Invest4Health

www.invest4health.eu

CALL: HORIZON-HLTH-2022-CARE-08

DELIVERABLE No: D2.3 Proposal for outcome measures and study designs to assess the real-world impact of smart capacitating investment (SCI)

* Dissemination Level:	PU
** Deliverable Type:	R
Due delivery date	31 December 2024
Work Package:	WP2, Task 2.3
Actual delivery date:	20 December 2024
Editors:	Maureen Rutten-van Mölken, Apostolos Tsiachristas, Bengt Stavenow, Jolanda van Vliet, Anant Jani, Lucas Goosens, Balázs Babarczy, Beatriz Salgado Pardal, Shirin Bartholdsson, Gareth Davies, Alison Maassen, Shabs Rajasekharan, Yannick Becerra, Karol Puscus, Rositsa Koleva-Kolarova, Vidya Oruganti, and Joanna Lane
Contributors:	Rhiannon Tudor Edwards, Jacob Davies, Sofie Roberts, Holly Whiteley on behalf of the I4H Consortium

* **Dissemination Level**:	*PU= Public* *CO= Confidential, only for members of the Consortium (including the Commission services)* *EU-REST= Classified information: RESTRAINT UE* *EU-CONF= Classified Information: CONFIDENTIEL UE* *EU-SECRET= Classified Information: SECRET UE* *(based on Council Decision 2013/488/EU - EU Classified Information (EUCI) Decision)*
** **Deliverable Type**:	*R= Document, DEM= Demonstrator, pilot, prototype, DEC= Website, patent filling, videos, etc., OTHER, ETHICS= Ethics requirement*

DISCLAIMER

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or [European Health and Digital Executive Agency (HADEA)]. Neither the European Union nor the granting authority can be held responsible for them.

Deliverable Summary

The challenge: Finding novel sources and mechanisms of finance for health promotion and disease prevention to promote population health requires evaluation of both the effectiveness and cost-effectiveness of interventions themselves and the evaluation of the performance of the smart capacitating investment (SCI) model, including the underpinning business and governance models.

Aim: Here we present a proposal for outcome measures and study designs to assess the real-world impact of SCI in the four regional Invest4Health (I4H) testbeds (Hywel Dda, Wales, UK; Skåne, Sweden; Galicia, Spain, and North Rhine Westphalia, Germany). It is to be noted that the timing and scope of the Invest4Health study preclude operationalisation of these proposals during the lifetime of the I4H project. Deliverable 2.3 was specifically bounded to propose the designs which would form part of the future bid for research funding to enable operationalisation.

Methods: Our proposal for outcome measures and study designs is informed by:

- Collective I4H consortium knowledge of evaluating complex interventions.
- Existing I4H outputs, including the SCI conceptual framework (T2.1) and SCI readiness self-assessment tool (T2.2) developed in Work Package (WP) 2.
- Informed by WP3 contribution such as multi-stakeholder business and governance models and the design of collaborative spaces and data management platforms.
- Expert interviews with investment experts (WP4) a review of the measurement of outcomes across multiple sectors (T4.3) undertaken in WP4.
- Close liaison with the four regional testbeds.

Results: Outcome measure and study design options for the four testbeds can be seen in Tables 5, 7, 9 and 11. Table 1 below presents an overview summary of proposed study designs and outcome measures for the four regional testbeds in the I4H project.

Table 1. Overview of the proposed study designs and outcome measures for the four I4H testbeds.

	Hywel Dda (Wales)	**Skåne (Sweden)**	**Galicia (Spain)**	**North Rhine Westphalia (Germany)**
Study Design	Intervention matched control pre-and post-intervention, systems-level randomised controlled trial, utilising a natural experiment design.	Systems level case study/ natural experiment.	Pilot patient-level randomised controlled trial with control and intervention arms.	Matched control experiment.
Outcome measures	Uptake of existing interventions while utilising SCI backed social prescribing compared to standard practice, changes in user quality of life and change in	Effectiveness of cross-sectoral/multi-stakeholder governance and pooling of resource, proportion of pooled resources invested by	Reduction in primary and secondary healthcare contacts for disease progression monitoring,	Absenteeism, productivity, reduction in conditions where sedentary behaviour and poor home

	biomarkers for Type 2 diabetes post-intervention.	each collaborating body.	improvement in quality of life.	working environment are risk factors.

Alongside the above proposals we also present a series of questions designed to help guide outcome choice in future SCI evaluations.

Discussion: This deliverable provides the foundations from which the operationalisation of proposed testbed evaluations can be materialised. This deliverable draws upon existing work in this project, the published literature, interviews with expert stakeholders in the space and through discussions with testbeds to assess what is rigorous, acceptable, and feasible in each respective testbed.

List of Authors

Partner	Authors
BU	Rhiannon Tudor Edwards, Jacob Davies, Sofie Roberts, Holly Whiteley
EUR	Maureen Rutten-van Mölken, Lucas Goossens
SRI	Balázs Babarczy
OU	Apostolos Tsiachristas, Rositsa Koleva-Kolarova, Anant Jani
RISE	Shirin Bartholdsson
IESE	Shabs Rajasekharan
ACIS	Beatriz Salgado Pardal
RWTH	Yannick Becerra, Karol Puscus
IS	Bengt Stavenow, Jolanda van Vliet
HDdUHB	Gareth Davies
EHN	Alison Maassen
HCN	Joanna Lane

Document History

Date	Version	Editors	Status
18/09/24	1.1	HW	First draft
10/09/24	1.2	JD, SAR, HW	Working draft
14/11/24	1.3	RTE, JD, SAR	Working draft
22/11/24	1.4	RTE	Working draft
27/11/24	1.5	HW, JD, SAR	Working draft
28/11/24	1.6	JD, SAR, RTE	Working draft
18/12/24	1.7	JD, SAR, RTE	Finalised version sent to BS for upload

Glossary

Acronym	Meaning
CA	Consortium Agreement. The management and regulatory framework from which the team works.
DALY	Disability Adjusted Life Year. The sum of years of potential life lost due to premature death and the years of productive life lost due to disability compared to a standardised life expectancy. One DALY represents the loss of the equivalent of one year of full health.
GDP	Gross Domestic Product. The total monetary or market value of all the finished goods and services produced within a country's borders in a specific period. It functions as a comprehensive scorecard of a given country's economic health.
GO Lab	Government Outcomes Lab, Oxford University.
MCDA	Multi Criteria Decision Analysis. MCDA is an evaluative tool for multiple conflicting criteria. It is a way of helping decision-makers rationally choose between multiple options where there are several conflicting objectives.
MRC	Medical Research Council.
NHS	National Health Service.
OECD	Organisation for Economic Co-operation and Development, an intergovernmental organisation with 38 member countries.
QALY	Quality Adjusted Life Year. It is the academic standard for measuring how well all different kinds of medical treatments lengthen and/or improve patients' lives. The metric has served as a fundamental component of cost-effectiveness analysis for more than 30 years.
ROI	Return on Investment. ROI measures the amount of return on a project relative to its cost. It is increasingly being used to evaluate the value for money of public health interventions
SCI	Smart Capacitating Investment. Task 2.3 uses the definition of SCI as constituting less conventional approaches to investment, financial or non-financial in health promotion and disease prevention initiatives that create returns somewhere along a spectrum between financial benefit on the one hand and the enhancement of the capacity or capability of individuals or communities to engage in healthier behaviour on the other.
SELFIE project	*Sustainable intEgrated care modeLs for multi-morbidity: delivery, FInancing and performancE* (SELFIE) project, (European Union's Horizon 2020 research and innovation programme, grant agreement number 634288).
SROI	Social Return on Investment. SROI is a framework for measuring and accounting for this much broader concept of value by incorporating social, environmental and economic costs and benefits.
Testbed	Regions in West Wales (UK), Skåne (Sweden), Galicia (Spain), and North Rhine Westphalia (Germany).
UK	United Kingdom.
WHO	World Health Organization.
WP	Work Package.

