



Samenvatting onderzoeksrapport

EEN EFFECTONDERZOEK NAAR DE ANDER LEVEN INTERVENTIE
OP BASIS VAN DE SF-36 VRAGENLIJST

Tim Choy

18-10-2018

Inhoudsopgave

1. Inleiding	2
2. Onderzoeksopzet.....	2
3. Steekproef en generaliseerbaarheid	4
4. Kenmerken onderzoeksgroep	6
5. Uitkomsten op de voor- en nametingen	7
6. Effectiviteit van de Ander Leven interventie.....	10
7. Percentage deelnemers dat duurzaam herstelt.....	15
8. Conclusies.....	16
 Bijlage: Bedrijfsprofiel Evidento	 17

1. Inleiding

Ander Leven past een interventie toe op mensen met (chronische) lichamelijke en geestelijke klachten. De interventie bestaat uit een consult, een driedaagse training en een verdieping/terugkomdag. Daarnaast is er een professioneel ondersteuningsproces bestaande uit intensieve nazorg (bijvoorbeeld tot aan de verdieping/terugkomdag altijd kunnen bellen) en reguliere nazorg (bijvoorbeeld een jaar na de training contact kunnen opnemen), een zelfregulerend ondersteuningsproces (o.a. buddy-systeem, Ander Leven CD en dagboek bijhouden) en een zelfregulerend helingsproces: het (thuis) zelfstandig lopen van AC-processen.

Sinds 2011 verzamelt Ander Leven data over de gezondheidstoestand c.q. het welzijn van cliënten voor aanvang en na toepassing van de interventie aan de hand van de SF/RAND-36 vragenlijst. Ander Leven wil namelijk weten in hoeverre de interventie effectief is: *verbetert de lichamelijke en geestelijke gezondheid na de interventie en is het resultaat blijvend?* Door antwoord op deze vragen te krijgen kan Ander Leven uitspraken doen over de effectiviteit van haar interventie en belanghebbenden daarover informeren.

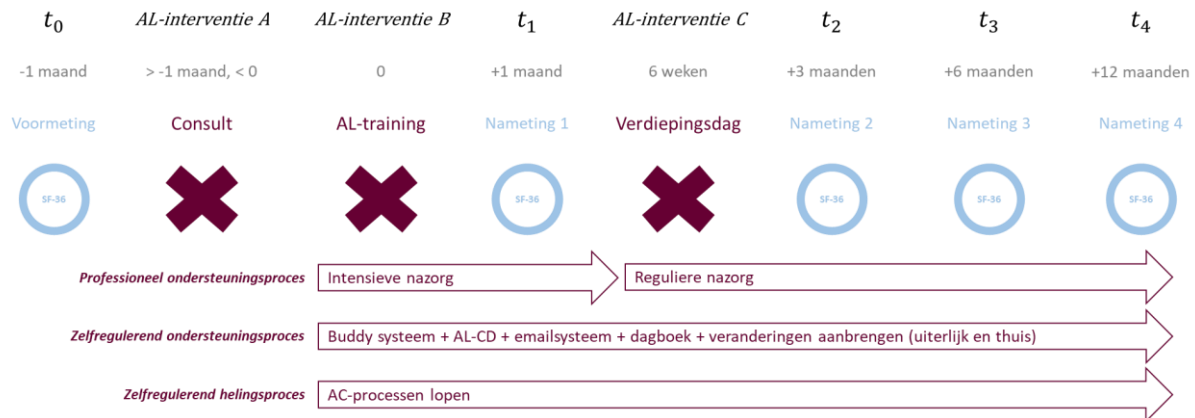
Ander Leven heeft in het voorjaar van 2017 een onafhankelijk wetenschappelijk onderzoeksbureau (Evidento) ingeschakeld om effectonderzoek naar de interventie te doen. De uitkomsten van het onderzoek geven meer uitsluitsel over de effectiviteit van de Ander Leven interventie en in hoeverre deze als 'evidence based' kan worden aangemerkt.

2. Onderzoeksopzet

De onderzoeksopzet is bepaald aan de hand van hoe Ander Leven mensen selecteert om deel te nemen aan de interventie, bij wie gegevens zijn verzameld, op welke tijdstippen gegevens zijn verzameld en welke gegevens zijn verzameld. Dit gaf een aantal mogelijkheden om adequaat effectonderzoek te kunnen doen, maar ook een aantal beperkingen.

Deelnemers met lichamelijke en/of geestelijke klachten, waarop de interventie is toegepast, zijn 'select' geselecteerd op basis van een (telefonisch) intake gesprek. De deelnemers zijn niet 'aselect' geselecteerd, waardoor niet elke aanmelding evenveel kans heeft om de interventie te ondergaan. Deelnemers moeten dus aan bepaalde (motivatie gedreven) voorwaarden voldoen. Dit is een eerste beperking van het onderzoek, omdat niet bekend is hoe de interventie uitwerkt op mensen met lichamelijke en/of geestelijke klachten die niet geselecteerd zijn.

In dit onderzoek is sprake van een "interventiegroep" waarop de Ander Leven interventie is toegepast en bestaat uit alle groepen c.q. deelnemers die tussen 2011 en 2017 de interventie hebben ondergaan. Op basis van deze groep moet bepaald worden in hoeverre de interventie effectief is gebleken. Dit gebeurt door de meting op de gezondheidstoestand van de deelnemers voor aanvang van de interventie te vergelijken met metingen op de gezondheidstoestand van de deelnemers na toepassing van (onderdelen van de) interventie. De voormeting vindt plaats voor het consult en de driedaagse training. De nametingen vinden vervolgens plaats na 1, 3, 6 en 12 maanden (tussen de eerste en tweede nameting zit de verdieping/terugkomdag). De figuur op de volgende pagina geeft een visuele representatie van de meetpunten in relatie tot de toepassing van interventieonderdelen.



In dit onderzoek is geen controlegroep waarmee de interventiegroep vergeleken kan worden. Een controlegroep is een overeenkomstige groep van mensen met lichamelijke en/of geestelijke klachten die de interventie niet krijgt toegepast en die gelijktijdig met de interventiegroep de vragenlijst invult. Indien een effect van de interventie zichtbaar zou worden bij de interventiegroep en niet bij de controlegroep, zou er ook met meer zekerheid vastgesteld kunnen worden dat de interventie effect heeft. Maar door praktische en ethische redenen was het niet mogelijk om een controlegroep te installeren. Dit is een tweede beperking van het onderzoek.

De lichamelijke en geestelijke gezondheidstoestand van de interventiegroep is gemeten aan de hand van de SF/RAND-36 vragenlijst. Dit is een wetenschappelijk gevalideerde en betrouwbare vragenlijst en stelt bij respondenten de volgende aspecten vast:

- Algemene gezondheid
- Lichamelijke gezondheid
- Rolbeperkingen n.a.v. lichamelijke problemen
- Lichaamspijn
- Geestelijke gezondheid
- Rolbeperkingen n.a.v. geestelijke (emotionele) problemen
- Vitaliteit/energie
- Sociaal functioneren

De meetschaal van deze variabelen loopt van 1 tot 100, met als uitgangspunt: hoe hoger de score, hoe beter de gezondheid c.q. het welzijn. Daarnaast meet deze vragenlijst twee geaggregeerde aspecten: lichamelijk welzijn en geestelijk welzijn. Deze meetschaal kent een standaardisatiewaarde van 50, welke als normering geldt voor het landelijk gemiddelde van een westerse bevolking (een score onder de 50 is lager dan het landelijk gemiddelde en een score boven de 50 is hoger dan het landelijk gemiddelde).

Deelnemers zijn niet bevraagd op de mate waarin zij gebruik maken van het professionele en zelfregulerend ondersteuningsproces, alsmede in hoeverre zij het zelfregulerend helingsproces (thuis) uitvoeren. Daardoor is het niet mogelijk om te achterhalen wat de invloed van deze processen op de gezondheidstoestand heeft. Dit is een derde beperking van het onderzoek.

Om binnen de beperkingen van het onderzoek de effectiviteit van de interventie vast te stellen zal er een statistisch significant verschil moeten zijn tussen de voormeting en de vier nametingen binnen de interventiegroep: de lichamelijke en geestelijke gezondheidstoestand van de deelnemers moet na de interventie dusdanig verbeterd zijn dat deze verbetering – statistisch gezien – niet op toeval

berust. Daarbij komt dat de vier nametingen min of meer gelijk moeten blijven om aan te kunnen geven dat de interventie, over een periode van 12 maanden, blijvend effectief is.

