informatie over waarom een interventie niet werkt en kunnen leiden tot betere implementatiestrategieën. Tegelijk mag een interventie niet eindeloos worden verdergezet wanneer herhaalde evaluaties geen effect aantonen, en moeten schadelijke interventies onmiddellijk worden stopgezet. Een interventie staat daarbij nooit op zichzelf: een interventie bereikt soms niet het gewenste effect omwille van bredere sociale en flankerende factoren, zoals sociale normen of een gebrek aan maatschappelijk draagvlak. In dat geval volstaat bijsturing niet, maar is er nood aan een bredere investering in onderliggende factoren.

Beleidsplannen moeten ruimte laten voor aanpassing op basis van voortschrijdend inzicht. In beheers-overeenkomsten is dit een structurele uitdaging: plannen liggen lang op voorhand vast, waardoor men minder snel kan inspelen op nieuwe inzichten. Wanneer een interventie goed werkt en men wil opschalen, moet dit vertaald kunnen worden naar jaarplannen en verschuivingen in acties (bv. een actie schrappen en meer inzetten op een andere actie). Het is ook een knelpunt dat tijdelijke projectmiddelen vaak niet worden verduurzaamd. Interventies die goed werken, moeten duurzaam gefinancierd en uitgevoerd kunnen worden.

Tot slot benadrukten deelnemers dat het identificeren van "effectieve interventies" zorgvuldigheid vereist: RCT-bewijs mag niet de enige maatstaf zijn, internationale evidentie moet steeds contextueel geïnterpreteerd worden, en er moet naast bewezen interventies ruimte blijven voor veelbelovende innovaties. Onderzoek naar **werkzame componenten** van programma's is daarbij essentieel, zeker omdat in de praktijk vaak "light versies" van interventiepakketten worden gebruikt.

2.1.3 Leercyclus

Een **lerend netwerk** heeft volgens de werkgroep enkel een meerwaarde wanneer er duidelijke doelstellingen, goede coördinatie en voldoende interne expertise bij de deelnemers aanwezig is. Eerdere ervaringen, zoals de vroegere werkgroep impactmeting, tonen dat een netwerk pas werkt wanneer er voldoende engagement is en deelnemers vergelijkbare rollen en competenties hebben. Hoewel organisaties vandaag vaak inspiratie zoeken in internationale voorbeelden, werd het als waardevoller beschouwd om ook te leren van **andere Vlaamse organisaties**, via uitwisseling rond gedeelde vragen en praktijken.

Het lerend netwerk is nu al deels aanwezig via **netwerkgroepen en sectoroverleg van Gezond Leven**. Het is interessanter om het thema 'evaluatie' hierin te versterken in plaats van nieuwe overleg-

structuren op te zetten. Afhankelijk van de noden in het werkveld of vragen van partners kan dan worden ingezoomd op specifieke thema's. Andere medewerkers of externe experts kunnen aansluiten wanneer relevant. Aanvullende vormen van kennisdeling, zoals een symposium of congres, werden interessant bevonden, al werd een digitaal platform als minder prioritair ingeschat vanwege de hoge onderhoudslast en kosten.

2.1.4 Beleidsimpact

De werkgroep benadrukte dat **populatiedata** een cruciale basis vormen voor het inschatten van beleidsimpact, met de Gezondheidsenquête en de Preventiebarometer als belangrijkste referentiepunten. Tegelijk blijft de representativiteit van doelgroepen een vraagteken, wat de interpretatie van effecten op populatieniveau bemoeilijkt.

Hoewel een **gedeeld uitkomstenkader** voor interventies op het eerste gezicht aantrekkelijk lijkt, achten de deelnemers dit vooralsnog **niet haalbaar** gezien de grote diversiteit in de preventiesector. Een **gedeeld kader voor implementatie-uitkomsten**, bijvoorbeeld gebaseerd op het RE-AIM-model, lijkt op termijn haalbaarder en relevanter. Dit kan inzicht bieden in het bereik van de sector als geheel (bv. welke doelgroepen worden onvoldoende bereikt?). Verder werd benadrukt dat **systeendenken noodzakelijk** is bij het begrijpen van beleidsimpact: interventies staan nooit op zichzelf, maar worden beïnvloed door bredere sociale, economische en organisatorische factoren, inclusief federale regelgeving en flankerende acties in andere domeinen.

2.1.5 Financiering van effectevaluatie

Nadenken over financieringsmechanismen voor effectevaluatie is volgens de werkgroep nauwelijks mogelijk zonder duidelijkheid over beschikbare middelen: zonder extra budget leidt elke keuze tot verschuivingen die andere kerntaken onder druk zetten. Ook middelen poolen wordt niet als realistisch gezien. Een percentage vrijhouden voor evaluatie lijkt dan het meest aangewezen. Een percentage voor evaluatie biedt dan de meeste zekerheid en autonomie, terwijl bijkomende middelen ruimte creëren voor grotere effectstudies en samenwerking met onderzoekers.

Een **evaluatiefonds** kan de kwaliteit van effectevaluaties versterken en samenwerking stimuleren, maar vraagt extra organisatie, governance en middelen, en biedt geen garantie dat ingediende voorstellen geselecteerd worden. **Gerichte oproepen** naar onderzoekers zijn eenvoudiger, maar zijn niet vernieuwend: uitbesteding is nu ook al mogelijk, maar alleen haalbaar wanneer organisaties intern over voldoende middelen beschikken. Een **vast evaluatiebudget binnen de eigen werking**, wat interne expertise vereist, is vooral voor kleinere organisaties moeilijk haalbaar. Een **tweesporenbeleid**, waarbij interne middelen dienen voor de permanente evaluatie en kleinere effectmetingen, en externe middelen worden ingezet voor robuust effectonderzoek, werd gezien als een veelbelovende middenweg. Tegelijk werd benadrukt dat nauwe samenwerking tussen organisaties en onderzoekers grote voordelen biedt, maar enkel werkt wanneer onafhankelijkheid gewaarborgd blijft. Tot slot werd nagedacht over **alternatieve financieringsbronnen**, zoals een verschuiving van het curatie- naar het preventiebudget, belastingen op schadelijke producten en social impact bonds, al gaan sommige opties gepaard met ethische of praktische beperkingen.

2.1.6 Centraal aanspreekpunt, advies en ondersteuning

De behoefte aan een centraal aanspreekpunt voor evaluatie hangt volgens de werkgroep sterk af van de verwachtingen die aan organisaties worden opgelegd. Wanneer preventieorganisaties zelf verantwoordelijk worden voor het uitvoeren van effectevaluaties, ontstaat er duidelijk nood aan **advies, kwaliteitsbewaking en methodologische ondersteuning**. Als de sector daarentegen sterker inzet op structurele samenwerking met onderzoeksinstellingen, wordt die expertise al deels ingebed via onderzoekspartners en is een afzonderlijk aanspreekpunt minder noodzakelijk. Indien een centraal aanspreekpunt wordt ontwikkeld, verdient **inbedding in bestaande structuren** de voorkeur. Het opzetten van een centrale databank met evaluatie-instrumenten wordt niet prioritair geacht, omdat instrumentkeuze sterk afhangt van thema en context. De grootste meerwaarde ligt volgens de groep in **strategische functies**, zoals het ontwikkelen van een overkoepelende evaluatiestrategie en governance binnen de preventiesector, zodat de aanpak die in dit traject werd ontwikkeld ook in de toekomst duurzaam kan worden verankerd.

2.2 Werkgroep met actoren uit onderzoek en beleid

De tweede bijeenkomst met de werkgroep met actoren uit onderzoek en beleid verliep digitaal, vond plaats op donderdag 19 november 2025 en telde vijf deelnemers.

2.2.1 Algemene opmerkingen en bedenkingen

De werkgroep vindt dat de verwachtingen rond meer en robuustere effectevaluaties begrijpelijk zijn, maar dat dit in de huidige context voor veel organisaties nog een **zeer ambitieus** is. Er is nood aan een **groeipad en plan van aanpak**. Voor de overheid is het daarbij aangewezen om na te gaan of de beoogde stappen effectief gezet worden en de sector in de juiste richting evolueert, bijvoorbeeld aan de hand van een jaarlijkse evaluatie. Daarbij kan worden nagegaan in welke mate de sector vooruitgang boekt op verschillende gebieden zoals kennisdeling, effectmeting en het gebruik van resultaten. De werkgroep waarschuwde ook voor een te enge focus op het evalueren van interventies op zich en het de-contextualiseren van interventies. De impact van interventies hangt immers sterk af van **uitvoerende organisaties en intermediairs**, die mee bepalen of een interventie werkt.

Als inspiratie werd gewezen op de aanpak van de databank *Kwaliteitsvolle Praktijken* van Opgroeien, waar interventies worden beschreven en beoordeeld op kwaliteit in plaats van louter effectiviteit. Een gelijkaardige aanpak zou ook in de preventiesector waardevol kunnen zijn, bijvoorbeeld door op de website van preventiemethodieken systematisch een **Theory of Change** op te nemen die helder maakt hoe een methodiek geacht wordt te werken en welke doelen zij nastreeft.

Tot slot werd gewezen op een structureel probleem: de **return on investment van preventie** vloeit niet terug naar het preventiebudget, waardoor evaluaties niet kunnen worden gefinancierd via de winst die de sector oplevert.

2.2.2 Nood en meerwaarde van effectevaluatie

2.2.2.1 *Wanneer is effectevaluatie nodig?*

De werkgroep gaf aan dat effectevaluatie in de meeste situaties zinvol is, maar altijd afgestemd moet worden op de **doelen en noden van de evaluatie**. Bijvoorbeeld, wanneer sterke internationale evidentie al aantoont dat een interventie werkt, is het niet nodig om dezelfde uitkomsten opnieuw te meten, maar blijft het wel belangrijk te evalueren of en hoe de interventie functioneert in de Vlaamse context. Bovendien vragen interventies die vanuit het buitenland worden overgenomen doorgaans een **adaptatieproces**, waarna moet worden bekeken of een nieuwe evaluatie nodig is. Zo kan worden nagegaan of aanpassingen leiden tot veranderingen in haalbaarheid, werking of effect.

2.2.2.2 *Prioritering van interventies*

Bij de prioritering van interventies benadrukte de werkgroep dat nuance cruciaal is en dat kleinschalige initiatieven niet systematisch uit de boot mogen vallen. Effectevaluatie wordt waardevoller naarmate een interventie aan meer prioriteringscriteria voldoet, maar geen enkel criterium sluit evaluatie volledig uit. Interventies die inspelen op **gezondheidsongelijkheden** en interventies met **opschalingspotentieel** kregen de meeste stemmen, gevolgd door **veelbelovende interventies** uit pilootstudies en interventies die in andere contexten al **effectief bleken**. Ook interventies die aansluiten bij **gezondheidsdoelstellingen** en **beleidsrelevante inzichten** bieden werden als prioritair beschouwd.

2.2.3 Gebruik van evaluatieresultaten en leercyclus

Het **overkoepelend ontsluiten van bestaande evaluatiekennis** is volgens de werkgroep essentieel om dubbel werk te vermijden en beter gebruik te maken van beschikbare inzichten. Er bestaan al erg veel kennisreviews, richtlijnen en leidraden voor bepaalde thema's, welke een goed startpunt vormen. We moeten vaker over het muurtje kijken in plaats van telkens nieuwe interventies te ontwikkelen: welke aanpak werkt al en welke kernelementen zijn nodig, en kunnen we die (mits aanpassing) ook inzetten voor andere doelgroepen of in andere contexten?

Een inspirerend voorbeeld is de Nederlandse databank effectieve interventies in de jeugdhulp. Het Nederlands Jeugdinstituut is op een gegeven moment meer gaan focussen op het overkoepelend (niet enkel interventiegericht) ontsluiten van kennis en toegankelijk maken van wat al geweten is, o.a. op werkzame ingrediënten van interventies, literatuuronderzoeken en reviews...

Daarnaast pleiten deelnemers voor meer aandacht voor **de-implementatie**: interventies die herhaaldelijk geen effect aantonen of zelfs schadelijk blijken, moeten kunnen worden afgebouwd. Het is belangrijk vast te leggen welke maatregelen volgen wanneer interventies herhaaldelijk ineffectief blijken. Een belangrijke nuance is dat er ruimte voor innovatie moet blijven. Dat betekent dat ook veelbelovende interventies, waarvan de effectiviteit nog moet worden aangetoond, een kans moeten kunnen krijgen. In de Vlaamse preventiesector is dit echter niet vanzelfsprekend: subsidies worden doorgaans niet toegekend op het niveau van afzonderlijke interventies, maar maken deel uit van vijfjarige beheersovereenkomsten. Een meer logisch model zou zijn dat, in overleg met het departement, kan worden beslist om een interventie stop te zetten of uit te faseren, zodat er middelen vrijkomen voor nieuwe initiatieven. Stopzetting kan daarnaast ook om andere redenen nodig zijn, bijvoorbeeld door het stellen van prioriteiten. Soms zijn meerdere interventies wel effectief, maar moeten er toch keuzes worden gemaakt.

Daarbij is het cruciaal om **te begrijpen waarom interventies al dan niet werken**, en om de focus te leggen op leren en verbeteren. Een interventie met tegenvallende resultaten kan nog wel potentieel hebben mits aanpassingen. Dit moet duidelijk zijn voor organisaties, zodat zij geen terughoudendheid tegenover evaluatie ontwikkelen uit vrees voor negatieve resultaten.

Tot slot benadrukte de werkgroep dat structureel **overleg tussen onderzoekers, partnerorganisaties en het Departement Zorg** noodzakelijk is. Zonder vaste contactmomenten worden belangrijke beslissingen over de noodzaak, aanpak en opvolging van evaluaties te versnipperd genomen, waardoor waardevolle inzichten onvoldoende worden benut.

2.2.4 Beleidsimpact

De werkgroep benadrukte dat een effectieve beoordeling van beleidsimpact nood heeft aan een **duidelijke visie, een uitgewerkt plan en een systemische benadering**. Op clusterniveau is het waardevol om meer in te zetten op samenhangende evaluaties. Hiervoor bieden de nieuwe clusters in de gezondheidsdoelstellingen kansen om **geclusterde evaluaties** uit te voeren en samen te werken rond gedeelde indicatoren. Op sectorniveau is er behoefte aan een duidelijk plan voor de **evaluatie van de volledige gezondheidsdoelstelling**. Daarnaast wezen deelnemers op de noodzaak om het principe *Health in All Policies* sterker in de praktijk te brengen: beleidsimpact wordt immers beïnvloed door factoren uit diverse beleidsdomeinen, zoals welzijn, werk, cultuur en economie.

2.2.5 Financiering van effectevaluatie

De werkgroep waardeerde het bredere en creatievere denken over financiering van effectevaluatie, maar wees erop dat elk mechanisme pas zinvol is wanneer voldoende middelen beschikbaar zijn. Een **evaluatiefonds** kan kwaliteitsvolle evaluaties stimuleren en samenwerking met onderzoekers versterken, maar brengt ook risico's mee, zoals mogelijke concurrentie tussen organisaties en een ongelijk speelveld voor partnerorganisaties zonder onderzoeks-inbedding. Daarnaast vraagt zo'n fonds extra governance en administratie. Een inspirerend voorbeeld is het ZonMW uit Nederland: dit fonds financiert praktijkgericht en wetenschappelijk verantwoord onderzoek waarbij een organisatie in samenwerking met een onderzoekspartner een project kan indienen.

De groep benadrukte het belang van **onafhankelijkheid en kwaliteit van evaluatoren**, omdat organisaties die hun eigen interventies ontwikkelen tegelijk niet de meest geschikte partij zijn om hun effectiviteit te onderzoeken. Uitbesteding via onderzoeksoproepen kan werken, maar is momenteel vaak financieel onhaalbaar. Een werkbare tussenoplossing kan liggen in **meer transparantie in evaluatieplannen** en, op termijn, het stimuleren dat partnerorganisaties wetenschappelijk verankerd zijn via structurele samenwerkingen met onderzoeksinstellingen.

2.2.6 Centraal aanspreekpunt, advies en ondersteuning

De werkgroep benadrukte dat ondersteuningstaken die in dit traject werden opgenomen, idealiter structureel worden verankerd. Een centraal aanspreekpunt (bv. als steunpunt) kan daarbij een belangrijke rol spelen, op voorwaarde dat het wordt ondergebracht bij een **neutrale partij** die de preventiesector goed kent en verbinding heeft met zowel het Departement Zorg als relevante onderzoeksgroepen. Dit aanspreekpunt kan ook binnen een bestaande organisatie of binnen het departement worden gepositioneerd, maar moet voldoende tijd, mandaat en middelen krijgen om effectief te functioneren. Ook wanneer er sterk wordt ingezet op meer samenwerking met onderzoeks-

instellingen, blijft extra versterking nodig. Dergelijke samenwerking met onderzoekers vergt eveneens voorbereidende stappen, zoals bepalen wat moet worden geëvalueerd en wat hiervoor nodig is.

Tot slot onderstreepte de werkgroep het belang van **opleidingen** rond evaluatie- en implementatievaardigheden. Hoewel er vandaag al externe opleidingen bestaan, zou een aanbod dat specifiek is afgestemd op de preventiesector, en waar mogelijk wordt verzorgd door experts uit het veld, een belangrijke meerwaarde bieden.

Hoofdstuk 13

Leidraad voor effectevaluatie in de Vlaamse preventiesector

Dit laatste hoofdstuk presenteert de leidraad voor effectevaluatie in de Vlaamse preventiesector. Deze leidraad vormt het resultaat van een zorgvuldig proces waarin inzichten uit de literatuur (hoofdstukken 3 tot en met 11) en de rijke input van twee werkgroepen met actoren uit praktijk, beleid en onderzoek (hoofdstuk 12) werden samengebracht. Ze biedt een **realistisch, haalbaar en gefaseerd kader** dat rekening houdt met de uiteenlopende contexten binnen de sector en met de randvoorwaarden die in de werkgroepen naar voren kwamen. De leidraad is bedoeld als **praktisch instrument** voor zowel organisaties als beleidsmakers om gefundeerde keuzes te maken over wanneer, waarom en hoe effectevaluatie wordt uitgevoerd, met aandacht voor proportionaliteit, kwaliteitsborging en leerprocesgericht werken.

1 Waarom deze leidraad?

Hoewel effectevaluatie formeel verankerd is in beheersovereenkomsten en subsidieaanvragen, blijft het vinden van een werkbaar evenwicht tussen wetenschappelijke validiteit en praktische haalbaarheid een uitdaging in de preventiesector. In de praktijk ligt de nadruk meestal op procesindicatoren, zoals bereik en tevredenheid, terwijl effectevaluaties slechts beperkt plaatsvinden. Daarnaast wordt er vaak gebruik gemaakt van methoden met beperkte bewijskracht. Dit bemoeilijkt een gefundeerde beleidsmatige evaluatie en bijsturing van bestaande interventies.

Kwaliteitsvolle effectevaluatie is essentieel en mag niet worden beschouwd als een randactiviteit. Het vereist een bewuste en structurele investering in tijd, expertise en middelen. Dit vraagt om een grondige voorbereiding, doordachte methodologische keuzes en continue inspanningen tijdens het volledige traject van ontwikkeling tot implementatie en opschaling van de interventie. In een context van beperkte middelen is het belangrijk om efficiënt te werk te gaan en te streven naar de best mogelijke evidentie binnen wat haalbaar is.

Deze leidraad wil kwaliteitsvolle effectevaluatie een vaste plaats geven binnen het preventieve gezondheidsbeleid. Het ondersteunt beleidsmakers bij het creëren van de juiste randvoorwaarden voor organisaties in de preventiesector en helpt deze organisaties om op een onderbouwde en haalbare manier de effectiviteit van hun interventies te evalueren via een gefaseerde aanpak. Op termijn draagt dit bij aan beter gefundeerde beleidsbeslissingen en aan een sterker preventief beleid dat de gezondheid van de Vlaming bevordert.

2 Evaluatie als geïntegreerd geheel: effect, proces en impact

Effectevaluatie is cruciaal om beleid en praktijk op wetenschappelijk onderbouwde wijze te ondersteunen, maar kan niet los worden gezien van procesevaluatie en beleidsimpactevaluatie. Een interventie is overigens zelden meteen klaar voor een effectevaluatie: eerst moet duidelijk zijn of deze implementeerbaar is en uitgevoerd wordt zoals bedoeld. Daarom is een geïntegreerde evaluatieaanpak nodig.

Effectevaluatie is een weldoordachte afweging

Een interventie is effectief als deze aantoonbaar bijdraagt aan het behalen van de beoogde doelstellingen. In de context van gezondheidsbevordering betekent dit dat de interventie leidt tot gunstige veranderingen in gedrag, gezondheid of relevante determinanten van gedrag of gezondheid bij de doelgroep(en). Het geeft antwoord op de vraag: *werkt de interventie, en in welke mate zijn effecten eraan toe te schrijven?* Daarbij is het van belang dat de waargenomen effecten overtuigend kunnen worden toegeschreven aan de interventie zelf, en niet aan externe factoren of toevallige veranderingen. Dit vraagt om robuust en zorgvuldig opgezet onderzoek, bijvoorbeeld met een geschikte controlegroep, herhaalde metingen of andere methoden die vertekening beperken.

Effectevaluatie kan op verschillende niveaus worden uitgevoerd, afhankelijk van de interventie- en evaluatiedoelen. De nadruk ligt doorgaans op de einddoelgroep, gezien daar de beoogde middellange en langetermijneffecten tot uiting moeten komen. Veranderingen op het niveau van intermediairs vormen echter vaak een noodzakelijke schakel in het realiseren van effecten bij de einddoelgroep. Daarom kan effectmeting zich in een eerste fase richten op intermediairs. Daarnaast kan gekozen worden voor een gelaagde effectevaluatie, waarbij zowel veranderingen bij intermediairs als bij de einddoelgroep systematisch in kaart worden gebracht.

Goed uitgevoerd effectiviteitsonderzoek is essentieel voor het nemen van onderbouwde beslissingen over de voortzetting, opschaling of afbouw van een interventie. Tegelijk vraagt effectevaluatie een aanzienlijke investering in tijd, expertise en middelen. Daarom is het belangrijk om steeds een bewuste afweging te maken van de verwachte baten van de meerwaarde die de evaluatie oplevert tegenover de kosten die ermee gepaard gaan.

Procesevaluatie als voorbereiding én aanvulling

Procesevaluatie gaat vaak vooraf aan effectevaluatie en blijft ook tijdens en na de effectevaluatie relevant. Procesevaluatie onderzoekt doelgericht de uitvoerbaarheid en processen van een interventie in de praktijk: de implementatie van de interventie, het werkingsproces en de invloed van de context.

In een vroege fase biedt procesevaluatie vooral inzicht in de haalbaarheid en uitvoering van de interventie: Wat maakt dat de interventie al dan niet haalbaar is? Wordt de interventie uitgevoerd zoals bedoeld? Bereikt het de juiste doelgroepen? Begrijpen uitvoerders de werkwijze goed? Welke drempels zijn er voor uitvoering of deelname, op niveau van intermediairs en/of de einddoelgroep? Welke contextfactoren beïnvloeden de uitvoering?

Bij een stabiele en brede implementatie verschuift de focus van procesevaluatie naar monitoring, kwaliteitsbehoud, structurele inbedding en duurzaamheid. Tijdens een effectevaluatie helpt de procesevaluatie verklaren hoe en waarom effecten optreden of uitblijven: Zijn er verschillen in uitvoering tussen locaties of uitvoerders die verschillen in effect verklaren? Via welke processen of werkzame elementen leidt de interventie tot effect? Door proces- en effectevaluatie te combineren, leren we wat werkt, voor wie en onder welke omstandigheden. Procesevaluatie levert bovendien input voor bijsturing tijdens de implementatie.

Beleidsimpactevaluatie: verder kijken dan directe effecten

Impact gaat verder dan de directe effectiviteit van afzonderlijke interventies en omvat bredere, duurzame veranderingen op maatschappelijk niveau. Beleidsimpactevaluatie onderzoekt hoe de gezamenlijke inspanningen van diverse actoren – zoals overheden, partnerorganisaties, organisaties met terreinwerking, Gezondheidsmakers, intermediairs en burgers – en verschillende beleidsmaatregelen samen bijdragen aan

gezondheid, welbevinden en gezondheidsongelijkheid. Dit perspectief is essentieel om te begrijpen of inspanningen op lange termijn renderen, en waar bijsturing nodig is om duurzame gezondheidswinst te realiseren.

3 Op welke interventies is de leidraad van toepassing?

De leidraad richt zich op interventies die deel uitmaken van het Vlaamse preventieve gezondheidsbeleid die de gezondheid van de Vlaming versterken, beschermen en behouden door mensen in staat te stellen actief invloed uit te oefenen op gezondheidsdeterminanten. Het gaat om **gezondheidsbevorderende interventies** en **primaire preventie** (gericht op het voorkomen van gezondheidsproblemen door het aanpakken van risicofactoren en oorzaken), zowel universeel (gericht op de hele bevolking) als selectief (gericht op groepen met verhoogde kwetsbaarheid).

Deze interventies bestaan uit een **samenhangend geheel van activiteiten, materialen en diensten**, uitgevoerd binnen een sociaal systeem met als doel de gezondheid, het gezondheidsgedrag of de determinanten van gedrag en gezondheid te veranderen. Vaak worden meerdere preventiestrategieën gecombineerd, zoals educatie, omgevingsinterventies, afspraken en regels of het aanbieden van preventieve zorg en begeleiding. Ze vinden doorgaans plaats in publieke settings zoals scholen, werkplekken, buurten of gemeenschappen.

Gezondheidsbevorderende interventies zijn vaak **complex van aard**. Een interventie is complex wanneer één of meer van volgende kenmerken aanwezig zijn: bestaan uit meerdere samenhangende onderdelen, gericht op groepen of uitgevoerd op meerdere niveaus, meerdere beoogde uitkomsten, kunnen evolueren door feedback en leerprocessen, en de effectiviteit is afhankelijk van de context en het gedrag van uitvoerders en deelnemers. Hoe moeilijker het is om precies te bepalen welke onderdelen de interventie doen werken en hoe die met elkaar samenhangen, hoe complexer de interventie.

Niet in scope van deze leidraad zijn losse materialen (folders, affiches), eenmalige vormingen of consulten, en de reguliere werking van organisaties in hun geheel.

Programma's zoals vaccinatie en bevolkingsonderzoeken vallen binnen de scope, maar enkel indien de nadruk ligt op de evaluatie van interventies die deelname hieraan bevorderen en dus gedrag beïnvloeden. De klinische effectiviteit van de programma's zelf valt buiten de scope.

4 Hoe kwam deze leidraad tot stand?

Deze leidraad werd ontwikkeld in opdracht van het Departement Zorg door onderzoekers van LUCAS KU Leuven. Het project combineerde een literatuurstudie met praktijkinput via twee werkgroepen, waarin zowel actoren uit de preventiesector (zoals partner- en terreinorganisaties) als actoren uit beleid en onderzoek (zoals academici en Sciensano) vertegenwoordigd waren. Een stuurgroep van het Departement Zorg volgde het proces op. Het onderzoek werd uitgevoerd tussen april 2025 tot december 2025.

5 Kernprincipes van iedere evaluatie in de preventiesector

Een kwaliteitsvolle evaluatie steunt niet alleen op methodologische keuzes, maar ook op een aantal **fundamentele principes die richting geven aan het hele proces**. Deze principes zorgen ervoor dat evaluaties niet enkel data opleveren, maar ook bijdragen aan rechtvaardigheid, transparantie en bruikbare inzichten voor beleid en praktijk. Ze vormen het kader waarbinnen evaluatoren werken, ongeacht de aard of omvang van de interventie. Volgende **vijf kernprincipe** zijn in deze context van belang:

Participatie: Evaluaties zijn sterker wanneer ze gedragen worden door alle betrokkenen. Daarom is het aangewezen om de doelgroep, intermediairs en relevante stakeholders actief te betrekken bij de opzet, uitvoering en interpretatie. Dit integreert diverse perspectieven en vergroot het draagvlak. Daarbij is het evenwel belangrijk om op machtsverhoudingen te letten door ruimte voor gelijkwaardige inbreng te creëren.