Table of Contents

1. Introduction

1.1. Background

The Invest4Health (I4H) project is developing and evaluating Smart Capacitating Investment (SCI) models in practical settings in four European testbeds with a view to create a roadmap for large-scale implementation that will help transform public health financing and other type of enablers related to capabilities and capacity for better health and well-being. Preceding work in the project has laid the theoretical foundation for SCI (e.g. T2.1 *Scoping review and realist synthesis of existing SCI examples* and T2.2 *SCIROCCO tool to self-assess implementation readiness*). This deliverable (D2.3 *Proposal for outcome measures and study designs to assess the real-world impact of smart capacitating investment [SCI]*) focuses on and builds upon these theoretical foundations and knowledge by proposing study designs and outcome measures for assessing the real-world impact of the SCI in each testbed.

1.2. Aim

The key aim of this deliverable (2.3) is to propose appropriate study designs and outcome measures for the pragmatic evaluation (testing) of SCI models in four regional I4H testbeds (Hywel Dda, Wales, UK; Skåne, Sweden; Galicia, Spain, and North Rhine Westphalia, Germany). Our approach to evaluation and outcome measurement in testbeds is pragmatic in the sense of a realistic, practical approach rather than wholly theory driven. In the development of this deliverable, we explored the evaluative frameworks, study designs and outcome measures available and suitable for assessing the real-world impact of SCI. We then applied our findings to the four regional I4H testbeds and we present the most appropriate study designs and outcome measures for their specific contexts in this report.

This deliverable (D2.3), alongside T2.1 and T2.2, helps set the scene for developing the guidance for performing cross-sectoral evaluation of SCI in Work Package (WP) 4. Such a comprehensive evaluation approach is vital in capturing the spill-over effects of investment to and interactions between different government sectors such as education, housing, transport, the built and natural environment, as well as the health sector.

It is to be noted that the timing and scope of the Invest4Health study preclude operationalisation of these proposals during the lifetime of the I4H project. Deliverable 2.3 was specifically bounded to propose the designs which would form part of the future bid for research funding to enable operationalisation.

2. Methods

This section gives a brief overview of the approach and methods we used to develop our proposal for outcome measures and study designs to evaluate the impact of SCI in the four regional I4H testbeds. Our extended methods section (Section 4) delves into our extensive wider exploration into available evaluative frameworks, study designs and outcome measures suitable for SCI evaluation in more detail. Table 2 below presents an overview of the steps taken in the development of the proposed evaluation plans with our testbed partners.

Table 2. Steps taken in developing testbed evaluation plans

	Describing intervention	**Determining counterfactual**	**Determining identification strategy**	**Selecting relevant outcome measures**	**Choosing method of analysis**	**Data collection**
Steps taken	Iterative development of logic models and decision trees with testbeds; exploration and discussion at dedicated 'deep dives' with each testbed.	Current care provision was mapped in each testbed prior to SCI implementation to understand and identify respective counterfactuals in each case.	Assumptions concerning data and exogeneous factors that may influence results were explored in testbed 'deep dives' after mapping current contexts and landscapes via logic models and decision trees.	Selected outcome measures reflect findings from testbed logic models of targeted short-, medium-, and long-term outcomes, as well as findings from scoping reviews in T2.1 and T4.3 of evaluation examples.	Methods of analysis were chosen after iterative discussions of testbed data availability and feasibility of study designs. Also informed by scoping reviews in T2.1 and T4.3.	Data collection methods described following discussions with testbeds on feasibility and availability of data. Further informed by preceding work informing the method of analysis.

This deliverable naturally follows the preceding deliverables, T2.1 and T2.2, and incorporates findings from interviews with investment experts (T4.1 or T4.2), a scoping review of frameworks and outcome measurements used in evaluations of complex, multi- and cross-sectoral healthcare interventions (T4.3), and close liaison with the four regional I4H testbeds (T2.3).

We recognise that evaluation and outcome measurement in the context of SCI involves the evaluation of a complex intervention that consists of multiple components, such as the intervention itself and the innovative way the intervention is financed. It may not always be possible to disentangle the impact of those components on every outcome measure that will be proposed.

In our exploration of available evaluation approaches and study designs for SCI evaluation, we collated existing I4H consortium knowledge and relevant literature in the domains of evaluative frameworks for complex, multi-sectoral interventions and study designs, for example from published frameworks such as the updated Medical Research Council (MRC) framework for developing and evaluating complex interventions (Skivington et al., 2021), and applied them in the novel context of SCI.

2.1: Reviews of outcome measures for complex, multi-sectoral health interventions and SCI examples

In this section we provide further details as to the research that informed the development of the proposed testbed assessment proformas (see section 4). A scoping review of empirical evaluations of healthcare interventions that incorporated outcomes and costs in sectors beyond the healthcare sector (T4.3) provides an overview of underlying evaluation frameworks and their operationalisation. The review identified 38 unique publications from which information regarding evaluation frameworks employed, sectors evaluated, how outcomes were measured and monetised, how resource use in different sectors was valued and monetised, and what methods were employed to measure outcomes and costs.

Identified cross-sectoral evaluations primarily concerned curative healthcare interventions, rather than disease preventing/health promoting interventions. However, the multi-sectoral nature of the evaluations builds upon and is complementary to the scoping review of SCI examples (T2.1). The scoping review of SCI examples compiled outcomes across health and well-being and non-health outcomes utilised in their evaluations. Further outcomes for evaluating the SCI mechanism are captured from expert (T4.1/T4.2) and stakeholder interviews.

3 Evaluative frameworks for SCI

SCIs by their nature (as described in the protocol paper *'A protocol for mobilising novel finance models for collaborative health promotion and disease prevention initiatives: Taking a Smart Capacitating Investment (SCI) approach in the Invest4Health project'*) are complex and typically **cross-sectoral**. Evaluations of their real-world impact must be considerate of this. Our testbed SCIs meet several of the MRC criteria for complex interventions. Namely, they are in themselves complex, multi-faceted interventions targeting health promotion/disease prevention in their respective contexts. These interventions are supported by innovative SCI investment and business models, often pooling resources (both financial and non-financial) from across sectors and stakeholders. In some cases, these stakeholders may have competing or different perspectives on the success of an SCI (be that monetary or social gains achieved). Further, all testbeds operate within healthcare systems that are currently focused on the distribution of curative rather than preventative healthcare. With the SCI testbed interventions being health promoting/disease preventing, this conflicting interaction with their contexts can add further complexity. This section gives an overview of how to approach evaluating complex interventions across sectors.

