Learning from integrated eCare practice and promoting deployment in European regions

BeyondSilos

D6.1 EVALUATION FRAMEWORK

WP6 Pilot evaluation

Version 1.0, 13th August 2014

The BeyondSilos project is co-funded by the European Commission within the ICT Policy Support Programme of the Competitiveness and Innovation Framework Programme (CIP).
Grant Agreement No.: 621069

cip competitiveness and innovation framework programme 2007–2013

Document information

Abstract

This document contains the scientific protocol for the BeyondSilos project.

Organisation responsible

Region of Southern Denmark (RSD)

Authors

Daugbjerg, Signe (RSD)
Da Col, Paolo (HIM)

Contributing partners

Ciccarelli, Michele (Campania)
Faubel, Raquel (LAFE)
Fernandes, Adriano (Amadora)
Georgieva, Lidia (CPRH)
Goncalves, Rute (Amadora)
Hobson, Penny (HSCNI)
Hurtado, Mayte (HIM)
Iaccarino, Guido (Campania)
Lewis, Leo (HIM)
Melle, Christian (Kinzigtal)
Meyer, Ingo (Emperica)
Müller, Sonja (Emperica)
Piera, Jordi (BSA)
Oates, John (HIM)
Roth, Monica (Kinzigtal)

Delivery date

15th August 2014

Dissemination level

P	Public

Version history

Version	Date	Changes made	By
0.1	3rd July 2014		
0.2	25th July 2014		SD
0.3	27th July 2014		PS
0.4	11th August 2014		SD
1.0	13th August 2014	Version for issue	JO

Outstanding issues

Parts of the evaluation framework will need further updating as data collection is finalised.

Filename

D6.1 v1.0 BS Evaluation framework

Statement of originality

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Executive summary

Deliverable D6.1 is the scientific protocol for the BeyondSilos project. The protocol presents descriptions of the relevant information for carrying out an evaluation of ICT supported integrated health and social care.

The protocol is based on the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement [1]. It presents the background of the evaluation, objectives, methodologies used for selection of participants, data collection, data management, statistics, monitoring and ethics. The protocol describes the evaluations of the new pilot sites organisational models along with the overall evaluation of BeyondSilos project.

Table of Contents

1 Introduction

1.1 Purpose of this document

WP6 for BeyondSilos includes detailing and finalising the methodology for the pilot evaluation. This scientific protocol will ensure that the evaluation data collection during the delivery of integrated health and social care (both in the comparison phase and the new care phase) is carried out according to a common methodology across all pilot sites.

BeyondSilos is collaborating closely with two other projects, CareWell and SmartCare. The three projects strive to create synergy and coherence between the methodologies used in the evaluation framework for the projects to allow comparison of results between the three projects.

However, the three projects are different, and the evaluation framework is adapted to the specific needs of each of them. Moreover, the evaluation of each project will be performed independently from the other two projects, but ensuring that the lessons learned within each project will be transferred to the others.

This evaluation framework constitutes D6.1. However, throughout the text, when referring to the current document, it will be termed a protocol as opposed to an evaluation framework. The document will be reissued as further details are agreed.

1.2 Structure of document

The following issues will be covered in the protocol:

- Section 2 provides background information, the rationale and the objectives of the project.
- Section 3 describes the methodology including the study design, the setting, the participants, the eligibility criteria, the variables, the indicators and comparators, and the statistical methods.
- Section 4 covers approvals from ethical committees, authorship guidelines including scientific dissemination strategy.

1.3 Glossary

Abbreviation	Full name
ACG	Adjusted Clinical Groups
ANOVA	Analysis of variance
CCI	Charlson Comorbidity Index
CHF	Chronic Heart Failure
CI	Confidence Intervals
COPD	Chronic Obstructive Pulmonary Disease
CRF	Case Report Form
CV	Curriculum Vitae
DES	Discrete Event Simulation
EHR	Electronic Healthcare Record

Abbreviation	Full name
EPR	Electronic Patient Records
EU	European Union
GP	General Practitioner
HADS	Hospital anxiety and depression scale
HTA	Health Technology Assessment
IC	Integrated Care
ICD9	International Classification of Disease, 9th edition,
ICT	Information Communication Technology
IHC	Integrated Health Care
ISPOR	The International Society for Pharmacoeconomics and Outcomes Research
MAST	Model for ASsessment of Telemedicine applications
NHS	National Health Service
OR	Odds Ration
PhD	Academic Degree of Doctor of Philosophy
PSP	Policy Support Programme
Renewing Health	REgioNs of Europe WorkINg toGether for HEALTH
STROBE	Strengthening the Reporting of Observational studies in Epidemiology
TM	Telemedicine
WHO	World Health Organisation
WP	Work Package