Equity: Elke evaluatie houdt rekening met gezondheidsongelijkheden. Dit betekent dat er inspanningen worden geleverd om kwetsbare doelgroepen te includeren en participatieve, rechtvaardige benaderingen te hanteren. Bij elke stap wordt nagegaan hoe beslissingen bijdragen aan het bevorderen van gelijkheid. Waar mogelijk worden verschillen tussen groepen inzake bereik en effectiviteit onderzocht, om vast te stellen of interventies bijdragen aan meer equity en gezondheidsrechtvaardigheid.

Transparantie: Het evaluatieproces en de resultaten moeten inzichtelijk en controleerbaar zijn. Dit betekent dat keuzes rond onderzoeksdesign, meetinstrumenten, steekproef en analyse, inclusief beperkingen duidelijk gedocumenteerd moeten worden. Rapporten moeten begrijpelijk zijn en, waar mogelijk, publiek toegankelijk.

Gebruik van inzichten voor verbetering: Evaluatie is geen eindpunt, maar in de eerste plaats een middel om te leren en verbeteren. Van bij de start moet duidelijk zijn hoe resultaten benut zullen worden voor beleid en praktijk. Het is daarom aangewezen om concrete aanbevelingen en acties te formuleren die richting geven aan bijsturing en verantwoording.

Systeemdenken: Interventies staan nooit op zichzelf, maar maken deel uit van een complex systeem waarin context, actoren en middelen elkaar beïnvloeden. Een kwaliteitsvolle evaluatie vraagt daarom om het identificeren van de belangrijke elementen in dit systeem, het bewust bepalen van de grenzen van de evaluatie, en het inbouwen van flexibiliteit om met dynamieken en onvoorspelbaarheid om te gaan. Bevindingen moeten daarom in een breder geheel geplaatst worden, zoals beleid, organisatieprocessen en maatschappelijke trends. Dit perspectief helpt om resultaten correct te interpreteren en om aanbevelingen te formuleren die aansluiten bij de complexe realiteit waarin interventies zich voordoen.

6 Een weloverwogen en transparante inzet van werkingsmiddelen voor kwaliteitsvolle evaluatie

Organisaties dragen zelf de verantwoordelijkheid om voldoende middelen te reserveren voor evaluatie en deze flexibel toe te passen, afgestemd op de aard en evaluatiebehoefte van hun interventies. In hun meerjarenplannen verduidelijken ze hoe budgetten voor zowel structurele evaluatie als effectevaluatie worden ingecalculeerd en op welke manier deze zullen worden ingezet. Dit helpt om de haalbaarheid, voortgang en kwaliteit van evaluaties op te volgen. Door vooraf keuzes te maken en deze duidelijk en onderbouwd te communiceren, wordt vermeden dat evaluaties ad hoc of onvoldoende onderbouwd

gebeuren. Transparantie in planning maakt opvolging door het Departement mogelijk en versterkt het vertrouwen in de resultaten.

Effectevaluatie vereist methodologische expertise. Sommige organisaties beschikken over interne onderzoekscapaciteit, vaak dankzij nauwe banden met onderzoeksinstellingen. Andere organisaties hebben minder interne expertise en moeten externe onderzoekers inschakelen voor meer gedegen effect-onderzoek.

Er bestaat niet één standaardaanpak voor het organiseren van effectevaluatie. Organisaties kunnen kiezen uit **verschillende scenario's**, afhankelijk van hun interne expertise en nood aan externe expertise, beschikbare middelen (budget, tijd, VTE) en de complexiteit van de interventie en beoogde evaluatie.

De drie meest voorkomende opties zijn:

- **Interne inzet:** interne medewerkers voeren de evaluatie uit.
- **Externe expertise:** externe onderzoekers voeren de evaluatie uit.
- **Samenwerking intern-extern:** een gedeelde aanpak waarbij interne en externe evaluatoren samenwerken.

Elk scenario heeft eigen voordelen en uitdagingen. Het is belangrijk om deze opties zorgvuldig af te wegen. Binnen deze variabiliteit aan samenwerkingen zijn er enkele rollen voor beleid, evaluatoren (medewerkers en/of externe onderzoekers die de evaluatie opzetten en/of uitvoeren), organisaties in de preventie-sectoren en actoren met terreinwerking (zie *Bijlage 1*).

Hoewel neutraliteit en transparantie essentieel zijn voor een geloofwaardige effectevaluatie, is volledige onafhankelijkheid niet altijd haalbaar of wenselijk. In de praktijk zijn ontwikkelaars, implementatoren en evaluatoren vaak nauw met elkaar verbonden. Daarom is het belangrijk om vooraf een evaluatieplan en rolverdeling vast te leggen, zodat belangenconflicten worden beperkt en de geloofwaardigheid van de evaluatie gewaarborgd blijft.

6.1 Scenario 1: Interne inzet (intern VTE inschakelen)

Interne medewerkers worden (gedeeltelijk) vrijgesteld om de evaluatie op te zetten en uit te voeren. Dit scenario is vooral geschikt voor organisaties die ingebed zijn in een academische context of beschikken over sterke interne onderzoeksexpertise.

Voordelen:

- **Praktijkkennis:** interne medewerkers hebben een grondige kennis van de interventie, doelgroep en context.
- **Continuïteit:** korte lijnen en directe afstemming tussen implementatie en evaluatie worden gewaarborgd.
- **Kostenefficiënt:** er zijn geen extra kosten voor externe onderzoekers.

Uitdagingen:

- **Tijdsinvestering:** als medewerkers ook andere kerntaken hebben, kan het vrijmaken van tijd voor de evaluatie andere kerntaken onder druk zetten.
- **Risico op bias:** wanneer dezelfde medewerkers zowel de interventie ontwikkelen en/of implementeren als evalueren, kan dit de objectiviteit en geloofwaardigheid van de evaluatie verminderen.

- **Toegang tot ethisch comité en statistische ondersteuning voor organisaties zonder inbedding in een onderzoeksinstelling:** organisaties zonder link met een onderzoeksinstelling hebben vaak geen directe toegang tot een erkend ethisch comité, noch tot de methodologische en administratieve ondersteuning die onderzoeksinstellingen doorgaans voorzien voor onderzoekers. Dit kan het verkrijgen van ethische goedkeuring bemoeilijken en mogelijk de kwaliteit van het onderzoek beïnvloeden. Punt 7.4.6 in deze leidraad verduidelijkt wanneer het vragen van ethische goedkeuring nodig is.

Randvoorwaarden:

- **Beschikbare onderzoekscapaciteit:** minstens één interne medewerker beschikt over methodologische expertise in evaluatieonderzoek. Interne medewerkers krijgen expliciet tijd en een duidelijk mandaat om de evaluatie uit te voeren.
- **Duidelijke rolafbakening:** waar mogelijk wordt een onderscheid gemaakt tussen interne medewerkers die betrokken zijn bij implementatie en interne medewerkers die instaan voor evaluatie, om bias te beperken.

6.2 Scenario 2: Inschakelen van externe expertise (bv. onderzoekers uit onderzoeksinstellingen)

Externe onderzoekers met gespecialiseerde methodologische expertise worden aangesteld om de effectevaluatie op te zetten en uit te voeren. Dit zijn gewoonlijk onderzoekers uit onderzoeksinstellingen zoals universiteiten en hogescholen.

Voordelen:

- **Methodologische expertise:** externe onderzoekers van onderzoeksinstellingen hebben sterke kennis van onderzoeksdesigns en data-analyse en kunnen rekenen op statistische en methodologische ondersteuning (bv. bij designkeuzes) vanuit de onderzoeksinstelling.
- **Toegang tot ethisch comité:** externe onderzoekers van onderzoeksinstellingen hebben toegang tot een ethisch comité voor het verkrijgen van ethische goedkeuring.
- **Objectiviteit:** externe onderzoekers zijn onafhankelijk van de implementatie en de ontwikkeling van de interventie wat het risico op bias vermindert en de geloofwaardigheid verhoogt.
- **Efficiëntie:** interne medewerkers kunnen zich focussen op de implementatie van de interventie.

Uitdagingen:

- **Kosten:** het inschakelen van externe expertise is duurder.
- **Afstand tot de praktijk:** externe onderzoekers zijn vaak minder vertrouwd met de interventie, doelgroep of context.

Randvoorwaarden:

- **Duidelijke opdrachtschrijving:** voor de start van de evaluatie zijn de doelstellingen, onderzoeksvragen, timing en verantwoordelijkheden expliciet en eenduidig vastgelegd.
- **Goede afstemming:** er is een regelmatige en doorlopende afstemming tussen de externe onderzoekers en het interne team met terugkoppelingen over het verloop van het onderzoek.

6.3 Scenario 3: Samenwerking tussen interne en externe evaluatoren

Interne en externe evaluatoren werken samen met een duidelijke rolverdeling. Interne medewerkers nemen vaak voorbereidende en uitvoerende taken op zich (bv. dataverzameling, contact met doelgroep), terwijl externe onderzoekers instaan voor het onderzoeksdesign, de analyses en methodologische ondersteuning.

- **Sterktes combineren:** interne praktijkkennis wordt aangevuld met externe onderzoeksexpertise waar nodig.
- **Kostenefficiënt:** gerichte inzet van externe onderzoekers drukt de kosten.
- **Capaciteitsopbouw:** interne medewerkers doen ervaring op met evaluatie, wat toekomstige evaluaties versterkt.

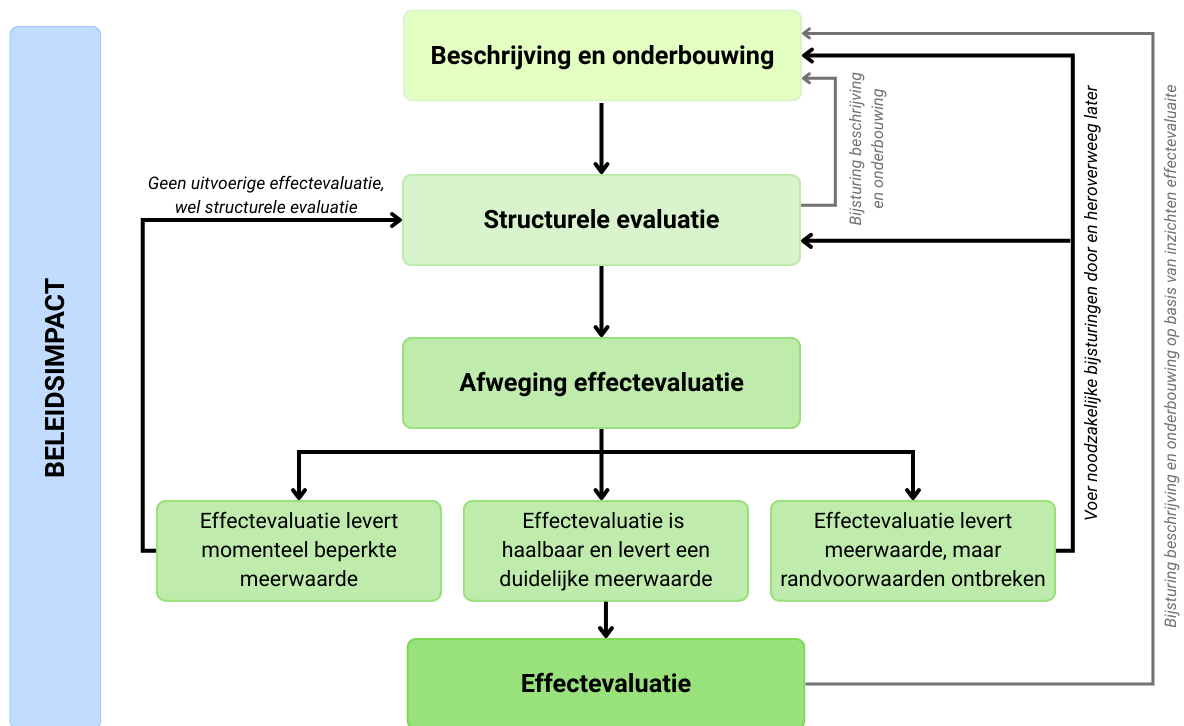
Randvoorwaarden:

- **Heldere rol- en taakverdeling:** taken en rollen van interne medewerkers en externe onderzoekers zijn vooraf duidelijk vastgelegd.
- **Structurele overlegmomenten:** er zijn regelmatige en geplande overlegmomenten om afstemming, voortgang en tijdige bijsturing te waarborgen.
- **Capaciteit aan beide zijden:** zowel het interne teams als de externe partner beschikken over voldoende tijd, middelen en expertise om hun rol effectief op te nemen.
- **Vertrouwen en transparantie:** er is wederzijds vertrouwen en openheid over de beperkingen, belangen en mogelijke spanningsvelden (bv. tussen praktijk en onderzoek).
- **Afspraken over eigenaarschap:** afspraken over data, rapportage, publicatie en gebruik van resultaten zijn vooraf expliciet vastgelegd in een samenwerkingsovereenkomst.

7 Een evaluatieproces in vier fasen

Het evaluatieproces van interventies verloopt **gefaseerd** om kwaliteit en haalbaarheid te waarborgen, maar is **iteratief** van aard. Elke fase bouwt voort op de vorige en biedt ruimte voor **bijsturing**. Inzichten uit latere fasen voeden steeds de eerste fase, waarin de beschrijving en onderbouwing van de interventie worden aangescherpt. Structurele evaluatie met een focus op procesevaluatie wordt doorlopend uitgevoerd, terwijl aan robuuste effectevaluatie een expliciete afweging voorafgaat. Dit verloop wordt weergegeven in Figuur 26.

Figuur 26 Verloop en iteratief karakter van de vier fasen in het evaluatieproces.



De vier fasen zijn:

- 1) **Beschrijving en onderbouwing van de interventie:** Leg stevige basis voor evaluatie door de interventie duidelijk te beschrijven en onderliggende aannames en uitgangspunten expliciet te maken.
- 2) **Structurele evaluatie en bijsturing:** Volg de uitvoering van de interventie systematisch op en voer procesevaluaties uit om de kwaliteit, uitvoerbaarheid en aanvaardbaarheid te verbeteren. Monitor uitkomsten op een pragmatische manier om te leren en bij te sturen, zonder conclusies te trekken over causale effecten.
- 3) **Beoordeel of een robuuste effectevaluatie aan de orde is:** Maak een afweging of een robuuste effectevaluatie een duidelijke meerwaarde biedt en haalbaar is. Op basis van deze afweging wordt bepaald of wordt overgegaan naar fase 4.
- 4) **Effectevaluatie:** Onderzoek met een vooraf gepland en methodologisch onderbouwd onderzoeksopzet of en in welke mate de interventie effecten veroorzaakt.

Tabel 12 vat de vier fasen overzichtelijk samen.

Tabel 12 Evaluatieproces in vier fasen.

Fasen	1. Beschrijving en onderbouwing	2. Structurele evaluatie en bijsturing	3. Afweging effectevaluatie	4. Effectevaluatie
Doel	Basis leggen voor evaluatie	Opvolging, bijsturing en verbeteren van uitvoering, kwaliteit en haalbaarheid	Bepalen of een meer robuuste effectevaluatie zinvol en haalbaar is	Op een meer robuuste manier onderzoeken of interventie werkt en de beoogde doelstellingen behaalt
Wat gebeurt er?	Narratieve beschrijving en bestaande onderbouwing, Theory of Change en Logic Model opstellen	Evaluatie plannen (eventueel evaluatieplan opstellen); procesevaluatie (focus op implementatie, context en mechanismen); pragmatische evaluatie van determinanten en uitkomsten van gedrag en gezondheid	Beoordelen van de stabiliteit van de interventie; afwegen van de kosten en baten; beslismoment	Evaluatieplan opstellen; effectmeting uitvoeren; analyse; aanbevelingen formuleren <i>Effectevaluatie kan met of zonder parallelle procesevaluatie plaatsvinden.</i>
Timing	Bij ontwikkeling of voorafgaand aan evaluatie, bij belangrijke bijstellingen	Doorlopend, periodiek	Periodiek	Afhankelijk van de beslissing van fase 3
Uitvoerder	Organisatie	Organisatie	Organisatie in afstemming met beleid	Evaluatoren (interne en/of externe onderzoekers)
Voorbeelden van uitkomsten	n.v.t.	Haalbaarheid, bereik, adoptie, getrouwheid van uitvoering, uitval, tevredenheid van deelnemers en uitvoerders, vroege veranderingen in determinanten en uitkomsten, contextfactoren	n.v.t.	Veranderingen in gedrag (bv. roken, bewegen, voeding), gezondheid (bv. mentaal welzijn, BMI, stress), en determinanten van gedrag en gezondheid (bv. kennis, sociale steun, attitude, motivatie)

Naast deze vier fasen kan op **beleidsniveau een impactevaluatie** plaatsvinden, afgestemd op beleidscycli zoals de evaluatie van de gezondheidsdoelstellingen. Deze kijkt naar duurzame effecten op populatieniveau en maatschappelijke trends (zie 8.3 *Beleidsimpact op systeem- en populatieniveau*).

7.1 Fase 1: Beschrijving en onderbouwing van de interventie

De eerste fase legt een solide basis voor alle latere evaluaties door (1) de interventie en bijhorende onderbouwing narratief te beschrijven, (2) een Theory of Change en (3) een Logic Model uit te werken. Het doel van deze fase is om bij de betrokkenen een **gedeeld begrip** creëren van wat de interventie inhoudt, waarom ze nodig is, hoe ze naar verwachting werkt en in welke context ze wordt uitgevoerd.

Deze fase is **dynamisch** en wordt doorheen het evaluatieproces bijgestuurd op basis van nieuwe inzichten.

Het is aanbevolen om het resultaat van deze fase te **publiceren in de catalogus preventiemethodieken** (preventiemethodieken.be). Door interventies op een meer uniforme manier te beschrijven, ontstaat een duidelijk en vergelijkbaar overzicht dat de zichtbaarheid, transparantie en consistentie van het preventieve aanbod vergroot.

7.1.1 Narratieve beschrijving en wetenschappelijke onderbouwing

Zorg voor een heldere beschrijving van de interventie en de wetenschappelijke onderbouwing waarop deze gebaseerd is.

Beschrijf in duidelijke, begrijpelijke taal de volgende elementen:

- **Korte omschrijving in duidelijke taal**
- **Doel:** Wat wil de interventie **bereiken**?
- **Doelgroep(en):** Wie is de intermediaire doelgroep en de einddoelgroep?
- **Context en setting:** Waar en onder welke omstandigheden wordt de interventie uitgevoerd (bv. school, werkplek, buurt)?
- **Aanpak, activiteiten en materialen:** Hoe gaat de interventie concreet te werk? Welke preventiestrategieën (= hoe de gezondheidsdoelstelling zal bereikt worden i.e., educatie, omgevingsinterventies, beleid via afspraken en regels en zorg en begeleiding), implementatiestrategieën (= hoe de preventiestrategieën worden uitgevoerd en geborgd bv. training en ondersteuning van uitvoerders, voorzien van materialen en tools, monitoring en feedback...), activiteiten en materialen worden ingezet?

Onderbouw de interventie met bewijs en randvoorwaarden op basis van een gestructureerde analyse die een antwoord geeft op volgende vragen:

- **Theoretisch kader:** Welke theorie of welke theoretische concepten ondersteunen de interventie?
- **Wetenschappelijke literatuur en richtlijnen:** Welke relevante wetenschappelijke studies en aanbevelingen uit de internationale literatuur ondersteunen de interventie?
- **Eerdere evaluaties en praktijkervaringen:** Wat kan men leren van eventuele eerdere evaluaties (bv. pilootstudies) van de interventie? Wat is al geweten uit praktijkervaringen?
- **Randvoorwaarden:** Welke voorwaarden (bv. middelen, competenties, organisatorische condities) zijn nodig voor een succesvolle implementatie?

Een **inspirerend voorbeeld voor praktijkbeschrijvingen** is de Databank kwaliteitsvolle praktijken van Opgroeien²⁹. Deze databank bundelt praktijkvoorbeelden uit Vlaanderen en Brussel van en voor professionals in het jeugd- en gezinsbeleid. Ingediende praktijkbeschrijvingen worden **beoordeeld door een commissie** van wetenschappers en praktijkprofessionals, die het niveau van de beschrijving inschat op basis van drie niveaus (goed beschreven, theoretisch onderbouwd, wetenschappelijk onderzocht) zonder de effectiviteit van de praktijk zelf te beoordelen.

7.1.2 Theory of Change

Het gebruiken of opstellen van een **Theory of Change of verandertheorie** van een interventie is een essentieel onderdeel van elke evaluatie. De Theory of Change bouwt voort op de vorige stap en beschrijft hoe en waarom een interventie naar verwachting leidt tot de gewenste verandering. De focus ligt op de bredere context, causale verbanden, aannames en externe factoren die invloed hebben op het succes van een interventie. De Theory of Change vormt dus ook een theoretische toets: heeft de interventie, vanuit conceptueel oogpunt, alles in zich om de beoogde effecten te realiseren?

²⁹ <https://www.opgroeien.be/kennis/databank-kwaliteitsvolle-praktijken>

Waarom is een Theory of Change belangrijk?

Een Theory of Change is naast een vorm van theoretische evaluatie ook een praktisch instrument voor verdere evaluatie. Door vooraf helder te beschrijven hoe en waarom een interventie naar verwachting werkt, wordt de veronderstelde effectketen zichtbaar: van activiteiten van de interventie, via veranderingen bij intermediairs, tot beoogde effecten bij de einddoelgroep. Dit maakt het mogelijk om latere evaluaties gericht op te zetten en ondersteunt het bepalen van de uitkomsten en mechanismen die moeten worden getoetst op één of meerdere niveaus. Zo vormt de Theory of Change het fundament voor een transparante evaluatieaanpak.

Gebruik een Theory of Change om:

- Een **gedeeld begrip** te creëren rond hoe en waarom de interventie leidt tot veranderingen
- **Veranderpaden en aannames** inzichtelijk te maken
- **Contextuele invloeden** te verduidelijken
- De **rationale** ('het waarom') achter een interventie te verantwoorden

Maak aan de hand van een Theory of Change de volgende elementen expliciet:

- **Impact:** uiteindelijke maatschappelijke verandering
- **Uitkomsten:**
 - Korte termijn: determinanten van gedrag of praktijkvoering zoals kennis, attitudes, vaardigheden
 - Middellange termijn: gedragingen en praktijkvoering
 - Lange termijn: duurzame veranderingen richting impact
- **Veranderpaden:** de opeenvolging van stappen van activiteit tot impact
- **Rationale:** verklaren het verband tussen verschillende stappen
- **Aannames:** voorwaarden waarvan wordt verondersteld dat ze waar zijn
- **Randvoorwaarden:** contextfactoren die buiten de directe controle van de interventie liggen
- **Activiteiten en preventiestrategieën:** wat de interventie doet, componenten van de interventie

Bij een Theory of Change gaat het niet om het volledig in detail beschrijven van alle elementen, maar om het creëren van een **plausibel, testbaar en bruikbaar kader** dat inzicht geeft in hoe de activiteiten en preventiestrategieën in de interventie leiden tot impact. De focus ligt op de stappen die cruciaal zijn voor het behalen van het doel. Daarbij is het nuttig om een **grens van verantwoordelijkheid** duidelijk te maken: de hoogste uitkomst waarvoor de interventie verantwoordelijk kan worden gehouden; alles boven dit plafond valt buiten de directe invloedssfeer.

Aanbeveling: Koppel lange termijndoelen aan Vlaamse gezondheidsdoelstellingen. Een dergelijke afstemming is belangrijk voor beleidsevaluatie, omdat het toelaat om interventies te situeren binnen het bredere preventieve gezondheidsbeleid. Waar mogelijk worden ook de interventiegebonden uitkomst-indicatoren afgestemd op populatie-indicatoren uit beschikbare databronnen, al zijn afwijkingen hierop mogelijk wanneer alternatieve indicatoren beter aansluiten bij de interventie.

Bijlage 2 geeft een voorbeeld van een Theory of Change en een overzicht van verschillende bronnen. Deze bronnen maken eveneens duidelijk dat er niet één standaardmodel bestaat. De structuur en het niveau van detail van een Theory of Change worden in belangrijke mate bepaald door de complexiteit en aard van interventie.

7.1.3 Logic Model

Een logic model is een **schematische weergave** van de onderliggende logica van een interventie. Het maakt inzichtelijk hoe beschikbare middelen (**inputs**) via geplande activiteiten (**activiteiten**) leiden tot concrete resultaten of producten (**outputs**) en uiteindelijk bijdragen aan gewenste effecten (**uitkomsten**) op korte, middellange en lange termijn.

Gebruik een logic model om:

- De evaluatie te plannen
- Indicatoren te selecteren
- Voortgang te monitoren

Een logic model wordt doorgaans voorgesteld als een horizontaal schema met opeenvolgende componenten:

- **Inputs:** middelen zoals budget, personeel en infrastructuur
- **Activiteiten:** acties die worden uitgevoerd, zoals trainingen, workshops, campagnes of dienstverlening
- **Outputs:** beoogde directe resultaten en producten van activiteiten, bijvoorbeeld het aantal sessies of bereikte personen
- **Uitkomsten:** beoogde effecten op verschillende niveaus:
 - **Korte termijn:** determinanten zoals kennis, attitudes, vaardigheden
 - **Middellange termijn:** gedrags- of praktijkveranderingen
 - **Lange termijn:** impact op gezondheid, welzijn of beleid

Bijlage 3 omvat een praktisch invulmodel van een logic model.

Theory of Change of logic model: wat is het belangrijkste verschil en wat komt eerst?

De **Theory of Change** biedt een onderbouwde en samenhangende uitleg van *hoe en waarom* verandering tot stand komt. Het kan op een flexibele en visuele manier worden weergegeven en is vooral geschikt om complexiteit, interacties, aannames en onzekerheden expliciet te maken. De focus ligt daarbij op het expliciteren van de onderliggende rationale: welke aannames worden gemaakt, welke mechanismen worden verondersteld en onder welke voorwaarden kan verandering verwacht worden.

Het **logic model** is een meer schematisch en lineair instrument dat inzicht geeft in de directe relatie tussen inputs/middelen, activiteiten, outputs en uitkomsten. In vergelijking met een Theory of Change is een logic model meer operationeel en concreet omdat het specifiek beschrijft wat er in de praktijk gebeurt en welke stappen worden gezet om tot resultaten te komen, vaak gekoppeld aan meetbare indicatoren.

Hoewel er dus wel degelijk verschillen zijn tussen de Theory of Change en het logic model, zijn ze **inhoudelijk verwant** en zijn beide bedoeld om de keten van verandering inzichtelijk te maken. In de literatuur worden de Theory of Change en het logic model soms door elkaar gebruikt of zelfs als synoniemen gepresenteerd. Daarnaast bestaan er **hybride vormen** waarin elementen van beide benaderingen worden gecombineerd. In sommige contexten kan dit een passende oplossing zijn, bijvoorbeeld wanneer de keten van verandering relatief lineair is, maar er toch behoefte bestaat aan explicitering van aannames en onderliggende mechanismen. Zo kan in een logic model ruimte worden gemaakt voor aannames, rationale of randvoorwaarden die de relatie tussen activiteiten en uitkomsten beïnvloeden. Deze combinatie kan helpen om zowel de

overzichtelijkheid en praktische bruikbaarheid van een logic model als de theoretische onderbouwing en reflectie die kenmerkend zijn voor een Theory of Change te behouden.

Wanneer zowel een Theory of Change als een logic model worden uitgewerkt, is vaak de vraag waarmee te beginnen. Bij **bestaande (goed gekende) en eenvoudige interventies of kleinschalige evaluaties** is het vaak efficiënter om te vertrekken vanuit een logic model, omdat dit snel duidelijkheid biedt over activiteiten, resultaten en indicatoren. Dit logic model kan worden verdiept met een Theory of Change om meer inzicht te krijgen in onderliggende rationale en aannames.

