Work Package 4

Development of an analytical framework to perform a comprehensive evaluation of integrated care programmes for multi-morbidity using Multi-Criteria Decision Analysis

Executive Summary

WP leader
Institute of Health Policy & Management, Erasmus University Rotterdam, The Netherlands
WP co-leader
University of Bergen (UiB), Norway
Authors
Fenna Leijten, Melinde Boland, Apostolos Tsiachristas, Maaike Hoedemakers, Nick Verbeek, Kamrul Islam, Jan Erik Askildsen, Antoinette de Bont, Roland Bal, Maureen Rutten-van Mölken
Date
January 31, 2017

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 634288. The content of this report reflects only the SELFIE groups' views and the European Commission is not liable for any use that may be made of the information contained herein.

1. Preface

This report constitutes the executive summary of the deliverable of Work Package (WP) 4 of the SELFIE project. In this WP we developed the methods to perform comprehensive evaluations of integrated care programmes for multi-morbidity using Multi-Criteria Decision Analysis (MCDA). The WP leader is the Institute of Health Policy & Management from Erasmus University Rotterdam in the Netherlands and the WP co-leader is the University of Bergen (UiB) in Norway.

The report is structured as follows. First, an overview of the SELFIE project is provided and an introduction to WP4 is given. Next, a summary of the two WP4 deliverable reports is given. In the first deliverable report we give an overview of the outcomes that will be included in the MCDA and we describe the methods used to identify and select them. The second deliverable report starts with a description of MCDA methods in general, followed by the MCDA methods that will be applied in SELFIE. At the end of this report we describe how this work feeds into the next WP in which the comprehensive evaluations will actually be carried out.

2. Introduction to SELFIE and WP4

The EU Horizon2020-funded SELFIE ('Sustainable Integrated Care Models for Multi-Morbidity: Delivery, Financing and Performance') project aims to improve person-centred care for persons with multi-morbidity by providing evidence on the impact of promising integrated chronic care (ICC) programmes and supporting financing/payment schemes on health- and wellbeing, experience with care, and cost outcomes (i.e., the Triple Aim). This four year research project is divided into nine work packages (WP) conducted by eight European partners: The Netherlands (coordinator), Austria, Croatia, Germany, Hungary, Norway, Spain, and the UK. SELFIE distinguishes itself from other research projects on integrated care and/or multi-morbidity by aiming to not only identify and describe promising integrated care programmes for multi-morbidity, but to evaluate them using an innovative approach: Multi-Criteria Decision Analyses (MCDA). MCDA is an umbrella term for a set of methods that aid decision-making when this is based on more than one criterion, whereby the relative impact that all criteria have on the decision-making process is made explicit. MCDA thus aims to improve transparency, accountability, and acceptability of the decision-making process by explicitly defining aims, or criteria.

In WP1 we first developed a conceptual framework to guide the further descriptive and evaluative work on integrated care for multi-morbidity in SELFIE. This conceptual framework includes six key components centred around the holistic understanding of an individual with multi-morbidity: (1) Service delivery (2) Leadership and governance, (3) Workforce, (4) Financing, (5) Technologies and medical products (6) Information and research. Subsequently, in WP1 and WP2, 17 promising integrated care programmes for persons with multi-morbidity were identified in the countries of the SELFIE consortium (2-3 per county) and extensively described, using both document analyses and interviews. This resulted in 17 'thick descriptions' that are being compared across countries (see also the executive summaries of WP1 and WP2). Preparations for the empirical evaluations using an MCDA framework were made in WP4.

Integrated care programmes are complex interventions consisting of a mixed package of interacting patient-, provider-, and organisational-oriented interventions that are tailored to

the context in which they are introduced and continuously improved as more experience is gained. They do not only aim to maximise health but also to improve wellbeing, experience of care and reduce costs. Therefore, traditional evaluation frameworks such as cost-utility analysis in which costs per Quality Adjusted Life Years (QALY) are calculated, were deemed insufficient. In contrast, using MCDA has the purpose of ensuring that these evaluations are broad enough to incorporate the wide range of different outcomes of these programmes.