2 Background and rationale

The protocol will evaluate the impact of the new organisational models developed in the framework of the BeyondSilos pilot service in order to provide ICT supported integrated health and social care to elderly patients. The evaluation will be performed covering the needs of the different principal stakeholders, such as end users (care recipients), informal carers, formal care staff / professionals, managers, decision-makers and third-party payers. Evaluation of integrated health and social care service delivery processes (process evaluation) will improve the current scientifically based knowledge base on barriers and facilitators towards integrated care (IC) delivery. Beyond this, scientific knowledge will be generated on outcomes of IC service delivery from the perspective of all actors involved. Apart from generating a number of self-standing deliverables, this work package will directly feed into WP7 with a view to support further exploitation of project outcomes beyond the end of the project by relevant stakeholders, and wider dissemination during the project.

2.1 Main hypothesis

Integrated care will lead to a more personalised and coordinated care, improve outcomes for elderly patients, deliver more effective care and support, and provide more cost efficient health and social services.

2.2 Objectives

The overall aim of the evaluation carried out in BeyondSilos is to identify the differences introduced by implementing ICT supported integrated care in different domains according to the MAST evaluation framework [2], including safety, clinical and social outcomes, resource use and cost of care, user/carer experience and organisational changes. The focus of the evaluation will be the impact of integration and changing organisational models on elderly patients.

3 Methods

The evaluation will be conducted using the MAST multi-dimensional evaluation methodology adapted to the needs of the BeyondSilos project, focusing on ICT supported integrated care, in accordance with the recommendations of the ISPOR Good Research Practice Task Force on Prospective Observational Studies [3] and the STROBE statement [1].

MAST is based on Health Technology Assessment (HTA), and has been successfully validated in the ICT PSP Type A project Renewing Health. It is encountering an increasing level of success among organisations involved in trials of complex interventions such as those piloted in United4Health, SmartCare and CareWell, because it fills a gap which has been widely felt in this area.

MAST was developed under contract with the European Commission (MethoTelemed project) by a multinational team led by the Odense University Hospital, which is participating in United4Health as part of the South Denmark Regional Partnership. The same team, which developed and validated MAST, will be in charge of the evaluation of BeyondSilos. MAST includes assessment of the outcomes of telemedicine applications divided into the following seven domains:

1) Health problem and characteristics of the application.
2) Safety.
3) Clinical effectiveness.
4) Patient perspectives.
5) Economic aspects.
6) Organisational aspects.
7) Socio-cultural, ethical and legal aspects.

For further description of the MAST domains, please see Kidholm et al 2012 [2].

3.1 Study design

The aim of the evaluation is to quantify the relationship between ICT supported integrated care services to elderly patients and specific outcomes 12-18 months after the deployment of the new organisational models. The most appropriate study design for the evaluation is the cohort-study (prospective observational study), given that random allocation is not possible.

The strengths of this study design are mainly the collection of real-life data about impact on effectiveness, costs and organisation (structure and processes) which allows the identification of barriers and facilitators for a wider service implementation. Furthermore, the long follow-up period allows for registering and monitoring long-term health effects and other outcomes, while the large sample size allows for stratification analysis and identification of patient subgroups that benefit most from the intervention.

In addition, from an ethical perspective, if the new care is proved effective, it should be offered to all potential users in need of integrated health and social care. This type of study design will assess the real-life effectiveness of the trialled services with a high degree of external validity and generalisability of the results. Due to inclusion of patients from many European countries, this study will be able to provide a valid estimate of the expected impact of the new organisational models in other regions of Europe.

3.2 Setting

All settings that are in any way relevant to the provision of health and social care are included. Therefore, out-of-hospital (community) services as well as hospitals, GPs' offices, community nurses, and any type of care practitioners, users' homes and volunteer service providers' offices will be engaged. Participants will be enrolled and the evaluation will be conducted at the following seven pilot sites.