Bij **nieuwe, grote of complexe interventies** is het aangewezen om te starten met een Theory of Change, zodat eerst het bredere kader, de aannames en de verschillende verandervaden helder worden. Op basis hiervan kan vervolgens een logic model worden afgeleid als vereenvoudigde en operationele weergave.

7.2 Fase 2: Structurele evaluatie en bijsturing

Evaluatie is een continu proces dat nauw verweven is met de implementatie van een interventie. In deze fase ligt de nadruk op procesevaluatie: het onderzoeken hoe de interventie wordt uitgevoerd, welke contextfactoren een rol spelen en of de beoogde werkingsmechanismen in gang worden gezet. Deze inzichten zijn cruciaal om tijdig bij te sturen, de kwaliteit van de uitvoering te verbeteren en de interventie voor te bereiden op een latere effectevaluatie.

Een goede planning is belangrijk bij procesevaluatie. Door vooraf te bepalen wat, wanneer en hoe wordt geëvalueerd, wordt evaluatie een integraal onderdeel van de uitvoering van de interventie. Indien gewenst, en met name bij meer grondige of omvangrijke procesevaluaties, kan hiervoor een evaluatieplan worden opgesteld. Dit evaluatieplan beschrijft onder meer welke aspecten van het proces worden geëvalueerd, wie betrokken wordt bij de evaluatie, hoe de gegevens worden verzameld en geanalyseerd en hoe de resultaten worden benut. Het is zinvol om het evaluatieplan voor de effectevaluatie af te stemmen op dat voor de procesevaluatie, om samenhang tussen beide evaluatievormen te creëren (zie *Fase 4: Effectevaluatie volgens een duidelijk evaluatieplan*). Het is eveneens een goede praktijk om een evaluatierapport met de resultaten van meer grondige of omvangrijke procesevaluaties te delen, waardoor kruisbestuiving kan ontstaan in de sector en de transparantie van evaluaties verhoogt (zie *Gebruik en disseminatie van resultaten*).

Procesevaluatie gebeurt periodiek en richt zich op de uitvoering en haalbaarheid, maar ook vroege signalen van verandering in gedrag en gedragsdeterminanten kunnen worden opgenomen. Dit geeft een indicatie of de interventie werkt zoals bedoeld en of de onderliggende mechanismen geactiveerd worden, zonder dat dit een grondige effectevaluatie vormt.

7.2.1 Uitvoering en procesevaluatie: continu en proportioneel

Bij een procesevaluatie staat niet de vraag *of* een interventie werkt centraal, maar *hoe* deze in de praktijk wordt uitgevoerd en ervaren. Procesevaluatie is een essentieel onderdeel van kwaliteitsvolle implementatie en vormt vaak de brug naar effectevaluatie. Het onderzoekt systematisch of een interventie wordt uitgevoerd zoals gepland en helpt begrijpen hoe, waarom en onder welke omstandigheden een interventie werkt. Dit geeft inzicht in wat goed loopt, waar knelpunten zitten en of de beoogde werkingsmechanismen worden geactiveerd.

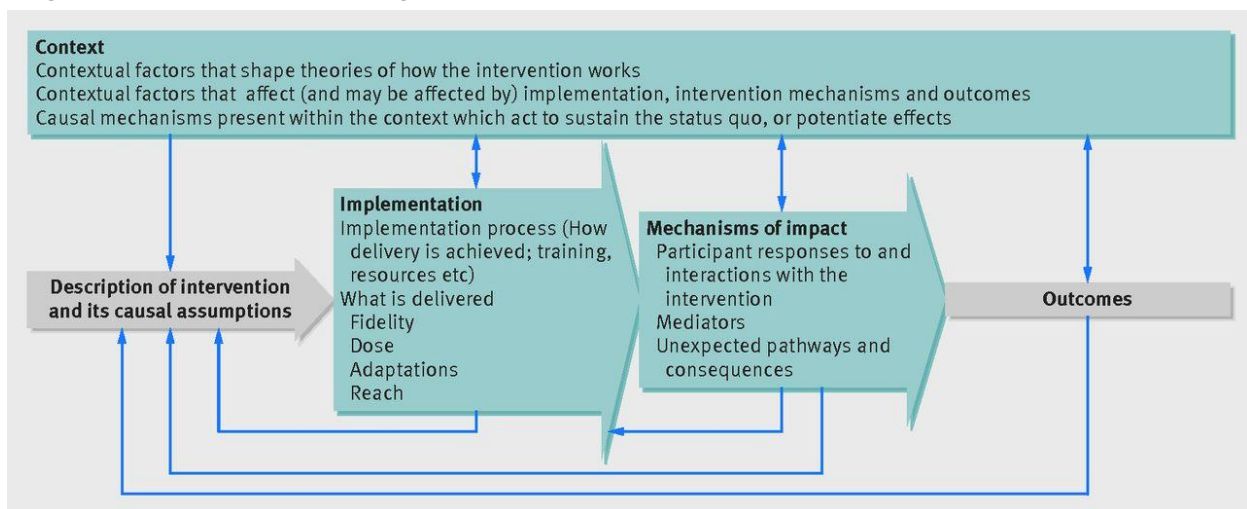
Waaruit bestaat een procesevaluatie?

Procesuitkomsten geven inzicht in de mate waarin een interventie volgens plan is geïmplementeerd en welke factoren het verloop van het proces beïnvloeden. Procesuitkomsten richten zich onder andere op aspecten zoals bereik, adoptie door implementatoren, getrouwheid aan het implementatieplan, acceptatie van de doelgroep en uitvoerbaarheid in de praktijk. Een overzicht van vaak gebruikte procesuitkomsten is opgenomen in *Bijlage 4*.

Binnen het raamwerk van de Medical Research Council (MRC) bestaat procesevaluatie uit drie kernonderdelen (zie Figuur 28, Moore et al., 2015):

- **Implementatie:** Hoe is de interventie uitgevoerd? (bv. getrouwheid, intensiteit, bereik, adaptaties)
- **Context:** Welke contextfactoren beïnvloeden uitvoering en uitkomsten? (bv. organisatie, beleid, sociale en economische omstandigheden)
- **Mechanismen van verandering:** Welke processen verklaren waarom de interventie werkt? (bv. psychologische, sociale en gedragsmatige mechanismen zoals kennis, attitude, sociale steun)

Figuur 27 Procesevaluatie volgens het MRC-framework (70).



Wanneer is een procesevaluatie aan de orde?

Procesevaluatie is zinvol in de vroege fases van implementatie en op geregelde tijdstippen tijdens de looptijd van een interventie. Aanbevolen momenten zijn:

- Bij start of pilot van een nieuwe interventie
- Bij opschaling of belangrijke aanpassingen
- Periodiek bij lopende interventies (bv. jaarlijks)

Bij de stabiele implementatie van een interventie ligt de focus in de structurele evaluatie op monitoring, kwaliteit en duurzaamheid. Procesevaluatie kan dan inzicht bieden in de mate waarin de kwaliteit van uitvoering behouden blijft, in hoeverre de interventie structureel is ingebed in organisaties en beleid, en welke aanpassingen in de praktijk plaatsvinden als gevolg van veranderende contexten. Daarnaast kan procesevaluatie laten zien hoe de interventie verantwoord kan worden opgeschaald of geadapteerd naar andere settings of doelgroepen en waar mogelijkheden liggen voor continue verbetering.

Wie staat in voor procesevaluatie?

Procesevaluaties zonder parallelle effectevaluatie worden doorgaans uitgevoerd door interne medewerkers van preventieorganisaties (partnerorganisaties, organisaties met terreinwerking, Gezondheidsmakers), omdat de focus ligt op implementatie en kwaliteitsverbetering.

Wanneer een procesevaluatie onderdeel is van een effectevaluatie, is het aangewezen om (ook) externe onderzoekers te betrekken om de neutraliteit en de methodologische kwaliteit te waarborgen.

Belangrijke aandachtspunten bij procesevaluatie:

- **Houd evaluatie proportioneel en afgestemd op het doel:** Het is niet in elke fase noodzakelijk of haalbaar om alle procesonderdelen (implementatie, context, mechanismen) en alle mogelijke procesuitkomsten uitgebreid te onderzoeken. Stem de procesevaluatie af op de fase van ontwikkeling en op de doelen van dat moment, bijvoorbeeld:
 - Monitoren (doorlopend): volg uitvoering, adoptie en bereik op regelmatige basis op.
 - Optimaliseren: focus op knelpunten en randvoorwaarden.
 - Verklaren: analyseer mechanismen en contextfactoren.
- **Gebruik van een kader voor implementatie- of procesevaluatie:** Een evaluatiekader zorgt voor consistentie en helpt om systematisch inzicht te krijgen in de implementatie van de interventie. *Bijlage 5* biedt een overzicht van enkele veelgebruikte evaluatiekaders die relevant zijn voor implementatieonderzoek en ruimere procesevaluatie van interventies in de preventiesector. Kies een kader dat past bij de fase van de interventie en het doel van de evaluatie:
 - **CFIR, Proctor en IRLM** voor implementatieonderzoek.
 - **RE-AIM** combineert implementatie en (vroege) effectuitkomsten.
 - **MRC-framework:** aanbevolen voor complexe interventies, vaak in combinatie met effectevaluatie.
- **Combineer methoden:** Procesevaluatie vraagt idealiter om een mix van kwalitatieve en kwantitatieve methoden. Deze combinatie versterkt de betrouwbaarheid van de bevindingen.
 - Kwalitatief: interviews, focusgroepen en observaties geven meer inzicht in ervaringen en context.
 - Kwantitatief: vragenlijsten, monitoring en checklists maken patronen zichtbaar en meten uitvoering.
- **Identificeer randvoorwaarden voor goede uitvoering:** Procesevaluatie biedt inzicht in wat nodig is om een interventie succesvol uit te voeren, welke risico's aandacht vragen en waarmee men rekening moet houden bij toekomstige implementaties. Beschrijf randvoorwaarden zoals: competenties en training van uitvoerders, draagvlak binnen de organisatie, samenwerking en communicatie met partners, beschikbare middelen, personeel, tijd en materialen. Deze informatie helpt bij opschaling en reproduceerbaarheid.
- **Documenteer adaptaties:** Interventies worden vaak aangepast tijdens de implementatie (bv. inhoud, intensiteit of duur, materialen...). Leg deze wijzigingen vast, inclusief de reden voor de aanpassing. Gebruik eventueel een 'fidelity checklist' (getrouwheidsvragenlijst) om na te gaan waar implementatoren zaken aanpassen en of kerncomponenten behouden blijven. Dit is cruciaal om te begrijpen hoe en waarom een interventie al dan niet werkt in de praktijk.
- **Onderzoek verandering bij de intermediaire doelgroep:** Wanneer een interventie ook veranderingen bij intermediairs beoogt, is het relevant om deze expliciet te evalueren. Dit kan gaan om veranderingen in kennis, vaardigheden of gedrag. Dergelijke intermediaire veranderingen vormen vaak een noodzakelijke voorwaarde om effecten bij de einddoelgroep te realiseren. Het blijft echter belangrijk om het onderscheid tussen proces- en effectevaluatie scherp te houden. Wanneer je vooral wil nagaan of intermediairs de interventie correct toepassen of uitvoeren, gaat het om een procesevaluatie. Enkel wanneer de interventie zelf ook gericht is op intermediairs (bv. de interventie omvat een training,

coaching of ondersteunende component voor intermediairs) is het zinvol om de effecten op intermediair niveau te meten als onderdeel van de effectevaluatie.

- **Onderzoek vroege signalen van verandering bij de einddoelgroep:** Hoewel de structurele evaluatie geen robuuste effectevaluatie omvat, kan men in deze fase wel verkennen of beoogde mechanismen worden geactiveerd en tussentijdse veranderingen worden gerealiseerd bij de einddoelgroep. Meet haalbare determinanten zoals kennis, motivatie, bewustwording of zelfeffectiviteit. Indien relevant kunnen op een meer pragmatische manier (i.e. zonder controlegroepen) ook voorlopige signalen van gedragsverandering worden meegenomen. Dit geeft een indicatie van potentieel effect. Het doel blijft echter verkennen hoe de interventie werkt, niet om conclusies te trekken over effectiviteit.

7.2.2 Bijsturing op basis van procesevaluatie

Procesevaluatie levert cruciale inzichten in wat in de praktijk haalbaar is, welke drempels optreden en welke mechanismen onvoldoende geactiveerd worden. Deze bevindingen moeten worden vertaald naar concrete aanbevelingen voor verbetering. Bijsturing kan betrekking hebben op zowel de interventie zelf als de theoretische onderbouwing.

Praktische en inhoudelijke wijzigingen van de interventie zijn vaak nodig om de uitvoerbaarheid en effectiviteit te vergroten. Voorbeelden: inkorten van sessies, verduidelijken van materialen of handleidingen, wijzigingen m.b.t. het benaderen van de doelgroep, het toevoegen of weglaten van interventie-componenten. Logischerwijs wordt dan ook de beschrijving van de interventie aangepast.

Procesevaluatie kan ook aanleiding geven om de Theory of Change te herzien. Dit gebeurt door het bijstellen van aannames die niet blijken te kloppen of het aanpassen van de rationale en bijhorende mechanismen om beter te verklaren hoe de interventie naar verwachting tot verandering leidt.

Door deze bijstellingen systematisch te documenteren versterkt de transparantie en reproduceerbaarheid van het evaluatieproces.

7.3 Fase 3: Beoordeel of grondige effectevaluatie zinvol en haalbaar is

Voordat een effectevaluatie wordt opgestart, is het cruciaal om na te gaan of een effectevaluatie een duidelijke meerwaarde oplevert en de interventie klaar is om geëvalueerd te worden. Deze afweging voorkomt dat te vroege, onrealistische of overbodige effectstudies worden uitgevoerd.

7.3.1 Levert een grondige effectevaluatie een duidelijke meerwaarde, gegeven de nodige investering?

Een effectevaluatie is enkel gerechtvaardigd wanneer het, gegeven de investering in tijd en middelen die nodig is voor effectevaluatie, een duidelijke meerwaarde biedt voor praktijk, beleid of kennisontwikkeling. Dit vraagt om een zorgvuldige analyse van bestaand bewijs en de potentiële opbrengst van een nieuwe studie.

Beoordeel bestaand bewijs op kwaliteit en overdraagbaarheid

Analyseer het beschikbare bewijs voor de interventie. **Beoordeel de kwaliteit van eerdere studies** m.b.t. methodologische sterkte (bv. gebruik van controlegroepen, randomisatie, longitudinale onderzoeksopzet), robuustheid en consistentie van resultaten, en conclusies uit literatuuronderzoeken en systematische reviews.

Vervolgens is het cruciaal om **na te gaan in welke mate deze bevindingen overdraagbaar zijn** naar de Vlaamse context en de specifieke implementatiesetting. Factoren zoals vergelijkbaarheid van de doelpopulatie, socioculturele kenmerken, organisatiecondities, zorgsystemen en regelgeving spelen hierbij een rol.

Aanpassingen aan de interventie zijn vaak nodig om de uitvoerbaarheid te garanderen, maar mogen de **kernmechanismen niet aantasten**. Het is essentieel dat de werkzame elementen van de interventie behouden blijven. Gebruik de Theory of Change als referentiepunt: hoe meer werkzame elementen van de oorspronkelijke interventie worden gewijzigd, hoe groter de nood aan (her)evaluatie.

Voorbeeld: Stel dat de effectiviteit van een onderzochte leefstijlinterventie afhankelijk is van wekelijkse begeleiding door een gemotiveerde en goed getrainde onderzoeker, en van kerncomponenten zoals motiverende gespreksvoering of sociale steun. Wanneer de interventie in een andere context wordt toegepast, kan niet zomaar worden aangenomen dat dezelfde effectiviteit wordt bereikt als de interventie minder intensief is, wordt uitgevoerd door intern personeel dat niet getraind of minder gemotiveerd is, in een setting die onder druk staat, of wanneer kerncomponenten van de interventie worden weggelaten.

Handige bronnen om de kwaliteit en overdraagbaarheid van het bestaand bewijs te beoordelen en te bepalen of de (her-)evaluatie van een interventie nodig is³⁰:

- ADAPT-richtlijn (114)
- Nederlandse richtlijn passend bewijs voor preventie (135)

Afweging: meerwaarde van een effectevaluatie voor praktijk en beleid

Een effectevaluatie moet een duidelijke meerwaarde hebben. Omdat tijd en middelen beperkt zijn, is het noodzakelijk om gerichte keuzes te maken. Dit vraagt om een afweging tussen de inspanning en kosten van evaluatie (bv. tijdsinvestering, belasting voor evaluatoren en deelnemers, benodigde middelen) en de verwachte opbrengst aan kennis en praktijkrelevantie.

Interventies komen meer in aanmerking voor een effectevaluatie wanneer meerdere van onderstaande criteria gelden:

- De evaluatie kan de inhoud of uitvoering van de interventie verbeteren.
- De evaluatie ondersteunt belangrijke beleidsbeslissingen.
- Bestaand bewijs is beperkt, inconsistent of van lage kwaliteit.
- De evaluatie levert nieuwe inzichten op, bv. over werkzame elementen van de interventie of duurzaamheid van effecten.
- De evaluatie kan aantonen of en hoe de interventie gezondheidsongelijkheden kan verminderen.
- De evaluatie kan aantonen of een effectieve interventie overdraagbaar is naar andere contexten.
- Er bestaat risico op negatieve of ongewenste effecten die de evaluatie kan identificeren.

³⁰ Aanbeveling voor onderzoek en beleid: Ontwikkel een gestandaardiseerde methode om te beoordelen in welke mate interventies uit internationale literatuur of andere contexten vertaald en verantwoord ingezet kunnen worden in de Vlaamse preventiesector, en om te bepalen in welke mate (her)evaluatie aangewezen is op basis van de mate van vergelijkbaarheid en de doorgevoerde adaptaties.

7.3.2 Is de interventie klaar voor een grondige effectevaluatie?

Zelfs wanneer een effectevaluatie noodzakelijk en zinvol lijkt, moet worden nagegaan of de interventie **evalueerbaar** is. Dit betekent dat er voldoende voorwaarden aanwezig zijn om een betrouwbare en betekenisvolle evaluatie uit te voeren. De voorgaande fasen leveren vaak al belangrijke informatie om deze inschatting te maken.

Kernvragen om de evalueerbaarheid na te gaan:

- Is de interventie duidelijk beschreven en reproduceerbaar (doelen, doelgroep, activiteiten, verwachte effecten)?
- Is er een plausibele Theory of Change beschikbaar?
- Wordt de interventie voldoende geïmplementeerd zoals bedoeld?
- Zijn er geschikte bestaande gegevens beschikbaar of kunnen relevante gegevens op een haalbare manier verzameld worden?
- Zijn stakeholders bereid en in staat om mee te werken aan de evaluatie en resultaten te gebruiken?

Belang van intermediairs en implementatoren: Intermediairs en implementatoren spelen een cruciale rol in de dataverzameling bij de einddoelgroep. Het is daarom belangrijk om vooraf af te stemmen welke vormen van effectevaluatie bij de einddoelgroep zij als haalbaar zien en welke randvoorwaarden daarbij gelden.

Dynamisch proces: Evalueerbaarheid is geen statisch gegeven. De omstandigheden kunnen veranderen, waardoor regelmatige herevaluatie aanbevolen is.

Bijlage 6 bevat een beoordelingsinstrument dat kan helpen om te bepalen of de interventie evalueerbaar is.

7.3.3 Besluitvorming of een grondige effectevaluatie haalbaar en de investering waard is

Fase 3 resulteert in een besluit over de vraag of een effectevaluatie haalbaar, zinvol en de investering waard is. Dit besluit komt tot stand via een **reflectief overleg tussen de betrokkenen van de interventie en het beleid**, waarbij wordt afgewogen of de verwachte inzichten opwegen tegen de benodigde middelen.

Het proces leidt tot drie mogelijke uitkomsten. Bij twijfel kan het besluit op elk moment opnieuw worden herzien, zodat de afweging opnieuw kan worden gemaakt zodra randvoorwaarden of informatie veranderen.

Uitkomst 1: Effectevaluatie is haalbaar en levert een duidelijke meerwaarde op

→ Ga door naar Fase 4

- De evaluatie is praktisch uitvoerbaar.
- Er is een duidelijke reden om effectevaluatie uit te voeren.
- De baten van de evaluatie (kennis, beleidsondersteuning, praktijkverbetering) rechtvaardigen de inzet van tijd en middelen.

Uitkomst 2: Effectevaluatie levert momenteel beperkte meerwaarde op

→ Geen uitvoerige effectstudie, wel structurele evaluatie.

- Er is reeds voldoende overtuigend en overdraagbaar bewijs beschikbaar.
- Aanvullende effectevaluatie zal waarschijnlijk weinig nieuwe inzichten opleveren.

- De inspanning en middelen die nodig zijn, wegen momenteel niet op tegen de verwachte baten.
- Structurele evaluatie wordt wel uitgevoerd.
- Merk op: Als er vanuit de structurele evaluatie signalen zijn dat de interventie minder effectief is dan aangenomen op basis van beschikbaar bewijs (bv. door uitblijven van veranderingen, veranderde contextfactoren) kan de beslissing steeds worden herbekeken.

Uitkomst 3: Effectevaluatie levert duidelijke meerwaarde maar randvoorwaarden ontbreken

→ Eerst bijsturen en later heroverwegen.

- Randvoorwaarden ontbreken (bv. stabiele en reproduceerbare implementatie, toegang tot gegevens, gedeeld begrip tussen betrokkenen).
- Een effectevaluatie zou op dit moment weinig betrouwbare inzichten opleveren, waardoor de investeringskosten niet gerechtvaardigd zijn.
- Noodzakelijke bijstellingen moeten eerst worden doorgevoerd, daarna kan de afweging opnieuw worden gemaakt.

7.4 Fase 4: Effectevaluatie volgens een duidelijk evaluatieplan

Een effectevaluatie richt zich op de vraag **of en in welke mate een interventie bijdraagt aan de beoogde effecten**. Soms volstaat een op zichzelf staande effectmeting, bijvoorbeeld wanneer de kernvraag enkel betrekking heeft op de effectiviteit, los van de invloed van context of uitvoering. In de meeste praktijksituaties bestaat echter nood aan inzicht in zowel de effectiviteit **als** de rol van uitvoerings- en contextfactoren, of in de onderliggende werkzame mechanismen.

Daarom is het aangewezen een effectevaluatie te **combineren met een procesevaluatie**. Deze combinatie helpt verklaren *hoe, voor wie en onder welke omstandigheden* effecten tot stand komen. Ze maakt de resultaten bovendien beter interpreteerbaar en bruikbaar voor praktijk en beleid.

Elke grondige effectevaluatie vertrekt vanuit een **duidelijk evaluatieplan** waarin gemaakte keuzes transparant worden vastgelegd. Dit evaluatieplan wordt voorafgaand aan de uitvoering gedeeld en besproken met het beleid. Eventuele afwijkingen tijdens de uitvoering worden eveneens gemotiveerd en afgestemd.

Het evaluatieplan bevat minimaal de volgende kernonderdelen:

- 1) Nood en context
- 2) Partnerschappen en participatie
- 3) Evaluatiedoelstellingen en onderzoeksvragen
- 4) Onderzoeksopzet
- 5) Dataverzameling en verwerking
- 6) Ethische en juridische aspecten
- 7) Gebruik en disseminatie van resultaten
- 8) Planning en middelen

In de volgende paragrafen worden deze onderdelen kort toegelicht, met bijhorende aanbevelingen voor een kwaliteitsvolle uitvoering.

7.4.1 Nood en context

De nood aan een grondige effectevaluatie vloeit voort uit het besluit van Fase 3.

Het evaluatieplan start met het beantwoorden van twee kernvragen:

- **Kennisbasis:** Wat is er al bekend uit literatuur, praktijk en eerdere evaluaties? (o.b.v. Fase 1)
- **Kennishiaten:** Welke vragen zijn nog onbeantwoord en moeten verder onderzocht worden?

7.4.2 Partnerschappen en participatie

Het evaluatieplan beschrijft hoe samenwerking en participatie worden georganiseerd en geborgd door volgende vragen te beantwoorden:

- Wie zijn de rechtstreeks betrokken partners (bv. partnerorganisatie, terreinorganisaties, Gezondheidsmakers, implementatiepartners, onderzoekers)?
- Wie neemt welke rol en verantwoordelijkheid op in de evaluatie?
- Hoe wordt samenwerking gestructureerd (bv. stuurgroep, afsprakenkader)?
- Welke andere stakeholders (bv. verenigingen, doelgroepvertegenwoordigers, intermediairs) worden betrokken, en op welke momenten (bv. onderzoeksdesign, dataverzameling, interpretatie van resultaten...)?

Houd vanaf de start rekening met de noden en prioriteiten van stakeholders.

Betrek bij de ontwikkeling van het evaluatieplan stakeholders die een goed zicht hebben op de haalbaarheid van de evaluatie in de praktijk, zoals de begrijpelijkheid en uitvoerbaarheid van de evaluatie-instrumenten en methoden.

Breng eveneens in kaart wat stakeholders belangrijk vinden in de evaluatie. Dit helpt om een evaluatie te ontwerpen die gedragen is en betekenisvolle resultaten oplevert. Verzamel informatie zoals:

- Welke resultaten vinden stakeholders het belangrijkst?
- Welke evaluatievragen willen zij beantwoord zien en waarom?
- Welke datatypen (kwalitatief, kwantitatief) hebben hun voorkeur?
- Hoe definiëren zij succes van de interventie?
- Hoe willen zij de resultaten gebruiken?

Leg rollen en verantwoordelijkheden vooraf duidelijk vast.

Dit voorkomt misverstanden en maakt het eenvoudiger om problemen tijdig te signaleren. Controleer periodiek of de rolverdeling nog effectief is en pas deze aan indien nodig.

Bijlage 1 geeft een algemeen overzicht van de rolverdelingen en verantwoordelijkheden bij effectevaluatie.

Het invulblad in *bijlage 7* kan worden gebruikt om per partner of stakeholder de betrokkenheid in iedere stap vast te leggen.

7.4.3 Evaluatiedoelstellingen en onderzoeksvragen

Beschrijf in het evaluatieplan zowel de evaluatiedoelstellingen als de onderzoeksvragen.

Evaluatiedoelstellingen: Waarom evalueren we?

Voor de start van een effectevaluatie moet helder zijn waarom de evaluatie wordt uitgevoerd en waarvoor de resultaten zullen worden gebruikt. Belangrijk zijn vragen zoals:

- Wie zal de resultaten gebruiken?
- Welke informatie is nodig?
- Welke acties kunnen op basis van de bevindingen worden ondernomen?

Onderzoeksvragen: Wat evalueren we?

De onderzoeksvragen concretiseren het doel en bakenen de grenzen van de evaluatie af. Mogelijke onderzoeksvragen zijn:

- In welke mate bereikt de interventie de beoogde veranderingen (specifieer) bij de doelgroep?
- Voor welke subgroepen (bv. leeftijd, risicogroepen) is de interventie effectief of effectiever? (subgroepanalyse, differentiële effectiviteit)
- Blijven effecten behouden na afloop? (duurzaamheid, lange termijn effectiviteit)
- Hoe verhoudt de effectiviteit van interventie X zich tot interventie Y?

Sommige vragen vereisen een combinatie van proces- en effectevaluatie, zoals:

- Welke onderdelen of mechanismen dragen bij aan effectiviteit? (mediatie-analyse)
- Hoe hangt implementatietrouw samen met effectiviteit?
- Leidt een hogere deelname-intensiteit tot sterkere effecten? (dosis-respons relatie)

7.4.4 Onderzoeksopzet

Gebruik van een evaluatiekader

Beschrijf in het evaluatieplan of er een specifiek evaluatiekader wordt gebruikt. Zo ja, geef aan welk kader en motiveer de keuze: waarom is dit kader geschikt voor deze evaluatie. Zo nee, leg uit waarom er geen kader wordt toegepast en hoe de evaluatie desondanks gestructureerd zal worden.