The evaluation of complex interventions is not a new concept, with UK MRC guidance for the design and evaluation of complex interventions first published in 2000 and subsequent updates in 2008 and 2021 (Skivington et al., 2021; Craig et al., 2008; Campbell et al., 2000). The 2021 update provided significant input into our proposal for outcome measures and study designs to assess the real-world impact of the SCIs in each of our four testbeds.

In developing our proposal for outcome measures and study designs to assess the real-world impact of our SCIs, we draw upon the MRC complex intervention research framework (Figure 1). This research framework was employed in collaboration with our I4H testbeds to understand their respective elements. Section 3 presents an overview of each testbed. In Annex 5 we present amended frameworks for developing and evaluating complex interventions (amended Figure 1s) representative of each testbed's complex intervention. This co-production with the testbeds allowed for a more accurate understanding of their respective complexities and contexts, in turn facilitating more accurate proposals for suitable study designs and outcome measures in each case.

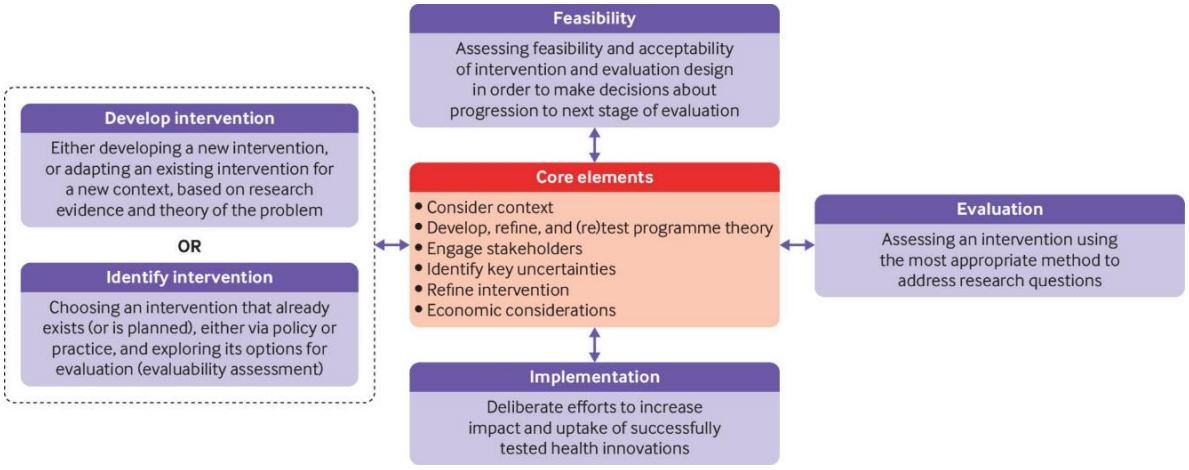


Figure 1: Framework for developing and evaluating complex interventions, re-used from (Skivington et al., 2021).

While being heavily informed by the MRC guidance update 2021, our proposals are also grounded in wider literature on evaluative frameworks in public health evaluations. Work by Moore and colleagues points to the need for process and outcomes evaluations in complex public health intervention research (Moore et al., 2013). The importance of understanding the contexts within which SCI-based prevention interventions exist and their interaction with one another complements the suggestions from the 2021 MRC guidance. The work adds that capturing how interventions are implemented in respective contexts can be captured through process evaluations. It further argues for combining process and outcomes evaluations to assess the 'fidelity and quality of implementation, clarify causal mechanisms and identify contextual factors associated with variation in outcomes' (Moore et al., 2013).

Noting the cross-sectoral nature of our SCI models and the likelihood of costs and outcomes falling on multiple sectors, we have taken learnings from the framework for economic evaluations when costs and effects fall on multiple sectors and decision makers (Walker et al., 2019). This framework formalises the capture of outcomes and costs further to the direct costs and outcomes to the healthcare sector. The capture of these wider outcomes will be compared with the findings from the preceding reviews in T2.1 and T4.3. The consideration of the perspectives in each testbed study design will also be informed by the descriptions in Walker et al. (2019).

Many influences of population health lie outside the health sector. Decision makers outside the health sector are primarily interested in delivering sector-specific outputs other than outputs relevant to health. Economic approaches to priority setting can help siloed sectors to consider the intersectoral outcomes of decisions (Weatherly et al., 2009). Economic evaluation can be employed to explicitly identify and measure all impacts and their values as inputs in total social welfare. This evidence evaluation can be translated into evidence-informed policy through prioritisation methodologies such as programme budgeting marginal analysis (PBMA) and multi-criteria decision analysis (MCDA) (Lawson et al., 2014). The I4H work programme utilises MCDA methodology, informed by the completion of O2.1-2.

We drew in particular upon the outputs of the previous EU-funded ''*Sustainable intEgrated care modeLs for multi-morbidity: delivery, FInancing and performancE*" (SELFIE) project, (European Union's Horizon 2020 research and innovation programme, grant agreement number 634288). The SELFIE project identified a range of study designs and statistical techniques suitable for assessing the causal impact of complex healthcare interventions, categorised into three groups: experimental designs, quasi-experimental designs and statistical techniques to improve

comparability between intervention and counterfactual, and observational designs (Table 3). These groups vary by the extent to which causality can be inferred. We recognise that the 'gold standard' study design is often cited to be experimental randomised controlled trials (RCTs), which minimise bias and better enable investigation of causal inference (Webber & Prouse, 2018). RCTs are not always practical nor possible in real-world contexts, however, and quasi-experimental or observational studies may be more feasible. Inferring causation can be more difficult for these types of study design that offer less control over conditions and exposure, but there are methods for strengthening casual inference (Craig et al., 2017).

The full list and description of the study designs identified in the SELFIE project are available in Annex 1. For each testbed we proposed an appropriate study design based on current understanding of the SCI to be evaluated and associated data availability.

Table 3. Hierarchy of study design/approaches for the evaluation of SCI by level of causality

Category	Study Designs/Approaches
Experimental designs	RCT Cluster-RCT
Quasi-experimental designs and statistical techniques to improve comparability between intervention and counterfactual	Stepped-Wedge Design Instrumental Variables Propensity Score Matching Regression Discontinuity Difference in Difference
Observational	Natural experiments Case Control Study Interrupted Time Series Analysis Cohort Study

In addition to drawing on existing knowledge and research, this deliverable is informed by several components and outputs from across I4H WP2–WP4. The infographic below (Figure 2) demonstrates the ways in which shared project learnings have supported the development of this deliverable.

Contextual information about each testbed was captured through close collaborative working with testbed partners and WP6 colleagues responsible for conducting interactive and exploratory workshops with testbeds. Problem trees and logic models for each testbed can be found in Annex 2. Completing the input from the testbed perspective are Task 2.2, the SCIROCCO self-assessment tool that has been incorporated as part of WP5 foresight exercises. We adopt the perspective in accordance with established understanding as the point of view adopted when deciding which types of costs and health benefits are to be included in an economic evaluation (York Health Economics Consortium, 2016).[1]

These exercises take testbeds through exploratory discussions of key stakeholders and to consider future implications of novel health promotion and disease preventing investment. This information informed the testbed proformas presented in the next section. Tasks 2.1 and 4.3

[1] Typical viewpoints are those of the patient, hospital/clinic, healthcare system or society. The societal perspective reflects a full range of social opportunity costs associated with different interventions. See: York Health Economics Consortium. (2016). *Perspective*. University of York. https://yhec.co.uk/glossary/perspective/

contributed published and grey literature evidence on SCI examples and cross-sectoral evaluations of healthcare interventions, respectively. This evidence contributed to the outcomes relating to the impact of the intervention itself and cross-sectoral outcomes. Finally, expert investor interviews conducted as part of WP4 and an expert interview with the Chief Strategy and Collaboration Officer in a National Health Service (NHS) Trust in the United Kingdom (UK) provided us with information on outcomes from the investor perspective that informed the outcomes evaluating the SCI mechanism.