- Northern Ireland
- Badalona
- Valencia
- Campania
- Amadora
- Kinzigtal
- Sofia

3.2.1 Dates and timetable

The enrolment will start on 1st September 2014 and will end 31st Marts 2015. However, for Northern Ireland there will be a short delay in the enrolment phase. In the enrolment period, the pilot sites will be expected to achieve the sample size they have declared and quoted in the Technical Annex. Any deviation from the declared number has to be reported with a proper explanation to the evaluation work package leader (WP6) and to the Pilot site preparation and operation leader (WP5). Early dropouts (within the first four months) should be replaced. The maximal duration of the follow-up will be 18 months, while the minimum will be 12 months. Data collection has to be completed before 31th August 2016, and the evaluation before 31st December 2016.

Table 1: Timetable

	Start enrolment	Expected finishing date of enrolling	Finish data collection	Data source for comparator group
Northern Ireland	1st November 2014	31st July 2015	31th August 2016	Nine month of comparison collection (usual care) followed by 12 month of new care
Badalona	1st September 2014	31st March 2015	31th August 2016	Parallel comparison groups
Valencia	1st September 2014	31st March 2015	31th August 2016	Six month of comparison collection (usual care) followed by 12 month of new care
Campania	1st September 2014	31st March 2015	31th August 2016	Parallel comparison groups
Amadora	1st September 2014	31st March 2015	31th August 2016	Six month of comparison collection (usual care) followed by 12 month of new care
Kinzigtal	1st September 2014	31st March 2015	31th August 2016	Parallel comparison groups from other geographical area
Sofia	1st September 2014	31st March 2015	31th August 2016	Parallel comparison groups

3.2.2 Data collection and management

A case report form (CRF) will be developed in Excel describing all the specific data that the pilot sites will have to collect from the participants and study settings. The CRF will specify level of variables, type of data (text or numbers), as well as validation rules, including minimum and maximum values. Each pilot site will be responsible for collecting their own data, and to cleaning data in line with the evaluation protocol (see section 3.9 Data handling) De-identified data from each pilot site will be uploaded to a central web-based database administered by Region of Southern Denmark. This will allow comparison of data between the different pilot sites. Each pilot site will have a separate log-in to access the database, and will be able to view its own data as well as aggregated data from all the pilot sites. The database will have daily back-up and secured data transfer. Internet connection is mandatory in order to access the central web-based database.

3.3 Study population

3.3.1 Eligibility criteria

Inclusion criteria for end users:

Participants eligible for the evaluation must comply with all of the following criteria:

- Age ≥65 years.
- Presence of health needs specified as:
  - Presence of heart failure, stroke, COPD or diabetes (diagnosed at hospital or at specialist visit) plus at least one additional chronic disease / condition included in the Charlson Comorbidity Index (CCI) [4].
- Presence of social needs based on Activities of Daily Living (ADL) or Instrumental Activities of Daily Living (IADL).
- Reasonable expectation of permanence in the BeyondSilos project for the whole data collection period (18 months).
- Informed consent, signed if necessary (by the subject or his/her delegate).
- Capability to handle ICT equipment / devices alone, or with the help from a delegate.
- Presence of good/reliable communication connection at home (internet, telephone or what is needed for the ICT connection).

Exclusion criteria for end users:

- Subjects who have been registered with an active cancer diagnosis and undergoing treatment, has undergone an organ transplant, or is undergoing dialysis prior to enrolment.
- Subjects in a terminal state.

3.3.2 Recruitment of study population

The set-up of all BeyondSilos pilot sites is cohort studies, which means that a group of people with similar characteristics will be followed over a period of time. In order to measure whether integrated health and social care has an effect, all pilot sites will provide both a group that will receive the new care and a comparator group that receives usual care. Potential participants are selected by screening electronic healthcare and social care records or/and the hospital / national databases and/or during long term condition annual reviews in the community setting. If necessary, candidates are informed about the

nature and the objectives of the evaluation. If a candidate passes the inclusion/exclusion criteria and signs the informed consent form, if necessary, they participate in the evaluation. Figure 1 summarises the flow of enrolment to clarify the steps of recruiting patients.

Screening (= subjects that in theory could be suitable; i.e. n=1000)

⬇

Eligible-candidates (=subjects that meet the inclusion/exclusion criteria; i.e = 200)

⬇

Enrolled subjects (=subjects that were evaluated and have accepted to participate giving a formal consent; i.e. n= 100).

Figure 1: Steps for the recruitment

It is relevant to underline three key issues in the recruitment of the study population:

- Deadlines for enrolment and follow-up period have to be followed (see section 3.2.1).
- The pilot sites are expected to replace early drop-outs, therefore there will be continuous attention to have available eligible candidate (thus maintaining the screening phase active).
- The recruitment can be performed by both healthcare practitioners and by social care professionals, who will be qualified to reach this goal.