Een evaluatiekader zorgt voor een systematische en transparante aanpak, vooral bij gecombineerde proces- en effectevaluatie. De keuze van het evaluatiekader is afhankelijk van de aard van de interventie, de context en de evaluatiedoelstellingen. Voorbeelden van veelgebruikte kaders voor effectevaluatie in combinatie met procesevaluatie zijn:

- **RE-AIM:** combineert op praktische wijze procesevaluatie met een effectuitkomst, maar is beperkt in het verklaren van effecten.
- **PRISM:** voegt contextfactoren toe die RE-AIM-uitkomsten verbeteren of belemmeren.
- **MRC-framework:** geschikt voor de evaluatie van complexe interventies door zowel theorie, effectevaluatie als procesevaluatie te integreren.
- **Realist Evaluation:** is een theoriegebaseerde benadering die onderzoekt wat, voor wie en onder welke omstandigheden werkt, waardoor uitkomsten worden gekoppeld aan verklarende factoren.

- **Contribution Analysis:** is eveneens een theoriegebaseerde benadering die onderzoekt in welke mate een interventie plausibel bijgedragen heeft aan waargenomen effecten, rekening houdend met andere factoren.

Bijlage 8 geeft een uitgebreider overzicht van bestaande evaluatiekaders die bruikbaar zijn voor effect-evaluatie in combinatie met procesevaluatie.

Kies een passend onderzoeksdesign

Beschrijf in het evaluatieplan het onderzoeksdesign dat wordt toegepast. Licht de methodologieën toe, zoals kwantitatieve, kwalitatieve of gemengde methoden, en motiveer waarom deze passend zijn voor de evaluatie.

Het kiezen van een geschikt onderzoeksdesign is cruciaal om betrouwbare en bruikbare resultaten te verkrijgen. Het design moet niet alleen methodologisch robuust zijn, maar ook haalbaar zijn binnen de context van de interventie.

Een goed onderzoeksdesign vindt een evenwicht tussen interne validiteit (de zekerheid dat een waargenomen effect aan de interventie toe te schrijven is) en externe validiteit (de mate waarin de bevindingen generaliseerbaar zijn buiten de onderzoekscontext). De effectiviteit van de interventie wordt daarom idealiter **in realistische settings onderzocht**.

Voor effectiviteitsonderzoek ligt de nadruk op **kwantitatief onderzoek**, zonder hierbij de meerwaarde van kwalitatief onderzoek te willen onderschatten. Tabel 13 geeft een overzicht van types kwantitatieve onderzoeksdesigns, inclusief hun kenmerken, voordelen en beperkingen. Dit helpt om een weloverwogen keuze te maken die past bij de evaluatiedoelstellingen, beschikbare middelen en praktische omstandigheden.

Tabel 13 Overzicht van types kwantitatieve onderzoeksdesigns.

Type design	Kenmerken	Voordelen	Beperkingen
Pre-experimenteel	Metingen zonder tegenfeitelijke situatie	Laagdrempelig, nuttig in vroege fase	Beperkte bewijskracht, geen causale uitspraken
Quasi-experimenteel	Tegenfeitelijke situatie (bv. controlegroep, herhaalde metingen) zonder randomisatie	Goede balans tussen haalbaarheid en validiteit, maakt uitspraken over effect mogelijk	Complexer, vereist meer middelen en planning
Experimenteel (RCT)	Randomisatie, controlegroep	Hoogste interne validiteit, overtuigend bewijs van causaliteit	Hoge kosten, gespecialiseerde expertise vereist, strikte standaardisering niet geschikt voor complexe contextafhankelijke interventies, beperkte externe validiteit

Een **tegenfeitelijke situatie** (*counterfactual*) in kwantitatief onderzoek is een hypothetisch scenario dat laat zien wat er zou zijn gebeurd als een interventie niet had plaatsgevonden. Door het verschil tussen de geobserveerde resultaten en de tegenfeitelijke situatie te bekijken, kunnen evaluatoren onderzoeken wat daadwerkelijk aan de interventie toe te schrijven is. Er zijn verschillende manieren om een tegenfeitelijke situatie te benaderen, zoals controlegroepen, synthetische controlemethoden en trendanalyses voor en na de interventie om het verwachte verloop zonder interventie te schatten.

Bijlage 9 biedt een uitgebreider overzicht van specifieke kwantitatieve onderzoeksdesigns.

Naast kwantitatieve effectmetingen zijn ook **kwalitatieve benaderingen** waardevol. Een **mixed-methods design** versterkt de **interpretatie van resultaten** en zorgt voor een **vollediger en betekenisvoller beeld van effectiviteit**. Het biedt inzicht in ervaringen en perspectieven van deelnemers, en in contextfactoren die van invloed zijn op de uitkomsten. Hierdoor kunnen (onverwachte) resultaten beter worden verklaard en kan de interventie gericht worden verbeterd. Kwalitatieve designs kunnen in alle fasen van een studie worden ingezet.

7.4.5 Dataverzameling en verwerking

Beschrijf in dit onderdeel hoe gegevens worden verzameld, verwerkt en geïnterpreteerd.

- Hoe worden de onderzoekseenheden afgebakend, geselecteerd en gerekruteerd?
- Welke indicatoren worden gebruikt?
- Welke databronnen en dataverzamelmethode worden gebruikt?
- Welke meetinstrumenten worden ingezet?
- Hoe worden de data geanalyseerd en geïnterpreteerd?

In de volgende paragrafen gaan we dieper in op elk van deze aspecten.

Onderzoekseenheden en rekrutering

De onderzoekseenheid is het niveau **waarop effecten worden gemeten**. Dit kan gaan om individuele deelnemers of om clusters zoals regio's, organisaties of populaties. De keuze hangt af van het doel van de interventie en het type effect dat onderzocht wordt.

Bij complexe interventies die via intermediairs of organisaties verlopen, kan een evaluatie **meerdere niveaus** van onderzoekseenheden en uitkomsten omvatten zoals de einddoelgroep (bv. leerlingen: mentaal welzijn), intermediairs (bv. leerkrachten: vaardigheden om mentaal welzijn te bespreken) en de organisatie (bv. schoolklimaat).

In het evaluatieplan worden volgende vragen beantwoord:

- Welke onderzoekseenheden worden onderzocht (bv. individuen, settings, organisaties, populaties)?
- Welke in- en exclusiecriteria gelden?
- Hoeveel onderzoekseenheden worden er naar verwachting opgenomen in de evaluatie? Licht toe hoe dit aantal bekomen wordt (bv. steekproefgrootte berekening).
- Hoe verloopt de rekrutering?

Indicatoren

Een kwaliteitsvolle evaluatie start met het zorgvuldig selecteren van indicatoren. Indicatoren bij effectevaluatie zijn **concrete, meetbare uitkomsten of tussenliggende factoren** die aantonen of de interventie werkt.

De gekozen indicatoren moeten **aansluiten bij zowel de doelstellingen van de evaluatie als het Logic Model** van de interventie. Beperk het aantal indicatoren tot wat nodig is om relevante inzichten te verkrijgen en de evaluatie uitvoerbaar te houden.

Maak bij de selectie van de indicatoren een onderscheid tussen **primaire indicatoren**, die rechtstreeks het hoofddoel van de interventie weerspiegelen, en **secundaire indicatoren**, die indirecte of ondersteunende effecten meten.

Waar mogelijk is het zinvol om indicatoren **af te stemmen op bestaande populatie-indicatoren en gezondheidsdoelstellingen**. Dit verhoogt de vergelijkbaarheid tussen interventies en maakt duidelijk hoe interventies bijdragen aan bredere beleidsdoelen.

Databronnen en dataverzamelmethode

Databronnen verwijzen naar de herkomst van de gegevens die in de evaluatie worden gebruikt, zoals: registers, bevolkingsonderzoeken, documenten of observaties.

De dataverzamelmethode beschrijven hoe deze gegevens worden verkregen, bijvoorbeeld via surveys, interviews, focusgroepen, observaties of documentanalyse.

Waar mogelijk is afstemming met **secundaire, reeds bestaande, databronnen** gewenst, omdat dit tijd en middelen bespaart en toegang biedt tot grotere populaties of langere tijdreeksen. Tegelijkertijd zijn dergelijke databronnen vaak niet ontworpen voor effectonderzoek, waardoor het volgen van dezelfde groep personen, de beschikbaarheid van voldoende meetmomenten en het toeschrijven van veranderingen aan de interventie beperkt kan zijn.

Vaak is (ook) **primaire dataverzameling** nodig, wanneer bestaande bronnen niet beschikbaar zijn of niet alle benodigde informatie bevatten. Het combineren van meerdere databronnen (= triangulatie) verhoogt de betrouwbaarheid van de evaluatie.

Meetinstrumenten

Meetinstrumenten zijn de **concrete tools** waarmee gegevens worden verzameld, zoals vragenlijsten, schalen, logboeken, dagboeken of biometrische meetinstrumenten (bv. activity trackers of weegschalen).

Het gebruik van **gevalideerde meetinstrumenten** verdient de voorkeur, omdat deze getest zijn op betrouwbaarheid en validiteit. Dit verhoogt de validiteit van de gegevens, de vergelijkbaarheid met andere studies en de generaliseerbaarheid van de resultaten.

In sommige situaties is het wenselijk om **aangepaste of zelfontwikkelde instrumenten** te gebruiken, afgestemd op de specifieke doelgroep en context. Het is een sterkte wanneer de doelgroep zelf betrokken wordt bij de ontwikkeling van het meetinstrument, waarbij ze de (gewenste) impact definiëren op basis van ervaringskennis. Zelfontwikkelde instrumenten kunnen, indien gevraagd, gedeeld worden binnen de sector om kennisuitwisseling te bevorderen.

Afstemming met gezondheidsdoelstellingen op populatieniveau kan worden gerealiseerd door gebruik te maken van dezelfde meetinstrumenten (bv. dezelfde schaal om mentaal welzijn te meten) of operationalisering, mits deze geschikt zijn voor de context en doelgroep van de interventie.

Data-analyse

Data-analyse omvat het verwerken van de verzamelde gegevens, het toepassen van statistische of kwalitatieve analysemethoden, het interpreteren van de resultaten en het trekken van conclusies over de effecten en werkingsmechanismen van de interventie.

Bij het plannen van de analyse is het essentieel om vooraf vast te leggen hoe de data worden gestructureerd en gecodeerd, welke analysemethoden worden toegepast en of subanalyses op specifieke subgroepen

nodig zijn. Daarnaast moet duidelijk zijn wie de analyse zal uitvoeren en of de vereiste expertise aanwezig is.

Afhankelijk van het type data kunnen verschillende **analysemethoden** worden ingezet, zoals descriptieve en toetsende statistiek, regressiemodellen, mixed-effects modellen (bijvoorbeeld deelnemers binnen clusters) en thematische analyse voor kwalitatief onderzoek.

Bijlage 10 biedt een invulmodel van een evaluatiematrix dat evaluatoren kan ondersteunen om de evaluatievragen te koppelen aan de centrale concepten, indicatoren, databronnen, dataverzamelingsmethoden, meetmomenten en analysemethoden.

Interpretatie

Het evaluatieplan beschrijft ook hoe belanghebbenden betrokken worden bij het **interpretatieproces**. Interpretatie is het moment waarop betekenis wordt gegeven aan de resultaten. Dit gebeurt idealiter in samenwerking met stakeholders, zodat sterktes, beperkingen en implicaties van de bevindingen in de context van de interventie besproken kunnen worden. Belangrijke reflectievragen zijn: hoe verhouden de resultaten zich tot de verwachtingen, en welke aanpassingen of verbeteringen zijn nodig op basis van de bevindingen?

Het analyse- en interpretatieproces mondt uit in **conclusies** die duidelijk maken wat werkt, wat verbeterd kan worden en hoe de resultaten aansluiten bij de oorspronkelijke doelstellingen.

7.4.6 Ethische en juridische aspecten

Bij het opstellen van een evaluatieplan is het essentieel om aandacht te besteden aan ethische en juridische aspecten. Evaluaties moeten respectvol en veilig worden uitgevoerd voor alle betrokkenen. Dit houdt in dat er geanticipeerd wordt op mogelijke risico's en nadelige effecten, dat privacy en gegevensbescherming worden gewaarborgd en dat relevante wettelijke en sectorale normen strikt worden nageleefd.

Het evaluatieplan beantwoordt daarom onder meer de volgende vragen:

- Zijn er ethische uitdagingen verbonden aan de evaluatie?
- Brengt de interventie of de evaluatie mogelijke risico's of bijwerkingen met zich mee, en hoe worden deze gemonitord of onderzocht?
- Wordt geïnformeerde toestemming gevraagd? Zo ja, hoe wordt dit georganiseerd? Zo nee, wat is de motivatie om dit niet te doen?
- Welke persoonsgegevens worden verzameld, met welk doel, en op basis van welke AVG-rechtsgrond(en)?
- Hoe worden de persoonsgegevens verwerkt, bewaard en beveiligd, en wie heeft toegang tot deze gegevens?
- Is goedkeuring door een ethisch comité vereist? Zo nee, waarom niet?

Elke evaluatie moet voldoen aan de geldende **ethische en juridische standaarden**. Dit betekent onder meer dat deelnemers in de mate van het mogelijke **geïnformeerde toestemming** geven, dat gegevens **vertrouwelijk** worden behandeld en dat deelname geen schade mag berokkenen. De ethische code voor

en door de Vlaamse sector preventieve gezondheid³¹ en de Europese Gedragscode voor Wetenschappelijke Integriteit³² geven hieraan richting.

Elk onderzoek dat prospectief persoonlijke gegevens verzamelt, moet bovendien in overeenstemming zijn met **de Algemene Verordening Gegevensbescherming (AVG), ofwel de General Data Protection Regulation (GDPR)**³³. De AVG/GDPR onderscheidt zes rechtsgronden voor de rechtmatige verzameling en verwerking van persoonsgegevens waarbij de rechtsgronden toestemming, algemeen belang en gerechtvaardigd belang het vaakst gebruikt worden in onderzoek.

Daarnaast is de **vrijwillige en geïnformeerde toestemming van deelnemers** voorafgaand aan hun deelname aan het onderzoek essentieel voor een ethisch verantwoorde evaluatie. Dit betekent dat deelnemers volledig, transparant en begrijpelijk worden geïnformeerd over het doel, de opzet, de procedures en de mogelijke gevolgen van het onderzoek en hun deelname daaraan.

Voor onderzoek dat valt onder de **Belgische wet van 7 mei 2004 inzake experimenten op de menselijke persoon** geldt dat een aanvraag bij een erkend medisch ethisch comité verplicht is en dat de hoofdonderzoeker een gezondheidszorgberoep beoefent. De Experimentenwet omschrijft een experiment als “elke op de menselijke persoon uitgevoerde proef, studie of onderzoek, met het oog op de ontwikkeling van de kennis eigen aan de uitoefening van de gezondheidszorgberoepen”.

Volgende studies vallen niet onder de experimentenwet:

- Onderzoek enkel gericht op praktijkverbetering, niet met het oog op de ontwikkeling van de kennis eigen aan de uitoefening van de gezondheidszorgberoepen (bv. een bevraging bij artsen over de toepassing van richtlijnen in hun dagelijkse praktijk)
- Bevestigingen zuiver in het kader van kwaliteitsmetingen of interne werkorganisatie
- Retrospectieve studies op basis van gegevens uit het verleden die zich in bestaande dossiers, of registraties bevinden
- Onderzoek dat enkel gebruik maakt van anonieme gegevens.

Niet iedere gezondheidsbevorderende interventies is bedoeld om kennis eigen aan de uitoefening van gezondheidszorgberoepen te ontwikkelen, en vaak is een ethische goedkeuring dus niet wettelijk verplicht. Maar, ook wanneer een aanvraag bij een ethisch comité niet wettelijk verplicht is, kan dit zinvol zijn omdat het de bescherming en privacy van deelnemers en de ethische kwaliteit van het onderzoek waarborgt. Bovendien helpt de beoordeling om mogelijke risico's vroegtijdig te identificeren en verhoogt een positief advies het vertrouwen naar deelnemers toe³⁴.

7.4.7 Gebruik en disseminatie van resultaten

Het evaluatieplan beschrijft hoe de resultaten zullen worden benut en gedeeld. Transparante communicatie van evaluatieresultaten bevordert kennisdeling en leren binnen en tussen organisaties. Dit verhoogt de bruikbaarheid van evaluaties en stimuleert evidence-informed beleid en praktijk.

³¹ <https://www.gezondleven.be/files/Gezond-Leven-ethische-code.pdf>

³² ALLEA (2023) Europese Gedragscode voor Wetenschappelijke Integriteit: 2023 Herziene Uitgave (Nederlands). Berlijn. DOI 10.26356/ECOC-Dutch

³³ <https://www.gegevensbeschermingsautoriteit.be/professioneel/avg>

³⁴ Belgische universiteiten beschikken ook over niet-medische ethische adviescommissies voor onderzoek binnen de gedrags-, humane en/of sociale wetenschappen. Voor dergelijk onderzoek gelden vaak instellingsgebonden verplichtingen tot het verkrijgen van ethische goedkeuring, ook wanneer er geen wettelijke verplichting bestaat op federaal niveau.

Bij het opstellen van het plan is het belangrijk om de volgende vragen te beantwoorden:

- Welke output wordt voorzien?
- Welke implicaties hebben de resultaten voor interventieverbetering, praktijk of beleid?
- Hoe zullen de resultaten en inzichten worden gedeeld?

Hefbomen ter bevordering van kennisdeling

Publicatie van resultaten van evaluaties: Maak resultaten toegankelijk via de website van de organisatie en via platforms zoals preventiemethodieken.be. Dit verhoogt transparantie richting beleid en stakeholders, stimuleert kennisdeling en maakt leren van elkaar mogelijk.

Standaardformat voor evaluatierapporten: Het gebruik van een uniform format bevordert consistentie en volledigheid, vergemakkelijkt de vergelijking van resultaten tussen interventies en helpt bij het identificeren van goede praktijken.

Minimale onderdelen van een evaluatierapport

Het rapport wordt opgesteld door de betrokken evaluatoren en bevat minimaal:

- **Algemene informatie:** Titel, betrokkenen (auteurs, evaluatoren), onderzoeksperiode en contact-informatie.
- **Samenvatting:** Kern van de evaluatievragen, methoden, belangrijkste bevindingen en aanbevelingen (± 300 woorden).
- **Doel en evaluatievragen:** Beschrijving van de interventie, doel van de evaluatie en onderzoeksvragen (± 2 pagina's).
- **Design en methoden:** Beschrijving van het onderzoeksdesign, de gebruikte dataverzamelmethode, meetinstrumenten en analysetechnieken, inclusief onderbouwing van de aanpak en sterktes en beperkingen (± 2 pagina's).
- **Resultaten en interpretatie:** Een overzicht van de resultaten met duidelijke antwoorden op de onderzoeksvragen en een interpretatie van de bevindingen (± 5 pagina's, afhankelijk van de omvang van het onderzoek).
- **Conclusies en aanbevelingen:** Een synthese van de belangrijkste conclusies, gevolgd door concrete aanbevelingen en acties. Deze kunnen betrekking hebben op interventieverbetering, aanpassing van de implementatie, wijziging van de Theory of Change, versterking van randvoorwaarden, opschaling of afbouw/stopzetting van de interventie, suggesties voor beleid of vervolgonderzoek (± 4 pagina's).
- **Bijlagen (indien nodig):** Bijvoorbeeld, gedetailleerde beschrijvingen van de rolverdeling, gebruikte instrumenten (zoals vragenlijst of leidraad voor kwalitatief onderzoek), gebruikte materialen voor communicatie en rekrutering, gedetailleerde statistische output.

Het aangegeven aantal pagina's per onderdeel dient als richtlijn en kan variëren afhankelijk van de omvang en complexiteit van het onderzoek.

7.4.8 Planning en middelen

Het evaluatieplan bevat een duidelijke tijdsplanning en een overzicht van de benodigde middelen. Dit helpt om het proces realistisch en haalbaar te maken.

Planning

Beschrijf de duur van de evaluatie en geef een globale tijdslijn van het evaluatieproces. Indien relevant, onderscheid verschillende fases (bijvoorbeeld voorbereiding, dataverzameling, analyse, rapportering).

Budget en middelen

Geef een overzicht van het beschikbare budget en hoe dit wordt ingezet. Denk hierbij aan personeelskosten (VTE, aantal dagen, profielen), externe expertise, kosten voor dataverzameling en materialen. Vermeld ook welk bijkomend budget zou toelaten om de evaluatie te versterken.

7.4.9 Bijsturing op basis van effectevaluatie

Wanneer resultaten afwijken van de oorspronkelijke verwachtingen, kan dit aanleiding geven tot een herziening van de Theory of Change of aanpassingen in de interventie.

Belangrijk om te benadrukken: zonder een parallelle procesevaluatie is het moeilijk te begrijpen waarom een effect wel of niet is opgetreden. Hierdoor wordt het aanpassen van de Theory of Change minder gericht en riskeert men foutieve conclusies.

Bijsturing van de interventie en/of Theory of Change is zinvol wanneer:

- De aannames en rationale relevant en plausibel blijven, maar nog nood hebben aan verfijning of aanpassing.
- Contextuele factoren en drempels beïnvloedbaar zijn.
- Aanpassingen aan de interventie praktisch uitvoerbaar en proportioneel zijn.

Overwegingen om de interventie grondig te herzien, af te bouwen of te stoppen:

- Verschillende implementatiestrategieën en mechanismen voor verandering zijn uitgetoetst, maar verwachte effecten blijven uit.
- De kosten of de complexiteit van benodigde aanpassingen zijn niet proportioneel ten opzichte van het verwachte effect.
- Contextfactoren blijken structureel en niet-beïnvloedbaar, waardoor effectieve implementatie vrijwel onhaalbaar is.

8 Aanbevelingen voor het beleid

Het versterken van effectevaluaties in de preventiesector vraagt meer dan methodologische verfijning op organisatieniveau. Het vereist een beleidskader dat realistische verwachtingen combineert met gerichte ondersteuning en een lerende evaluatiecultuur. Onderstaande aanbevelingen richten zich op beleidskeuzes die organisaties in staat stellen om beter te evalueren en zo meer maatschappelijke impact te realiseren.

8.1 Versterk de wisselwerking tussen beleid, praktijk en onderzoek

Een nauwe samenwerking tussen beleidsmakers, de preventiesector en onderzoekers is cruciaal om ervoor te zorgen dat effectevaluaties zowel maatschappelijk relevant als wetenschappelijk verantwoord zijn.

Aanbeveling 1: Stimuleer kennisdeling- en ontsluiting

De preventiesector genereert waardevolle praktijkkennis en inzichten uit evaluaties, maar deze blijven vandaag vaak impliciet of lokaal verankerd. Om de preventiesector te versterken en kruisbestuiving tussen organisaties te stimuleren, kan beleid inzetten op een structurele aanpak voor kennisdeling en ontsluiting.

Belangrijke hefboomen daarbij zijn:

- **Publicatie van evaluatierapporten:** Maak evaluatierapporten toegankelijk via platforms zoals preventiemethodieken.be. Hierin worden ook de gebruikte onderzoeksmethoden en meet-instrumenten opgenomen. Dit bevordert transparantie en kruisbestuiving.
- **Lerend netwerk voor evaluatie in de preventiesector:** Faciliteer een netwerk waarin actoren uit de preventiesector kennis uitwisselen over verschillende evaluatievormen en methoden, afgestemd op actuele vragen uit de sector.
- **Gedeelde verantwoordelijkheid met een duidelijk mandaat:** Hoewel kennisdeling een gezamenlijke opdracht is, is het belangrijk om één organisatie of contactpersoon aan te wijzen met een formeel mandaat om kennisdeling te coördineren, samenwerking te stimuleren en continuïteit te bewaken.

Aanbeveling 2: Richt een onafhankelijk expertiseforum op

Om de kwaliteit en onafhankelijkheid van effectevaluaties te waarborgen, is het zinvol om een **onafhankelijk expertiseforum of consortium** op te richten, samengesteld uit onderzoekers en praktijk- en beleidsvertegenwoordigers. Dit forum vormt een structureel platform waar praktijkervaring, wetenschappelijke expertise en beleidsprioriteiten samenkomen. De leden engageren zich om jaarlijks een aantal evaluaties te ondersteunen, beoordelen en prioriteren waar nodig. Hun rol kan onder meer bestaan uit:

- Ondersteuning van de besluitvorming voor effectevaluatie in fase 3
- Methodologisch advies inzake opties bij effectevaluatie
- Feedback op evaluatieplannen wat betreft kwaliteit en haalbaarheid
- Analyse van evaluatierapporten en voorstellen van acties op basis van nieuwe inzichten
- Het stellen van algemene beleidsprioriteiten en een onderzoeksagenda in de sector

Aanbeveling 3: Zet in op een duidelijke en gedeelde onderzoeksagenda voor de preventiesector

Beleidsprioriteiten geven richting aan onderzoek en zorgen ervoor dat het maatschappelijk en beleidsmatig relevant is, terwijl onderzoeksresultaten het beleid onderbouwen en ondersteuning bieden bij het nemen van beslissingen. Beleidsmakers kunnen hierin een sturende rol opnemen door samen met de preventiesector en onderzoekers een **gedeelde onderzoeksagenda** op te stellen. Deze agenda brengt kennishiaten in kaart en bepaalt welke gezondheidsthema's en interventies prioritair zijn voor verder onderzoek over organisaties heen. Zo draagt onderzoek sterker bij aan beleidsontwikkeling, in plaats van dat het louter voor verantwoording gebruikt wordt. Om deze agenda duurzaam te verankeren, kunnen de vastgelegde prioriteiten worden opgenomen in de **meerjarenplanning en beheersovereenkomsten** van organisaties in de preventiesector.

Aanbeveling 4: Faciliteer structurele verbinding tussen de preventiesector en onderzoeksinstituten

Samenwerking tussen organisaties uit de preventiesector en onderzoeksinstituten rond effectevaluatie ontstaat vandaag ad hoc, wat kan leiden tot ongelijke toegang tot onderzoeksexpertise. Het structureel versterken van de samenwerking tussen de preventiesector en onderzoeksinstituten verhoogt zowel de methodologische kwaliteit als de onafhankelijkheid van effectonderzoek. Het beleid kan deze verbinding structureel als volgt versterken:

- Voorzie een **toegankelijk overzicht van onderzoekers en onderzoeksgroepen** met expertise in gezondheidsbevordering en preventie.
- Ondersteun **netwerking en matchmaking-initiatieven** die organisaties en onderzoekers samenbrengen op basis van inhoudelijke en/of methodologische noden.

8.2 Kwaliteitsvolle effectevaluaties vragen voldoende middelen en tijd

Effectevaluaties in de preventiesector vragen realistische randvoorwaarden. Zonder voldoende middelen en tijd lopen evaluaties het risico om oppervlakkig of onbetrouwbaar te zijn. Beleidsmakers spelen een sleutelrol in het voorzien van voldoende middelen voor evaluatie.

Aanbeveling 5: Voorzie realistische financiering voor organisaties in de preventiesector

Effectevaluatie is methodologisch, organisatorisch en financieel veeleisend. Zonder voldoende middelen is kwaliteitsvol onderzoek nauwelijks haalbaar, wat kan resulteren in ondermaatse designs of symbolische metingen zonder echte bewijskracht. De **verantwoordelijkheid voor financiering is gedeeld** bij zowel de overheid als de organisaties zelf: beleidsmatig is het essentieel dat de overheid voldoende middelen beschikbaar stelt voor organisaties waarvan verwacht wordt dat ze de effectiviteit van hun interventies evalueren, terwijl organisaties binnen hun mogelijkheden deze middelen gericht dienen in te zetten om dit onderzoek goed uit te voeren.

Aanbeveling 6: Voorzie voldoende tijd en stel realistische verwachtingen

Preventieve effecten tonen zich zelden op korte termijn. Middellange en langetermijneffecten worden pas zichtbaar na meerdere jaren. Goed verwachttingsmanagement is daarom cruciaal. Beleidsmakers kunnen dit ondersteunen door tussentijdse veranderingen in determinanten te erkennen als relevante vooruitgang en een voldoende lange tijdshorizon te voorzien voor de effectevaluatie van gezondheidsbevorderende interventies. Het gebruik van een Theory of Change helpt om deze verwachte veranderingen richting effectiviteit zichtbaar te maken.