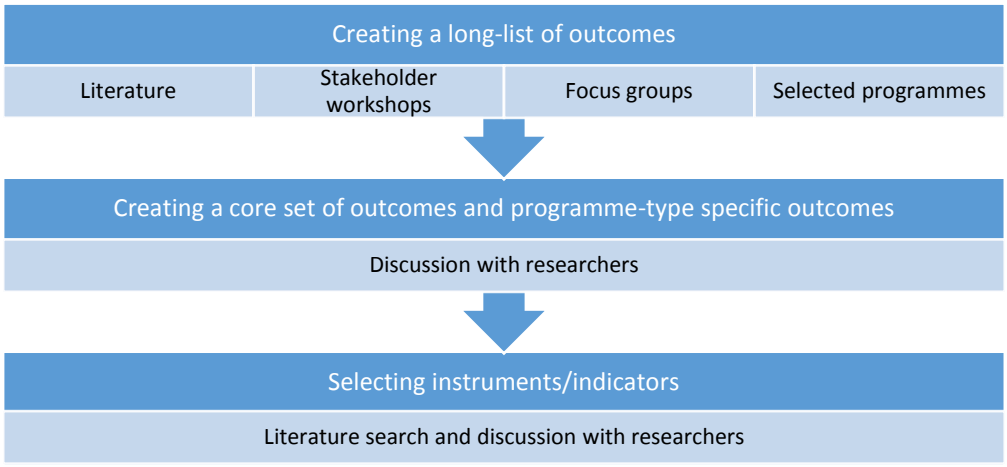
The work done in WP4 is divided into two deliverables (i.e., two reports). In the first report we defined a set of outcomes for which we will **measure performance** of the promising integrated care programmes and for which we will **elicit weights**. In the second report, we created an MCDA evaluation framework by selecting an MCDA method, constructing study designs to measure the performance of the 17 promising programmes, and preparing for the weight-elicitation study whereby the **importance** of the various outcomes will be determined. With this MCDA framework we will, in WP5, conduct 17 comprehensive evaluations in which integrated care programmes are compared to usual care or a control group.

The results of the two deliverables will be summarised below.

3. WP4 Deliverable Report 1: Selecting and defining outcomes for the evaluation

In the first deliverable report of WP4, entitled "*Outcomes and indicators in integrated care for persons with multi-morbidity",* we selected and defined a set of outcomes that are specifically relevant for the empirical evaluation of integrated care programmes for persons with multi-morbidity. Secondly, we selected instruments or indicators to measure these outcomes. A great number of outcomes and instruments exist to measure integrated care, and we witnessed a tremendous growth in new types of measures such as patient reported outcomes (PROMs) and patient reported experience measures (PREMs). Although multiple criteria, or outcomes, can be incorporated in an MCDA, feasibility in collecting performance information on these and the need to avoid cognitive overload for respondents in the weight-elicitation study forced us to select the most relevant ones. **Figure I** provides an overview of the steps in this selection process.

Figure I: *Steps to develop the list of outcomes measures and indicators included in the MCDA: a core set and a programme-type specific set*



We used four main sources to create an initial 'long-list' of outcome measures: (1) literature review, (2) stakeholder workshops, (3) focus groups in individuals with multi-morbidity, and (3) a review of outcomes currently used in the 17 programmes selected for evaluation in SELFIE.

Key (grey) **literature** was used to explore innovative definitions and measures of health and integrated care. We focused in particular on outcomes to measure integrated care according to the Triple Aim (health- and well-being, experience of care, and costs). In scientific literature, we identified outcomes that were measured in integrated care programmes specifically targeting individuals with multi-morbidity.

Alongside the literature review, each country in the SELFIE consortium organised a **workshop with national stakeholders** that represented five stakeholder groups, the 5Ps: Patients, Partners (i.e., informal caregivers), Professionals, Payers, and Policy makers. During these workshops the stakeholders were asked to name and define what would make them reimburse, participate in, offer, or implement an integrated care programme for multi-morbidity. The outcomes that resulted from these workshops were added to those found in the literature.