3.3.3 Comparator group

To take appropriate account of particular national/regional circumstances the rules for selecting a comparator group can differ between pilot sites. Most pilot sites will have a parallel running comparator group that receives usual care (option 1, Figure 2). However, some pilot sites plan to enrol subjects in a two phase-observation period, where phase 1 is the comparator phase (usual care), and phase 2 is the new care phase (option 2, Figure 3).

Option 1: All enrolled subjects (n = 100) will be assigned to one of the two groups. Both groups will be followed in parallel over time.

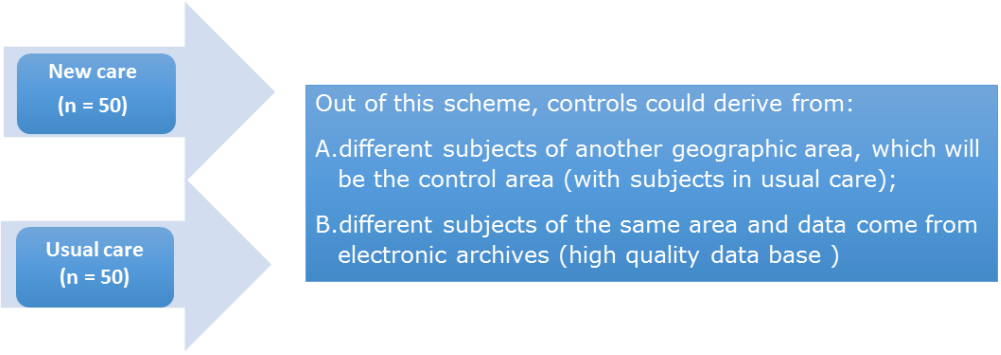

Figure 2: Composition of the comparator group - Option 1 parallel group

Option 2: All enrolled subjects (n = 100) enter into a two phases-observation period of 18 months, where phase I is the comparator phase (usual care) and phase II is the new care phase.

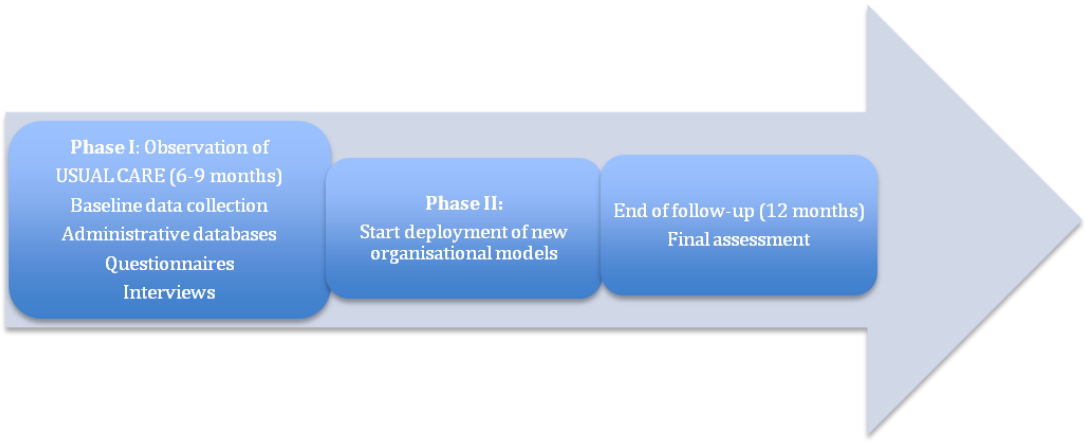

Figure 3: Composition of the comparator group - Option 2 phased group

3.4 Variables

All outcome metrics, timing and explanation for variables, along with identification of variable, analysis metric, time point and explanation for inclusion of each variable are presented in Table 2 below. In addition, the table indicates whether each variable can be included on a voluntary basis (V) by pilot sites, or if they are required to collect data (mandatory, M).

The mandatory variables are defined by study aims and objectives, and will be used in the final analyses of the study.

The variables / indicators of interest cover the following domains:

1. Overall service effectiveness and specific outcome metrics:
    - 1.a Disease specific health status metrics.
    - 1.b Generic health related / functional quality of life.
    - 1.c Psychological metrics.
2. Safety.
3. End user / client / carer perspectives.
4. Economic measures.
5. Organisational impact measures.
6. Possible confounders / control variables.