8.3 Valueer de beleidsimpact op systeem- en populatieniveau

Beleidsimpactevaluatie in de preventiesector gaat verder dan het optellen van resultaten van afzonderlijke interventies. Beleidskaders, financieringsmechanismen en regelgeving sturen een complex systeem van actoren, interventies en contexten. Het succes van beleid manifesteert zich daarom vooral in **verschuivingen op systeemniveau** (zoals samenwerking, toegankelijkheid en aanbod) en in **veranderingen op populatieniveau**.

Preventieve effecten zijn vaak minder zichtbaar omdat ze problemen voorkomen voordat ze optreden. Het aantonen van deze impact is uitdagend, maar juist daarom van groot belang om het bestaansrecht van een preventief beleid te onderbouwen.

Aanbeveling 7: Gebruik de gezondheidsdoelstellingen als ankerpunt voor beleidsimpact

Gezondheidsdoelstellingen op populatieniveau vormen het centrale referentiepunt voor beleidsimpactevaluatie. Het opvolgen ervan gebeurt via periodieke monitoring van de prevalentie van gezondheids- en gedragsuitkomsten, determinanten en verschillen tussen bevolkingsgroepen.

Interventiedata en populatiedata hebben elk een eigen rol. Het is belangrijk om maatschappelijke trends **niet te reduceren tot het effect van één beleidsmaatregel of interventie**, omdat gezondheid ook beïnvloed wordt door bredere maatschappelijke ontwikkelingen in andere sectoren zoals economische omstandigheden, werk en onderwijs. Hoewel directe causale effecten van interventies op populatieniveau moeilijk vast te stellen zijn, kunnen interventie-evaluaties wel bijdragen aan een **plausibele verklaring** van geobserveerde trends. Methoden zoals contribution analysis zijn bruikbaar in deze context.

Aanbeveling 8: Ontwikkel een Theory of Change op sectorniveau

Een **Theory of Change op sectorniveau** beschrijft hoe beleid via beleidshefbomen en een samenhangende set van interventies bijdraagt aan gezondheidsdoelstellingen. In tegenstelling tot de Theory of Change voor een interventie is een Theory of Change voor het beleid globaler en meer gericht op systeemmechanismen, structurele randvoorwaarden en duurzame impact.

Interventies worden hierin gepositioneerd als acties, gekoppeld aan beleidsdoelen. Naast een overkoepelende Theory of Change kunnen ook **thematische en/of settinggerichte Theory of Change's** worden ontwikkeld (bv. rond mentaal welzijn, beweging of voeding in scholen, werk en gemeenten). Deze thematische benadering laat toe om beleidsdoelen, interventies en specifieke partners te positioneren in

het geheel. Een volgende stap is dan om deze thematische Theory of Change's met elkaar te verbinden, gezien preventiethema's elkaar vaak versterken.

Aanbeveling 9: Werk toe naar een systeem voor collectieve impactmeting

Preventie gebeurt door meerdere acties tegelijk, uitgevoerd door verschillende organisaties in diverse settings. **Collectieve impact** is een samenwerkingsaanpak waarbij meerdere actoren zich richten op complexe maatschappelijke problemen met gedeelde doelen. De aanpak steunt op een gemeenschappelijke agenda, gedeelde meting, elkaar versterkende activiteiten, continue communicatie en een coördinerende organisatie.

Om te begrijpen wat deze gezamenlijke inspanningen opleveren, is een aanpak nodig die verder kijkt dan losse interventies. **Collectieve impactmeting** richt zich op het in kaart brengen van de gezamenlijke bijdrage van verschillende actoren aan brede maatschappelijke doelen, zoals gezondheidsdoelstellingen. Het vaststellen van causale verbanden is in complexe systemen nauwelijks haalbaar: het evalueren van collectieve impact moet dus vooral dienen als instrument voor leren en bijsturen, met ruimte voor experimenteren, reflectie en feedback.

Omdat betrokken organisaties verschillende informatiebehoeften hebben, is het beter om meerdere, elkaar aanvullende evaluatievormen te gebruiken in plaats van één uniform evaluatiekader. Monitoring en meting moeten daarbij pragmatisch worden toegepast: gedeelde indicatoren kunnen helpen bij afstemming, maar moeten eenvoudig en haalbaar zijn om administratieve lasten te beperken. Daarnaast is het belangrijk om ook onbedoelde effecten mee te nemen, zodat het beeld van de gezamenlijke werking volledig is.

In plaats van uitsluitend te focussen op het aantonen van causale effecten van individuele initiatieven, kan het beleid samen met de preventiesector geleidelijk toewerken naar een systeem van collectieve impactmeting door **per thema gedeelde doelen en een beperkte set kernindicatoren vast te leggen**. Dit vergroot het inzicht in gezamenlijke resultaten en hoe dit bijdraagt aan de gezondheidsdoelstellingen, voorkomt versnippering en stimuleert samenwerking.

Collectieve impactmeting vereist **duidelijke coördinatie en intensieve afstemming**. Dit vraagt om een langdurig traject met de preventiesector en experts. Als eerste stap kunnen pilots op thema- of sectorniveau worden opgezet (bv. mentaal welzijn op school) waarin gedeelde doelstellingen en indicatoren in een afgebakende context worden getest en verfijnd.

8.4 Evolueren naar beleid dat effectieve interventies identificeert en versterkt

Effectieve preventieve interventies ontstaan niet vanzelf. Ze vragen om een beleid dat ruimte biedt voor leren, verbeteren en opschalen. In plaats van te verwachten dat interventies onmiddellijk bewezen 'effectief' zijn, is een groeipad nodig waarin evidentie stapsgewijs wordt opgebouwd. Tegelijk moet beleid op lange termijn inzetten op het versterken van interventies die potentieel of aantoonbaar effectief zijn, en werken aan een transparant beoordelingskader. De volgende aanbevelingen schetsen hoe dit kan worden gerealiseerd.

Aanbeveling 10: Voorzie een groeipad voor de onderbouwing van interventies

Hanteer een **realistische en toekomstgerichte visie op effectiviteitsbewijs**. Verwacht niet dat interventies onmiddellijk als ‘effectief’ kunnen worden beoordeeld, maar zorg voor een groeipad waarin evidentie stapsgewijs wordt opgebouwd.

In een eerste fase ligt de focus op haalbare evaluaties die vooral inzicht geven in uitvoerbaarheid, kwaliteit van implementatie en succes- en faalfactoren. Pre-experimentele en pragmatische onderzoeksdesigns zijn in deze fase geschikt om eerste signalen van effectiviteit te detecteren, aannames uit de Theory of Change te toetsen en bijsturing mogelijk te maken.

Tegelijk moeten evaluatoren inschatten wanneer een interventie voldoende rijp is om over te gaan naar meer robuuste effectevaluatie. Zodra haalbaar, wordt het gebruik van onderzoeksopzetten met een tegenfeitelijke vergelijking sterk aanbevolen.

Merk hierbij op dat starten van nul niet nodig is: evaluatoren benutten internationale evidentie over bewezen interventies en gaan na in welke mate deze te vertalen is naar de Vlaamse context en de specifieke context van de setting waarin de interventie geïmplementeerd wordt. Op basis hiervan volgt een proportionele (her)evaluatie van de haalbaarheid en effectiviteit.

Aanbeveling 11: Ontwikkel een beoordelingskader voor preventieve interventies met verschillende niveaus van onderbouwing

Werk op termijn toe naar een **beoordelingskader dat erkent dat interventies zich op verschillende niveaus van onderbouwing bevinden**. Een gelaagd systeem maakt het mogelijk om interventies passend bij hun ontwikkelingsfase te ondersteunen, bijvoorbeeld via niveaus zoals goed beschreven, theoretisch onderbouwd en wetenschappelijk onderzocht of effectief.

Een dergelijk kader stimuleert kwaliteitsvolle praktijkontwikkeling door duidelijk te maken waar interventies staan en welke vervolgstappen realistisch zijn. Dit vraagt om helderheid over wat als ‘passend bewijs’ wordt beschouwd per niveau, evenals om versterking van evaluatiepraktijken binnen de sector. Daarna kan in een breed gedragen traject een realistisch beoordelingskader worden ontwikkeld.

Internationale voorbeelden tonen aan dat een gelaagd beoordelingskader een krachtige hefboom kan zijn voor kennisdeling, transparantie en kwaliteitsverbetering in de preventiesector (bijvoorbeeld: Databank Erkende Leefstijlinterventies van het Nederlandse Loket Gezond Leven³⁵).

³⁵ Zie <https://www.loketgezondleven.nl/leefstijlinterventies/erkenningstraject>

Aanbeveling 12: Ontwikkel een beleidssysteem dat effectieve interventies versterkt

Op langere termijn is het wenselijk om te evolueren naar een beleidssysteem dat actief inzet op het identificeren en versterken van interventies die potentieel of aantoonbaar effectief zijn. Dit vraagt om gedragen keuzes en samenwerking tussen beleid, kennisinstellingen en actoren in de preventiesector.

Mogelijke richtlijnen voor zo een systeem zijn:

- **Gerichte ondersteuning van veelbelovende interventies:** Interventies met sterke onderbouwing of positieve voorlopige resultaten verdienen extra ondersteuning om verder te ontwikkelen en evalueren.
- **Verduurzaming van bewezen effectieve interventies:** Wanneer interventies effectief zijn, moet beleid inzetten op schaalvergroting, structurele verankering en financiering via de beheers-overeenkomsten.
- **Primair inzetten op verbeteren van interventies:** Niet elke interventie levert onmiddellijk het gewenste effect op. Voordat wordt besloten tot stopzetting, moet ruimte worden geboden om de interventie, implementatiefactoren, context en andere flankerende factoren te verbeteren en de interventie opnieuw te onderzoeken.
- **Stopzetten van schadelijke of herhaaldelijk ineffectieve interventies:** Bij aanwijzingen van schadelijke effecten of herhaaldelijke ineffectiviteit is stopzetten noodzakelijk om middelen vrij te maken voor nieuwe beter onderbouwde interventies.

9 Slotwoord

De aanbevelingen schetsen een ambitieus toekomstbeeld en bieden richting voor de verdere ontwikkeling van een sterke evaluatiecultuur in de preventiesector. Ze zijn bedoeld als stip op de horizon, niet als checklist die onmiddellijk en integraal moet worden uitgevoerd. De leidraad zal dan ook geleidelijk worden uitgerold: het kan niet in één keer volledig worden waargemaakt en voor bepaalde lopende trajecten komt het te laat om nog volledig toegepast te worden. In de toekomst moet de leidraad echter steeds meer richtinggevend zijn.

Verandering is complex en vraagt tijd. Structurele vernieuwing is een geleidelijk proces dat inzet op lange termijn en voortdurende dialoog tussen beleid, praktijk en onderzoek vereist. In die rol zal beleid ook mechanismen creëren om de opvolging van de leidraad te monitoren en te ondersteunen. Sommige aanbevelingen zijn al vertaald naar concrete voorstellen tot actie, maar bij elke actie is het belangrijk om te **beoordelen wat realiseerbaar is** met de beschikbare middelen en context, of dat er alternatieve wegen zijn om het onderliggende doel te bereiken.

De leidraad is een startpunt van verandering, geen eindpunt. Verdere afstemming en samenwerking met de sector zijn wellicht nodig om de leidraad effectief te verankeren in de praktijk en om evaluatie structureel te integreren in de werking van organisaties. Indien dit nodig blijkt, kunnen bepaalde aspecten van de leidraad worden bijgestuurd of kunnen aanvullende randvoorwaarden worden gecreëerd die de processen versterken. Een langetermijnvisie is daarbij cruciaal: door continu te leren, bij te sturen en te investeren in gezamenlijke expertise kan de sector steeds beter evalueren en zo bijdragen aan duurzame gezondheidswinst.

10 Bijlagen in leidraad

10.1 Overzicht: Rolverdelingen en verantwoordelijkheden bij effectevaluatie

De aard van de samenwerking hangt sterk af van het type en de fase waarin de interventie zich bevindt, de complexiteit van de evaluatie, de beschikbare expertise en middelen, en de specifieke onderzoeksvraag. Binnen deze variabiliteit zijn er enkele terugkerende rollen die richting geven aan een efficiënte taakverdeling:

Actor	Rollen en verantwoordelijkheden bij effectevaluatie
Beleid	• Creëert de randvoorwaarden voor kwaliteitsvolle evaluaties (middelen, ondersteuning). • Bepaalt de onderzoeksagenda en prioriteiten afgestemd op beleidsnoden en -doelstellingen. • Levert input bij methodologische keuzes van onderzoekers of organisaties waar nodig. • Stimuleert samenwerking en kennisdeling • Gebruikt evaluatieresultaten bij beleidsontwikkeling, programmasturing en financieringsbeslissingen.
Onderzoekers (intern aan organisatie of extern)	• Houden rekening met beleidsnoden en indicatoren die relevant zijn voor besluitvorming • Zorgen voor de planning en uitvoering van evaluaties, rekening houdend met methodologische kwaliteit, validiteit en objectiviteit. • Analyseren en interpreteren data. • Rapporteren helder en transparant de resultaten. • Formuleren onderbouwde conclusies op basis van de bevindingen.
Partnerorganisaties, organisaties met terreinwerking, Gezondheidsmakers	• Houden rekening met beleidsnoden en indicatoren die relevant zijn voor besluitvorming • Brengen praktijkkennis en context in (bv. wat is geschikt voor de doelgroep/setting). • Faciliteren toegang tot deelnemers en dataverzameling indien nodig. • Denken mee over interpretatie van resultaten • Gebruiken de resultaten om interventie te optimaliseren.

Andere actoren met terreinwerking en intermediairs	• Brengen praktijkkennis en context in (bv. wat is geschikt voor de doelgroep/setting). • Faciliteren toegang tot deelnemers en dataverzameling indien nodig.
---	--

10.2 Voorbeelden van een uitgewerkte Theory of Change en bronnen ter ondersteuning bij de ontwikkeling

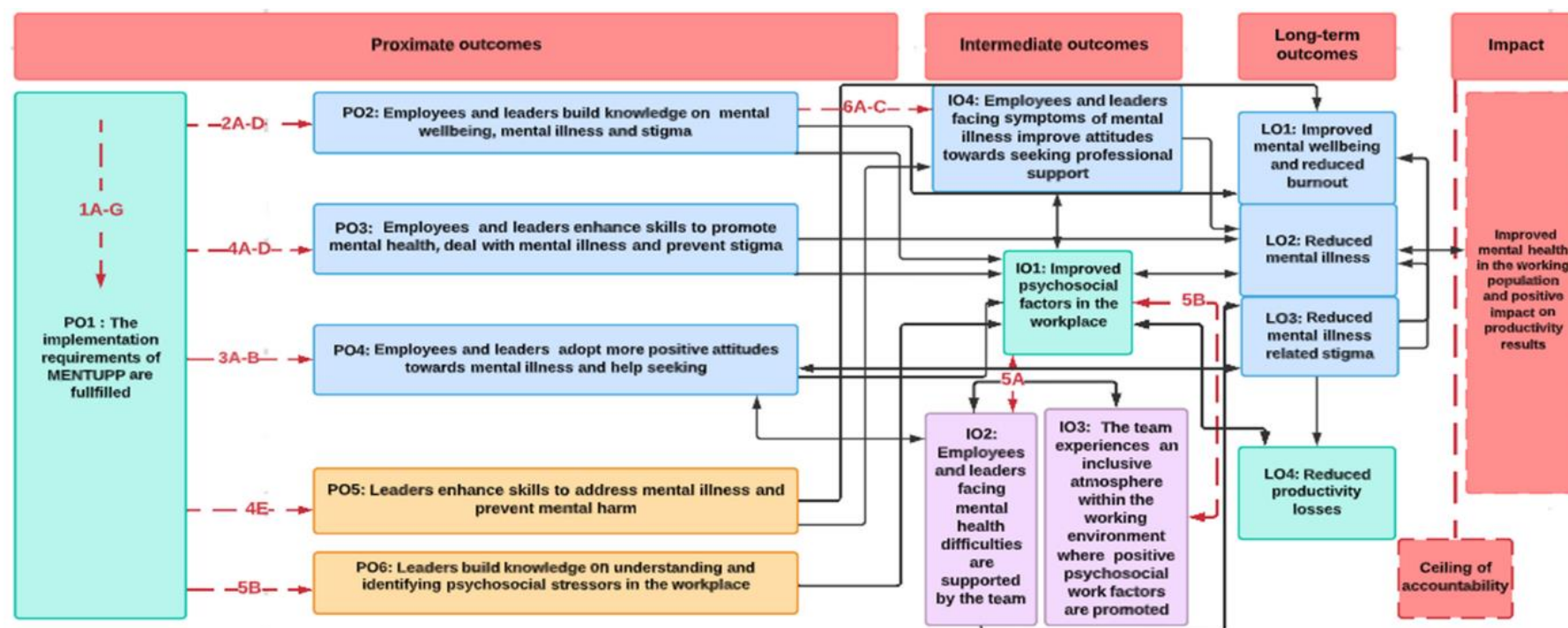
Voorbeeld: Theory of Change van de MENTUPP-interventie

Een eerste voorbeeld van hoe een Theory of Change in de praktijk wordt toegepast binnen een complexe gezondheidsbevorderende interventie, is te vinden in het Europese MENTUPP-project. Dit project richt zich op het bevorderen van mentale gezondheid in Kleine en Middelgrote Ondernemingen (KMO's). De Theory of Change werd ontwikkeld volgens een participatief proces met experts en partners uit negen landen, en maakt de causale verbanden tussen interventieactiviteiten en verwachte uitkomsten inzichtelijk.

De Theory of Change maakt onderscheid tussen directe en indirecte effecten van interventiecomponenten. Zo tonen rode pijlen in het schema de directe causale verbanden aan tussen interventiecomponenten en resultaten, terwijl zwarte pijlen de indirecte effecten aanduiden. De aannames en rationale zijn niet opgenomen in de visual, maar in aparte tabellen in het artikel.

Tot slot werd de Theory of Change gebruikt als basis voor het evaluatieproces: per causale keten werden indicatoren ontwikkeld (zoals attitudes ten opzichte van hulp zoeken en sociale steun van collega's) die met gevalideerde meetinstrumenten werden gemeten. Dit maakt het mogelijk om hypothesen te toetsen over hoe veranderingen op korte en middellange termijn bijdragen aan langetermijneffecten en de uiteindelijke impact.

Bron: Tsantila, F., Coppens, E., De Witte, H. *et al.* Developing a framework for evaluation: a Theory of Change for complex workplace mental health interventions. *BMC Public Health* 23, 1171 (2023). <https://doi.org/10.1186/s12889-023-16092-x>



Levels of change	Intervention components	4.Enhancing skills
SME level	1.Implementation requirements	4A.Generic module: toolkit-mental wellbeing and stress management for employees and leaders
Employee & leaders level	1A. Initiate communication to recruit SME	4B.Generic module:interactive learning activities -depression and anxiety for employees and leaders
Leaders only level	1B.Secure commitment from senior management	4C.Sector-specific module:understanding stigmatisation of mental health issues in the workplace for employees and leaders
Team level	1C.Assessment to meet the the minimum requirements	4D.Generic module:challenging stigma for people with mental health issues for employees and leaders
	1D.Conduct introductory session	4E.Generic module: toolkit- how to address mental illness in the workplace for leaders
	1E. Inform about what the implementation entails for the SME	5.Improving psychosocial factors
	1F.Invite participants to join the MENTUPP Hub platform	5A.Generic module:toolkit- supporting each other at work for employees and leaders
	1G.Establish a pilot palnning group to promote engagement	5B.Sector-specific module: toolkit-creating mentally healthy workplaces for leaders
	2.Building knowledge	6.Arranging additional support
	2A.Generic module: understanding depression for employees and leaders	6A.iFightDepression tool (optional)
	2B.Generic module: undersanding anxiety for employees and leaders	6B.iFightDepression website
	2C.Generic module: test your stigma for employees and leaders	6C. National helplines contact details
	2D.Sector-specific module: understanding mental wellbeing, stress and burnout for employees and leaders	
	3.Adopting more positive attitudes	
	3A.Generic module: understanding stigma and discrimination for employees and leaders	
	3B.Generic module:challenging stigma for employees and leaders	

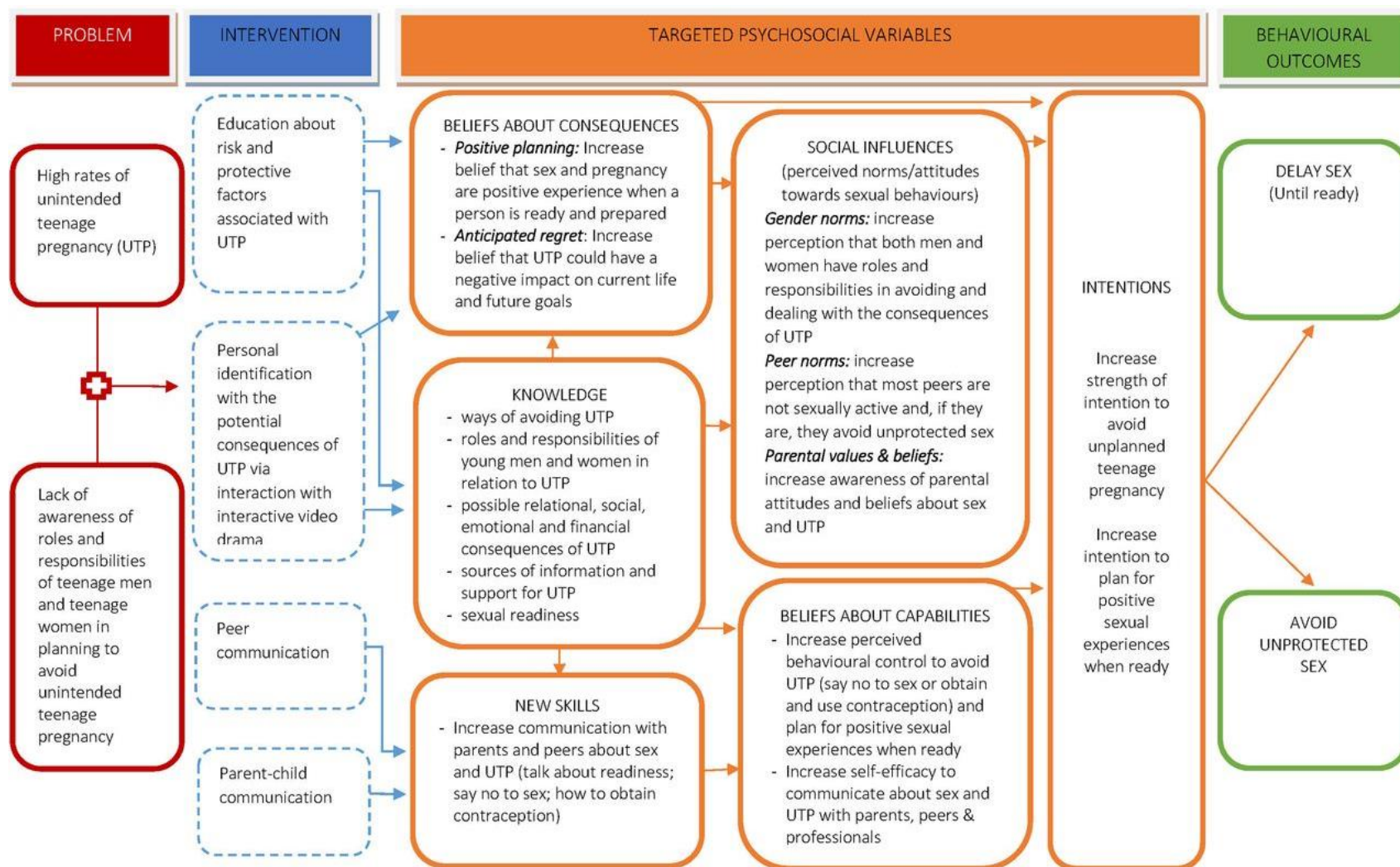
Voorbeeld: If I Were Jack interventie voor seksuele opvoeding door leerkrachten

De interventie If I Were Jack is een door leerkrachten uitgevoerde interventie voor relationele en seksuele opvoeding, gericht op het voorkomen van onbedoelde tienerzwangerschappen (136). De Theory of Change voor werd ontwikkeld op basis van theoretische fundamenteën (Theory of Planned Behaviour), bestaande evidentie en de psychosociale mechanismen die de interventie beoogt te beïnvloeden.

De Theory of Change beschrijft een causale keten die de interventiecomponenten verbindt met veranderingen in psychosociale variabelen en, uiteindelijk, met gedraguitkomsten. De Theory of Change dient als kader voor de hypothesen en voor de selectie van uitkomstmaten (zoals kennis, attitudes, communicatievaardigheden en intenties). Deze uitkomsten worden gemeten om te beoordelen of de door de veronderstelde mechanismen daadwerkelijk zijn beïnvloed.

Merk op dat in de Theory of Change geen expliciet geformuleerde aannames zijn opgenomen. Wel wordt in het artikel een inhoudelijke rationale voor de veronderstelde causale verbanden beschreven.

Bron: Lohan, M., Aventin, Á., Clarke, M., Curran, R. M., Maguire, L., Hunter, R., McDowell, C., McDaid, L., Young, H., White, J., Fletcher, A., French, R., Bonell, C., Bailey, J. V., & O'Hare, L. (2018). JACK trial protocol: a phase III multicentre cluster randomised controlled trial of a school-based relationship and sexuality education intervention focusing on young male perspectives. *BMJ open*, 8(7), e022128. <https://doi.org/10.1136/bmjopen-2018-022128>



Enkele helpende bronnen en bijkomende voorbeelden bij het ontwikkelen van een Theory of Change zijn:

- World Health Organization. (2025). How to develop an evidence-informed theory of change for health: WHO technical guidance. <https://www.who.int/publications/i/item/9789240097063>
- The Evaluation Taskforce (UK Cabinet Office). Module 2: Developing a Theory of Change. In: Resources from the Evaluation Task Force Evaluation Academy training programme (2025). <https://www.gov.uk/government/publications/etf-evaluation-academy-20-resources>
- The Evaluation Support Team (Government Analysis Function). The Theory of Change Process – Guidance for Outcome Delivery Plans <https://analysisfunction.civilservice.gov.uk/policy-store/the-analysis-function-theory-of-change-toolkit/>
- James Noble (2019). Theory of change in ten steps. <https://www.thinknpc.org/resource-hub/ten-steps/>
- Max Stanford (2024). Running a theory of change workshop: A quick reference for workshop facilitators. <https://www.eif.org.uk/resource/running-a-theory-of-change-workshop-a-quick-reference-for-workshop-facilitators>
- De Silva, M., Lee, L., & Ryan, G. (2014). Using Theory of Change in the development, implementation and evaluation of complex health interventions. *London: the centre for global mental health at the London school of hygiene & tropical medicine and the mental health innovation network.* https://www.mhinnovation.net/sites/default/files/content/document/MHIN%20ToC%20guidelines_May_2015_0.pdf
- Er bestaan verschillende [software](#) die de visualisatie van een Theory of Change kunnen ondersteunen

10.3 Tool: Invulmodel Logic Model

De organisaties en evaluatoren kunnen deze tabel gebruiken als invulmodel om na te denken over de samenhang tussen **geplande inputs en activiteiten**, en de **beoogde outputs en uitkomsten** van een interventie: *“Als we deze inputs hebben, kunnen we deze activiteiten uitvoeren. Als we deze activiteiten uitvoeren, leveren we deze outputs en verwachten we deze uitkomsten.”* Neem concreet meetbare indicatoren op in het model.

Inputs	Activiteiten	Outputs	Uitkomsten (korte termijn)	Uitkomsten (middellange termijn)	Impact (lange termijn)
Geld Tijd Middelen ...	Interventies Voorlichting Campagne Workshops ...	Bereik Aantal sessies Waardering Uitvoering Samenwerking ...	Kennis Motivatie Bewustwording ...	Gedrag Praktijkvoering ...	Gezondheid Maatschappij ...