Om dat vast te kunnen stellen zijn de voor- en nametingen vergeleken bij deelnemers die tijdens elke meting de vragenlijst volledig hebben ingevuld. Dit betekent dat in elke meting dezelfde deelnemers zitten (in onderzoekstermen worden dit *gepaarde* metingen of groepen genoemd). Dit maakt het onderzoek krachtiger, omdat een eventueel geconstateerd verschil tussen de voor- en nametingen dan minder snel toegerekend kan worden aan mogelijke verschillen in de karakteristieken van de deelnemers (denk bijvoorbeeld aan veel meer oudere deelnemers in de voormeting dan in de nameting, waardoor 'ouderdom' het verschil zou kunnen verklaren). Dat betekent ook dat een geconstateerd verschil tussen voor- en nametingen beter aan de interventie kan worden toegewezen.

Het design kan in onderzoekstermen worden aangeduid als een *non-gerandomiseerde longitudinale cohort studie* (Katz, 2010). Non-gerandomiseerd wil zeggen dat de mensen met (chronische) lichamelijke en/of geestelijke klachten die zich aanmelden bij Ander Leven niet evenveel kans hebben om de interventie te ondergaan en dat dit op willekeurige wijze is gebeurd (deelnemers zijn specifiek geselecteerd). Een longitudinale cohort studie betekent dat een groep mensen met een bepaald gemeenschappelijk kenmerk (in dit geval: lichamelijke en/of geestelijke klachten) in de tijd wordt gevolgd c.q. in de tijd wordt gemeten.

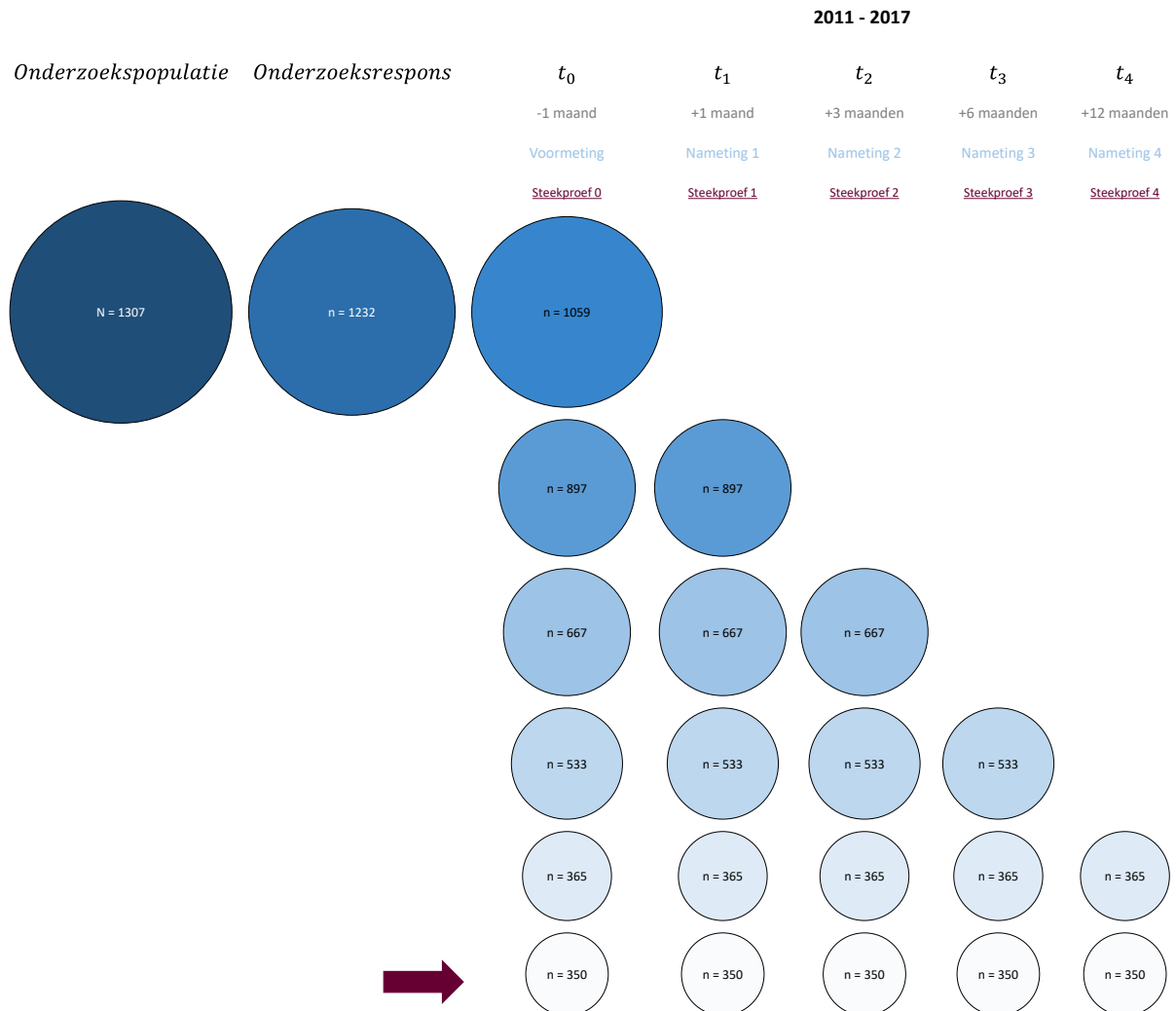
Het onderzoek doet dus uitspraak over het effect van het consult en de driedaagse training tezamen (verschil tussen de voormeting en eerste nameting), de verdieping/terugkomdag (verschil tussen de eerste nameting en de tweede nameting) en in hoeverre het effect stand houdt (verschil tussen de vier nametingen, ten opzichte van de voormeting). De veronderstelling van de ontwikkelaar(s) is dat een geconstateerd effect met name aan de driedaagse training is toe te wijzen en dat alle type deelnemers daar profijt van hebben.

Om daar uitspraken over te kunnen doen is het verschil tussen de voor- en nametingen ook gecontroleerd op karakteristieken van de deelnemers, o.a. geslacht, leeftijd, nationaliteit, inkomenssituatie en type klacht. Met andere woorden, als de interventie voor iedereen effectief is, dan moet het verschil tussen voor- en nametingen hetzelfde zijn voor bijvoorbeeld mannen en vrouwen, ouderen en jongeren, Nederlanders en Belgen, armen en rijken en mensen met lichamelijke en geestelijke klachten.

3. Steekproef en generaliseerbaarheid

Om uitspraken te kunnen doen op welke populatie van mensen de Ander Leven interventie effect sorteert, is het van belang om meer inzicht te krijgen in de doelpopulatie, onderzoekspopulatie en kenmerken van deelnemers in de steekproef. De doelpopulatie van Ander Leven zijn mensen in Nederland en Vlaanderen met (chronische) lichamelijke en/of geestelijke klachten die voldoen aan de selectievoorwaarden van Ander Leven. De onderzoekspopulatie bestaat uit geselecteerde deelnemers uit de doelpopulatie waarop de Ander Leven interventie, in de periode januari 2011 t/m november 2017, is toegepast (in totaal 1307 mensen). De steekproeven zijn deelnemers uit de onderzoekspopulatie die vijf keer (1 voormeting en 4 nametingen) de SF/RAND-36 vragenlijst volledig hebben ingevuld (in beginsel 365 mensen, maar na het opschonen van de data en het verwijderen van uitzonderlijke gevallen 350 mensen).

In onderstaande figuur staat de onderzoekspopulatie en de steekproefomvang bij elke meting visueel weergegeven. De steekproeven worden steeds kleiner, omdat niet iedereen de vragenlijst bij elke meting (volledig) invult. De vraag waarom deelnemers dat niet hebben gedaan en in hoeverre hun gezondheidstoestand daarmee te maken heeft, kan niet worden beantwoord en is een vierde beperking van het onderzoek (er is bijvoorbeeld geen non-respons onderzoek uitgevoerd, waarbij gevraagd wordt waarom deelnemers de vragenlijst niet invullen).



Op basis van de steekproef kan een uitspraak worden gedaan over de totale onderzoekspopulatie. Er zijn geen gegevens bekend over de omvang van de doelpopulatie, waardoor er geen uitspraken gedaan kunnen worden over de generaliseerbaarheid van de steekproeven naar de doelpopulatie.

Ruim 23% van de onderzoekspopulatie heeft de vragenlijst op elke meting volledig ingevuld. Maar statistisch zegt dit nog weinig over de representativiteit van de steekproef. De generaliseerbaarheid van de steekproeven hangt af van de steekproefomvang in relatie tot de omvang van de onderzoekspopulatie en wordt bepaald aan de hand van een *betrouwbaarheidsmarge* en een *foutmarge*.