Table 2: Outcome, metrics, timing and explanation for variables

Measurement	Respondent / target group	Level of data	Level of detail	M/V	Preferred collection method	Timing of measurement	Reason
1. Overall service effectiveness and specific outcome measures							
Number of hospitalisations	Patient	Individual level	Number	M	Administrative databases, EHR	Baseline / end	Total number of hospitalisations is 1) easy to establish (was there a contact or not), and 2) it is available in all sites
Total days hospitalised							Total days hospitalised is 1) easy to establish, and 2) it is available in all sites
Number of emergency department visits							Total number of ED visits is 1) easy to establish (was there a contact or not), and 2) it is available in all sites
Number of re-hospitalisations within 30 days							Total number of re-hospitalisations is 1) easy to establish (was there a contact or not), and 2) it is available in all sites
Number of contacts, healthcare services	Patient / carer	Individual level	Number	M	Administrative databases, EHR	Baseline / end	Total number of face to face contacts (planned and unplanned) is 1) easy to establish (was there a contact or not), and 2) it is available in all sites
Unplanned contacts, healthcare services				V			Unplanned contacts is chosen because it is 1) easy to establish (was there an unplanned contact or not), and 2) it reflects both the aim of the interventions in clinical terms, but also safety issues, organisational and economic aspects. At each site, the exact meaning and operationalisation of this outcome measure needs to be defined.

Measurement	Respondent / target group	Level of data	Level of detail	M/V	Preferred collection method	Timing of measurement	Reason
Number of contacts, social care services	Patient / carer	Individual level	Number	M	Administrative databases, EHR	Baseline / end	Total number of contacts is 1) easy to establish (was there a contact or not), and 2) it is available in all sites
Unplanned contacts, social care services				V			Unplanned contacts is chosen because it is 1) easy to establish (was there an unplanned contact or not), and 2) it reflects both the aim of the interventions in clinical terms, but also safety issues, organisational and economic aspects. At each site, the exact meaning and operationalisation of this outcome measure needs to be defined.
Number of contacts, volunteer sector services				M, if relevant in setting			Total number of contacts is 1) easy to establish (was there a contact or not), and 2) it is available in all sites
Unplanned contacts, volunteer sector services				V			Unplanned contacts is chosen because it is 1) easy to establish (was there an unplanned contact or not), and 2) it reflects both the aim of the interventions in clinical terms, but also safety issues, organisational and economic aspects. At each site, the exact meaning and operationalisation of this outcome measure needs to be defined.
1.a Disease specific health status measures							
Blood pressure	Patient	Individual level	Number	V	Administrative databases, EHR, clinical assessment	Baseline / end	Indicator for health status
Heart rate							Indicator for health status
Weight							Indicator for health status
Oxygen saturation							Indicator for health status (only for patients with COPD or CHF)

Measurement	Respondent / target group	Level of data	Level of detail	M/V	Preferred collection method	Timing of measurement	Reason
Blood glucose							Indicator for health status (diabetics only)
HbA1c	Patient	Individual level	Number	V	Administrative databases, EHR, clinical assessment	Baseline / end	Indicator for health status (diabetics only)
Status/severity of primary condition			Scale or number				Predictor of health outcome
1.b Generic health related / functional quality of life							
Charlson Comorbidity Index (CCI)	Patient / carer	Individual level	Scale	M	Clinical measurement	Baseline / exit	Indicator for health status
Barthel index	Patient / carer	Individual level	Scale	V	Clinical measurement	Baseline / exit	Indicator for health status
SF 36 v2	Patient	Individual level	Scale	V	Questionnaire or interview	Baseline / exit	Might be affected by the intervention
1.c Psychological measures							
Anxiety and depression according to HADS	Patient	Individual level	Number	V	Questionnaire or interview	Baseline / exit	HADS is used to determine the levels of anxiety and depression in end users. It is a 14-item scale. Seven of the items relate to anxiety, and seven related to depression.
Depression according to GDS	Patient	Individual level	Number	V	Questionnaire or interview	Baseline / exit	Geriatric Depression Scale-15 (GDS-15) is a short, 15-item instrument specifically designed to assess depression in geriatric populations. Its items require a yes/no response. GDS was first introduced by Yesavage et al. in 1983, and the short form (GDS-15) was developed by Sheikh and Yesavage in 1986.