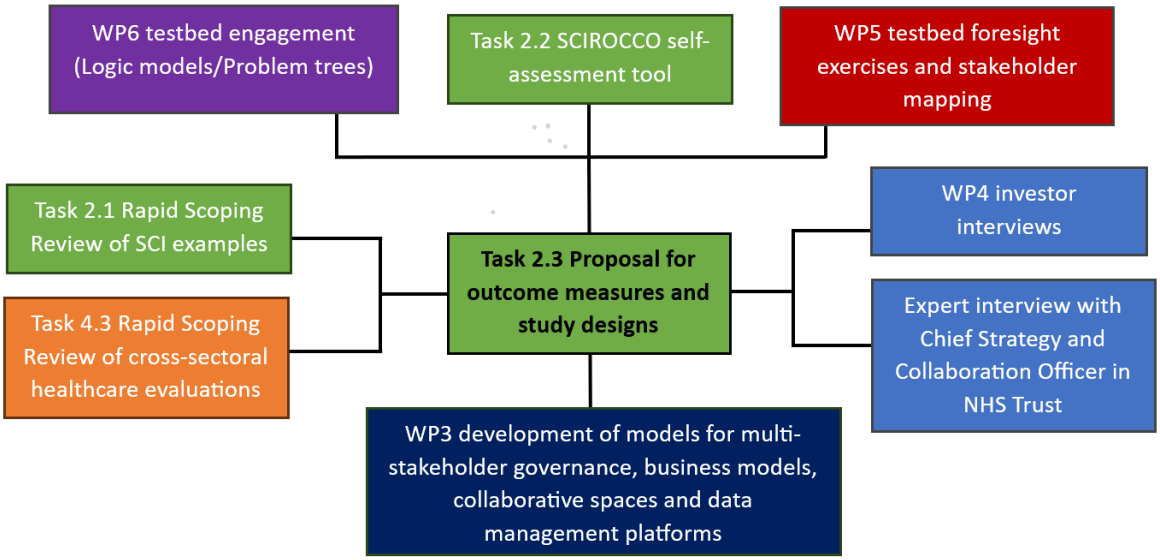


Figure 2. Overview of I4H tasks and outputs informing and contributing to D2.3.

4. Proposed study designs and outcome measures for I4H regional testbeds

Each of the four regional I4H testbeds are introduced below with their key features noted, along with our proposal for appropriate study designs and outcomes measures.

4.1 Hywel Dda, West Wales, UK

4.1.1 Context and intervention

Hywel Dda University Health Board (HDUHB) serves 388,000 citizens across three counties in southwest Wales. It is made up of largely rural communities, with dense population centres in two university towns. Key health challenges in the testbed region include: an ageing population; increase in prevalence of people living with more than one long-term illness; social isolation and loneliness; physical inactivity and unhealthy eating. The health promoting/disease preventing challenge targeted in this testbed is the prevention of type 2 diabetes and increasing physical and mental health and well-being problems.

The intervention consists of combining multiple existing programmes into one initiative that offers the services on an integrated digital platform. The programmes include a food-based diabetes prevention programme, a national exercise referral scheme and social prescribing. The All Wales Diabetes Prevention Programme (AWDPP) offers targeted consultation, information and support to those identified to be at risk of developing type 2 diabetes. The Wales National Exercise Referral Scheme (NERS) is a tailored, supervised exercise programme for 16 weeks that aims to improve the health and wellbeing of sedentary adults in Wales. The scheme is typically for

adults referred by NHS health professionals. Social prescribing is a method of connecting people to non-medical support to address healthcare and other unmet needs, typically through signposting to ongoing activities in the local community/third sector.

4.1.2 Innovative investment model

The intervention evaluated in the Hywel Dda testbed takes an approach to re-purposing existing health board funds, with a small contribution from service users through a novel alignment of existing services (NERS and AWDPP), facilitated by social prescribing.

The SCI targets the health promotion and disease prevention of type 2 diabetes and those with risk factors for developing the condition.

Table 4: Overview of Hywel Dda testbed characteristics

Characteristic	Description
Context and intervention	**Health challenge being addressed:** Type 2 diabetes. 8.4% of all adults aged 17+ in the Hywel Dda testbed region are living with a form of diabetes. **Description of intervention:** The intervention aims to reduce population-level incidence and prevalence of Type-2 diabetes by identifying and supporting those people at risk of developing Type 2 diabetes occurring (pre-diabetic) The intervention combines existing the existing AWDPP and NERS with local social prescribing services, community hubs and an online intervention delivery platform (Actif anywhere).
Sectors involved	**Primary sectors:** Health and Social Care, Community, Third sector. **Other sectors potentially impacted:** Transport, Environment, Employment.
Stakeholders involved	Health Board, Local Authorities in West Wales, Primary care services, social prescribers, NERS staff, community centres, leisure centres, fitness instructors and third sector organisations delivering socially prescribed interventions, general public.
Outcomes of interest	**Short term:** Increased participation of people at risk of developing type 2 diabetes (HP/DP intervention outcome); increased participation of people at risk of developing type 2 diabetes from hard to reach groups (HP/DP intervention outcome); higher adherence rates of people on the programme compared to NERS only (SCI process outcome); increased awareness of healthy diet and exercise (HP/DP intervention outcome); increased social networks, trust and civic engagement (cross-sectoral outcome); improved staff skills on the integrated approach (SCI process outcome). **Medium term:** Improved physical health (HP/DP intervention outcome); increased well-being (HP/DP intervention outcome); improved mental and social health (HP/DP intervention outcome); reduction in Body Mass Index (BMI) (HP/DP intervention outcome); reduction in glycated haemoglobin (hbA1c index) (HP/DP intervention outcome). **Long term:** Increased equality and reduced health inequalities (cross-sectoral outcome); improved quality of life (HP/DP intervention outcome); population level change in risk (cross-sectoral outcome); reduced hospital admissions; reduced medication use; reduced GP visits (HP/DP intervention outcome); reduced health service utilisation (HP/DP intervention outcome); increased retention of staff (SCI process outcome). increased cost-effectiveness and the sustainability of services only (SCI process outcome).
SCI Model	Resources pooled in a public-public partnership with a small fee-for-service contribution from participants.

Investor(s)	Health Board, Local Authority, fee-for-service contribution from participants. The investment from these stakeholders funds intervention costs such as staffing, equipment, room hire and training.
Expected Return on Investment	Financial (individual contribution to service to recoup some of initial health board expenditure in the short term, reductions in hospital admissions to reduce long-term expenditure) and social/health (improved physical and mental health, improvements in population health).
Data availability	A dataset containing biomarkers (hbA1c) and lifestyle (e.g., smoking, exercising) measurements for all individuals on the programme is actively being collected. The Wales testbed is currently working through the information governance process to make the data available for evaluation purposes.