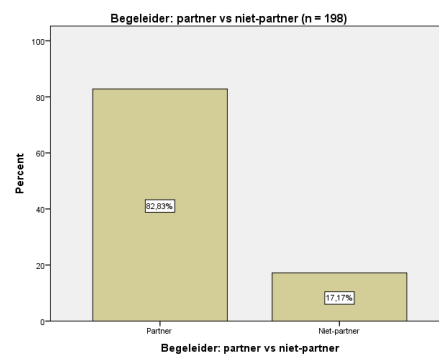
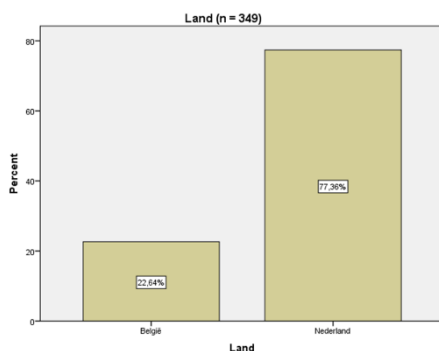
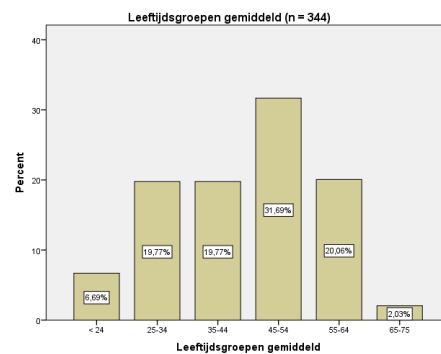
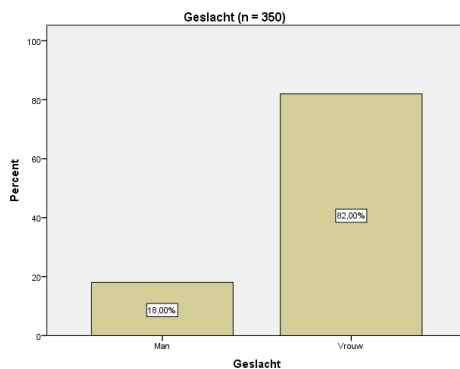
In de regel zijn de onderzoeksresultaten uit de steekproef te generaliseren naar de onderzoekspopulatie bij een betrouwbaarheidsmarge van 95% en een foutmarge van 5%. Dat wil zeggen dat als uit de onderzoekspopulatie 100 steekproeven worden getrokken met dezelfde steekproefomvang er in 95 gevallen overeenkomstige gemiddelde scores uitkomen (die niet ver van elkaar afwijken) en in 5 gevallen de gemiddelde scores – gegeven het toeval – veel af zullen wijken.

Om deze percentages aan te kunnen houden moet de steekproefomvang, ten opzichte van de omvang van de onderzoekspopulatie, een bepaalde grootte hebben. Dit wordt vastgesteld aan de hand van een statistische procedure. Deze wijst uit dat een steekproefomvang van 350 deelnemers ten opzichte van een onderzoekspopulatie van 1307 deelnemers, een foutmarge van 4,48% oplevert als het betrouwbaarheidsmarge op 95% wordt gezet. De foutmarge blijft dus onder de 5%, waardoor de steekproeven in dit onderzoek representatief zijn voor de onderzoekspopulatie en de onderzoeksresultaten gelden voor alle deelnemers aan de Ander Leven interventie.

4. Kenmerken onderzoeksgroep

In veel gevallen konden administratieve gegevens over de deelnemers uit de interventiegroep aan het onderzoek worden toegevoegd, waardoor meer informatie beschikbaar kwam over de kenmerken van de interventiegroep. De belangrijkste staan hieronder weergegeven:

- De meeste deelnemers zijn vrouw (82%).
- Het merendeel van de deelnemers bevindt zich in de leeftijd van 25 tot 65 jaar, waarbij mensen in de leeftijd van 45 tot 55 het sterkst vertegenwoordigd zijn.
- Vrijwel alle deelnemers komen uit Nederland (77%) en Vlaanderen (23%).
- De begeleiders van de deelnemers tijdens de training zijn voornamelijk de partners (83%) ten opzichte van niet-partners (17%).



De type lichamelijke en/of geestelijke klachten van de deelnemers verschillen sterk van elkaar. Om een beter overzicht te maken zijn de specifieke klachten geaggregeerd naar een abstracter niveau (zie onderstaande tabellen). Geaggregeerde lichamelijke klachten die bij meer dan 30 deelnemers geconstateerd zijn, hebben betrekking op:

- het spijsverteringsstelsel (n = 36)
- het zenuwstelsel (n = 41)
- het circulatiestelsel (n = 30)
- SOLK (n = 101)
- een combinatie van diverse klachten (n = 125)

Statistics																		
	Ademhalingsstelsel	Bloedvatenstelsel	Botstelsel	Celstelsel	Gewrichtstelsel	Hormoonstelsel	Huidstelsel	Lymfevaten/immuunstelsel	Sensorisch stelsel	Spiersstelsel	Spijsverteringsstelsel	Urinewegstelsel	Voorplantingsstelsel	Zenuwstelsel	Motorisch stelsel	Circulatiestelsel	SOLK*	Diversen
N Valid	336	336	336	336	336	336	336	336	336	336	336	336	336	336	336	336	336	336
Missing	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14
Aanwezig	3	12	7	7	3	6	8	7	4	9	36	0	1	41	19	30	101	125
Afwezig	333	324	329	329	333	330	328	329	332	327	300	336	335	295	317	306	235	211

* Spasmodie, CVS, FM, ME, etc.

Geaggregeerde geestelijke klachten die bij meer dan 30 deelnemers geconstateerd zijn, hebben betrekking op:

- Angst- en paniekstoornissen (n = 50)
- Persoonlijkheidsstoornissen (n = 46)
- Slaapstoornissen (n = 37)
- Stemningsstoornissen (n = 36)
- Stress (n = 108)
- Vitaliteit (n = 74)

Statistics												
	Angst- en paniekstoornis	Emotioneel probleem	Gedragprobleem	Obsessief-compulsieve stoornis	Ontwikkelingsstoornis	Persoonlijkheidsstoornis	Slaapstoornis	Somatische stoornis	Stemmingsstoornis	Stress	Trauma stoornis	Verslavingsstoornis
N Valid	336	336	336	336	336	336	336	336	336	336	336	336
Missing	14	14	14	14	14	14	14	14	14	14	14	14
Aanwezig	50	2	0	1	0	46	37	0	36	108	7	1
Afwezig	286	334	336	335	336	290	299	336	300	228	329	335

Van de mensen in de steekproeven hebben in totaal 130 deelnemers lichamelijke klachten, 190 deelnemers geestelijke klachten en 125 deelnemers zowel lichamelijke als geestelijke klachten.

5. Uitkomsten op de voor- en nametingen

Voormeting

De gemiddelde scores op de SF/RAND-36 variabelen tijdens de voormeting, op een schaal van 1-100, zijn als volgt:

- Algemene gezondheid ($\bar{x}$ = 44)
- Lichamelijke gezondheid ($\bar{x}$ = 70)
- Rolbeperkingen n.a.v. lichamelijke problemen ($\bar{x}$ = 26)
- Lichaamspijn ($\bar{x}$ = 62)
- Geestelijke gezondheid ($\bar{x}$ = 53)
- Rolbeperkingen n.a.v. geestelijke (emotionele) problemen ($\bar{x}$ = 46)
- Vitaliteit/energie ($\bar{x}$ = 35)
- Sociaal functioneren ($\bar{x}$ = 44)
- Lichamelijk welzijn ($\bar{x}$ = 43)

- Geestelijk welzijn ($\bar{x} = 33$)

Opvallend genoeg scoort de lichamelijke gezondheid hoog (gemiddeld 70 punten) en rolbeperkingen naar aanleiding van lichamelijke problemen erg laag (gemiddeld 26 punten). Dit verschil is waarschijnlijk te verklaren door de type lichamelijke klachten van de deelnemers en het soort van vragen dat gesteld is om de lichamelijke gezondheid te meten. Veel deelnemers hebben chronische lichamelijke klachten (bijvoorbeeld een prikkelbare darm of continue vermoeidheid), maar kunnen over het algemeen nog goed traplopen, afstanden lopen en zware objecten tillen. De SF/RAND-36 vragenlijst gaat in op dergelijke aspecten om de lichamelijke gezondheid vast te stellen en niet op chronische lichamelijke en/of geestelijke klachten.