In order to zoom into and learn more about the importance of person-centred integrated care for individuals with multi-morbidity, each country in the SELFIE consortium organised a **focus group**. In total 58 individuals with multi-morbidity attended these focus groups, in which they discussed what it means to them to be in 'good health' and how they define a good care process. Outcomes mentioned in the focus groups were again added to the list of candidate outcomes.

Lastly, we reviewed the 17 **promising integrated care programmes** included as case studies in SELFIE as a source of information on possible outcomes: what are the goals of these programmes, what are they already measuring, and what do they find important to measure?

An abundance of outcomes was obtained from these four sources, many with large conceptual overlap. We clustered the outcomes into higher-level concepts and categorised them into the Triple Aim. There was considerable agreement between the outcomes mentioned by the various stakeholders. There was general consensus that we should focus on patient-reported outcome measures and patient-reported experience measured to extend the frequently used structural indicators or indicators of the adherence to

programme-components that are extracted from routine organisational and system-level databases. Furthermore, an overarching theme was that evaluations of integrated care for multi-morbidity should go beyond traditional clinical health outcomes, and should focus more broadly on well-being. Further, it was felt that in the selection process the outcomes that were frequently mentioned by the persons with multi-morbidity in the focus groups should be leading. The group of persons with multi-morbidity is in and of itself complex and varied; they deal with many different health- and social problems that may interact with one-another. For these reason, in SELFIE, we incorporate **'social relationships and participation'**, **'enjoyment of life'** and **'resilience'** as health- and well-being outcomes alongside the more traditional outcomes of **'physical functioning'** and **'psychological wellbeing'** as these apply to persons with all different types of disease- and problem-combinations. Persons with multi-morbidity often deal with care providers from different sectors, with a high risk of fragmentation and repetition. Thus we identified **'person-centeredness'** and **'continuity of care'** as the two key elements of experience of care that should be included in evaluations. Moreover, we also include **'total health- and social care costs'** in SELFIE, to capture the care utilisation in different sectors.

The above mentioned outcomes were also selected because they were found relevant and applicable across all 17 integrated care programmes. The discussion of these outcomes amongst the SELFIE researchers revealed the need to add important programme-specific outcomes to the MCDA. This resulted in the decision to construct a) a core set of outcomes to be included in all 17 programme evaluations, and b) programme-type specific outcomes. Defining outcomes that are relevant across multiple programmes is important because one of the SELFIE aims is to develop a reusable MCDA, where criteria-weights can be used again by others who want to monitor different integrated care programmes, facilitated by an online tool.

To keep the weight elicitation study feasible, the 17 integrated care programmes were categorised into four types of programmes: population health management programmes, programmes targeting frail elderly, palliative care and oncological programmes, and programmes targeting persons with problems in multiple life domains. For each type of programme a list of outcomes was defined. The core set of outcomes includes the eight

mentioned above; they are presented and defined in **Table I.** An overview of the core set and the programme-type specific outcomes are presented in **Table II**. Both the core set of outcomes and the programme-type specific outcomes will be included in the MCDA. This implies that we will elicit weights for both of them.

The outcomes in **Tables I** and **II** were defined at a conceptual level in order to allow the use of different instruments or indicators to measure a particular outcome-concept. The reason that we permit the use of different instruments and indicators to measure a particular concept is that some programmes have already been measuring certain outcomes for years, and this retrospective data is of great value. In the cases when data collection still needs to be set up, we make recommendations for instruments or indicators that best operationalise the outcomes in SELFIE.

Table I: *Core set of outcomes*

Health / well-being
Physical functioning *"Acceptable physical health and being able to do daily activities without needing assistance"* **Psychological well-being** *"Absence of stress, worrying, listlessness, anxiety, and feeling down"* **Social relationships & participation** *"Having meaningful connections with others as desired"* **Enjoyment of life** *"Having pleasure and happiness in life"* **Resilience** *"The ability to recover from or adjust to difficulties and to restore ones equilibrium"*
Experience with care
Person-centeredness *"Care that care matches an individual's needs, capabilities and preferences and jointly making informed decisions"* **Continuity of care** *"Good collaboration, smooth transitions between caregivers, and no waste of time"*
Costs
Total health and social care costs *"Total health and social care costs per participant"*