Table 5. Proposal for study designs and outcome measures to assess the real-world impact of smart capacitating investment (SCI) in Hywel Dda UHB, Wales

Study design and outcome measures proposal	
Study design	Intervention matched control pre-and post-intervention, systems-level randomised controlled trial, utilising a natural experiment design. Recommended time horizon is a minimum of 6-to-12 months.
Target Population	People identified as being at risk of pre-diabetes/developing type 2 diabetes by their local GP or NHS health professional.
Intervention	Two Local Authorities (LA) in the region will be utilising the SCI, while one will not. The LA not using SCI will act as comparator.
Comparator	The comparator LA not implementing the SCI will not offer the connected NERS or AWDPP interventions facilitated through social prescribing as is the case in the LAs implementing the SCI.
Perspective of analysis	**Primary perspective:** NHS Health and social care perspective (see section 3). **Other:** Wider societal perspective (see section 3).
Outcome measures	**1) Process outcomes** Engagement and adherence rates; participation in traditionally hard to reach demographics; retention rates; health service utilisation; proportion of total delivery costs covered by fee-for-service participants; staff retention rates. **2) Population-health outcomes** Change in population risk profile (change in prevalence of pre-diabetic population); reductions in type 2 diabetes related complications (e.g., heart attack, diabetic retinopathy, stroke, amputations). **3) Clinical outcomes** Reduction in Body Mass Index (BMI); reduction in hbA1c index. **4) Patient-reported Outcomes** Changes in mental and physical health; self-reported quality of life. **5) Resource-utilisation outcomes** Reduction in GP visits, reduction in hospital admissions, reductions in medication use and prescriptions; sustainability of service provision.

	6) Broader Societal Outcomes Increased strength of social networks, trust, and civic engagement; improved staff skills on delivering the integrated approach.

4.2 Galicia, Spain

4.2.1 Context and Intervention

Geographic concentration of the population in Galicia is very low, with many sparsely populated rural areas. Increasing demands on the health and social care system and rurality make health and social care coverage difficult particularly for patients living with rare and life-limiting conditions. The prevention of increase in chronicity, in waiting lists for chronically ill patients, the decrease in the quality of life of the population because of their state of health, the prevention of frailty and the improvement in the control of the disease by the patient are key aspects in ensuring a more efficient and sustainable healthcare system for patients with rare and life-limiting conditions.

This testbed uses cystic fibrosis as an exploratory case study of the roll out of telemedicine technologies. This is a small group of chronically ill patients (around 120 CF patients in Galicia) which will allow us to carry out a pilot project that can later be scaled up to other groups of patients or pathologies.

The intervention evaluated in the Galician testbed takes an approach to overcoming the difficulty of health and social care coverage in a rural region by providing home monitoring equipment to Cystic Fibrosis (CF) patients. This home monitoring equipment enables remote monitoring of CF patients and removes the need to make frequent hospital trips for regular follow-ups. The data obtained with this equipment is linked to the teleassistance platform (TELEA) and to the patient's single electronic medical record. This intervention also aims to empower CF patients to manage and take control over their conditions. TELEA was created in 2012 through an innovative public procurement process with European funds. It is currently integrated into the SERGAS system, integrated with the single electronic health record and is continually updated. SERGAS is responsible for the management and maintenance of this tool.

4.2.2 Innovative Investment Model

The equipment needed to make this telemonitoring possible (e.g. hand-held spirometers, cnoga glucometer, thermometer, weight scales, activity bracelet) is expensive and the Galician Health Service does not cover this cost. That's why we aim to find new finance models that can enable CF patients to have these devices and be included in the telemonitoring program.

Table 6. Overview of Galicia testbed characteristics

Characteristic	Description
Context and intervention	**Health challenge being addressed:** Cystic Fibrosis (CF), a rare, life-limiting condition requiring self-management and family/carer participation. **Description of intervention:** Monitoring of routine health data through wearables/devices that may detect other symptoms that are not easily detectable (e.g., associated pathologies such as mental health, nutrition), These wearables provide live information of existing CF patients to better understand and manage their condition. This timely health information can facilitate where necessary, hospital care at home. Purchase of wearables facilitated through SCI involving CF charity investment.
Sectors involved	**Primary sectors:** Health and Social Care, Private Organisations, Households, Third sector (CF charity). **Other sectors potentially impacted:** Transport, Environment.
Stakeholders involved	Galician healthcare provider (SERGAS), including healthcare professionals, TELEA maintenance and IT professionals; CF Association (patients, carers, families).
Outcomes of interest	**Short-term:** Patient adherence in monitoring, reduce the number and frequency of exacerbations including the need for hospitalisation, improve patient awareness of their own care, improve staff knowledge, increase access to specialist care. **Medium-term:** Reduce the number of (face-to-face) consultations and hospital admissions, reduce risks (cross-infection and multi-resistant bacteria), reduce the number of spirometries done at the hospital by doing them at home, patient empowerment. **Long-term:** Less saturation in health system and more time to attend to those people with major complications, improvement in patient quality of life, avoid cross-infection and be able to predict that a patient is going to have a future exacerbation.
SCI Model	Public-third sector partnership (public healthcare sector working with a Cystic Fibrosis charity). Exploration of investment from philanthropic investor ongoing as of time of deliverable submission.
Investor(s)	Third sector cystic fibrosis charity helping to obtain home monitoring equipment; potential philanthropic donor who would buy the equipment. TELEA was financed by EU regional development funding. The TELEA contractor holds intellectual and industrial property rights over the intervention, however a proportion of revenue is received by SERGAS.
Expected Return on Investment	Social/health (improvement in measurement and monitoring of patients living with CF; improving access to healthcare in a rural region, especially those living in the hardest to reach areas) and financial (long-term reductions in primary and secondary care contacts to monitor conditions resulting in cost savings to the health system).
Data availability	Health measurement data from the wearable is currently manually transcribed to the TELEA platform. Linking of wearable data directly to the platform is being explored to improve data reliability and accuracy.

Table 7. Study design and outcome measure proforma: Galicia, Spain

Study design and outcome measures proposal	
Study design	Pilot patient-level RCT with control and intervention arms. Minimum 6/12-month time horizon of study.
Target Population	Individuals living with Cystic Fibrosis in the Galician region.
Intervention	Intervention arm receiving and using the TELEA home monitoring equipment.
Comparator	Control arm not receiving the equipment.
Perspective of analysis	**Primary perspective:** Health and social care perspective. **Other:** Wider societal perspective.
Outcome measures	**1) Process outcomes** Patience adherence to wearing monitoring devices; improve patient awareness of their own care needs; patient empowerment. **2) Population-health outcomes** Reduce risks (cross-infection and multi-resistant bacteria) associated with visiting hospitals for check-ups; avoid cross-infection. **3) Clinical outcomes** Reduced frequency and incidence of exacerbations; increase ability to predict that a patient is going to have a future exacerbation. **4) Patient-reported Outcomes** Improvement in CF patient quality of life; improvement in mental and physical health. **5) Resource-utilisation outcomes** Reduce the number of (face-to-face) consultations and hospital admissions; reduce the number of spirometries done at the hospital by doing them at home; decreased saturation in health system and more time to attend to those people with major complications. **6) Broader Societal Outcomes** Improved staff skills and knowledge of remote monitoring.