10.4 Overzicht: Definities en voorbeelden van belangrijke procesuitkomsten

Deze tabel geeft een niet-exhaustief overzicht van mogelijke procesuitkomsten, hun definities en voorbeelden van indicatoren. Het is **niet de bedoeling, noch mogelijk**, om alle uitkomsten tegelijkertijd op te nemen. Organisaties bepalen zelf welke uitkomsten relevant zijn voor de interventie en welke op dat moment praktisch haalbaar en waardevol zijn om te meten. Het overzicht is dus bedoeld als **inspiratie**, niet als verplicht kader.

Merk op: uitkomsten en indicatoren op determinant- en effectniveau zoals kennis, attitude, vaardigheden, bewustzijn, gedragingen en gezondheid zijn hierin niet opgenomen. Onderstaande procesuitkomsten focussen op *hoe* een interventie wordt uitgevoerd en wordt ervaren, eerder dan *wat* ze oplevert.

De vermelde procesuitkomsten richten zich daarnaast op hoe een interventie wordt uitgevoerd en ervaren. Uitkomsten en indicatoren op determinant- of effectniveau, zoals kennis, attitude, vaardigheden, bewustzijn, gedragingen en gezondheid, verschillen sterk per thema en type interventie en zijn daarom niet opgenomen.

Procesuitkomst	Definitie	Voorbeelden van indicatoren
Bereik	De mate waarin de beoogde doelgroep wordt bereikt door de interventie.	Aantal deelnemers, percentage van doelgroep dat deelneemt, spreiding over demografische groepen
Adoptie	De mate waarin implementatoren of organisaties de interventie daadwerkelijk toepassen.	Aantal organisaties dat de interventie gebruikt, percentage implementatoren dat de interventie toepast in praktijk
Getrouwheid (Fidelity)	De mate waarin de interventie wordt uitgevoerd zoals ontworpen of gepland.	Aantal activiteiten uitgevoerd versus gepland volgens protocol, score op getrouwheidsvragenlijst, redenen voor aanpassingen
Intensiteit / Dosis	De hoeveelheid of frequentie van de interventie die deelnemers ontvangen.	Aantal sessies per deelnemer, gemiddelde duur van interventie-componenten
Haalbaarheid / uitvoerbaarheid	De mate waarin de interventie praktisch uitvoerbaar is binnen de bestaande middelen en context.	Tijd en middelen benodigd versus beschikbaar, drempels bij de implementatie, feedback van uitvoerders over de uitvoerbaarheid
Samenwerking	De kwaliteit en effectiviteit van samenwerking tussen betrokken partijen.	Aantal overlegmomenten, beoordeling van samenwerking
Tevredenheid	De mate waarin deelnemers of implementatoren tevreden zijn met de interventie of het evaluatieproces.	Gemiddelde tevredenheidsscore, percentage deelnemers dat tevreden is, gerapporteerde zaken waarover men wel of niet tevreden is

Acceptatie	De mate waarin de doelgroep of implementatoren bereid zijn de interventie te gebruiken of eraan deel te nemen.	Aandeel deelnemers dat de interventie nuttig vindt, aandeel organisaties met intentie tot voortzetting, redenen voor deelname of stopzetting
Context	Contextfactoren die de implementatie en effecten van de interventie beïnvloeden.	Organisatorische steun of drempels, beleidsmatige factoren, beschikbaar personeel
Duurzaamheid van implementatie	De mate waarin de interventie behouden blijft na de implementatieperiode.	Aandeel organisaties dat de interventie na 1 jaar nog uitvoert, mate van integratie in standaardpraktijk
Uitval	De mate waarin deelnemers of organisaties stoppen of niet volledig deelnemen aan de interventie.	Aandeel deelnemers of organisaties dat voortijdig stopt, redenen voor uitval

10.5 Overzicht: Evaluatiekaders bruikbaar voor implementatie- of procesevaluatie

De onderstaande tabel biedt een overzicht van enkele evaluatiekaders en hun kenmerken. De kaders in de onderstaande tabel zijn specifiek geselecteerd omdat ze relevant zijn voor **implementatieonderzoek en ruimere procesevaluatie** van interventies in de preventiesector. Het overzicht kan dienen als hulpmiddel om te bepalen welk kader het meest geschikt is voor een specifieke interventie, fase of evaluatiedoel. Bij de keuze spelen factoren mee zoals de complexiteit van de interventie, beschikbare middelen, gewenste diepgang en de vraag of de focus ligt op implementatie, context, mechanismen of een combinatie met effectevaluatie.

Sommige kaders zijn specifiek bedoeld voor onderzoek met een focus op de implementatie (o.a. CFIR, Proctor, IRLM). Het MRC-framework is het meest geschikt bij grondige evaluatie van complexe interventies, waarbij de procesevaluatie samen kan gaan met effectevaluatie. RE-AIM is geschikt voor meer pragmatische evaluatie waarbij in een vroege fase al effectuitkomsten worden onderzocht.

Evaluatiekader	Focus	Kernvraag	Aanbevelingen van onderzoekers
MRC-gids voor procesevaluatie van complexe interventies (70)	Implementatie, context en mechanismen van verandering	Hoe en waarom werkt een interventie, welke mechanismen en contextfactoren beïnvloeden resultaten, en hoe verloopt de uitvoering?	Aanbevolen voor de integratie van procesevaluatie en effectevaluatie bij complexe interventies (i.e. in het ruimere MRC-framework) Stimuleert het gebruik van mixed-methods onderzoeksmethoden Helpt verklaren waarom een interventie werkt of niet Flexibel inzetbaar afhankelijk van noden gedurende een pilootstudie, opschaling en continue evaluatie. Kan methodologisch veeleisend zijn.
RE-AIM (64)	Implementatie en evaluatie van interventies	Wat is het bereik, effect, de adoptie, implementatie en duurzaamheid van de interventie in de praktijk?	Hoge gebruiksvriendelijkheid Diverse voorbeelden en tools beschikbaar³⁶ Breed toepasbaar, goed afgestemd op preventiecontext Richt zich zowel op individuele als organisatorische aspecten van implementatie. Minder sterk in verklaren waarom een interventie werkt

³⁶ Zie <https://re-aim.org/>

PRISM (Practical, Robust Implementation and Sustainability Model) (137)	Uitbreiding van RE-AIM, met nadruk op context en duurzaamheid	Hoe beïnvloeden contextfactoren de RE-AIM uitkomsten?	Geschikter voor complexe interventies en diverse contexten dan RE-AIM Sterker in verklaren waarom een interventie werkt dan RE-AIM Complexer en tijdrovender dan RE-AIM
CFIR (Consolidated Framework for Implementation Research) (79, 80)	Identificeert en meet determinanten van implementatiesucces	Welke factoren in de context, implementatieproces, organisatie, individuen en de interventie beïnvloeden de implementatie?	Nuttig voor koppeling tussen implementatiefactoren en aparte effectevaluatie. Diverse tools en templates beschikbaar Complex raamwerk, vereist aanzienlijke middelen en tijd Enkel factoren gerelateerd aan implementatie en context, geen mechanismen van verandering
Proctor's Implementation Outcomes (83)	Definieert acht uitkomstmaten van implementatie	Wat zijn de acceptatie, toepassing/adoptie, geschiktheid, haalbaarheid, getrouwheid, kosten, bereik en duurzaamheid van de interventie?	Nuttig voor eenvoudige evaluatie van enkel de implementatie (bv. in vroege fase) Beperkt in verklaren van mechanismen van verandering of contextinvloeden.
Implementation Research Logic Model (IRLM) (85)	Integreert determinanten (CFIR), implementatie-strategieën en implementatie-uitkomsten (Proctor) in een Logic Model	Hoe hangen determinanten, strategieën en uitkomsten in het implementatieonderzoek samen?	Voor een volledige implementatie-evaluatie met koppeling tussen interventie, strategie, context, mechanismen en uitkomsten. Laat toe om implementatie-uitkomsten te verbinden met uitkomsten op niveau van diensten en gebruikers. Complex kader, kan overladen raken met variabelen en pijlen

10.6 Tool: Evalueerbaarheidsbeoordeling

Dit beoordelingsinstrument kan helpen om te bepalen of de interventie evalueerbaar is. Geef aan in welke mate elke stelling waar is. De reflectie op basis van de ingevulde tabel bepaalt mee de beslissing over wel of niet effecten te evalueren, en:

- **Zo ja:** welke vorm van evaluatie passend is en welke voorbereidingen nodig zijn.
- **Zo nee:** welke aanpassingen eerst nodig zijn voordat een evaluatie kan worden uitgevoerd.

Bron : CDC Program Evaluation Framework (55)

Evaluatiecomponent	Evalueerbaarheid	Helemaal niet waar	Deels waar	Waar	N.v.t.
Doel en Theory of Change	De Theory of Change is logisch en begrijpelijk.	☐	☐	☐	☐
Doel en Theory of Change	Er is gedeeld begrip tussen betrokkenen over de kernelementen en de context van de interventie	☐	☐	☐	☐
Waarschijnlijkheid van de interventie	De verwachtingen van de interventie zijn realistisch.	☐	☐	☐	☐
Toegankelijkheid van data	Bestaande data zijn valide, betrouwbaar en beschikbaar.	☐	☐	☐	☐
Toegankelijkheid van data	Nieuwe data kunnen op een haalbare manier worden verzameld.	☐	☐	☐	☐
Toegankelijkheid van data	Er zijn systemen aanwezig om indicatoren op te volgen.	☐	☐	☐	☐
Gereedheid voor gedegen evaluatie	Er zijn voldoende middelen (personeel, financiering...) om de evaluatie uit te voeren of deze kunnen voorzien worden.	☐	☐	☐	☐
Gereedheid voor gedegen evaluatie	De timing van de evaluatie sluit aan bij de looptijd van de interventie	☐	☐	☐	☐

10.7 Tool: Invulmodel voor betrokkenheid van partners en stakeholders

Vul per partner / stakeholder in in welke mate deze betrokken zal zijn bij de verschillende stappen van het evaluatieproces. Deze matrix kan worden aangepast naargelang het specifieke evaluatieontwerp.

Partner / Stakeholder	Evaluatie-focus	Onderzoeksdesign	Data-verzameling	Interpretatie van resultaten	Gebruik van resultaten	Andere (licht toe)

10.8 Overzicht: Evaluatiekaders bruikbaar voor effectevaluatie in combinatie met procesevaluatie

Evaluatiekader	Focus	Kernvraag	Aanbevelingen van onderzoekers
RE-AIM (64)	Implementatie en evaluatie van interventies	Wat is het bereik, effect, de adoptie, implementatie en duurzaamheid van de interventie in de praktijk?	- Hoge gebruiksvriendelijkheid - Diverse voorbeelden en tools beschikbaar - Breed toepasbaar, goed afgestemd op preventiecontext - Richt zich zowel op individuele als organisatorische aspecten van implementatie. Minder sterk in verklaren waarom een interventie werkt
PRISM (Practical, Robust Implementation and Sustainability Model) (137)	Uitbreiding van RE-AIM, met nadruk op context en duurzaamheid	Hoe beïnvloeden contextfactoren de RE-AIM uitkomsten?	- Geschikter voor complexe interventies en diverse contexten dan RE-AIM - Sterker in verklaren waarom een interventie werkt dan RE-AIM - Complexer en tijdrovender dan RE-AIM
MRC-framework (15)	Ontwikkeling, implementatie en evaluatie van complexe interventies (i.e., uitgevoerd en geëvalueerd op verschillende niveaus), benadrukt het belang van theoretische onderbouwing, procesevaluatie, contextanalyse en effectmeting.	Hoe kan een complexe interventie stap voor stap worden ontwikkeld, geïmplementeerd en geëvalueerd?	Zeer sterk voor complexe interventies in gezondheidsbevordering. Combineert theoriegebaseerde evaluatie, procesevaluatie en effectevaluatie Ondersteunt mixed-methods evaluatie. Complex raamwerk, vereist aanzienlijke middelen en tijd indien richtlijn nauwkeurig wordt gevolgd
Realist Evaluation (73)	Theoriegedreven evaluatie van interventies op basis van Context-Mechanisme-Outcome verbanden	Wat werkt, voor wie, in welke context, en waarom?	Maakt inzichtelijk waarom resultaten verschillen per context. Zeer geschikt voor complexe interventies waarin context een grote rol speelt De aanpak is complex en tijdsintensief

			Minder voor harde effectmeting, meer verklarend en gericht op mechanismen
Contribution analysis (35)	Evaluatie van bijdragen van een interventie aan uitkomsten door theorie, data en context te combineren	Hoe, waarom en voor wie resultaten tot stand zijn gekomen, en in welke mate heeft de interventie daar waarschijnlijk aan bijgedragen?	Sterk in theoriegedreven evaluatie van complexe interventies, wanneer kwantitatieve metingen niet haalbaar zijn Vereist toegang tot verschillende databronnen om verschillende aannames te kunnen toetsen Niet geschikt om effecten te kwantificeren

10.9 Overzicht: Kwantitatieve onderzoeksdesigns

Het gebruik van een beslisboom om een onderzoeksdesign te selecteren wordt als minder geschikt beschouwd, vanwege de vele contextuele overwegingen en randvoorwaarden, zoals beschikbare middelen, expertise en data. De keuze voor een onderzoeksdesign zou daarom altijd afgestemd moeten zijn op de beschikbare evaluatiecapaciteit, het doel en type interventie, en de specifieke context.

Tegelijkertijd is een overzicht van mogelijke designs wel waardevol, omdat het evaluatoren helpt een weloverwogen keuze te maken. Het overzicht illustreert dat er een breed scala aan designs bestaat die een balans bieden tussen methodologische robuustheid en praktische haalbaarheid.

In het uitgebreide onderzoeksrapport worden de verschillende designs met hun toepasbaarheid en voor- en nadelen meer in detail toegelicht.

Design	Controlegroep?	Randomisatie?	Beschrijving
RCT	Ja	Ja	Willekeurige toewijzing van deelnemers op individueel niveau aan interventie- of controlegroep.
Cluster-RCT	Ja	Ja	Willekeurige toewijzing van groepen (bv. scholen, regio's) aan interventie- of controlegroep.
Cross-over design	Ja	Van volgorde: Ja: experimenteel Nee: QE	Alle deelnemers ontvangen zowel de interventie als controleconditie, maar in een verschillende volgorde.
Stepped-wedge design	Ja	Van tijdstip: Ja: experimenteel Nee: QE	Alle clusters beginnen in controleconditie en ontvangen de interventie op verschillende tijdstippen.
Natural experiment	Ja	Nee	Interventie wordt niet willekeurig toegewezen; natuurlijke omstandigheden bepalen de groepen.
Synthetic control	Ja	Nee	De controlegroep wordt samengesteld uit een combinatie van meerdere groepen. Het gemiddelde van deze controlegroep wordt vergeleken met de uitkomsten van de interventiegroep.
Regressie discontinuïteit	Ja	Nee	Groepen worden bepaald door drempelwaarde op een continue variabele (bv. leeftijd). De uitkomst na de interventie wordt vergeleken bij mensen net onder en boven de drempel.
Interrupted time series	Nee	Nee	Meerdere metingen in één groep vóór en na de interventie laten trends zien, deze trends worden gebruikt om het verwachte verloop zonder interventie te reconstrueren.

Gecontroleerde pretest-posttest (<i>difference-in-difference analyse</i>)	Ja	Nee	Vergelijking van voor- en nameting in interventie- en controlegroep, met focus op verschillen in veranderingen tussen groepen.
Gecontroleerde enkelvoudige posttest studie ('static group comparison design')	Ja	Nee	Alleen vergelijking van nameting tussen interventiegroep en controlegroep
Pretest-posttest in één groep	Nee	Nee	Vergelijking van voor- en nameting bij interventiegroep
Enkelvoudige posttest in één groep	Nee	Nee	Alleen nameting bij interventiegroep.

10.10 Tool: Evaluatiematrix

De evaluatiematrix kan evaluatoren helpen bij het overzichtelijk maken van hoe verschillende evaluatievragen beantwoord zullen worden. In de matrix worden de belangrijkste concepten of constructen van de evaluatie gekoppeld aan specifieke evaluatievragen, indicatoren en de bijbehorende manier van kwantificeren. Daarnaast worden de databronnen en dataverzamelingsmethoden opgenomen, samen met de timing van dataverzameling en de geplande analysemethoden.

Door op deze manier alle elementen systematisch in kaart te brengen, ontstaat een **duidelijk overzicht van wat er wordt gemeten, hoe dit gebeurt en wanneer de gegevens beschikbaar zijn**. Dit ondersteunt de planning en uitvoering van de evaluatie, en zorgt voor **transparantie** zowel voor het evaluatieteam als voor beleid en andere betrokken stakeholders.

Concept / Construct	Evaluatievragen	Indicator(en)	Databron	Methode en meet-instrument	Timing van dataverzameling	Data-analyse
Welke concepten/constructen ga je evalueren?	Wat wil je concreet weten?	Welke data heb je nodig om de evaluatievragen te beantwoorden?	Waar komt de data vandaan?	Hoe ga je de data verzamelen? Welk meet-instrument wordt gebruikt?	Wanneer wordt de data verzameld?	Hoe ga je de data analyseren en interpreteren?
Bv. Gezonde voedingsgewoonten	Verandert de consumptie van groenten en fruit bij leerlingen na de interventie?	Gemiddeld aantal porties groenten en fruit per dag	Leerlingen van deelnemende scholen	Online vragenlijst	Voor de interventie (T0), direct na interventie (T1), 3 maanden na interventie (T2)	Berekening van gemiddelde porties, vergelijking T0–T1–T2, paired t-tests
...						

11 Referenties leidraad

- Adviescommissie richtlijn passend bewijs voor preventie. (2025). *Passend bewijs voor preventie*. Retrieved from <https://open.overheid.nl/documenten/551ff7b5-c686-4956-9c0e-a36dca3f4f3c/file>
- Damschroder, L. J., Reardon, C. M., Widerquist, M. A. O., & Lowery, J. (2022). The updated Consolidated Framework for Implementation Research based on user feedback. *Implementation Science*, 17(1), 75. <https://doi.org/10.1186/s13012-022-01245-0>
- Holtrop, J. S., Estabrooks, P. A., Gaglio, B., Harden, S. M., Kessler, R. S., King, D. K.,...Glasgow, R. E. (2021). Understanding and applying the RE-AIM framework: Clarifications and resources. *J Clin Transl Sci*, 5(1), e126. <https://doi.org/10.1017/cts.2021.789>
- Kidder, D. P., Fierro, L. A., Luna, E., Salvaggio, H., McWhorter, A., Bowen, S. A.,...Young, K. (2024). CDC Program Evaluation Framework, 2024. *MMWR Recomm Rep*, 73(6), 1-37. <https://doi.org/10.15585/mmwr.rr7306a1>
- Lohan, M., Aventin, Á., Clarke, M., Curran, R. M., Maguire, L., Hunter, R.,...O'Hare, L. (2018). JACK trial protocol: a phase III multicentre cluster randomised controlled trial of a school-based relationship and sexuality education intervention focusing on young male perspectives. *BMJ Open*, 8(7), e022128. <https://doi.org/10.1136/bmjopen-2018-022128>
- Mayne, J. (2008). Contribution analysis: An approach to exploring cause and effect. *Institutional Learning and Change (ILAC) Initiative, ILAC Briefs*, 16.
- Moore, G., Campbell, M., Copeland, L., Craig, P., Movsisyan, A., Hoddinott, P.,...Evans, R. (2021). Adapting interventions to new contexts-the ADAPT guidance. *Bmj*, 374, n1679. <https://doi.org/10.1136/bmj.n1679>
- Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W.,...Baird, J. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *Bmj*, 350, h1258. <https://doi.org/10.1136/bmj.h1258>
- Pawson, R., & Tilley, N. (1997). An introduction to scientific realist evaluation. In *Evaluation for the 21st century: A handbook*. (pp. 405-418). Sage Publications, Inc. <https://doi.org/10.4135/9781483348896.n29>
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A.,...Hensley, M. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health*, 38(2), 65-76. <https://doi.org/10.1007/s10488-010-0319-7>
- Reardon, C. M., Damschroder, L. J., Ashcraft, L. E., Kerins, C., Bachrach, R. L., Nevedal, A. L.,...Rogal, S. (2025). The Consolidated Framework for Implementation Research (CFIR) User Guide: a five-step guide for conducting implementation research using the framework. *Implementation Science*, 20(1), 39. <https://doi.org/10.1186/s13012-025-01450-7>
- Rebekah Gomes, C. B., Fort, M., Maw, A., McCreight, M., Borsika Rabin, E. R., Studts, C.,...Glasgow, R. (2023). A guidebook to the pragmatic and iterative use of the practical, robust implementation and sustainability model (PRISM) and reach, effectiveness, adoption, implementation, maintenance framework (RE-AIM) for planning, implementation, and sustainment. *The Colorado Implementation Science Center for Cancer Control*.
- Skivington, K., Matthews, L., Simpson, S. A., Craig, P., Baird, J., Blazeby, J. M.,...Moore, L. (2021). A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *Bmj*, 374, n2061. <https://doi.org/10.1136/bmj.n2061>
- Smith, J. D., Li, D. H., & Rafferty, M. R. (2020). The Implementation Research Logic Model: a method for planning, executing, reporting, and synthesizing implementation projects. *Implementation Science*, 15(1), 84. <https://doi.org/10.1186/s13012-020-01041-8>

Bijlagen onderzoeksrapport

Bijlage 1: Voorbereidingen bijeenkomsten van werkgroepen

1. Voorbereiding en leidraad bijeenkomst 1

1.1 Inleidende vragen

- *Werkgroep stakeholders preventiesector:* Wat is jullie ervaring met het evalueren van gezondheidsbevorderende/preventieve interventies?
- *Werkgroep onderzoek en beleid:* Wat is jullie ervaring met het uitvoeren, opzetten, begeleiden of beoordelen van evaluaties van gezondheidsbevorderende interventies?
- *Beide werkgroepen:* Wat wordt er precies geëvalueerd (bv. effectiviteit, bereik, proces), met welke methoden, en door wie?
- *Werkgroep stakeholders preventiesector?* Hoe wordt evaluatie ervaren binnen de organisatie? Wordt het als iets positief, moeilijk, waardevol, tijdrovend of als een verplichting ervaren?

1.2 Prioritaire randvoorwaarden en hefbomen

Voor deze vraag wordt een vooraf opgestelde lijst met randvoorwaarden en hefbomen overlopen. De deelnemers kunnen deze lijst aanvullen met eigen ideeën via post-its. Vervolgens prioriteert iedere deelnemer individueel de belangrijkste hefbomen/randvoorwaarden (zowel uit lijst als toegevoegde) met behulp van vijf stickers. Ze mogen deze vrij verdelen over de hefbomen/randvoorwaarden (waarbij men bv. meerdere stickers op één randvoorwaarde kan kleven om hieraan meer gewicht te geven).

De resultaten van de oefening worden in groep besproken: wat valt op, waar is er consensus, waarover lopen de meningen uiteen?

- Wat zijn volgens jullie de **meest belangrijke voorwaarden en hefbomen** voor een haalbare en kwaliteitsvolle effect- en impactevaluatie van interventies? = Wat is essentieel om effect- en impactevaluatie te doen slagen? Welke elementen verdienen prioriteit? Waar zou het beleid of een leidraad volgens jou vooral op moeten inzetten?

Vooraf opgestelde lijst met randvoorwaarden en hefbomen:

- **Heldere beleidsvisie** over wie waarvoor verantwoordelijk is
- **Financiële middelen** expliciet gereserveerd voor effect- en impactevaluatie in projectbudgetten
- **Tijdsinvestering** ingebouwd in de planning van teams en projecten
- **Vertrouwen en draagvlak** voor effect- en impactevaluatie binnen de organisatiecultuur
- **Praktische training of coaching** rond effect- en impactevaluatie
- **Ondersteuning** door onderzoekers, beleidsmakers en financiers in het **ontwikkelen van meetbare indicatoren en goede evaluatieplannen** die doelgericht inzicht geven in de effectiviteit en mechanismen van interventies.
- **Interne wetenschappelijk medewerker** met expertise in effect- en impactevaluatie
- **Directe samenwerking en afspraken met externe onderzoekers** (bv. academici)

- Effect- en impactevaluatie opnemen als verplicht onderdeel van interventies
- **Methodologische richtlijnen, evaluatiekaders en beslissingsondersteunende tools**
- Erkenning van **flexibele/alternatieve evaluatiebenaderingen** en methoden
- **Infrastructuur** voor data-analyse, monitoringtools en digitale platformen ter ondersteuning van effect- en impactevaluatie
- Het **in kaart brengen van organisaties die effect- en impactevaluatie succesvol integreren in hun programmaontwikkeling**, om inzicht te krijgen in de kenmerken, strategieën en werkwijzen die helpen om obstakels bij evaluatie te overwinnen.
- **Meerjarenperspectief voorzien**, zodat effect- en impactevaluatie niet ad hoc gebeurt maar structureel en over langere termijn wordt opgevolgd.
- **Systematisch gebruik van evaluatieresultaten** in beleid, praktijk en de verdere ontwikkeling van interventies.
- ... (zelf aan te vullen via post-its)

- Wat **helpt of ondersteunt jullie vandaag** al bij de effect- en impactevaluatie van gezondheidsbevorderende interventies? Welke **bestaande praktijken, ondersteuning of tools werken goed**?
- Welke concrete hefboomen, ondersteuningsvormen of tools **ontbreken vandaag nog of zijn onvoldoende aanwezig**?

1.3 Evaluatiekaders

- In welke mate zijn jullie vertrouwd met het gebruiken van een **verandertheorie (Theory of Change)** voor het evalueren van interventies, en wat zijn jullie ervaringen hiermee (bv. wat werd als helpend ervaren, waar liepen jullie tegenaan ...)?
- Welke **evaluatiekaders** kennen en/of gebruiken jullie, en wat zijn jullie ervaringen hiermee (bv. wat werd als helpend ervaren, waar liepen jullie tegenaan ...)? (enkele voorbeelden: RE-AIM, Realist Evaluation, MRC-framework, CFIR...)

Stellingen (akkoord/niet akkoord + groepsdiscussie):

- "Het gebruik van een evaluatiekader is onmisbaar om een effect- en impactevaluatie systematisch en transparant te laten verlopen."
- "Evaluatiekaders moeten steeds zowel de effectiviteit als de implementatie meten."
- "De meeste evaluatiekaders zijn te complex voor gebruik door medewerkers van organisaties in de preventiesector."
- "Het zou helpen als de leidraad één uniform en helder evaluatiekader aanreikt dat iedereen gebruikt."

1.4 Onderzoeksdesign

- Hebben jullie reeds ervaring met (quasi-)experimenteel onderzoek, d.w.z. met controlegroep en/of randomisatie?

Zo nee, waarom niet?

Zo ja, hoe wordt dit ervaren?

- Is het haalbaar en/of wenselijk om meer gecontroleerd en/of gerandomiseerd onderzoek (bv. RCT's) uit te voeren naar gezondheidsbevorderende interventies in Vlaanderen?
- Is ondersteuning bij de keuze van het onderzoeksdesign wenselijk (bv. via een tool)?
- Is het wenselijk om per gezondheidsthema een set van minimaal te meten effectindicatoren te voorzien (bv. met oog op afstemming met indicatoren uit gezondheidsdoelstellingen)? Waaraan dienen deze indicatoren te voldoen?

1.5 Rolverdeling en samenwerkingsmogelijkheden tussen organisaties en externe onderzoekers/ evaluatoren

- Wat zijn randvoorwaarden voor een goede samenwerking (co-creatie, tijd, overlegstructuren...)?
- Welke samenwerkingsverbanden bestaan er momenteel tussen onderzoeksinstellingen en partnerorganisaties voor de evaluatie van gezondheidsbevorderende interventies? Hoe kan het Departement Zorg zulke samenwerkingen faciliteren en stimuleren?
- Wanneer is samenwerking met externe onderzoekers wenselijk of noodzakelijk?
- Wat is een goede rolverdeling tussen betrokken actoren bij de evaluatie van interventies, per evaluatiestap? In een tabel noteren de deelnemers welke actoren een sleutelfunctie opnemen en eventueel bijhorende rollen/taken. De uitkomsten worden per groep besproken per evaluatiestap om tot een gezamenlijke invulling te komen.