Nametingen

De gemiddelde scores op de SF/RAND-36 variabelen op de nametingen, genomen over de metingen van 1, 3, 6 en 12 maanden, zijn als volgt:

- Algemene gezondheid ($\bar{x} = 76$): stijging van 32 punten
- Lichamelijke gezondheid ($\bar{x} = 85$): stijging van 15 punten
- Rolbeperkingen n.a.v. lichamelijke problemen ($\bar{x} = 74$): stijging van 48 punten
- Lichaamspijn ($\bar{x} = 80$): stijging van 18 punten
- Geestelijke gezondheid ($\bar{x} = 77$): stijging van 24 punten
- Rolbeperkingen n.a.v. geestelijke (emotionele) problemen ($\bar{x} = 84$): stijging van 38 punten
- Vitaliteit/energie ($\bar{x} = 70$): stijging van 35 punten
- Sociaal functioneren ($\bar{x} = 82$): stijging van 38 punten
- Lichamelijk welzijn (agg.) ($\bar{x} = 52$): stijging van 9 punten
- Geestelijk welzijn (agg.) ($\bar{x} = 49$): stijging van 16 punten

SF/RAND-36	Voormeting (gemid.)	Nametingen (gemid.)	Vershil voor- en nametingen
Algemene gezondheid	44	76	32
Lichamelijke gezondheid	70	85	15
Rolbeperkingen n.a.v. lichamelijke problemen	26	74	48
Lichaamspijn	62	80	18
Geestelijke gezondheid	53	77	24
Rolbeperkingen n.a.v. geestelijke (emotionele) problemen	46	84	38
Vitaliteit/energie	35	70	35
Sociaal functioneren	44	82	38
Lichamelijk welzijn	43	52	9
Geestelijk welzijn	33	49	16

Uit de onderzoeksresultaten komt naar voren dat de gezondheidstoestand van de deelnemers na de Ander Leven interventie evident verbetert. De grootste verbetering is waarneembaar op Rolbeperkingen n.a.v. lichamelijke problemen met maar liefst 48 punten. De gezondheidstoestand op Algemene gezondheid, Rolbeperkingen n.a.v. geestelijke (emotionele) problemen, Vitaliteit/energie en Sociaal functioneren verbetert ook zeer sterk. De Lichamelijke gezondheid, Lichaamspijn en de Geestelijke gezondheid zijn ook verbeterd, maar doordat deze relatief wat hoger scoorden in de voormeting is het verschil in toename relatief wat lager. De verbetering op de geaggregeerde variabelen Lichamelijk welzijn en Geestelijk welzijn zijn relatief wat kleiner, maar dat komt door het rekenmodel voor geaggregeerde scores.

Metingen in perspectief

De gemiddelde scores op de voor- en nametingen krijgen meer betekenis door deze af te zetten tegen normscores van representatieve bevolkingsonderzoeken. In 1991 is een landelijk bevolkingsonderzoek in de Verenigde Staten uitgevoerd met de SF/RAND-36 vragenlijst op basis van

een steekproef (2474 respondenten) uit de totale populatie volwassen Amerikanen. Deze steekproef wordt als representatief beschouwd en kan als normering gebruikt worden om scores met elkaar te vergelijken. Om meer inzicht te krijgen in de lichamelijke en geestelijke gezondheidsproblematiek voor aanvang van de Ander Leven interventie, zijn de gemiddelde scores in de voormeting uitgesplitst naar lichamelijke en geestelijke klachten en vergeleken met normeringen van twee lichamelijke klachten (rugpijn en dermatitis) en een geestelijke klacht (klinische depressie) uit het Amerikaanse bevolkingsonderzoek (zie onderstaande tabel).

SF/RAND-36	Algemeen	Voormeting		Rugpijn	Normeringen (VS)		
		Lichamelijk	Geestelijk		Dermatitis	Clinische depressie	
Algemene gezondheid	44	42	48	58	61	53	
Lichamelijke gezondheid	70	73	76	66	75	72	
Rolbeperkingen n.a.v. lichamelijke problemen	26	21	36	47	57	44	
Lichaamspijn	62	59	67	59	67	59	
Geestelijke gezondheid	53	54	50	75	76	46	
Rolbeperkingen n.a.v. geestelijke problemen	46	52	38	71	79	39	
Vitaliteit/energie	35	36	38	52	57	40	
Sociaal functioneren	44	47	48	81	84	57	

Uit het overzicht komen overeenkomsten en verschillen naar voren, waarbij het lastig is om een duidelijk patroon op te maken. Opvallend is dat deelnemers die alleen lichamelijke klachten hebben op Algemene gezondheid en op Rolbeperkingen n.a.v. lichamelijke problemen evident lager scoren dan mensen met rugpijn of dermatitis. Ook valt het op dat deelnemers die alleen lichamelijke klachten hebben beduidend lager scoren op alle geestelijke variabelen ten opzichte van mensen met rugpijn en dermatitis. De scores van deelnemers met alleen geestelijke klachten trekken meer naar de scores toe van mensen met een klinische depressie.

Een andere normering zijn de geaggregeerde lichamelijke en geestelijke welzijnscores. Het Amerikaanse bevolkingsonderzoek gaf een gemiddeld lichamelijk en geestelijk welzijnsniveau van 50 punten aan, met een standaarddeviatie van 10 (gemiddeld wijken mensen 10 punten boven en 10 punten onder deze score af). Deze 50-puntenscore is in de praktijk als standaardisatiescore aangenomen waarmee lichamelijke en geestelijke welzijnscores uit andere SF/RAND-36 onderzoeken vergeleken kunnen worden. In de regel geldt hoe lager of hoger de scores van 50 afwijken, hoe slechter of beter het betreffende welzijnsniveau van de onderzochte populatie.

Verder geeft een Brits bevolkingsonderzoek uit 1999 (op een steekproef uit 1991) aan dat mensen met een langdurige chronische ziekte op lichamelijk welzijn gemiddeld 48 punten en op geestelijk welzijn gemiddeld 45 punten scoren (n = 3256). Aangezien deze groep overeen komt met de deelnemers aan de Ander Leven interventie, kunnen deze scores als vergelijkende indicatoren worden gebruikt.

De interventiegroep scoorde 43 punten op lichamelijk welzijn en 33 punten op geestelijk welzijn. Dit geeft aan dat deelnemers aan de Ander Leven interventie voor aanvang laag scoren op lichamelijk welzijn en met name laag scoren op hun geestelijk welzijn. Dit betekent dat de deelnemers in hun beleving serieuze lichamelijke en/of geestelijke klachten ervaren.

De gemiddelde scores op de nametingen tezamen kunnen vergeleken worden met normeringen uit bevolkingsonderzoeken onder volwassenen om meer inzicht te krijgen in hoeverre de lichamelijke en/of geestelijke gezondheidstoestand na toepassing van de Ander Leven interventie normaliseert. De scores zijn vergeleken met het eerder genoemde Amerikaanse bevolkingsonderzoek in 1991 (n = 2474) en het Brits bevolkingsonderzoek in 1999 (n = 8889), alsmede met een Canadees bevolkingsonderzoek in 2000 (n = 9423) en een Nederlands bevolkingsonderzoek in de gemeente Emmen in 1992 (n = 1063) (zie de tabel op de volgende pagina).

SF/RAND-36	Nameting	Normeringen			Gemeente Emmen
	Interventiegroep	Verenigde Staten	Verenigd Koninkrijk	Canada	
Algemene gezondheid	76	72	71	77	73
Lichamelijke gezondheid	85	84	88	86	87
Rolbeperkingen n.a.v. lichamelijke problemen	74	81	87	82	79
Lichaamspijn	80	75	79	76	80
Geestelijke gezondheid	77	75	72	78	77
Rolbeperkingen n.a.v. geestelijke (emotionele) problemen	84	81	86	84	84
Vitaliteit/energie	70	61	58	66	67
Sociaal functioneren	82	83	83	86	87
Lichamelijk welzijn	52	50	50*	51	n.b.
Geestelijk welzijn	49	50	50*	52	n.b.

* gemiddelde score tussen Britse mannen en vrouwen

Uit het overzicht komt naar voren dat de gemiddelde scores op de nametingen min of meer overeenkomen met de bevolkingsonderzoeken, met twee uitzonderingen daargelaten (waarbij een verschil van meer dan 10 punten is geconstateerd).