Table II: *Overview of core set and programme-type specific outcomes in SELFIE*

Outcomes for integrated care for persons with multi-morbidity					
	Core set outcomes	**Programme-type specific outcomes**			
		Population health management	*Frail elderly*	*Palliative and oncology*	*Problems in multiple life domains*
Health & well-being	Physical functioning	Activation & engagement	Autonomy	Mortality	Self-sufficiency
	Psychological well-being			Pain and other symptoms	
	Social participation/ relationships				
	Resilience				
	Enjoyment of life				
Experience	Person-centeredness		Burden of medication	Compassionate care	
	Continuity of care		Burden of informal caregiving	Timely access to care	
				Preferred place of death	
				Burden of informal caregiving	
Costs	Total health- and social care costs	Ambulatory care sensitive hospital admissions	Living at home		Justice costs
		Hospital re-admissions	Falls leading to hospital admissions		

4. WP4 Deliverable Report 2: Creating an MCDA evaluation framework

In the WP4 Deliverable Report 2, entitled *"MCDA framework"*, the reason why MCDA was chosen as the evaluative framework in SELFIE is explained. As described above, this broad evaluation framework allows for multiple outcomes (in MCDA-terms 'criteria') to be included in the evaluation, and weighs these from specific perspectives. There are different MCDA methods, which can broadly be divided into the value-based, outranking, and goal- or reference point methods. Each is briefly described below.

In **value-based methods**, the aim is to assign values to alternatives and construct a value function. In SELFIE the alternatives are the integrated care programme and its comparator. In the commonly used value-based method Multi-Attribute Utility Theory (MAUT), a single overall value is created. The performance of each alternative, on all criteria, is determined. Separately, the importance of the criteria needs to be determined. For each alternative, the weighted performance on each criterion is aggregated into an overall value score. This overall value score is compared between the integrated care programme and its comparator.

In **outranking methods**, pairwise comparisons are made of the performance of all alternatives on all criteria. In the simplest case, if we compare the performance of alternatives on all criteria and one alternative scores better on all, then this is the preferred alternative. In less simple cases, patterns of dominance between alternatives are studied to reach a decision about the preferred alternative. For this method the performance of alternatives on criteria needs to be known, as well as the weights for these criteria.

In **goal** or **reference point methods**, alternative care programmes are compared by calculating the weighted deviations from a priori set goals. This method requires a specification of desirable levels of performance for each criterion.

Given that in SELFIE we aim to compare each integrated care programme to its comparator and not to rank all 17 programmes, and considering the theoretical foundations of all

methods, we have opted for MAUT methods to be applied in the MCDA. The seven steps commonly undertaken in MCDAs, and MAUT specifically, are as follows:

1. Establish the decision-context
2. Identify and structure criteria
3. Determine performance on criteria
4. Weight-elicitation
5. Creating a global score
6. Sensitivity analyses
7. Examine results

The work done in earlier SELFIE WPs has helped us understand and establish the **decision-context** of integrated care for multi-morbidity (step 1). For most of the 17 case studies in SELFIE, the decision pertains to reimbursement, continuation, extension, and/or wider implementation of the integrated care programme. Step 2 was described in **WP4 Deliverable Report 1** "*Outcomes and indicators in integrated care for persons with multi-morbidity".*

To determine the performance on the criteria, step 3, empirical evaluations for all 17 promising programmes are being set up. In order to be able to attribute effects to the intervention, performance will be repeatedly assessed in both the integrated care group as well as a comparator group. Each SELFIE partner selected the most appropriate study design for their evaluation and started working on a study protocol to make this explicit.

There are different methods to elicit weights (step 4), including: direct ranking, visual analogue scales, point allocation, analytical hierarchy process, swing weighting, and **discrete choice experiments** (DCE). In a DCE, choice sets with scenarios are presented to stakeholders. The scenarios consist of various alternatives (e.g., care programmes) that systematically differ on performance criteria (i.e., outcomes, attributes). Stakeholders are asked which scenario they prefer. Hereafter, weights for each criterion can be statistically derived on the basis of the likelihood that one scenario, with specific criteria performance, is preferred over another. The stakeholders are forced to make trade-offs in criteria and take the full range of potential performance into account. DCE was selected as the method to

elicit weights, in all eight SELFIE partner countries, from the 5P perspectives. Due to the number of different outcomes/criteria, it is not possible to conduct a DCE for the core set and the programme-type specific outcomes. For this reason, a DCE will only be used to elicit weights for the core set of outcomes.