Measurement	Respondent / target group	Level of data	Level of detail	M/V	Preferred collection method	Timing of measurement	Reason
Isolation according to Perceived Isolation Questionnaire	Patient	Individual level	Number	V	Questionnaire or interview	Baseline / exit	Previous research has identified a wide range of indicators of social isolation that pose health risks, including living alone, having a small social network, infrequent participation in social activities, and feelings of loneliness. However, multiple forms of isolation are rarely studied together, making it difficult to determine which aspects of isolation are most harmful to health. Cornwell and Waite (2009) used population-based data from the National Social Life, Health, and Aging Project to generate questions combining multiple indicators of social isolation into scales assessing social disconnectedness (e.g., small social network, infrequent participation in social activities) and perceived isolation (e.g., loneliness, perceived lack of social support). These questions can be ascribed numerical values so that, when repeated, they provide a way for people to self-rate whether they are more or less socially disconnected and isolated from others than at the previous time of measurement
Anxiety	Patient / carer	Individual level	Scale	V	Questionnaire or interview	Baseline / end	Indicator for health status

Measurement	Respondent / target group	Level of data	Level of detail	M/V	Preferred collection method	Timing of measurement	Reason
Carer burden according to ZBI (short version)	Carers	Individual level	Number	V	Questionnaire or interview	Baseline / exit	The Zarit Burden Interview was developed to measure subjective burden among family carers of adults with dementia. Items were generated based on clinical experience with family carers and previous research, resulting in a 22-item self-report inventory that examines the burden associated with functional or behavioural impairments and the home care situation. Most researchers use the 22-item version of ZBI. However, the length of the instrument may be a deterrent to its use in clinical and research environments. Bédard et al produced a short version consisting of 12 items, with results comparable to the full version. Cronbach's α for the 12-item version is 0.88.
Carer burden according to CADI-CASI-CAMI suite	Carers	Individual level	Number	V	Questionnaire or interview	Baseline / exit	Carers are also assessed for difficulties, satisfaction and management in caring using the CADI-CASI-CAMI suite. This suite is a collection of three instruments used to assess family carers' perceptions of difficulty, satisfaction and management (coping strategies). The Carer Assessment of Difficulty Index (CADI) is a 30-item index, and contains a series of statements which carers have made about the difficulties they face. Carers are asked to tick the box next to each statement that most applies to them from the following options: 'this does not apply to me', 'not stressful', 'stressful', and 'very stressful'. The Carer Assessment of Satisfaction Index (CASI) is also a 30-item index, and contains a series of statements about the satisfaction of carers experience. The Carer Assessment of Management Index (CAMI) is a 38-item questionnaire and contains a series of statements about the coping strategies used by family carers.

Measurement	Respondent / target group	Level of data	Level of detail	M/V	Preferred collection method	Timing of measurement	Reason
2. Safety							
Deaths	Patient	Individual level	Yes/no (dichotomous)	M	Administrative databases, EHR	End	Easy to establish, common as adverse outcome
3. End user / client / carer perspectives							
End user / client / carer empowerment	Citizen / client / carer	Individual level	Scale for each question	M	Questionnaire	Exit	Reflects the aim of BeyondSilos
End user / client / carer self-management				M	Questionnaire		Reflects the aim of BeyondSilos
End user / client /carer satisfaction				M	Questionnaire, IFIC		This would be based on the eCare Client Impact Survey developed in CommonWell and INDEPENDENT in response to a lack of instruments measuring impacts on older end-users and informal carers beyond clinical outcomes, and with particular focus on impacts occurring from combined social and health care.
End user perception of integration	End-users	Individual level		M	Questionnaire	Exit	?
End user level of independence				V			?
End user level of adherence to drug therapy				V			?

Measurement	Respondent / target group	Level of data	Level of detail	M/V	Preferred collection method	Timing of measurement	Reason
4. Economic measures							
Efforts related to service development & implementation	Patient / carer/ Service providers	Individual or organisational level	Number	M	Various	Exit Implementation and pilot phase	To support the design and implementation of viable and sustainable services. To produce supportive economic data for internal decision making processes. To allow for an overall, post-hoc assessment of socio-economic impacts.
Efforts related to service operation or use				M			
Equipment cost	Service providers	Organisational level	Number	M	Various	Implementation and pilot phase	As above.
Service effectiveness benefits				M			As above.
Service efficiency benefits				M			As above.
Revenue streams				M			As above.
Willingness to pay	Citizen / client / carer	Individual level	Scale	V	Questionnaire	Exit	Relevant if a service fee payable by end user / client /carer is considered to become part of the revenue model.
5. Organisational impact measures							
Impacts on staff	Service providers: staff members and key informants / decision makers	Organisational level	Scales, qualitative	M	Questionnaire or interview	Pilot end	Key measures to understand the organisational changes caused by the new service, as well as to get a better understanding of what was actually achieved through the integration of different service silos. Can also capture where staff members and organisational decision makers are (still) not satisfied with the result.
Impacts on organisations				M			
Service integration aspects				M			