Een eerste uitzondering is dat deelnemers, na toepassing van de Ander Leven interventie, lager scoren op hun rolbeperkingen n.a.v. lichamelijke problemen dan de bevolkingsonderzoeken (m.n. ten opzichte van het Verenigd Koninkrijk). Dit kan mogelijk verklaard worden doordat de deelnemers die bij Ander Leven binnenkomen uitzonderlijk sterke klachten ervaren op hun rolbeperkingen n.a.v. lichamelijke problemen (deze scores zijn erg laag). Aan de andere kant kan het ook betekenen dat deelnemers, ondanks de Ander Leven interventie, nog steeds wat meer rolbeperkingen n.a.v. hun lichamelijke problemen ervaren dan een 'gemiddelde' westerling.

Een tweede uitzondering is dat deelnemers, na toepassing van de Ander Leven interventie, wat hoger scoren op hun vitaliteit/energie vergeleken met de normscores uit de bevolkingsonderzoeken (m.n. ten opzichte van het Verenigd Koninkrijk, alsook de Verenigde Staten: landen waar de werkdruk en stress over het algemeen hoog kunnen zijn). Een waarschijnlijke verklaring is dat veel deelnemers specifiek bij Ander Leven binnenkomen voor vitaliteits- en energiekachten, zoals moeheid en weinig puf hebben en dat de interventie ook specifiek voor deze doelgroep is ontworpen. In ieder geval voelen deelnemers zich na de Ander Leven interventie over het algemeen vitaler en energiever dan een 'gemiddeld' Amerikaans of Brits staatsburger (de verschillen met iemand uit Canada of Emmen zijn kleiner; plekken in de wereld waar de werkdruk en stress waarschijnlijk ook wat lager liggen).

De interventiegroep scoort op de geaggregeerde lichamelijke welzijnsscore 2 punten boven de standaardisatiescore van 50 (dus beter dan gemiddeld). De geestelijke welzijnsscore zit slechts 1 punt onder standaardisatiescore van 50, maar kent – zoals eerder aangegeven – een enorme stijging van 16 punten ten opzichte van de voormeting.

Concluderend kan, over het algemeen, gesteld worden dat de Ander Leven interventie deelnemers, die serieuze lichamelijke en/of geestelijke gezondheidsklachten ervaren, min of meer 'terugbrengt' naar een normaal gezondheids- en welzijnsniveau.

6. Effectiviteit van de Ander Leven interventie

Effectiviteitsbepaling

Bovenstaande uitkomsten indiceren dat de Ander Leven effectief is, maar om dat met meer zekerheid vast te stellen zijn specifieke statistische procedures nodig. Op basis van het onderzoeksdesign, de karakteristieken van de variabelen en bepaalde statistische voorwaarden (o.a. in hoeverre de SF/RAND-36 variabelen een normale verdeling kennen) komen twee statistische tests naar voren om de effectiviteit te bepalen: de Friedman test en de 'Repeated Measures' (RM) test. De

eerste is een zogenoemde non-parametrische test en de tweede een parametrische test. Een parametrische test is sterker dan een non-parametrische test, maar deze moet voldoen aan bepaalde statistische voorwaarden. Dit onderzoek voldoet, min of meer, aan deze voorwaarden. Maar omdat er ruimte is voor discussie, is ervoor gekozen om ook de non-parametrische test uit te voeren. Dit biedt niet alleen meer zekerheid, maar geeft ook de mogelijkheid om de tests met elkaar te vergelijken en meer over de effectiviteit van de interventie te zeggen. Indien de twee testen allebei hetzelfde uitwijzen, kan met meer stelligheid een uitspraak over de interventie worden gedaan.

De testen geven uitsluitsel of er een significant verschil is tussen de voormeting ten opzichte van de nametingen. Binnen de RM test kan ook getoetst worden of de onderlinge nametingen significant van elkaar verschillen. Bij de Friedman test is dat niet mogelijk. Daarom wordt de Wilcoxon test uitgevoerd om deze verschillen non-parametrisch met elkaar te vergelijken.

Een significant verschil wil zeggen dat het verschil – statistisch gezien – niet meer op toeval berust en dus dat het verschil aan ‘iets’ kan worden toegewezen. De veronderstelling is dat met name de driedaagse training (in combinatie met het consult) een verklaring geeft voor een significant verschil tussen de voormeting en de eerste nameting. Mogelijk kan er nog een klein verschil optreden tussen de eerste nameting en de twee nameting door de verdieping/terugkomdag, maar over het algemeen wordt verwacht dat de nametingen op hetzelfde niveau zullen zitten. Als de voormeting significant verschilt van de nametingen en de nametingen nauwelijks van elkaar verschillen, geeft dit ook aan dat het effect van de interventie c.q. de driedaagse training stand houdt in de tijd (na 1, 3, 6 en 12 maanden).

Een significant verschil wordt vastgesteld als de nulhypothese wordt verworpen. De nulhypothese geeft de verwachting aan dat de scores tussen de voor- en nametingen gelijk zijn (of weinig van elkaar verschillen). Een aanname of verwerping van de nulhypothese wordt bepaald door het daadwerkelijk gemeten verschil tussen de voor- en nametingen af te zetten tegen de verwachting dat er geen/weinig verschil is. Indien het gemeten resultaat cijfermatig gezien veel afwijkt van de verwachting dat er geen/weinig verschil is, kan de nulhypothese verworpen worden. Dit gebeurt aan de hand van statistische procedures en formules die binnen de Friedman test en RM test van toepassing zijn. Hierbij speelt kansverdeling een grote rol: de kans om, als de nulhypothese waar is (in werkelijkheid is er geen verschil), per toeval een verschillscore te meten die evident kleiner of groter is. Dan wordt de nulhypothese ten onrechte verworpen en wordt er een foutieve uitspraak gedaan.

De kans op dit toeval kan via een statistische procedure verkleind worden door de zogenoemde toevalwaarde (ook wel p-waarde genoemd) binnen de testen zo laag mogelijk te stellen. De laagste toevalwaarde binnen de Friedman test is 1%, dat wil zeggen dat in 1 van de 100 gevallen de kans bestaat dat een verschilwaarde wordt gemeten die toevallig veel afwijkt, terwijl in de andere 99 gevallen geen afwijking wordt gevonden. De laagste toevalwaarde binnen de RM test is 0,1%, hetgeen wil zeggen dat in 1 van de 1000 gevallen de kans bestaat dat een verschil wordt gevonden die toevallig veel afwijkt. Deze toevalwaarden zijn binnen dit onderzoek toegepast, waardoor de strengste statistische procedures van toepassing zijn. Dus indien een verschilwaarde tussen de voor- en nametingen wordt gevonden die significant afwijkt van de verwachte waarde dat er geen/weinig verschil is (nulhypothese), kan met zeer grote zekerheid de nulhypothese worden verworpen en het verschil aan de Ander Leven interventie worden toegewezen (de kans dat de meting per toeval afwijkt is uitermate klein).

Hoe groot of hoe klein het cijfermatige verschil tussen het daadwerkelijk gemeten verschil en het verwachte verschil moet zijn om een uitspraak te doen c.q. de nulhypothese aan te nemen of te

verwerpen, hangt af van de statistische formules die binnen de specifieke test gehanteerd worden en de steekproefomvang (de formules houden rekening met een kleine of een grote steekproefomvang: het maakt verschil uit of er 10 of 100 respondenten in de metingen zitten). De gemeten toevalwaarde bij de verschijscore tussen de voor- en nametingen moet in de Friedman test in ieder geval lager dan 1% zijn en in de RM test lager dan 0,1%.

Daarnaast is de gekeken naar de zogenaamde 'effect size', oftewel de sterkte van het effect als er een significant verschil tussen de voor- en nametingen wordt vastgesteld. Dit is ook berekend aan de hand van statistische procedures en formules.

Ten slotte is via de RM test gecontroleerd of verschillen tussen de voor- en nametingen gerelateerd zijn aan bepaalde factoren, zoals geslacht, leeftijd, nationaliteit, type begeleider en type klacht (de toevalswaarde is hier op 5% gezet). Daarmee kan worden aangegeven in hoeverre deze factoren van invloed zijn op het geconstateerde verschil tussen de voor- en nametingen. Het zou bijvoorbeeld kunnen zijn dat de interventie meer effect heeft op vrouwen dan op mannen of meer op jongeren dan op ouderen. Indien dat kan worden uitgesloten, kan met meer zekerheid worden aangegeven dat de Ander Leven interventie – bij een significant verschil tussen de voor- en nametingen – op "iedereen" van toepassing is en hetzelfde (blijvende) effect heeft.