4.3 Region Skåne, Sweden

4.3.1 Context and Intervention

Skåne is the third most populous region of Sweden, home to 1.4 million people. Some of the key health challenges in the region are impaired mental health, obesity, and health inequalities. It is difficult to address these because of the fragmented health and social care system in Sweden along the children's health pathway for mental health. Testbed Skåne is aiming to reduce health inequalities in mental health amongst children through Early Coordinated Interventions (ECI).

4.3.2 Innovative Investment Model

The intervention evaluated in the Skåne testbed is the application of a SCI-based multi-stakeholder governance model or SCI to reinforce, further develop and increase the uptake of already existing Early Coordinated Interventions (ECI). This means health and social care organisations are working more closely together in a more strategic way, which includes the alignment and pooling of resource. These interventions have been in place in a number of municipalities in Sweden for some years, with the aim of improving mental and physical health of children as well as supporting improvement in educational attainment. Most of these interventions are supported locally through shared, pooled resources amongst public sector institutions.

Table 8. Overview of Skåne testbed characteristics

Characteristic	Description
Context and Intervention	**Health challenge being addressed:** mental health amongst children. **Description of intervention:** Increase the uptake of children accessing early coordinated interventions (ECIs) across Skåne (and other primary prevention and health promotion initiatives, many at universal and selective levels). This incentivises and reinforces prevention initiatives and protective factors for mental health among children.
Sectors involved	**Primary sectors:** Local and Regional Administration, Health and Social Care, Education, Households, Voluntary Organisations (NGOs). **Other sectors potentially impacted:** Private sector, University, Employment
Stakeholders involved	Municipalities, primary healthcare providers, schools, children and their families, national/regional administrations and authorities, NGOs. Civil society, police/ law enforcement authorities
Outcomes of interest	**Short term:** Understanding of how to combine a Skåne-shared support function with novel financing mechanism, meeting requests from municipalities of a systematic support that increases the shared readiness and capacity at the shared level, combines methods and financing in a holistic cross-sectoral implementation framework, improved implementation and dissemination capacity, improved coordination between region, municipalities and third sector organisations. **Medium term:** More effective coordination, reinforcement and scaling of interventions and services to children and youth in general (universal) and at risk of mental ill health and anti-social behaviour. Improved early detection, holistic offering of prevention and early intervention services to children and parents, increased rake up/acceptance of support from children/youth and families. **Long term:** Health equality. Improved mental wellbeing among children and young people. Improved school outcomes, improved mental health for risk groups, reduced inflow of children and youth to institutional care, reduced proportion of children and youth with criminal behaviour.
SCI Model	Co-governance of pooled resources on the regional shared level. Support by collaborative platform and improved shared data management. Existing local interventions like Getting It Right for Every Child (GIRFEC) act on a Social Impact Bond-like model. And/or a public-public partnership where resources are pooled across public sector and wider sector organisations. Resources pooled not clearly defined, but likely financial as well as physical and human capital (knowledge transfer, person hours).
Investor(s)	**(Current)** The healthcare landscape in the region is contained under one health area, covering both specialised and primary care. Existing ECI interventions in this area take the form of Social Impact Bond-like interventions and

	interventions based on different types of strategic partnerships between public and private partners. Specific temporary quality-improved funding from national administrations. **(Potential)** combined local/reginal/national funding, public/private combined funding. Shared resources more widely.
Expected Return on Investment	Social/health (improvements in educational attendance and attainment amongst most deprived children in Sweden).
Data availability	**(Current)** National dashboards of child education attainment in different regions. Repeated assessments of SCIROCCO. Regional surveys on children's and young people's self-experienced health. Health records in student health (managed by schools). Health promotion indicators per municipality. Quality registry for student health (partly available). **(Expected)** Aggregated data of existing and specific sources to visualize system characteristics (process data).

Table 9. Study design and outcome measure proforma: Skåne, Sweden.

Study design and outcome measures proposal	
Study design	Systems level case study/ natural experiment. Focusing on governance, inter-organisational collaboration, and shared goals, and shared/integrated resources. Pre-and-post comparator of natural experiment. 1 year time horizon to remain within project scope.
Target Population	Children and young people at risk of developing or who have mental health problems.
Intervention	The closer alignment of public bodies in Region Skåne to work collaboratively and pool resource to facilitate the implementation of ECI interventions that target children and young people in Sweden.
Comparator	Regions that have not pooled resources to implement an ECI.
Perspective of analysis	**Primary perspectives:** Capabilities and capacity in societal, health and social system perspectives.
Outcome measures	**1) Process outcomes** Engagement and willingness to engage of various public bodies; effectiveness of shared governance; proportion of pooled resources invested by each collaborating body. **2) Population-health outcomes** Population level changes in school attainment and attendance; population level changes in child and adolescent mental health. **3) Clinical outcomes** Not recorded. **4) Patient-reported outcomes** Self-reported measure of health and wellbeing gain in children and young people. **5) Resource-utilisation outcomes** Reductions of admissions to mental health care and facilities.

	6) Broader societal outcomes Indicators at system level: minimal requirements on school grades; value-based spare time activities (valued through financial value of leisure time/opportunity cost of leisure time).
Other measure(s) of SCI performance	**Governance**: established strategic and steering groups, performance of those **Business Model**: fund raising (private, national, innovation initiatives) **Data management:** increased use of data sources made available and accessible **Dispatching resources/competencies/knowledge**

4.4 North Rhine Westphalia, Germany

4.4.1. Context and Intervention

North Rhine-Westphalia is the most populous state of Germany, with a population of 18 million people. The state has a significant proportion of older adults, with a growing number of people reaching retirement age. Since the proportion of people that are active in the workforce is decreasing and the region is already short of staff, it is important to keep them as healthy as possible. Especially when people are working for longer. Major cities like Düsseldorf, Cologne, and Bonn tend to have higher average incomes and a higher standard of living. Musculoskeletal disorders accounted for more than one-fifth of absences from work (22%), in all of Germany in 2021, followed by mental disorders (12%), injuries (10%), respiratory diseases (10%), and diseases of the circulatory system and digestive organs (5% and 4%, respectively). Amongst home workers, common conditions arising from poor working conditions and equipment include back pain, carpel tunnel syndrome, repetitive strain injuries and mental health. These types of illnesses often result in long periods of absence from work.

4.4.2 Innovative Investment Model

The intervention evaluated in the North Rhine Westphalia testbed is a subscription-based fee-for-service model to access a bundle containing ergonomic home working equipment, information, and fitness classes. This equipment and classes are available as individual items, yet the bundle offers a comprehensive package at a small cost to the employer/employee through a subsidised contribution. The bundle can be built upon, scaled up and rolled out further. Issues arise when home offices are not fit for purpose (not ergonomically sufficient) and where working practices are too sedentary. Both are risk factors for future ill health. This intervention brings previously disparate aspects surrounding home working in a novel bundled package.