Measurement	Respondent / target group	Level of data	Level of detail	M/V	Preferred collection method	Timing of measurement	Reason
Mainstreaming potential and sustainability	Service providers: key informants / decision makers	Organisational level	Scales, qualitative	M	Questionnaire or interview	Pilot end	As above.
6. Possible confounders / control variables							
Year of birth	Citizen / client / carer	Individual level	YYYY	M	Registries or interview	Inclusion	Age is a strong predictor of any health outcome
Gender			Male/ female	M			Gender is very often related to health outcomes
Level of education			Categories (3)	M			Level of education is a strong predictor of any health outcome. Generally, it is said that one Euro given to education increases the level of health more than one Euro given to health care. Categories are important and have to be used in a similar way throughout pilots
Marital status			Categories	M			Marital status is a strong predictor of health outcomes. It is better to be married than being single. Categories are important and have to be used in a similar way throughout pilots

Measurement	Respondent / target group	Level of data	Level of detail	M/V	Preferred collection method	Timing of measurement	Reason
Ethnicity	Citizen / client / carer	Individual level	Categories	V	Questionnaire or interview	Inclusion	Ethnicity is strongly related to health outcomes
Work status			Categories	V			Work status is being recognised as a strong indicator of health outcome. It turns out that people belong to the social group in which they work rather than the one in which they are educated. Categories are important and have to be used in a similar way throughout pilots
People older than 18 living in household			Number	V			Indicator for the level of informal care received
Household income			Number	V			Necessary if willingness-to-pay is analysed.
Daily tobacco use	Citizen / client	Individual level	Dichotomous	V	Questionnaire or interview	Inclusion	Indicator for health status
Frequency of alcohol (12 months)			Categories	V			Indicator for health status
Height (CM)			Number	V			Indicator for health status
Weight (Kg)			Number	V			Indicator for health status
Co-morbidity (CCI)			ICD-10 codes	V			Indicator for health status, highly relevant for the usability of results after finishing pilots

3.5 Data source/measurement

See section 3.4 Table 2.

3.6 Bias

Some methodological issues should be considered when planning the evaluation, data collection and the analyses for the BeyondSilos project.

3.6.1 Information bias

Data will be collected from different sources, which includes administrative databases, questionnaires and interviews. Missing or inaccurate reporting in the administrative databases might occur. Therefore, each pilot site has to provide information on the quality of the databases used to collect information in order to assess the quality of the data. Missing or inaccurate information from the administrative databases are not expected to depend on the implementation of the ICT supported integrated healthcare (the exposure). Therefore, this possible non-differential misclassification should only have minor effect on the analyses.

Information collected from questionnaires and interviews may be influenced by recall bias. One pilot site has to collect data directly from the subjects for some of the primary outcomes, such as prior hospitalisations and contacts with healthcare services, and therefore may be especially effected by this form of bias. The magnitude of the effect of the bias on the study's result will be assessed and discussed in the analysis phase.

3.6.2 Selection bias

Differences in characteristics between subjects who consent to participate in the BeyondSilos project and those who decline may affect the external validity of the results of the new care. In order to address this issue, demographic characteristics of the included subjects will be compared with those who decline in order to examine for any systematic differences.

Pilot sites that are planning to enrol subjects in a two phase-observation period have to be aware of potential bias due to seasonal changes in the outcome measures. As an example, it is well established that more hospitalisations occurs in the winter period compared to the summer period, especially in the elderly population. It is therefore important that the pilot sites consider the seasonal calendar related effect when planning their comparison period.

3.7 Sample size

The number of patients which will be recruited and included in the evaluation of the project is:

- Pilot 1 - Northern Ireland: 300-450
- Pilot 2 – Badalona: 100
- Pilot 3 – Valencia: 150
- Pilot 4 - Campania: 50-100
- Pilot 5 - Amadora: 150
- Pilot 6 - Kinzigtal: 50
- Pilot 7 - Sofia: 100

In total, more than 900 patients will be included in the evaluation of the project.