Onderzoeksresultaten

De onderzoeksresultaten staan weergegeven in onderstaande tabel, waarbij de scores in de cellen de toevalwaarden (p-waarden) aangeven en de scores tussen haakjes de sterkte van het effect ('Effect size') als er een significant verschil is gevonden. Bij de non-parametrische testen moet de toevalwaarde kleiner zijn dan 1% (0.01) en bij de parametrische testen kleiner dan 0,1% (0.001) om een significant verschil aan te tonen. Bij een Wilcoxon test is het effect zwak bij scores rond de 10% (0.1), medium sterk bij scores rond de 30% (0.3) en groot bij scores rond de 50% (0.5) of hoger. Deze rangschikking geldt ook voor de 'within-subjects contrast' binnen de RM test. De sterkte van het effect bij de algemene RM test is echter anders: deze is zwak rond de 1% (0.01), medium sterk rond de 6% (0.06) en erg sterk bij 14% (0.14) of hoger.

	Non-parametrisch (significant verschil als $p < 0.01$)					Parametrisch (significant verschil als $p < 0.001$)				
	Friedman	Wilcoxon				RM	RM: within-subjects contrasts			
	t0-t4	t0-t1	t1-t2	t2-t3	t3-t4	t0-t4	t0-t1	t1-t2	t2-t3	t3-t4
Algemene gezondheid	0.000	0.000 (0.593)	0.516	0.616	0.235	0.000 (0.324)	0.000 (0.619)	0.549	0.779	0.392
Lichamelijke gezondheid	0.000	0.000 (0.528)	0.018	0.164	0.611	0.000 (0.283)	0.000 (0.434)	0.012	0.249	0.624
Rolbperkingen n.a.v. LG	0.000	0.000 (0.525)	0.068	0.781	0.410	0.000 (0.210)	0.000 (0.523)	0.056	0.867	0.361
Lichaamspijn	0.000	0.000 (0.450)	0.171	0.867	0.461	0.000 (0.139)	0.000 (0.358)	0.171	0.882	0.377
Geestelijke gezondheid	0.000	0.000 (0.573)	0.115	0.772	0.002	0.000 (0.279)	0.000 (0.558)	0.107	0.793	0.005
Rolbperkingen n.a.v. GG	0.000	0.000 (0.454)	0.071	0.848	0.057	0.000 (0.145)	0.000 (0.398)	0.083	0.802	0.057
Vitaliteit/energie	0.000	0.000 (0.607)	0.000 (0.147)	0.624	0.865	0.000 (0.353)	0.000 (0.693)	0.000 (0.113)	0.967	1.000
Sociaal functioneren	0.000	0.000 (0.582)	0.125	0.534	0.180	0.000 (0.310)	0.000 (0.602)	0.084	0.442	0.357
Lichamelijk welzijn (agg.)	0.000	0.000 (0.532)	0.115	0.452	0.388	0.000 (0.231)	0.000 (0.455)	0.351	0.572	0.228
Geestelijk welzijn (agg.)	0.000	0.000 (0.598)	0.033	0.987	0.013	0.000 (0.276)	0.000 (0.580)	0.003	0.853	0.014

In de cellen staan de toevalwaarden (p-waarden) weergegeven: bij een significant verschil staat de Effect size tussen haakjes vermeld

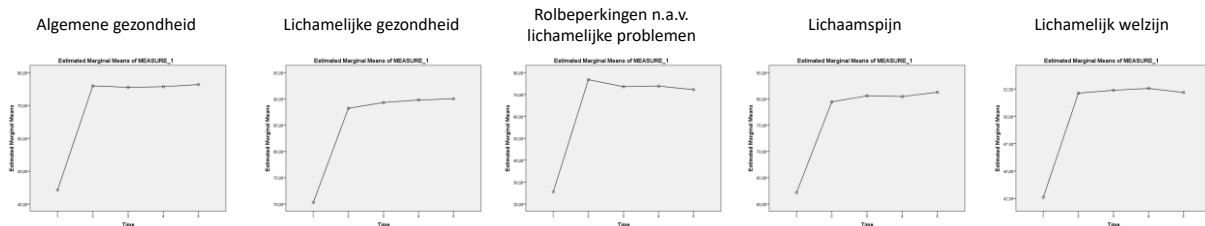
Zowel de niet-parametrische als de parametrische testen geven aan dat er een significant verschil is tussen de voormeting ten opzichte van de nametingen en tussen de voormeting en eerste nameting op alle SF/RAND-36 variabelen, alsmede dat de sterkte van het effect daarvan (zeer) groot is.

De nametingen op alle SF/RAND-36 variabelen verschillen niet significant van elkaar, behalve bij Vitaliteit/energie. Daar is een significant verschil gevonden tussen de eerste en tweede nameting. Dit impliceert dat de vitaliteit/energie na de training nog kan toenemen (bijvoorbeeld door thuis de AC-processen te lopen) en/of dat het geconstateerde verschil is toe te wijzen aan de verdieping/terugkomdag.

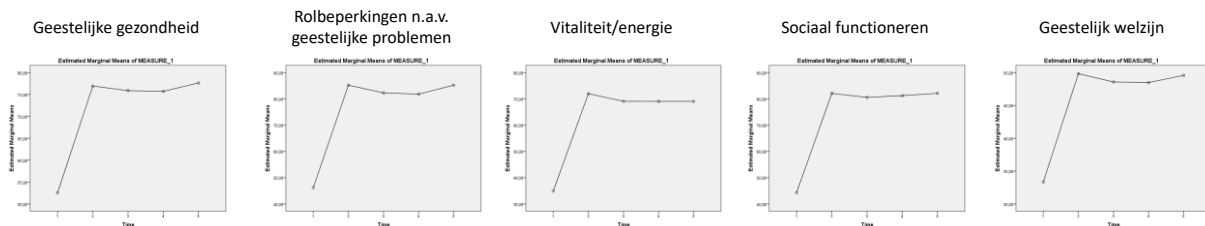
De uitkomsten geven aan dat, binnen de eerder genoemde onderzoekbeperkingen, met grote zekerheid kan worden aangenomen dat het verschil tussen de voor- en nametingen aan de Ander Leven interventie (m.n. de driedaagse training) kan worden toegewezen. Met andere woorden, de Ander Leven interventie heeft een sterk bewezen ('evidence based') effect opgeleverd.

In onderstaande grafieken wordt het verschil tussen de voor- en nametingen op de SF/RAND-36 variabelen visueel weergegeven. Er is een duidelijk verschil waarneembaar tussen de voormeting en de eerste nameting (de lijn gaat stijl omhoog) en er is relatief weinig verschil waarneembaar tussen de nametingen (de lijn blijft min of meer op hetzelfde niveau).

Lichamelijk

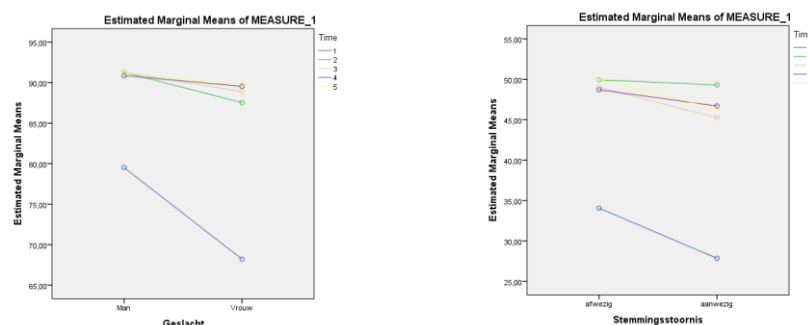


Geestelijk



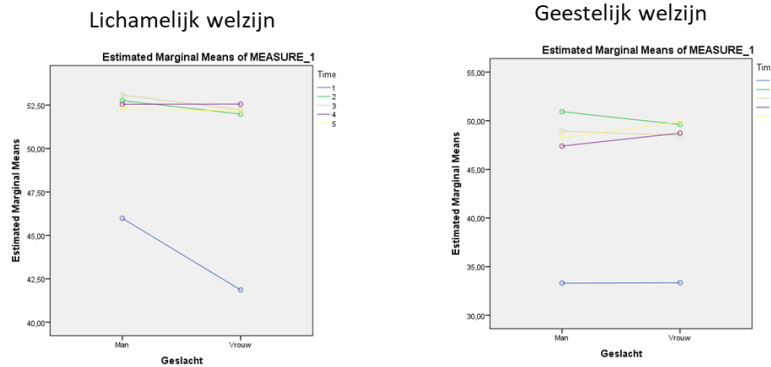
Verskil tussen factoren op de voor- en nametingen

Een significant verschil tussen de voormeting en eerste nameting is niet toe te wijzen aan een externe factor, zoals geslacht, leeftijd, nationaliteit, inkomenssituatie of type klacht. Er waren enkele uitzonderingen, maar het verschil tussen de factoren (bijvoorbeeld tussen man en vrouw of tussen mensen met of zonder stemmingsstoornis) is te verklaren doordat de scores op de voormeting veel van elkaar afwijken (de lichamelijke gezondheid van vrouwelijke deelnemers voor aanvang van de interventie scoort evident lager dan de lichamelijke gezondheid van mannelijke deelnemers). In onderstaande grafieken komen de verschillen op de voormeting c.q. baseline goed naar voren: het verschil tussen de voormeting (blauwe lijn) en de eerste nameting (groene lijn) is hoger door de baseline verschillen op de voormeting (vrouwen scoren op de voormeting bijvoorbeeld evident lager dan mannen).