Evaluatiestappen:

- Stakeholders betrekken tijdens het hele evaluatieproces
- Verandertheorie van interventie omschrijven
- Focus van evaluatie bepalen (selectie van uitkomst-indicatoren)
- Onderzoeksdesign en methodologie bepalen
- Dataverzameling uitvoeren
- Analyse van de data
- Interpretatie van resultaten
- Rapportage van evaluatie-uitkomsten
- Gebruik van resultaten om interventie te optimaliseren

2. Voorbereiding en leidraad bijeenkomst 2

2.1 Wanneer is effectevaluatie nodig en hoe kunnen evaluaties eventueel geprioriteerd worden?

Een effectevaluatie vraagt aanzienlijke **tijd, middelen en expertise**. Daarom is het essentieel om vooraf goed na te denken waarom, wanneer en hoe een effectonderzoek zinvol is. Evalueren mag nooit een

doel op zich zijn: een effectevaluatie is enkel waardevol wanneer de resultaten daadwerkelijk gebruikt zullen worden voor beleid, praktijk of de interventie. Daarom is het belangrijk om **duidelijke criteria te hanteren die aangeven wanneer effectevaluatie aangewezen is**.

1) Wanneer is een effectevaluatie nodig, en wanneer niet?

a) Welke bedenkingen hebben jullie bij deze overwegingen?

b) Aan welke andere overwegingen of situaties denken jullie nog waarin effectevaluatie (niet) wenselijk is?

Onderstaande tabel geeft enkele mogelijke overwegingen weer (gebaseerd op Ogilvie ea.) die kunnen helpen om te bepalen of effectevaluatie zinvol is. Deze zijn niet bedoeld als strikte checklist, maar helpen bij de reflectie en discussie over al dan niet over te gaan tot effectevaluatie.

Fase van interventie	In een vroege fase is het vaak zinvoller om procesvragen te stellen, zoals welke drempels er bestaan voor deelname en welke doelgroepen worden bereikt. Effectevaluatie komt meestal pas in een latere fase aan bod, wanneer de interventie stabiel is en haalbaar gebleken is.
Invloed op beleid	Een evaluatie is vooral zinvol als de resultaten kunnen bijdragen aan beleidsbeslissingen of strategische keuzes binnen de sector.
Omvang en bereik van impact	Interventies met een groot bereik of een potentieel groot effect verdienen vaker effectevaluatie. Toch kunnen ook kleinschalige of innovatieve projecten relevant zijn, zeker als ze schaalbaar zijn of inspelen op gezondheidsongelijkheid.
Bestaande evidentie	Niet elke interventie hoeft uitgebreid onderzocht te worden, zeker niet als de effectiviteit al voldoende bewezen is. Effectevaluatie is zinvol wanneer evidentie ontbreekt of er onduidelijkheden zijn, bv. over effecten in andere populaties, contexten, mechanismen, duurzaamheid...
Praktische overwegingen	De meerwaarde van effectevaluatie hangt ook af van de haalbaarheid van evaluatie: is het mogelijk om effecten te meten, voldoende deelnemers te betrekken, externe invloeden te controleren, en de nodige expertise in te zetten?

2) Individuele prioriteringsoefening Omdat er met de beschikbare middelen keuzes moeten gemaakt worden en niet elke interventie tegelijk geëvalueerd kan worden, is **prioritering** nodig. Welke interventies in de preventiesector verdienen prioriteit inzake (meer wetenschappelijke) evaluatie? (top 5)

- Kostelijke interventies
- Interventies met groot bereik
- Interventies met potentieel om op grote schaal uit te rollen
- Interventies met beperkte bestaande (internationale) evidentie
- Interventies die bewezen effectief zijn in andere contexten (landen/regio's/settings) en mogelijk overdraagbaar zijn

- Interventies met veelbelovende resultaten (bv. in kleinere studies of pilootstudies)
- Interventies met een duidelijke meerwaarde voor beleidsbeslissingen
- Interventies die aansluiten bij de gezondheidsdoelstellingen.
- Interventies die inspelen op hiaten in de preventiesector
- Interventies die inspelen op gezondheidsproblemen met grote ziektelast
- Interventies die inspelen op gezondheidsongelijkheden (bv. gericht op kwetsbare groepen)
- Andere

2.2 Gebruik van evaluatieresultaten

Effectevaluaties zijn pas waardevol als hun resultaten **worden gebruikt om beleid en praktijk te versterken**. Het doel is **leren, bijsturen en verbeteren van interventies en beleid** op basis van evidentie. Daarom is het belangrijk om na te denken over **hoe evaluatieresultaten effectief kunnen leiden tot veranderingen** en hoe er een **leercyclus** kan ontstaan binnen de preventiesector.

- 1) Hoe kunnen het beleid en de preventiesector ervoor zorgen dat resultaten van evaluaties daadwerkelijk worden gebruikt voor beleidskeuzes, bijsturing van interventies en het versterken van succesvolle praktijken?

Mogelijke pistes voor beter gebruik van evaluatieresultaten

- Bij evaluatieplannen al duidelijk maken **hoe de resultaten beleid, praktijk en de interventie zullen beïnvloeden** en wat de gevolgen van de resultaten zullen zijn
- **Effectieve interventies identificeren en erkennen:** Interventies worden beoordeeld en erkend aan de hand van vastgelegde criteria (rekening houdend met de complexiteit van onderzoek naar gezondheidsbevordering)
- **Effectieve interventies versterken:**
 - Bijkomende middelen voorzien voor impact te vergroten
 - Opschalen / breder uitrollen
 - Verduurzamen via o.a. structurele financiering of structurele inbedding
- **Formuleren van acties bij ongunstige resultaten:**
 - Bijsturen van de interventie en eventuele implementatiefactoren aanpakken
 - Stopzetten van interventies die schadelijk of herhaaldelijk ineffectief blijken, zodat middelen doelgerichter kunnen worden ingezet
- Formuleren van **concrete aanbevelingen en acties voor beleid, praktijk en interventieverbetering bij afronding** van de evaluatie
- Integratie van evaluatieresultaten in **beleidsprocessen**, zoals de opvolging van gezondheidsdoelstellingen, meerjarenplanning of beheersovereenkomsten.

2) Hoe kan beleid een leercyclus stimuleren?

Een **leercyclus** betekent dat evalueren, reflecteren, leren en bijsturen **structureel en herhaald** plaatsvinden. Het doel is om van de preventiesector een **lerend systeem** te maken, waarin kennis systematisch wordt opgebouwd en benut.

- Hoe zou een leercyclus concreet kunnen werken? Welke elementen lijken bijzonder waardevol?
- Welke andere initiatieven zouden een leercultuur met inzet op kennisdeling nog meer kunnen stimuleren?

Leercyclus en stimuleren van leercultuur

Mogelijke elementen leercyclus:

- Periodieke overlegmomenten met vertegenwoordigers van beleid, praktijk en onderzoek
- Bespreking van lessen uit recente evaluaties (wat werkte, wat niet, waarom?).
- Bepaling van nieuwe prioriteiten voor volgende evaluaties of verbetertrajecten
- Delen van inzichten (bv. via de Expertisecel of andere communicatiekanalen).

Mogelijke ondersteunende initiatieven:

- Ruimer lerend netwerk van evaluatoren in de preventiesector
- Een digitaal platform met toegankelijke rapporten en evaluatieresultaten
- Thematische netwerkdagen waar evaluaties worden voorgesteld en besproken.

2.3 Beleidsimpact

Naast het verbeteren van afzonderlijke interventies is het belangrijk dat ook de **beleidsimpact op systeemniveau** in kaart wordt gebracht. Denk hierbij aan de link tussen populatiedata, de gezondheidsdoelstellingen en de indicatoren en uitkomsten die gemeten worden in de interventie. Daarom is het belangrijk om reeds in de programmatheorie van een interventie aandacht te besteden aan de mogelijke impact op beleidsniveau: op welke gezondheidsdoelstelling(en) de interventie inspeelt en via welke populatiedata deze effecten kunnen worden gemeten.

- 1) Hoe kan de beleidsimpact op **systeem/sectorniveau beter worden geëvalueerd**?
- 2) Hoe kunnen we ervoor zorgen dat evaluaties van interventies bijdragen aan een **breder inzicht in de impact van het preventieve gezondheidsbeleid**?

2.4 Financiering van effectevaluatie

Tijdens de vorige werkgroepen werd het belang benadrukt van een meer gecoördineerde en duurzame **financiering van effectevaluaties** in de preventiesector. Beleid kan hierin een belangrijke rol spelen door de beschikbare budgetten meer **doelgericht te verdelen**. Op basis van eerdere input zijn enkele eerste **voorstellen van mogelijke financieringsmodellen** uitgewerkt.

Financiering van effectevaluatie

Voorzien van een systeem voor oproepen of calls vanuit preventiesector naar onderzoekers:

De overheid ontwikkelt een systeem voor oproepen of calls. Organisaties besteden een evaluatie uit (met eigen middelen) via een oproep, maar blijven actief betrokken in het proces. Onderzoekers kunnen een voorstel indienen in de vorm van een evaluatieplan.

1. *PO heeft nood aan evaluatie:*
 - *omschrijft interventie*
 - *geeft richtinggevend budget*
 - *omschrijft wat geëvalueerd dient te worden (o.a. welke uitkomsten, determinanten, subgroepen, mechanismes, context...), eventueel via specifieke evaluatieaanpak of onderzoeksdesign*
2. *Externe onderzoeksactor schrijft voorstel uit op basis van sjabloon/evaluatieplan*
3. *De PO selecteert een voorstel*
 - *eventueel in afstemming met beleid*
 - *eventueel op basis van vooraf vastgelegde beoordelingscriteria.*

Centraal Evaluatiefonds

De overheid beheert met een evaluatiefonds en lanceert regelmatig (bv. jaarlijks) oproepen voor effectonderzoek naar preventiemethodieken. Dit staat open voor PO's die een evaluatievoorstel uitschrijven: zowel zelf (met interne expertise), of samen met externe onderzoeksactoren

Eventueel kan er met verschillende sporen worden gewerkt, bv. een kleiner fonds voor evaluatie in een vroegere (pilot)fase, en een groter fonds voor meer wetenschappelijke evaluaties

1. *De PO met interne of externe onderzoekers schrijven een voorstel volgens een vast format waarin o.a. volgende zaken aan bod komen:*
 - *Beschrijving van interventie (o.a. doelgroep, setting, Theory of Change, fase van uitrol)*
 - *Nood aan evaluatie (wat is al wel en nog niet geweten)*
 - *Methodologie (o.a. design, indicatoren, steekproef, ethiek)*
 - *Samenwerking en governance (o.a. betrokkenen, rollen, borging van samenwerking, betrekking doelgroep en intermediairs)*
 - *Nodige budget en middelen*
 - *Planning en tijdslijn*
 - *Disseminatie en gebruik van resultaten (o.a. vertaling naar beleid en praktijk)*

2. Om te bepalen welke evaluatievoorstellen worden gefinancierd, is het belangrijk om te werken met **transparante en evenwichtige beoordelingscriteria**.

We stellen **twee complementaire dimensies** voor:

(a) **Kwaliteit van evaluatieplannen**

eventueel met meer gewicht voor belangrijke criteria of onderdelen (bv. nood aan evaluatie, methodologie)

idealiter door onafhankelijk expertpanel om de objectiviteit te versterken.

(b) **Strategische dimensie: beleidsmatig en maatschappelijk belang**

- het departement beoordeelt de mate waarin de interventie beleidsmatig en maatschappelijk prioritair is (bv. relevant voor gezondheidsdoelstellingen, kosten van interventie, bereik...)
- deze inschatting kan door departement zelf, eventueel met (beleids)experts

→ **door de dimensies te combineren ontstaat een evenwichtige prioritering:**

- Evaluaties met **hoge kwaliteit en hoge strategische relevantie** krijgen prioriteit.
- Evaluaties met **lage kwaliteit** maar hoge beleidsrelevantie kunnen worden herwerkt op basis van feedback.
- Evaluaties met **beperkte beleidsrelevantie** maar hoge kwaliteit kunnen later worden ingepland

- Vast percentage
- Verschillende opties:
 - - Vast percentage van de middelen in beheersovereenkomsten reserveren voor evaluatie per organisatie
 - - Vast percentage van budget voor project/interventie reserveren
 - - Andere?
- Combinaties van verschillende opties zijn eveneens mogelijk

- 1) Welke opties spreken het meeste aan?
- 2) Welke bezwaren of bedenkingen zijn er?
- 3) Zijn er nog andere opties om middelen efficiënter te verdelen?
- 4) Indien er gekozen wordt voor een Evaluatiefonds, hoe kan dit fonds op een duurzame en rechtvaardige manier worden gefinancierd?
- 5) Wat zijn de voor- en nadelen van de financieringsmodellen?

Hou bv. rekening met:

- Toegankelijkheid: zowel kleinere als grote organisaties met en zonder interne expertise hebben nood aan toegang tot middelen
- Flexibiliteit: inzet van middelen wanneer het nodig is
- Duurzaamheid: structurele financiering van evaluatie op lange termijn
- Samenwerking tussen onderzoek en praktijk als sterkte

2.5 Centraal advies voor evaluatie in de preventiesector

Uit de eerste bijeenkomsten van de werkgroepen kwam het voorstel om een **centraal aanspreekpunt en ondersteuningsstructuur voor evaluatie in de preventiesector** te ontwikkelen. Hieronder staan enkele mogelijke manieren waarop dit vorm kan krijgen en welke mogelijke taken deze structuur zou kunnen opnemen.

Centraal advies voor evaluatie in de preventiesector

Mogelijke taken:

- Ontwikkelen van een **overkoepelende strategie en governance ondersteunen** om de evaluatie van interventies systematisch en transparant te laten verlopen.
- **Adviesverlening en ondersteuning over de evaluatieaanpak** (o.a. verandertheorie, onderzoeksvragen, onderzoeksdesign, indicatoren, dataverzameling...), afgestemd op het doel en de aard van de interventie, de beschikbare middelen en expertise en de context.
- **Stimuleren van kennisdeling en uitwisseling**, zoals het faciliteren van netwerken en (online) platformen voor het delen van inzichten, ervaringen en geleerde lessen tussen organisaties, onderzoekers en beleidsmakers.
- **Opvolging van evaluaties** binnen de preventiesector
- **Voorzien van training en opleiding** rond evaluatievaardigheden.
- **Kwaliteitsbewaking**, bijvoorbeeld door toetsing van evaluatieplannen
- **Ondersteuning bij budgettering van evaluaties**, zoals het inschatten van realistische evaluatiekosten en het integreren van evaluatie in projectbudgetten.
- **Faciliteren van verbindingen en gerichte matchmaking** tussen organisaties en onderzoekers, op basis van hun expertise en thematische focus, om samenwerking rond de evaluatie van specifieke interventies te versterken
- **Administratieve en praktische ondersteuning** zoals het aanleveren van sjablonen voor het opzetten en rapporteren van evaluaties.
- **Beheer van een overzicht** of databank met evaluatie-instrumenten, tools, rapporten en andere ondersteunende documenten
- **Andere**

- 1) Hoe staan jullie tegenover het idee van een centraal aanspreekpunt en/of centrale ondersteuningsstructuur?
- 2) Hoe zou dit best vorm krijgen? Bv. verankering in bestaande structuur, beleid of organisatie; vereiste expertise of betrokkenheid...
- 3) Welke taken zouden volgens jullie het belangrijkste zijn voor zo'n structuur? Zijn er taken die minder prioritair zijn of die beter elders thuishoren?

Bijlage 2: Verwachte expertise inzake evaluatie van de Canadese overheid voor junior, intermediate en senior evaluatiespecialisten

Tabel 14 Verwachte expertise inzake evaluatie van de Canadese overheid voor junior, intermediate en senior evaluatiespecialisten (<https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/evaluation-government-canada/evaluation-competencies.html>)

Junior evaluator	Intermediaire evaluator	Senior evaluator
Geeft advies om credibiliteit en bruikbaarheid te waarborgen	Geeft advies om credibiliteit en bruikbaarheid te waarborgen	Bepaalt en controleert evaluatiedoelen, waarborgt de credibiliteit en bruikbaarheid
Onderzoekt en ondersteunt ontwikkeling van programmatheorie en identificatie van stakeholders	Ontwikkelt/valideert programmatheorie, koppelt die met resultatenframework van Departement (i.e., Departmental Results Framework) en andere programma's, identificeert stakeholders	Ontwikkelt/valideert programmatheorie, identificeert linken met resultatenframework van Departement (i.e., Departmental Results Framework) en andere programma's (ook in andere departementen), identificeert stakeholders
Ondersteunt bij formuleren van vragen passend bij doel	Stelt evaluatievragen voor	Stelt of valideert evaluatievragen en issues die aansluiten bij doelen
Helpt bepalen datavereisten en beschikbaarheid van bestaande data, signaleert hiaten	Bepaalt of verifieert datavereisten, beoordeelt kwaliteit, betrouwbaarheid en validiteit, identificeert hiaten	Adviseert management over kwaliteit, betrouwbaarheid en validiteit van data; wijst op hiaten
Ondersteunt ontwikkeling van aanpak, ontwerp en methoden van evaluatie, o.a. strategieën om datahiaten aan te pakken	Stelt aanpak, ontwerp en methoden van evaluatie voor, o.a. dataverzameling en analysetechnieken die aansluiten bij de middelen en risico's; en strategieën om datahiaten aan te pakken	Stelt aanpak, ontwerp en methoden van evaluatie voor, o.a. dataverzameling en analysetechnieken die aansluiten bij de middelen en risico's; en strategieën om datahiaten aan te pakken
Past technieken, instrumenten en protocollen voor dataverzameling toe	Ontwerpt instrumenten en protocollen, implementeert of monitort hun implementatie, beoordeelt validiteit en betrouwbaarheid van de data	Waarborgt kwaliteit en integriteit van instrumenten en protocollen; waarborgt validiteit en betrouwbaarheid van data; lost dataproblemen op

Analyseert data, ondersteunt identificatie van hiaten, stelt oplossingen voor	Analyseert en interpreteert data, identificeert hiaten, beoordeelt de evidentie en vult hiaten aan waar nodig	Waarborgt adequaatheid van de evidentie en nauwkeurigheid van data-interpretatie
Ondersteunt de identificatie en formulering van bevindingen en conclusies, formuleert aanbevelingen die logisch volgen uit conclusies	Identificeert bevindingen en beoordeelt attributie of bijdrage, formuleert conclusies. Stelt haalbare aanbevelingen voor, gericht op resultaatverbetering	Identificeert en valideert bevindingen en conclusies, beoordeelt attributie/bijdrage; legt verbanden met organisatiedoelen en overheidsprioriteiten. Zorgt dat aanbevelingen logisch en haalbaar zijn en aansluiten bij bevindingen, conclusies en strategische doelen
Ondersteunt voorbereiding van rapporten en presentaties	Bereidt rapporten en presentaties voor	Zorgt voor duidelijke en tijdige rapporten en presentaties

Bijlage 3: Matrix van Bourgois en Cousins inzake organisatorische evaluatiecapaciteit.

Tabel 15 Matrix van Bourgois en Cousins inzake organisatorische evaluatiecapaciteit (42).

Dimensie	Laag	Ontwikkeland	Gemiddeld	Voorbeeldig
Menselijke middelen	Onvoldoende bemand, geen opvolging, zwakke methoden	Gedeeltelijk bemand, losse loopbanen, beginnend vertrouwen	Volledig bemand, goede mix vaardigheden, interne leer-mogelijkheden	Optimale personeelsbezetting, sterk leiderschap, vertrouwen en innovatie
Organisatorische middelen	Geen budget, geen performantiegegevens, zwakke structuren	Ad hoc financiering, onvolledige data, steun vaak gebrekkig	Stabiel maar gedeeld budget, gedeeltelijk systeem, structuren in groei	Vast budget, volwaardig meetsysteem, sterke governance en ondersteuning
Evaluatieplanning & uitvoering	Geen plan, weinig samenwerking, zwakke netwerken	Ad hoc planning, beperkte consultants, weinig kennisdeling	Jaarplan, consistente methoden, matige netwerken	5-jaarplan, risico-gebaseerd, sterke samenwerking en externe steun
Evaluatiekennis	Weinig begrip, zwakke methodische kennis	Basiskennis, enkele vormingen, beperkte toepassing	Goede kennis, kan resultaten interpreteren	Sterke geletterdheid, geavanceerde methoden, systematisch gebruik
Besluitvorming	Resultaten nauwelijks gebruikt, weinig steun leiding	Sporadisch gebruik in planning, enige steun leiding	Regelmatig gebruikt in budget en ontwerp, leiding betrokken	Evaluatie stuurt strategie, volledig geïntegreerd
Leer- en verbeterkansen	Nauwelijks leren, geen verantwoording	Enkele verbeteringen, beperkte leercultuur	Programma's verbeteren, groeiende leercultuur	Continue verbetering, sterke leer- en verantwoordingscultuur

Bijlage 4: De originele *Evaluation Practice Analysis Survey* (EPAS).

Tabel 16 De originele *Evaluation Practice Analysis Survey* (EPAS). Van: Schwarzman, J., Bauman, A., Gabbe, B. J., Rissel, C., Shilton, T., & Smith, B. J. (2019). Understanding the factors that influence health promotion evaluation: The development and validation of the evaluation practice analysis survey. *Evaluation and program planning*, 74, 76–83.
<https://doi.org/10.1016/j.evalprogplan.2019.03.002>

Survey domeinen

Schalen

Evaluation practice

Self-reported evaluation practice

- a) Percentage of projects evaluated in the past two years at the formative level
- b) Percentage of projects evaluated in the past two years at the process level
- c) Percentage of projects evaluated in the past two years at the impact level
- d) Percentage of projects evaluated in the past two years at the outcome level

Individual level

Attitude and beliefs - Program and staff level benefits

I think that evaluation...

- a) Will help me understand my program
- b) Will inform the decisions I make about my program
- c) Is absolutely necessary to improve my program
- d) Should involve program participants in the evaluation process
- e) Is worth spending time on

Attitude and beliefs - Organisation and wider benefits

I think that evaluation...

- a) Will influence policy relevant to my program (e.g. stakeholder/partnership, local, state or national policy)
- b) Will help improve programs for people from minority and hard to reach populations
- c) Will justify funding for my program
- d) Will help to convince managers that changes are needed in my program
- e) Will inform changes in our internal processes and documentation

Skills for different levels of evaluation

- a) Formative evaluation (i.e. pretesting or piloting)
- b) Process evaluation (i.e. monitoring of implementation)
- c) Impact (i.e. achievement of objectives)
- d) Outcome evaluation (i.e. achievement of project goals)
- e) Economic evaluation (eg. cost benefit analysis)

Skills for different evaluation tasks

- a) Clearly stating goals and objectives for my project
- b) Deciding what questions to answer in an evaluation
- c) Choosing designs to evaluate project impacts or outcomes
- d) Developing data collection tools
- e) Identifying from whom to collect the information (i.e. a sampling strategy)

- f) Implementing or appraising ethical standards in evaluation
- g) Analysing quantitative evaluation data
- h) Analysing qualitative evaluation data
- i) Developing recommendations based on evaluation results
- j) Writing an evaluation report
- k) Disseminating findings to relevant stakeholders
- l) Managing an evaluation

Organisational level

Leadership for evaluation

The leaders in my organisation...

- a) Have a clear vision for evaluation
- b) Demonstrate a good understanding of program evaluation
- c) Are able to build good partnerships with those who can support evaluation when required
- d) Can successfully negotiate evaluation and reporting requirements with our funding source
- e) Ensure our evaluation findings are useful to decision makers
- f) Champion evaluation practice and use internally
- g) Have a strong desire to share our findings
- h) Lack interest in evaluation
- i) Advocate in support of our evaluation efforts when required

Organisational culture to create the right environment

My organisation...

- a) Encourages a learning environment
- b) Is an innovative organisation
- c) Uses evidence from evaluations to inform decision making
- d) Shares evaluation findings with other organisations
- e) Has a system that allows us to learn successful practices from other organisations
- f) Is worried about evaluation findings making the organisation look bad

Organisational culture for systems and structures

- a) Evaluation is an important part of our organisational culture
- b) Evaluation of our programs is part of our core business
- c) Everyone in the organisation has some responsibility to contribute to evaluation

My organisation...

- a) Embeds evaluation in everything we do
- b) Uses evaluation findings to inform the next phase of work
- c) Actively provides staff with opportunities to upskill in the area of evaluation

Systems and structures

- a) It is clear from my job description what is expected of me in terms of evaluation
- b) We have an evaluation framework in place for the health promotion work that we do
- c) We have a user friendly way to store and manage our program data (electronic, or paper based)
- d) It is easy to access evaluation expertise from within the organisation
- e) There is a process for seeking advice from external evaluation experts if needed
- f) We have regular meetings to facilitate shared understanding regarding the evaluation plan
- g) Our organisation has clear expectations and timeframes for reporting

Internal support

- a) There is adequate support within my organisation for the evaluation I'm involved in
- b) My organisation arranges structured support and training for evaluation as required
- c) Management is happy to support evaluation and allows time and resources for it
- d) We have a strong working relationship with people who can support our team in evaluation

External support

- a) I am comfortable contacting people outside of my organisation for evaluation support and advice
- b) Formal arrangements with other organisations support our evaluation practice
- c) I get evaluation support from a network of practitioners I am involved with

Tools for evaluation

- a) We have access to data collection instruments appropriate to our work
- b) We have a functioning database to collate information
- c) We have appropriate tools for evaluation involving hard to reach and vulnerable groups
- d) Our organisation provides us with the basic resources to support evaluation (e.g., computers, software, copying, administrative support)
- e) We can access a range of good templates for reporting on our evaluation

Time for evaluation

- a) We have time to engage with our community and stakeholders for evaluation
- b) We have adequate time for evaluation planning
- c) We have enough time to think and reflect on our work
- d) We have time to write up our findings to share them

Resources dedicated to evaluation

- a) Our organisation has long-term, dedicated financial support for our evaluation activities
- b) Evaluation funding is set aside as part of every program budget
- c) Even when there are budget constraints we still allocate funds to evaluation
- d) Staff are encouraged to spend time on program evaluation

Evaluation role in the organisation

The evaluation person or team...

- a) Champions evaluation of health promotion projects in our organisation.
- b) Supports the health promotion program staff to undertake elements of the evaluation
- c) Mentors staff members in evaluation practice
- d) Is involved in helping to plan and design our evaluations right from the beginning

Partnerships

- a) Partnerships allow us to attract additional resources for evaluation
- b) Involving our partners in evaluation provides additional skills for evaluation
- c) Our partners help to facilitate evaluation involving hard to reach population groups
- d) Our partnerships are valuable for obtaining data needed in our evaluation
- e) There are clear expectations about evaluation between our partners
- f) Involving our partners in evaluation of the programs is worth the effort

External Evaluators

We engage external evaluators...

- a) To ensure the findings are independent
- b) To ensure a high quality of evaluation
- c) When we don't have the required skills internally

University partnerships

- a) We work with university partners to help obtain funds for evaluation
- b) Our university partners understand how we work
- c) We have the opportunity to build evaluation capacity through our contacts with university partners

System level

Adequate funding for evaluation

We have sufficient funding for...