3.8 Statistical methods

Separately analyses will be performed for each pilot site, as well as some common analyses comparing the results between the pilot sites for primary outcomes and for some selected secondary outcomes.

3.8.1 Local pilot sites

The choice of method to analyse the data depends on:

- the type of data that are investigated (dichotomous, categorical or numerical);
- whether or not the data are normally distributed.

Simple comparisons of the distribution of data will be performed and presented in tables or histograms.

Depending on the distribution of the data, continuous outcome variables are planned to be analysed using multivariate ANOVA tests examining the difference between group means.

Binary outcome variables will be analysed using multiple logistic regression models estimating the Odds Ratio (OR) with proper confidence intervals (CI). Different models adjusting for age, sex and other possible confounding variables will be performed.

A final detailed strategy for analyses will be elaborated before analysing data.

3.8.2 Overall analyses

Meta-analyses are planned to summarise and compare the results for the primary and secondary outcomes for the different pilot sites. The results from the meta-analyses will be presented as tables along with graphs showing the forest plots. In order to assess the percentage of the total variation in estimated effects across the studies that is due to heterogeneity rather than to chance, the I^2 will be shown for all meta-analyses.

First, a meta-analysis including results from all the pilot sites will be performed for the primary outcome, and investigated for subgroup impacts based on similarities among populations. The relevance of this analysis will be discussed based on the level of heterogeneity presented in the meta-analysis. Next, the pilot sites that have similar populations in terms of disease, frailty or other factors will be analysed together in a meta-analysis.

A final detailed strategy for the meta-analyses will be elaborated before analysing data.

3.9 Data handling

A two step procedure will be performed in order to detect and handle errors in the data that might impact the study results:

- Step 1: All pilot sites have to perform the following data cleansing process before submitting the data to the central web-based database. All subjects with missing values, or values that are considered to be illegal or outliers, must be checked and compared to an alternative reliable data source if such is available. The correct value (the most plausible) should be included in the dataset. However, a note must be made about the alteration of the value.
- Step 2: The following data cleansing will be performed when the data has been collected in the central web-based database before performing the analyses.

Missing values

- If one subject has <50% missing values, the remaining values are allowed in analyses.
- Analyses that require some of the missing data will be run without the values, and reporting will present the total number of subjects in all analyses.

Outlier

- If a value is considered to be a realistic outlier, the value will remain unchanged. Sensitivity analysis will be carried out to assess the impact of the outliers.
- If a value is considered to be an unrealistic outlier, the value will be re-coded as missing.

Range check

- A value is considered illegal if it falls outside the min-max range of possible values, and will be re-coded as missing.

Categorical variables

- All observations must relate to the predefined categories, otherwise the value will be registered as missing.

4 Ethics and dissemination

4.1 Plans for seeking research approval

Whenever necessary, pilot sites will seek ethical approval in order to collect and evaluate patient data.

4.2 Authorship guidelines

Regarding scientific dissemination, the BeyondSilos project will agree on a process for authorship, acknowledgment of the project work and other supportive works, and sign off. This process will be mandatory for all publications conducted in the context of the BeyondSilos project or its data. The process for authorship will be decided at the next PCC meeting, and included in an updated version of the evaluation protocol.

5 Conclusion

This protocol contains the suggested evaluation framework for the BeyondSilos project, reviewed and agreed by the BeyondSilos evaluation group. The protocol will be finally approved at the next PCC meeting to be held in Lisbon, Portugal, on 16th-17th October, 2014. Parts of the evaluation framework will need further updating as data collection is finalised. Additional conclusions will be added parallel with the data collection and analyses.

6 References

1. Vandenbroucke JP, Elm E von, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Explanation and Elaboration. Epidemiology 2007;18: 805–835.

2. Kidholm K , Ekeland AG, Jensen LK, Rasmussen J, Pedersen CD, Bowes A, Flottorp SA, Bech M. A model for assessment of telemedicine applications: MAST. *International Journal of Technology Assessment in Health Care* 2012: 28 (1) 44-51.

3. Berger ML, Dreyer N, Anderson Fred, et al. Prospective Observational Studies to Assess Comparative Effectiveness: The ISPOR Good Research Practices Task Force Report. Value Health 2012;15:217-230.

4. Charlson, M. E., Pompei, P., Ales, K. L., & MacKenzie, C. R. (1987). A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of Chronic Diseases, 40(5), 373–383.