Table 10. Overview of North Rhine Westphalia testbed characteristic

Characteristic	Description
Context and intervention	**Health challenge being addressed:** Musculoskeletal issues affecting the working age population. Musculoskeletal conditions represent the largest proportion of ill-health related sickness absences in the region. **Description of intervention:** Subscription-based fee-for-service model to access a bundle containing ergonomic working equipment, information and fitness classes. The bundle can be built upon, scaled up and rolled out further. The intervention targets an improved ergonomic environment in the home office to prevent the development of associated conditions where sedentary behaviour and poor working conditions are risk factors.
Sectors involved	**Primary sectors:** Local government, private-for-profit organisations, households. **Other sectors potentially impacted:** Health and social care, Transport.
Stakeholders involved	Local authority, municipalities, employers, equipment providers and manufacturers, insurance providers, local physical activity providers/gyms and employees
Outcomes of interest	**Short term:** Improved ergonomic environment in the home office, employee satisfaction, raise awareness/education of ergonomic working practice and behaviour change. **Medium term:** Reduced absenteeism, improvement in employee health. **Long term**: Increased productivity and reduction in associated conditions where sedentary behaviour and poor working conditions are risk factors.
SCI Model	Bundled service provision funded through private-private partnerships and accessed via a fee-for-service model by employers and employees.
Investor(s)	Multiple funding sources included in the business case. Including a fee from the employee and employers for the subscription; investment from insurance companies, municipalities, and investors. This investment is used to fund the interventions included in the subscription package (software, hardware, fitness centres and activity groups)
Expected Return on Investment	Financial (fee-for-service contribution to access bundle will return some of initial investment over a period, long-term reductions in absenteeism and increasing productivity to have positive financial returns to employers).
Data availability	*Recording of data still to be confirmed by NRW testbed.*

Table 11. Study design and outcome measure proforma: Germany.

Study design and outcome measures proposal	
Study design	Matched control experiment across NRW region with intervention (receiving intervention) and control (not receiving intervention) arms. To run for a minimum of 12 months.
Target Population	Individuals who work from home/hybrid workers in the NRW region.
Intervention	Fee-for-service bundle containing ergonomic home working equipment, information, and fitness classes targeting reduction and prevention of injuries leading to time off work in home workers.
Comparator	Individuals who work from home but are not receiving the bundle subscription.

Perspective of analysis	**Primary perspective:** Employer perspective. **Other:** Societal perspective
Outcome measures	**1) Process outcomes** Employee satisfaction with working environment **2) Population-health outcomes** Reduction in prevalence of associated conditions where sedentary behaviour and poor working conditions are risk factors. **3) Clinical outcomes** Not recorded. **4) Patient-reported outcomes** Improved employee health and wellbeing. **5) Resource-utilisation outcomes** Reduction in healthcare utilisation through reduction in prevalence of associated conditions. **6) Broader societal outcomes** Increased productivity; reduced absenteeism; raise awareness of ergonomic working practice and influence behaviour change.

4.5 Overarching approach to sensitivity analysis

Sensitivity analysis is an important part of the evaluation process and gives valuable information to decision-makers about the robustness of their decision based on the findings of an economic evaluation, as well as the potential value of collecting more information before making a decision (York Health Economics Consortium, 2016). Our proposed recommendations for each of the study designs listed above is to apply sensitivity analysis to key assumptions. This determines how sensitive our results are to key assumptions, in order to deal with and quantity endogenous (local economy) and exogenous (national prevention policy landscape) uncertainty.

4.6. Overarching approach to addressing confounders and reducing bias

It is important for the testbeds to control for confounders to mitigate potential biases and produce more accurate and trustworthy results. For example, in the Germany testbed, there would need to be a control for individuals who are doing things on their own (i.e., attend own fitness classes or have purchased own ergonomic equipment) or have another source of support. In this case, and others, there would need to be a means of collecting such information from both the intervention and comparator group, if feasible to ensure biases are accounted for and dealt with in the analysis.

5. Outcome measures

In formulating our proposed outcome measures, study designs and evaluative frameworks, we draw upon evidence from the published and grey literature identified in the scoping reviews of T2.1 and T4.3. The findings of these reviews are presented thematically below.

T4.3 delivered a scoping review of existing frameworks for cross-sectoral evaluation of healthcare interventions and provided an understanding of how they have been applied and operationalised.

5.1 The components of SCI outcome measures

In developing our proposed outcome measures for evaluating the real-world impact of SCI, we considered the appropriate approaches to measurement (as illustrated in Figure 3). First, we will evaluate the outcomes of the disease prevention/health promoting activity on end users. This can take the form of outcomes such as health gain and quality of life. These outcomes most closely map to the outcomes collected as part of the scoping review of SCI examples in T2.1, but appropriate measures identified in T4.3 will also be presented here. Secondly, we will consider outcomes to evaluate the impact of the SCI mechanism itself (the success of the business model, governance, and ability to engage stakeholders/investors). For example, these outcomes are reflected in our discussions with expert stakeholders as part of WP4 and our case study with colleagues in NHS England. The image below (Figure 3) shows how we have attempted to demonstrate disentangling the impact of the intervention from the impact of the investment model by using different types of outcomes. However, it is important to note that the combination of the two components (intervention arm and investment arm) can impact many different categories of outcomes, including outcomes beyond the healthcare sector (for example broader societal benefits from higher productivity in paid or unpaid labour), and these are not represented in this image. A proposed evaluative checklist for the testbeds based on these two components is presented in Annex 4.

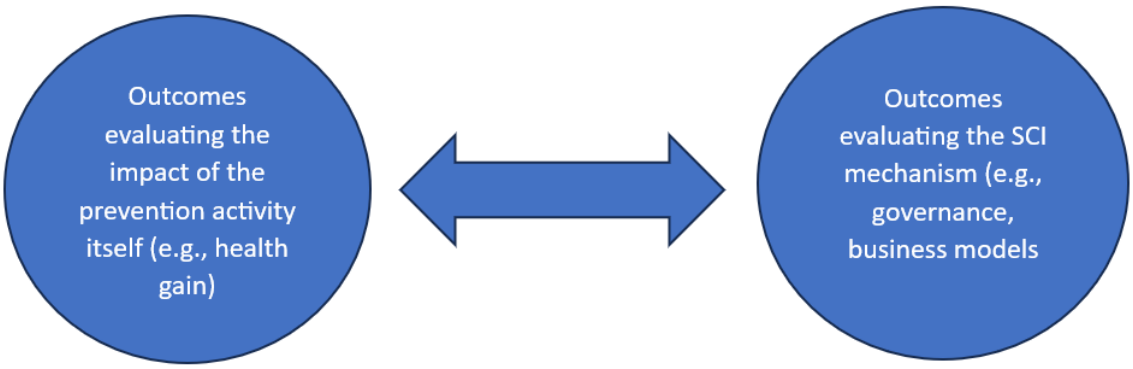


Figure 3. The components in evaluating the real-world impact of SCI

5.2 Outcomes evaluating the impact of the prevention activity

When presenting outcomes in the testbed proformas (Section 3), we listed outcomes. Our scoping review of SCI examples comprehensively extracted data from identified articles across several domains of interest, including outcomes. Outcomes evaluating the healthcare activities in the review of multi-sectoral evaluations were also collected. Information on outcomes, how they are measured, monetised and valuated can be found in Table 12 following the framework developed by Drummond and colleagues (Drummond et al., 2015).

Across included articles in T2.1, outcome measurements were poorly reported. Where reported, the most common outcome measures (Table 12) used to monitor the impact of the disease preventing/health promoting activity were general quality of life and well-being-related outcomes, behaviour-change-related outcomes (e.g., physical activity levels), process outcomes (e.g., number of participants, patient satisfaction) and multiple outcomes (mixture of health outcome and process measures) - mirroring the generic nature of the health and well-being goals targeted.