- a) Health promotion evaluation (overall)
- b) Formative evaluation (i.e. pretesting and piloting)
- c) Process evaluation (i.e. monitoring of implementation)
- d) Impact and/or outcome evaluation (i.e. achievement of objectives)
- e) Outcome evaluation (i.e. achievement of project goals)

Clear and realistic expectations

- a) There are clear evaluation requirements from our funding body
- b) There are clear expectations from our funding body about the level of investment required in evaluation
- c) The evaluation activities I engage in are consistent with funders' expectations
- d) The funding body has realistic expectations for evaluation given available resources

Control and flexibility

- a) We have good flexibility from our funding body to evaluate the way we think is most appropriate
- b) We have a high level of control over the indicators chosen to evaluate and report on
- c) There is sufficient time available to develop an evaluation plan that meets stakeholder requirements

Competing demands

- a) It is challenging to meet funding body requirements for evaluation due to the complex nature of our programs
- b) There is pressure to deliver the programs at the expense of quality evaluation
- c) Evaluation is cut or reduced when the budget is tight overall
- d) Our funding body is generally more concerned with health service delivery than health promotion or prevention
- e) There is tension between what is important for us to evaluate locally and the funding body's priorities for evaluation

Organisational influence in the system

- a) We've built political support for our health promotion or disease prevention work to continue and grow
- b) We go to extra lengths to secure funding for evaluation beyond the usual program funding
- c) We distil our evaluation findings into key messages that help us engage politicians
- d) We have had success in securing additional funds for program evaluation

Learning and sharing

- a) Our funding body communicates clearly about how they will use our evaluation
- b) Our funding body encourages us to share our evaluation findings
- c) Our funding body provides leadership for conducting evaluation
- d) The funding body supports learning from evaluation findings

Political decision making

- a) Changes in funding priorities have impacted on the evaluation of our programs
- b) Despite having evaluation findings available, decisions appear to be made based on political factors
- c) Despite demonstrating effectiveness, projects have not been refunded
- d) It's more important for us to evaluate using rigorous methods in the current political climate
- e) Evaluation findings that include an economic component are used more often by decision makers

Reporting facilitates evaluation

- a) I am confident to use the reporting templates that we have been provided with
 - b) Evaluation is expected as part of the funding arrangements
 - c) Our reporting requirements and key performance indicators drive our evaluation work
 - d) Our funding body provides a reporting template
-

Referenties

1. IOM (Institute of Medicine). Preventing psychological disorders in service members and their families: An assessment of programs. Washington, DC: The National Academies Press; 2014.
2. Caplan G, Grunebaum H. Perspectives on primary prevention. A review. Arch Gen Psychiatry. 1967;17(3):331–46.
3. Hage S, Romano J. Best Practices in Prevention. Thousand Oaks, California: SAGE Publications, Inc.; 2013. Available from: <https://sk.sagepub.com/book/mono/best-practices-in-prevention/toc>.
4. Gordon RS. An operational classification of disease prevention. Public Health Rep. 1983;98(2):107–9.
5. World Health Organization. Summary reports on proceedings minutes and final acts of the international health conference held in New York from 19 June to 22 July 1946. New York: WHO. 1948;35.
6. Marmot M, Friel S, Bell R, Houweling TA, Taylor S. Closing the gap in a generation: health equity through action on the social determinants of health. Lancet. 2008;372(9650):1661–9.
7. Kickbusch I. Health in all policies: where to from here? Health Promotion International. 2010;25(3):261–4.
8. Vlaamse Overheid. Decreet betreffende het preventieve gezondheidsbeleid. België: Vlaamse Overheid; 2003.
9. Vlaamse Overheid. Het preventielexicon. In: Zorg D, editor. België: Departement Zorg; 2025.
10. WHO. The Ottawa Charter for Health Promotion.; 1986.
11. Arango C, Diaz-Caneja CM, McGorry PD, Rapoport J, Sommer IE, Vorstman JA, et al. Preventive strategies for mental health. Lancet Psychiatry. 2018;5(7):591–604.
12. Loss J. Intervention concepts in prevention. In: Kirch W, editor. Encyclopedia of Public Health. Dordrecht: Springer Netherlands; 2008. p. 808–11.
13. Rootman I, Goodstadt M, Hyndman B, McQueen DV, Potvin L, Springett J, et al. Evaluation in health promotion: principles and perspectives. Copenhagen: World Health Organization. Regional Office for Europe; 2001 2001.
14. O'Cathain A, Croot L, Duncan E, Rousseau N, Sworn K, Turner KM, et al. Guidance on how to develop complex interventions to improve health and healthcare. BMJ Open. 2019;9(8):e029954.
15. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ. 2021;374:n2061.
16. Prochaska JJ, Prochaska JO. A Review of Multiple Health Behavior Change Interventions for Primary Prevention. Am J Lifestyle Med. 2011;5(3).
17. Zhang AL, Liu S, White BX, Liu XC, Durantini M, Chan MS, et al. Health-promotion interventions targeting multiple behaviors: A meta-analytic review of general and behavior-specific processes of change. Psychol Bull. 2024;150(7):798–838.
18. Litvak E, Dufour R, Leblanc É, Kaiser D, Mercure SA, Nguyen CT, et al. Making sense of what exactly public health does: a typology of public health interventions. Can J Public Health. 2020;111(1):65–71.
19. World Health Organization. Regional Office for Europe. Evaluating the impact of interventions addressing health behaviour: considerations and tools for policy-makers. 2024. Contract No.: WHO/EURO:2024-10200-49972-75147.
20. Bernet AC, Willens DE, Bauer MS. Effectiveness-implementation hybrid designs: implications for quality improvement science. Implementation Science. 2013;8(1):S2.
21. NASEM. Appendix C, Evaluation Types and Data Requirements. In: Wells KB, Amaro H, McHugh M, Stratton K, editors. Review of Four CARA Programs and Preparing for Future Evaluations. Washington (DC): The National Academies Press; 2023.

22. Cargo M, Mercer SL. The value and challenges of participatory research: strengthening its practice. *Annu Rev Public Health*. 2008;29:325–50.
23. Patton MQ, Campbell-Patton CE. *Utilization-Focused Evaluation*: SAGE Publications; 2021.
24. Akrami F, Zali A, Abbasi M, Majdzadeh R, Karimi A, Fadavi M, et al. An ethical framework for evaluation of public health plans: a systematic process for legitimate and fair decision-making. *Public Health*. 2018;164:30–8.
25. OECD. *Better Criteria for Better Evaluation: Revised Evaluation Criteria Definitions and Principles for Use*. Paris: OECD Publishing; 2019.
26. Puddy RW, Wilkins N, Centers for Disease Control and P, National Center for Injury Prevention and Control . Division of Violence Prevention ib. *Understanding evidence.: a guide to the continuum of evidence of effectiveness : Part 1: Best available research evidence*. Atlanta, GA: National Center for Injury Prevention and Control, Division of Violence Prevention; 2011.
27. Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am J Community Psychol*. 2008;41(3-4):327–50.
28. Klaic M, Kapp S, Hudson P, Chapman W, Denehy L, Story D, et al. Implementability of healthcare interventions: an overview of reviews and development of a conceptual framework. *Implementation Science*. 2022;17(1):10.
29. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ : British Medical Journal*. 2015;350:h1258.
30. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. *Med Care*. 2012;50(3):217–26.
31. Rossi P, Lipsey M, Freeman H. *Evaluation: A systematic approach*. 7 ed. United States of America: Sage Publications, Inc.; 2004.
32. Nutbeam D. Evaluating Health Promotion—Progress, Problems and solutions. *Health Promotion International*. 1998;13(1):27–44.
33. Bonell C, Jamal F, Melendez-Torres GJ, Cummins S. 'Dark logic': theorising the harmful consequences of public health interventions. *J Epidemiol Community Health*. 2015;69(1):95–8.
34. Lorenc T, Petticrew M, Welch V, Tugwell P. What types of interventions generate inequalities? Evidence from systematic reviews. *J Epidemiol Community Health*. 2013;67(2):190–3.
35. Mayne J. Contribution analysis: An approach to exploring cause and effect. *Institutional Learning and Change (ILAC) Initiative, ILAC Briefs*. 2008;16.
36. Lobo R, Petrich M, Burns SK. Supporting health promotion practitioners to undertake evaluation for program development. *BMC Public Health*. 2014;14(1):1315.
37. Greer SL, Vasev N, Jarman H, Wismar M, Figueras J. *European Observatory Policy Briefs. It's the governance, stupid! TAPIC: a governance framework to strengthen decision making and implementation*. Copenhagen (Denmark): European Observatory on Health Systems and Policies
- © World Health Organization 2019 (acting as the host organization for, and secretariat of, the European Observatory on Health Systems and Policies). 2019.
38. Schwarzman J, Bauman A, Gabbe BJ, Rissel C, Shilton T, Smith BJ. How practitioner, organisational and system-level factors act to influence health promotion evaluation capacity: Validation of a conceptual framework. *Evaluation and Program Planning*. 2022;91:102019.
39. Nielsen SB, Lemire S, Bourgeois I, Fierro LA. Capturing Evaluation Capacity: Findings from a Mapping of Evaluation Capacity Instruments. *Canadian Journal of Program Evaluation*. 2024;38(3):433–60.
40. Labin SN, Duffy JL, Meyers DC, Wandersman A, Lesesne CA. A Research Synthesis of the Evaluation Capacity Building Literature. *American Journal of Evaluation*. 2012;33(3):307–38.
41. Ignacio LA, Meigh N, Dixon S, Bateup S, Mickan S. Core competencies of health program evaluators: A scoping review. *Evaluation and Program Planning*. 2025;111:102604.

42. Bourgeois I, Cousins JB. Understanding Dimensions of Organizational Evaluation Capacity. *American Journal of Evaluation*. 2013;34(3):299–319.
43. Schwarzman J, Bauman A, Gabbe BJ, Rissel C, Shilton T, Smith BJ. Understanding the factors that influence health promotion evaluation: The development and validation of the evaluation practice analysis survey. *Evaluation and Program Planning*. 2019;74:76–83.
44. Bourgeois I, Maltais S. Translating Evaluation Policy Into Practice in Government Organizations. *Am J Eval*. 2023;44(3):353–73.
45. Chen H, Hailey D, Wang N, Yu P. A review of data quality assessment methods for public health information systems. *Int J Environ Res Public Health*. 2014;11(5):5170–207.
46. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*. 2016;3(1):160018.
47. Broholm-Holst M, Tjørnhøj-Thomsen T. Ethical reflexivity: A critical examination of ethics in public health intervention research. *Evaluation*. 2025;13563890251367001.
48. American Evaluation Association. Principles for Effective Use of Systems Thinking in Evaluation. 2018.
49. Agnello DM, Anand-Kumar V, An Q, de Boer J, Delfmann LR, Longworth GR, et al. Co-creation methods for public health research — characteristics, benefits, and challenges: a Health CASCADE scoping review. *BMC Medical Research Methodology*. 2025;25(1):60.
50. Movsisyan A, Arnold L, Copeland L, Evans R, Littlecott H, Moore G, et al. Adapting evidence-informed population health interventions for new contexts: a scoping review of current practice. *Health Research Policy and Systems*. 2021;19(1):13.
51. Pearce T, Maple M, Shakeshaft A, Wayland S, McKay K. What Is the Co-Creation of New Knowledge? A Content Analysis and Proposed Definition for Health Interventions. *Int J Environ Res Public Health*. 2020;17(7).
52. Porroche-Escudero A, Popay J. Reflections on Mainstreaming Health EquityHealth equity in a Large Research Collaboration: “If I can’t dance it is not my revolution”. In: Potvin L, Jourdan D, editors. *Global Handbook of Health Promotion Research, Vol 1: Mapping Health Promotion Research*. Cham: Springer International Publishing; 2022. p. 693–704.
53. Faruquzzaman O, Tse S, Cameron R, Chee K, Kirst M, Wilson C. Approaches to incorporating equity into program evaluation: A scoping review. *Evaluation and Program Planning*. 2025;112:102648.
54. Perkins AN, Nowakowski AXCH, Alva S, Rimal N, Sharer M, Davis N, et al. Expanding the Lens From Health Outcomes to Learn More About Key Stakeholders and the Structural Factors as a Pathway Toward Health Equity. *American Journal of Evaluation*. 2023;46(3):470–83.
55. Kidder DP, Fierro LA, Luna E, Salvaggio H, McWhorter A, Bowen SA, et al. CDC Program Evaluation Framework, 2024. *MMWR Recomm Rep*. 2024;73(6):1–37.
56. Funnell SC, Rogers PJ. *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*: Wiley; 2011.
57. United Nations Development Group. *Theory of Change*. 2017.
58. World Health Organization. *How to develop an evidence-informed theory of change for health: WHO technical guidance*. Geneva: World Health Organization; 2024 2024–12–31.
59. Breuer E, Lee L, De Silva M, Lund C. Using theory of change to design and evaluate public health interventions: a systematic review. *Implementation Science*. 2016;11(1):63.
60. WHO Centre for Health Development. *WHO guidance on research methods for health emergency and disaster risk management*. Kayano R, Murray V, Clarke M, Chan EYY, editors: World Health Organization; 2022.
61. Kellogg WK. *WK Kellogg Foundation Logic Model Development Guide*. Battle Creek, MI: WK Kellogg Foundation; 2004.
62. Lawlor ER, Cupples ME, Donnelly M, Tully MA. Promoting physical activity among community groups of older women in socio-economically disadvantaged areas: randomised feasibility study. *Trials*. 2019;20(1):234.

63. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health*. 1999;89(9):1322–7.
64. Holtrop JS, Estabrooks PA, Gaglio B, Harden SM, Kessler RS, King DK, et al. Understanding and applying the RE-AIM framework: Clarifications and resources. *J Clin Transl Sci*. 2021;5(1):e126.
65. Gaglio B, Shoup JA, Glasgow RE. The RE-AIM Framework: A Systematic Review of Use Over Time. *American Journal of Public Health*. 2013;103(6):e38–e46.
66. Feldstein AC, Glasgow RE. A practical, robust implementation and sustainability model (PRISM) for integrating research findings into practice. *Jt Comm J Qual Patient Saf*. 2008;34(4):228–43.
67. Rebekah Gomes CB, Fort M, Maw A, McCreight M, Borsika Rabin ER, Studts C, et al. A guidebook to the pragmatic and iterative use of the practical, robust implementation and sustainability model (PRISM) and reach, effectiveness, adoption, implementation, maintenance framework (RE-AIM) for planning, implementation, and sustainment. The Colorado Implementation Science Center for Cancer Control.
68. Rabin BA, Cakici J, Golden CA, Estabrooks PA, Glasgow RE, Gaglio B. A citation analysis and scoping systematic review of the operationalization of the Practical, Robust Implementation and Sustainability Model (PRISM). *Implement Sci*. 2022;17(1):62.
69. Jolles MP, Fort MP, Glasgow RE. Aligning the planning, development, and implementation of complex interventions to local contexts with an equity focus: application of the PRISM/RE-AIM Framework. *International Journal for Equity in Health*. 2024;23(1):41.
70. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*. 2015;350:h1258.
71. Westhorp G. *Realist Impact Evaluation: An Introduction*. 2014.
72. Louart S, Ridde V. Critical Realism for Health Promotion Evaluation. In: Jourdan D, Potvin L, editors. *Global Handbook of Health Promotion Research, Vol 3: Doing Health Promotion Research*. Cham: Springer International Publishing; 2023. p. 49–59.
73. Pawson R, Tilley N. An introduction to scientific realist evaluation. *Evaluation for the 21st century: A handbook*. Thousand Oaks, CA, US: Sage Publications, Inc; 1997. p. 405–18.
74. Malengreaux S, Doumont D, Scheen B, Van Durme T, Aujoulat I. Realist evaluation of health promotion interventions: a scoping review. *Health Promot Int*. 2022;37(5).
75. Wong G, Westhorp G, Manzano A, Greenhalgh J, Jagosh J, Greenhalgh T. RAMESES II reporting standards for realist evaluations. *BMC Medicine*. 2016;14(1):96.
76. Pommier J, Guével M-R, Jourdan D. Evaluation of health promotion in schools: a realistic evaluation approach using mixed methods. *BMC Public Health*. 2010;10(1):43.
77. Befani B, Mayne J. Process Tracing and Contribution Analysis: A Combined Approach to Generative Causal Inference for Impact Evaluation. *IDS Bulletin*. 2014;45(6):17–36.
78. Buetti D, Fitzgerald M, Barber C, Labelle PR, Bourgeois I, Aubry T, et al. The Use of Contribution Analysis in Evaluating Health Interventions: A Scoping Review. *Evaluation & the Health Professions*. 2024;48(3):363–73.
79. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. *Implementation Science*. 2022;17(1):75.
80. Reardon CM, Damschroder LJ, Ashcraft LE, Kerins C, Bachrach RL, Nevedal AL, et al. The Consolidated Framework for Implementation Research (CFIR) User Guide: a five-step guide for conducting implementation research using the framework. *Implementation Science*. 2025;20(1):39.
81. Damschroder LJ, Reardon CM, Opra Widerquist MA, Lowery J. Conceptualizing outcomes for use with the Consolidated Framework for Implementation Research (CFIR): the CFIR Outcomes Addendum. *Implementation Science*. 2022;17(1):7.
82. Waltz TJ, Powell BJ, Fernández ME, Abadie B, Damschroder LJ. Choosing implementation strategies to address contextual barriers: diversity in recommendations and future directions. *Implementation Science*. 2019;14(1):42.

83. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health*. 2011;38(2):65–76.
84. Proctor EK, Bunger AC, Lengnick-Hall R, Gerke DR, Martin JK, Phillips RJ, et al. Ten years of implementation outcomes research: a scoping review. *Implement Sci*. 2023;18(1):31.
85. Smith JD, Li DH, Rafferty MR. The Implementation Research Logic Model: a method for planning, executing, reporting, and synthesizing implementation projects. *Implementation Science*. 2020;15(1):84.
86. Green LW, Kreuter MW. *Health Promotion Planning: An Educational and Ecological Approach*: Mayfield Publishing Company; 1999.
87. Center for Community Health and Development. PRECEDE/PROCEED University of Kansas [Available from: <https://ctb.ku.edu/en/table-contents/overview/other-models-promoting-community-health-and-development/preceder-proceder/main>].
88. Kim J, Jang J, Kim B, Lee KH. Effect of the PRECEDE-PROCEED model on health programs: a systematic review and meta-analysis. *Systematic Reviews*. 2022;11(1):213.
89. Gielen AC, McDonald EM, Gary TL, Bone LR. Using the precece-proceed model to apply health behavior theories. *Health behavior and health education: Theory, research, and practice*, 4th ed. Hoboken, NJ, US: Jossey-Bass/Wiley; 2008. p. 407–33.
90. Fernandez ME, Ruiter RAC, Markham CM, Kok G. Intervention Mapping: Theory- and Evidence-Based Health Promotion Program Planning: Perspective and Examples. *Front Public Health*. 2019;7:209.
91. Doorn R, Massar K, Kok G. Gedragsverandering binnen organisaties: Kan Intervention Mapping een bijdrage leveren aan het ontwikkelen van effectieve interventies? *Gedrag en Organisatie*. 2018;31.
92. Coone A, Jongbloet J, van der Kreeft P. *Europees Preventiehandboek. Een leidraad voor wetenschappelijk onderbouwde preventie in de praktijk, vertaling van European Prevention Curriculum (EMCDDA. Politeia, editor2020.*
93. Marchand E, Stice E, Rohde P, Becker CB. Moving from efficacy to effectiveness trials in prevention research. *Behaviour Research and Therapy*. 2011;49(1):32–41.
94. Beets MW, von Klinggraeff L, Weaver RG, Armstrong B, Burkart S. Small studies, big decisions: the role of pilot/feasibility studies in incremental science and premature scale-up of behavioral interventions. *Pilot and Feasibility Studies*. 2021;7(1):173.
95. Leviton LC, Khan LK, Rog D, Dawkins N, Cotton D. Evaluability assessment to improve public health policies, programs, and practices. *Annu Rev Public Health*. 2010;31:213–33.
96. Ogilvie D, Cummins S, Petticrew M, White M, Jones A, Wheeler K. Assessing the evaluability of complex public health interventions: five questions for researchers, funders, and policymakers. *Milbank Q*. 2011;89(2):206–25.
97. Adviescommissie richtlijn passend bewijs voor preventie. Passend bewijs voor preventie. In: Ministerie van Volksgezondheid WeS, editor. 2025.
98. Coalitie Leefstijl in de Zorg. *Handreiking onderzoeksmethoden leefstijlinterventies in de zorg | Mogelijkheden voor het evalueren van effectiviteit*. 2025.
99. Tock WL, Maheu C, Johnson NA. Considerations of Control Conditions Designs in Randomized Controlled Trials of Exercise Interventions for Cancer Survivors. *Canadian Journal of Nursing Research*. 2022;54(4):377–91.
100. Stuart EA. Matching methods for causal inference: A review and a look forward. *Stat Sci*. 2010;25(1):1–21.
101. Handley MA, Lyles CR, McCulloch C, Cattamanchi A. Selecting and Improving Quasi-Experimental Designs in Effectiveness and Implementation Research. *Annu Rev Public Health*. 2018;39:5–25.
102. Calonico S, Jawadekar N, Kezios K, Zeki Al Hazzouri A. Regression discontinuity design studies: a guide for health researchers. *BMJ*. 2024;384:e072254.
103. Bouttell J, Popham F, Lewsey J, Robinson M, Craig P. Use of synthetic control methodology for evaluating public health interventions: a literature review. *The Lancet*. 2017;390:S26.

104. Levasseur M, Chaintré-Prieur A, Dubois MF, Maisonneuve C, Filiatrault J, Vassiliadis HM. Strengths, challenges, and strategies for implementing pragmatic multicenter randomized controlled trials (RCTs): example of the Personalized Citizen Assistance for Social Participation (APIC) trial. *Trials*. 2024;25(1):415.
105. Cabinet Office Behavioural Insights Team. Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials. Verenigd Koninkrijk; 2012.
106. Rosen L, Manor O, Engelhard D, Zucker D. In defense of the randomized controlled trial for health promotion research. *Am J Public Health*. 2006;96(7):1181–6.
107. Fetters MD, Molina-Azorin JF. Utilizing a Mixed Methods Approach for Conducting Interventional Evaluations. *Journal of Mixed Methods Research*. 2020;14(2):131–44.
108. Creswell JW, Plano Clark VL. Designing and conducting: mixed methods research. 2nd ed. Los Angeles: Sage; 2011.
109. UK Office for Health Improvement and Disparities. Outcome evaluation: evaluation in health and wellbeing <https://www.gov.uk/guidance/evaluation-in-health-and-wellbeing-outcome2018> [
110. World Health Organization - Eastern Mediterranean Region. Indicators <https://www.emro.who.int/child-health/research-and-evaluation/indicators/Type-of-indicators.html>2014 [Available from: <https://www.emro.who.int/child-health/research-and-evaluation/indicators/Type-of-indicators.html>].
111. Gottfredson DC, Cook TD, Gardner FE, Gorman-Smith D, Howe GW, Sandler IN, et al. Standards of Evidence for Efficacy, Effectiveness, and Scale-up Research in Prevention Science: Next Generation. *Prev Sci*. 2015;16(7):893–926.
112. Farrington DP, Gottfredson DC, Sherman LW, Welsh BC. The Maryland scientific methods scale. Evidence-based crime prevention: Routledge; 2003. p. 13–21.
113. Madaleno M, Waights S. Guide to scoring methods using the Maryland Scientific Methods Scale. London: What Works Centre for Local Economic Growth. 2015.
114. Moore G, Campbell M, Copeland L, Craig P, Movsisyan A, Hoddinott P, et al. Adapting interventions to new contexts-the ADAPT guidance. *Bmj*. 2021;374:n1679.
115. Bagnall A-M, Radley D, Jones R, Gately P, Nobles J, Van Dijk M, et al. Whole systems approaches to obesity and other complex public health challenges: a systematic review. *BMC Public Health*. 2019;19(1):8.
116. Michel J, Schneider K. Demystifying impact evaluation: an impact evaluation framework. *Front Epidemiol*. 2025;5:1460997.
117. New Philanthropic Capital (NPC). Developing theories of change for entire sectors 2016 [Available from: <https://www.thinknpc.org/blog/developing-theories-of-change-for-entire-sectors/>].
118. Panjwani S, Graves-Boswell T, Garney WR, Muraleetharan D, Spadine M, Flores S. Evaluating Collective Impact Initiatives: A Systematic Scoping Review. *American Journal of Evaluation*. 2023;44(3):406–23.
119. Cabaj M. Evaluating collective impact: Five simple rules. *The Philanthropist*. 2014;26(1).
120. NPC. Inspiring Impact: Catalysing shared measurement. The future of shared measurement - A guide to assessing individual sector readiness. 2014.
121. Lambeth Early Action Partnership (LEAP). Shared Measurement System – Learning Journey. 2025.
122. Atwal S, Schmider J, Buchberger B, Boshnakova A, Cook R, White A, et al. Prioritisation processes for programme implementation and evaluation in public health: A scoping review. *Front Public Health*. 2023;11:1106163.
123. Cowan K, Fulop NJ, Harshfield A, Ng PL, Ntouva A, Sidhu M, et al. Rapid prioritisation of topics for rapid evaluation: the case of innovations in adult social care and social work. *Health Research Policy and Systems*. 2021;19(1):34.
124. Downey LH, Peterson DJ, LeMenestrel S, & Leatherman J. The Systematic Screening and Assessment Method: An Introduction and Application. *The Journal of Extension*. 2015;53(2).
125. Leviton L, Gutman M. Overview and Rationale for the Systematic Screening and Assessment Method. *New Directions for Evaluation*. 2010;2010:7–31.

-
126. Piervil E, Wong L, Marshall KJ, Earl T, Leonard S, Waajid M, et al. Systematic Screening and Assessment of Hospital-Based Youth Violence Prevention Programs. *Health Promot Pract.* 2025;26(2):372–80.
127. Spilsbury MJ, Norgbey S, Battaglini C. Priority setting for evaluation: Developing a strategic evaluation portfolio. *Evaluation and Program Planning.* 2014;46:47–57.
128. La Brooy C, Kelaher M. The research–policy–deliberation nexus: a case study approach. *Health Research Policy and Systems.* 2017;15(1):75.
129. Murray L. Deliberative Research for Deliberative Policy Making: Creating and Recreating Evidence in Transport Policy. *Social Policy and Society.* 2011;10(4):459–70.
130. Samenwerkingsverband Erkenning van Interventies. *Erkenningssystematiek gezondheidsbevorderende interventies.*; 2024–2027.
131. Van Yperen T, Veerman J. *Zicht op effectiviteit: handboek voor praktijkgestuurd effectonderzoek in de jeugdzorg*; Eburon Uitgeverij BV; 2008.
132. Mitton C, Dionne F, Donaldson C. Managing healthcare budgets in times of austerity: the role of program budgeting and marginal analysis. *Appl Health Econ Health Policy.* 2014;12(2):95–102.
133. Gongora-Salazar P, Rocks S, Fahr P, Rivero-Arias O, Tsiachristas A. The Use of Multicriteria Decision Analysis to Support Decision Making in Healthcare: An Updated Systematic Literature Review. *Value in Health.* 2023;26(5):780–90.
134. Edwards RT, Charles JM, Thomas S, Bishop J, Cohen D, Groves S, et al. A national Programme Budgeting and Marginal Analysis (PBMA) of health improvement spending across Wales: disinvestment and reinvestment across the life course. *BMC Public Health.* 2014;14(1):837.
135. Adviescommissie richtlijn passend bewijs voor preventie. *Passend bewijs voor preventie.* In: Ministerie van Volksgezondheid WeS, editor. 2025.
136. Lohan M, Aventin Á, Clarke M, Curran RM, Maguire L, Hunter R, et al. JACK trial protocol: a phase III multicentre cluster randomised controlled trial of a school-based relationship and sexuality education intervention focusing on young male perspectives. *BMJ Open.* 2018;8(7):e022128.
137. Rebekah Gomes CB, Fort M, Maw A, McCreight M, Borsika Rabin ER, Studts C, et al. A guidebook to the pragmatic and iterative use of the practical, robust implementation and sustainability model (PRISM) and reach, effectiveness, adoption, implementation, maintenance framework (RE-AIM) for planning, implementation, and sustainment. The Colorado Implementation Science Center for Cancer Control. 2023.