The perceived 'next-best' method, **swing weighting**, will be used to elicit weights for the programme-type specific outcomes. Swing weighting is also a trade-off weighting method, in which the relative importance is determined on the basis of moving from the worst to best score on a scale. Specifically, the SMARTER method will be used, whereby a stakeholder is asked to pretend there is an alternative (e.g., care programme) that has the lowest possible scores on all criteria. The stakeholder then ranks which criteria would be selected first to swing from the worst to the best level. This is subsequently done for the remaining criteria. These ranks are then turned into weights using, for example, the rank ordered centroid method.

Table III illustrates the type of information that will be obtained in the MCDA. This includes the (standardised) performance scores of two alternative care programmes (e.g., integrated vs. usual) on a number of criteria, the weights of these criteria from the viewpoint of different stakeholder groups (S1 and S2), and the weighted aggregation. In the simplified example, in **Table III** the aggregated weight for resilience is calculated by multiplying the criteria weight of stakeholder 1 (0.30) or stakeholder 2 (0.15) with the standardised performance (0.74 for the integrated care programme and 0.67 for the usual care). When these weighted performance scores are summed across all criteria the overall value of a programme is obtained (step 5 of the MCDA).

Table III: *SELFIE example of information needed for an MCDA of integrated care programmes for multi-morbidity*

				Care alternatives		Weighted aggregation			
		Weight		*Integrated*	*Usual*	*Integrated*		*Usual*	
Triple aim	Criteria	S1	S2	Standardised performance*		S1	S2	S1	S2
Improved health	Resilience	.30	.20	0.74	0.67	0.22	0.15	0.20	0.13
	Physical functioning	.20	.15	0.68	0.73	0.14	0.10	0.15	0.11
Improved experience	Person-centeredness	.15	.05	0.80	0.60	0.12	0.04	0.09	0.03
	Continuity of care	.25	.05	0.77	0.63	0.19	0.04	0.16	0.03
Reduced costs	Health care costs	.05	.30	0.28	0.30	0.01	0.08	0.02	0.09
	Social care costs	.05	.25	0.24	0.34	0.01	0.06	0.02	0.09
					Sum	**0.69**	**0.47**	**0.64**	**0.48**

*Note: Aggregation on the basis of Multi-Attribute Utility Theory (MAUT). *Standardised performance based on relative standardisation. S1 = Stakeholder 1 (e.g., patient), S2 = Stakeholder 2 (e.g., payer).*

In step 6, sensitivity analyses will be done. This will include subgroup analyses, such as per gender, educational level or types of morbidities. Furthermore, we will conduct deterministic analyses, whereby certain criteria are excluded, as well as probabilistic analyses, in which uncertainty in weights and performance is modelled simultaneously. In step 7, the results will be examined. This will be done by the SELFIE researchers, but will also involve reflecting and interpreting the findings with representatives from the 5Ps in international and national stakeholder workshops.

In the **WP4 Deliverable Report 2** we extensively describe the background of MCDA methods, the seven steps undertaken in MAUT, MCDA and weight-elicitation choices in SELFIE, possible study designs, the weight-elicitation procedure, and the draft study designs to measure the performance of the 17 programmes.

5. Next steps

In the next phase of SELFIE research, each SELFIE partner will start data collection from the five defined stakeholders (5Ps) for the weight-elicitation. Parallel to this, the study designs for the empirical evaluations will be made definite and data collection in all 17 promising integrated care programmes will begin. Subsequently, the performances of the promising programmes on the (core) set of outcomes and the weights from the various stakeholders will be brought together in the proposed SELFIE framework, resulting in MCDAs of 17 promising integrated programmes for persons with multi-morbidity

In order to allow findings from the SELFIE study to be shared with others, an online MCDA tool will be developed that will allow others to also apply the criteria weights from the 5Ps to their own programme performance. The tool will stay available after the SELFIE research project has ended.