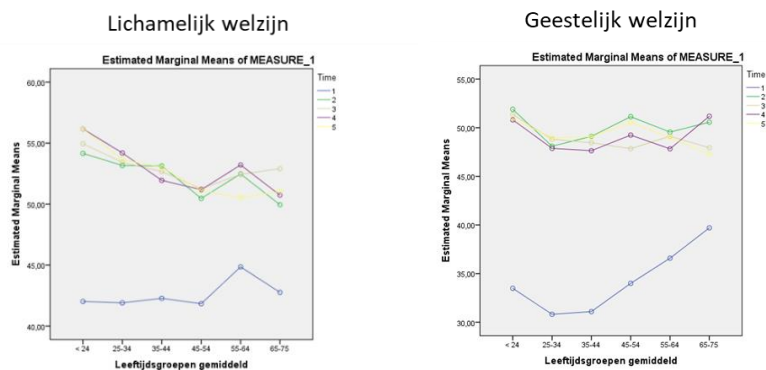


De verschillen op de voormetingen zijn in sommige gevallen opvallend. De belangrijkste staan hieronder weergegeven:

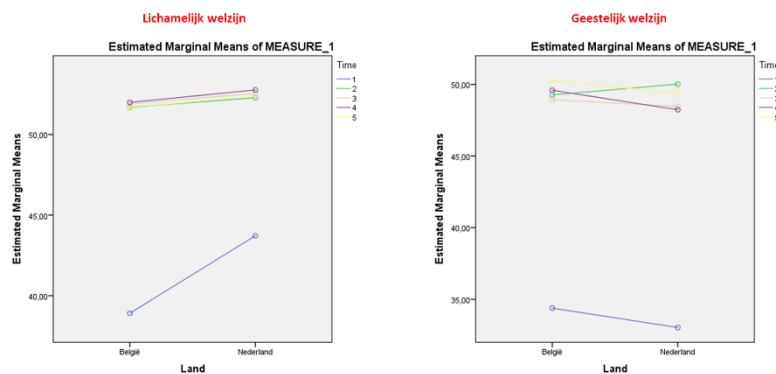
- vrouwen scoren voor aanvang van de training beduidend slechter op hun lichamelijk welzijn dan mannen (het geestelijk welzijn voor mannen en vrouwen is op de voormeting vrijwel gelijk).



- Mensen in de leeftijd van 55-75 jaar scoren voor aanvang van de training hoger op hun lichamelijk welzijn en met name hun geestelijk welzijn dan mensen die jonger zijn (18-54 jaar).



- Belgen scoren voor aanvang van de training wat slechter op hun lichamelijk welzijn dan Nederlanders (op geestelijk welzijn is dat vrijwel gelijk)

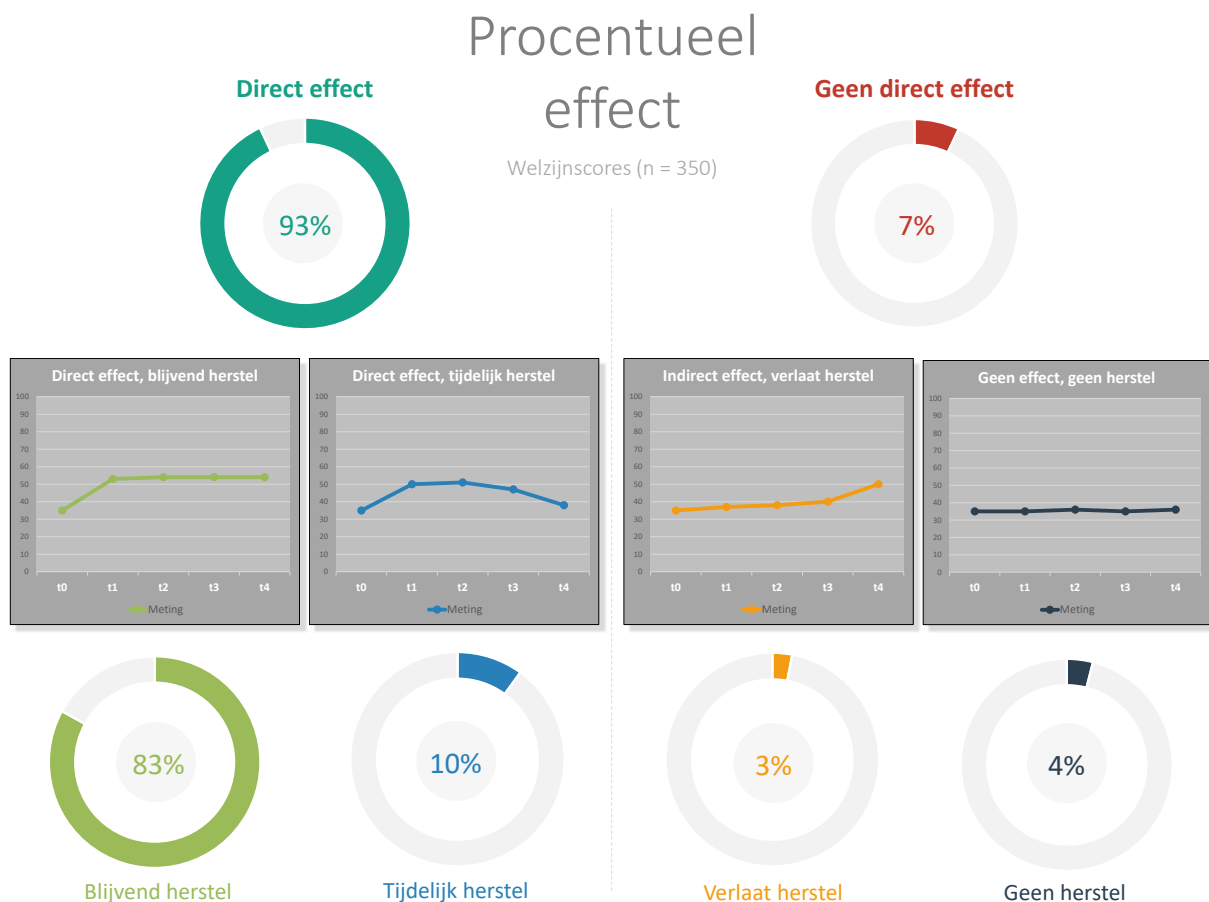


Naast externe factoren is ook gekeken naar karakteristieken van de Ander Leven Interventie zelf. De resultaten laten zien dat het niet uitmaakt welke hooftrainer voor de groep staat, hoe groot de groep is (al scoren kleine groepen wat lager dan grote groepen) en welk type begeleider de deelnemer

meeneemt (partner of geen-partner). De Ander Leven interventie heeft zich in de tijd doorontwikkeld. Hierbij kan worden aangegeven dat na toevoeging van een buddy kiezen, de dansavond, het begeleiden van partners in AC-processen en het vragenuurtje, de lichamelijke welzijnsscores wat hoger scoren en de nametingen vrijwel gelijk blijven.

7. Percentage deelnemers dat duurzaam herstelt

Uit de resultaten komt naar voren dat bij 93% van de deelnemers een direct trainingseffect optreedt. Dit effect is bij 83% van deze groep ook blijvend (duurzaam herstel). Bij 10% van deze groep treedt een tijdelijk herstel op en vallen na een jaar terug tot onder de kritische gezondheidsgrens. De training heeft op 7% van de deelnemers geen direct trainingseffect, maar bij 3% van deze groep treedt na een jaar alsnog een indirect effect en verlaat herstel op. Op 4% van de deelnemers heeft de training geen effect en vindt er ook geen herstel plaats. In onderstaande figuur worden de procentuele verschillen visueel weergegeven.



8. Conclusies

De uitkomsten van het onderzoek tonen aan dat de lichamelijke en geestelijke gezondheid van deelnemers na de Ander Leven interventie verbetert en dat het resultaat voor het overgrote deel van de deelnemers blijvend is.

In de periode 2011-2017 hebben in totaal 1307 mensen de Ander Leven de training gevolgd. Uit deze populatie is een onderzoeksgroep van 350 deelnemers naar voren gekomen die de vragenlijst op alle voor- en nametingen volledig heeft ingevuld. Dit vormt de steekproef om de effectiviteit van de interventie te evalueren. Statistische procedures wijzen uit dat de steekproefomvang groot genoeg is om de resultaten te generaliseren naar de totale populatie van deelnemers.

Op basis van de steekproef kan geconcludeerd worden dat de Ander Leven interventie zeer effectief is. De resultaten tonen aan dat de gemiddelde lichamelijke en geestelijke gezondheidstoestand van deelnemers voor aanvang van de Ander Leven interventie slecht scoort (deze liggen lager dan normeringen voor mensen met langdurige chronische lichamelijke en geestelijke klachten) en dat deze na toepassing van de interventie normaliseert naar een normale gezondheidstoestand (overeenkomstig met westerse samenlevingen). Verschillende statistische testen in dit onderzoek tonen allemaal aan dat het gemeten gezondheidsverschil tussen de voor- en nametingen niet op toeval berust en dat de sterkte van het effect zeer groot is. Dus de geconstateerde gezondheidsverbetering kan met grote zekerheid aan de Ander Leven interventie worden toegewezen. Ook houdt het effect stand in de tijd: de gemiddelde welzijnsscore van alle deelnemers in de steekproef blijft na 1, 3, 6 en 12 maanden op vrijwel hetzelfde genormaliseerde niveau.

Het geconstateerde effect van de interventie is voornamelijk van toepassing op de driedaagse training. De verdiepingsdag heeft over het algemeen weinig tot geen effect meer op de gezondheidstoestand van deelnemers, behalve voor de vitaliteit/energie van deelnemers. Deze verbetert nog significant na de verdiepingsdag. Ook is het opvallend dat het gemiddeld vitaliteitsniveau van deelnemers na de training hoger ligt dan het landelijk gemiddelde van een Westerse samenleving.

Het effect van de Ander Leven interventie hangt niet af van geslacht, leeftijd, nationaliteit, of inkomenssituatie. Het onderzoek geeft ook een indicatie af dat de gezondheidstoestand van deelnemers, onafhankelijk van hun type klacht, significant verbetert na toepassing van de Ander Leven interventie (vervolgonderzoek met grotere steekproeven moet hier meer uitsluitsel over geven).

Procentueel gezien ervaart 93% van de deelnemers een direct trainingseffect: een verbetering in hun lichamelijk en/of geestelijk welzijn. Bij 83% van deze groep is het effect blijvend en treedt duurzaam herstel op. Bij 10% van deze groep is het effect tijdelijk van aard. Van de 7% die geen direct trainingseffect ervaart, herstelt uiteindelijk 3% na een jaar alsnog na een gezond welzijnsniveau. Bij 4% van de deelnemers treedt geen effect en geen herstel op.

De onderzoeksresultaten in dit onderzoek wijzen uit dat de Ander Leven interventie effectief is, maar dit onderzoek kent ook een aantal beperkingen. Ten eerste zijn de deelnemers op basis van bepaalde criteria geselecteerd om deel te nemen aan de interventie. Er heeft geen aselechte selectie plaats gevonden: niet elke aanmelding heeft evenveel kans op de training te volgen. Daarom geldt de uitspraak alleen voor mensen met lichamelijke en geestelijke klachten die aan deze criteria voldoen, bijvoorbeeld voldoende gemotiveerd om de training te gaan volgen.

Ten tweede zijn deelnemers niet gerandomiseerd (op basis van toeval) aan de onderzoeksgroep toegewezen. In de steekproef zitten dus mogelijk deelnemers die op basis van een bepaalde motivatie de vragenlijst op alle tijdstippen hebben ingevuld (bijvoorbeeld deelnemers die meer tevreden of meer tijd hebben om de vragenlijst in te vullen).

Ten derde zijn deelnemers niet bevraagd op de mate waarin zij de aangeleerde technieken zelfstandig toepassen. Nu is bijvoorbeeld niet te controleren of deelnemers waarop de training geen effect heeft of waar geen blijvend herstel optreedt, te maken heeft met het al dan niet uitvoeren van de aangeleerde technieken. Of dat deelnemers die de technieken meer toepassen ook hoger scoren op hun welzijnsniveau.

Tot slot is de onderzoeksgroep niet vergeleken met een controlegroep (mensen met overeenkomstige klachten die de Ander Leven interventie niet ondergaan), waardoor minder uitsluitsel over het effect kan worden gegeven.

Bijlage: Bedrijfsprofiel Evidento

Evidento is een bureau dat gespecialiseerd is in het opzetten en uitvoeren van wetenschappelijk effectonderzoek en het ontwikkelen van evaluatiesystemen voor organisaties. Het bureau richt zich met name op organisaties in zorg en welzijn.

Organisaties voeren diverse processen, interventies en activiteiten uit, maar weten doorgaans niet wat deze uiteindelijk opleveren en in hoeverre de gewenste resultaten geboekt worden. Evidento helpt om inzicht te geven in de effectiviteit van processen, interventies en activiteiten (bijvoorbeeld in hoeverre cliënten naar aanleiding van een interventie vooruitgang boeken). Dit gebeurt aan de hand van wetenschappelijk effectonderzoek. Evidento verzamelt data, voert geavanceerde statistische analyses uit via SPSS en levert een onderzoeksverslag op. Met het verslag kunnen organisaties op een valide, betrouwbare en representatieve manier aantonen in hoeverre hun processen, interventies en activiteiten de gewenste effecten opleveren ('evidence based' zijn).

Door technologische ontwikkelingen wordt het voor organisaties mogelijk om op een gebruiksvriendelijke, kostenbesparende en doeltreffende manier *big data* te verzamelen en te analyseren. Evidento begeleidt dit proces en helpt organisaties een evaluatiesysteem te ontwikkelen via een Business Intelligence (BI) systeem. In dit systeem wordt het verzamelen en analyseren van data geautomatiseerd en worden de analyseresultaten weergegeven in een online dashboard waarop gebruikers kunnen inloggen. Het dashboard geeft gerichte stuurinformatie over wat precies te optimaliseren om tot betere resultaten te komen. Daarmee kunnen organisaties hun kwaliteit en succes op een systematische manier vergroten.

Evidento is een bureau dat sinds 1 juni 2016 staat ingeschreven bij de Kamer van Koophandel in Nederland onder de vlag van Maradata en is gevestigd aan de Keizersgracht 241 in Amsterdam. Tim Choy (1973) is de oprichter en eigenaar van Evidento. Daarnaast is hij verbonden aan Familiezorg Nederland, een platform om familiezorg ontwikkeling, onderzoek en onderwijs te stimuleren. In het verleden heeft hij Beleid, Communicatie en Organisatie (BCO) aan de Vrije Universiteit gestudeerd en promotieonderzoek aan de Universiteit van Tilburg verricht (onderzoek naar de effecten van evaluatiesystemen op de structuur en processen binnen organisaties en het menselijk gedrag). Hij heeft gewerkt voor verschillende universiteiten (Vrije Universiteit Amsterdam, Universiteit van Tilburg, Radboud Universiteit Nijmegen) en onderzoeksbureaus (TNS/NIPO, fluent Zorgadvies, Prismant).

Tim Choy is sociale wetenschapper en heeft zich verdiept in bepaalde onderwerpen binnen de organisatiekunde, sociologie, psychologie en statistiek. In de beginperiode van zijn carrière heeft hij zich voornamelijk gericht op welke evaluatiesystemen organisaties gebruiken, hoe ze werken en wat het oplevert. Gaandeweg is hij zich steeds meer gaan bezighouden met het zelf opzetten en uitvoeren van evaluatieonderzoek. Hij heeft zich met name gespecialiseerd in het evalueren van zorginterventies op effectiviteit. De afgelopen vijftien jaar heeft hij effectonderzoek verricht bij verschillende soorten zorgaanbieders in de Care sector (instellingen voor Geestelijke gezondheidszorg, Gehandicaptenzorg, Ouderenzorg en Jeugdzorg). In het bijzonder heeft hij veel effectonderzoek gedaan naar de Methode Familiezorg en welke invloed familiecommunicatie heeft op het stressniveau en het welzijn van mantelzorgers. Hij is bekwaam in het opstellen van onderzoeksdesigns voor effectonderzoek en het uitvoeren van geavanceerde statistische analyses in SPSS. Ook heeft hij aan de Radboud Universiteit Nijmegen een certificaat Geavanceerde Statistiek voor promovendi behaald.

Voor meer informatie neem contact op via: tim@evidento.nl of bel 06-13839493.