Table 12. Outcomes evaluating the impact of the prevention activity itself from T2.1

Outcome	Measurement	Monetisation/ valuation
Generic quality of life [Quality Adjusted Life Years (QALYs)/Disability Adjusted Life Years (DALYs)]	EuroQOL 5 dimensions (EQ-5D); General Anxiety Disorder-7 (GAD-7); Patient Health Questionnaire-9 (PHQ-9); Lifetable estimation; utility evaluation from the literature	Estimation using collected health resource use and valuation by unit costing; assumption of health resource use from literature; estimation of productivity impacts from data
Health and well-being	UCLA well-being scale; bespoke online well-being questionnaire; well-being self-assessment tool; Warwick-Edinburgh Mental Wellbeing Scale (WEMWEBS); Clinical Outcomes in Routine Evaluation-10 (CORE-10), Örebro Musculoskeletal Pain Screening Questionnaire ÖMPSQ 10), KINDL, Pediatric Quality of Life Inventory (PedSQL) and Satisfaction with Life Scale (SWLS), PAM score, and ICEpop CAPability measure for Adults (ICECAP-A).	Estimation using collected health resource use and valuation by unit costing; assumption and valuation of health resource use from literature; estimation of pupil attainment from the literature and costed by unit costing
Improved mental health	Self-designed survey based on EQ-5D, GAD-7, or Short Form Survey-12 (SF-12)	Not defined in literature
Independence, resilience and social isolation	TBS Resilience Outcomes Tool; semi structured interviews	Not defined in literature

Behavior change	Interviews; International Physical Activity Questionnaire (IPAQ)	Not defined in literature
Increase in physical activity	Accelerometers; logs of physical activity; International Physical Activity Questionnaire (IPAQ); GAD-7, IPAQ-short, PHQ-9	Estimation using collected health resource use and unit costing
Clinical outcomes (HbA1C, BMI)	Clinical measurement; interviews; clinical tests	Not defined in literature
Multiple outcomes (mixture of health outcome and process measures)	Surveys across multiple dimensions	Not defined in literature
Process Outcomes (Averted hospitalisations, cases, disease, and death; participants perception of activity)	Routine healthcare data on persons treated, appointments avoided; interviews; data on engagement with services; questionnaires	Estimation using collected health resource use and valuation by unit costing; assumption and valuation of health resource use from literature

5.3 Measurement and valuation/monetisation of outcomes evaluating the impact of the prevention activity itself

Of the outcomes evaluating the impact of the prevention activity itself identified in T2.1, most were measured through administering of a questionnaire to participants. The questionnaires completed differed by the outcome measure of interest. Interviews were used to measure a number of outcomes, in most cases in conjunction with some other form of measurement such as questionnaires or clinical assessments. Where multiple outcomes were grouped within an SCI example publication, these were often measured through surveys covering several outcome domains in one measure. Process outcomes (e.g., number of participants and patient satisfaction) were measured typically by routine data on contacts with the activity and wider healthcare resource contacts. Captured through published routine data or via questionnaire. Not all identified SCI examples were set within economic evaluations, so not all outcomes were valuated or monetised. To supplement this information, outcomes evaluating the impact of the health intervention identified in T4.3 have been combined in Table 12 to add greater depth to the findings.

The most common outcomes evaluating the impact of the health intervention identified in the review of multi-sectoral health evaluations of health interventions were the quality-adjusted life years (QALYs) or disability-adjusted life years (DALYs). QALYs and DALYs were typically measured through the EQ-5D questionnaire or estimated through lifetables and published utilities. QALYs and DALYs were monetised/valued through valuing collected health resource (or health resource

use estimates from the literature) by unit costing or estimating and valuing productivity impacts from QALY/DALY utility values.

The second most commonly utilised outcome was well-being. Well-being encompassed pain, suffering, stress, happiness, confidence, life control, relationships, maintaining dignity, or sense of self. Well-being outcomes were commonly measured through questionnaires, such as the CORE-10, ÖMPSQ 10, PAM score, and ICECAP-A. Well-being measures were monetised/valuated in a similar way to QALYs/DALYs by valuing collected health resource (or health resource use estimates from the literature) by unit costing. One valuation of well-being measurement estimated pupil attainment from the literature as a result of improved well-being and monetised these through unit costing.

Improved physical and mental health formed two other areas of outcome measure identified by the review. For physical health the outcomes were measured typically by questionnaires such as the GAD-7, IPAQ-short, International Physical Activity Questionnaire Item, and PHQ-9. Physical health outcomes were valued through collected health resource use and monetised through unit costing. Improved mental health had significant overlap with well-being measures, using similar measurement tools as the ones listed above as well as domain-specific quality of life tools for mental health (a self-designed survey based on EQ-5D, GAD-7/General Anxiety Disorder-7, or SF-12. No description of how mental health outcomes were monetised or valued was identified in the literature.

Outcomes regarding hospitalisations, cases, disease and deaths averted were measure through published statistics and within the literature. Such outcomes were valuated through collected health resource (or health resource use estimates from the literature) and monetised by unit costing.

5.4 Evaluating the SCI mechanism

As explored through T2.1, various models of SCI can be classified within a typology based on several categories such as their scope (primary, secondary, or tertiary prevention); their partnership (investors, commissioners, service providers, beneficiaries, and outcome assessors) and their investment contract (type and size of investment, whether expected return is financial or non-financial). These various models of SCI each possess unique characteristics. This inherent heterogeneity can pose difficulty in assigning outcome measures evaluating the SCI mechanism. The SCI mechanism can be defined as the achievement of pre-determined goals set out by all stakeholders, over an appropriate and pre-determined time period. In this section we draw upon evidence considering the governance, sustainability, and the perspectives of investors in SCIs from T2.1, T4.1/T4.2 and an expert interview with the Chief Strategy and Collaboration Officer at West Herts Teaching Hospitals NHS Trust. These elements are prerequisites for its successful implementation and should be monitored.

Table 13. Implementation features to monitor when evaluating the SCI mechanism.

Implementation features	Measured by
Service Sustainability	Financial sustainability (evergreen fund), organisational sustainability (staff retention)
Governance	Not yet defined
Scalability	Success of pilot study, evidence of scaling plan
Business model	Success of obtaining novel investment, interest in business model
Impact measurement	Health and well-being, social impact, typically part of an evaluation
Motivation for investor	Expected return to investor in contractual discussions: no returns, preferential return, non-preferential returns, non-financial investment, and other models
Investment ecosystem	Not yet defined
Regulatory issues	Political buy-in, health system architecture

5.5 Outcomes evaluating cross-sectoral impacts

Outcomes evaluating the healthcare activities in the review of multi-sectoral evaluations (T4.3) are presented below (Table 14). The table presents the various outcomes, the ways in which they were measured and their impact in terms of monetisation or valuation, following the framework of Drummond et al. (2015).

Table 14. Outcomes evaluating cross sectoral impacts.

Outcome measure	Measurement	Monetisation/valuation
Productivity (absenteeism, presenteeism, unpaid work and informal care)	Statistical data, including the Bureau of Labour Statistics, Census data, OECD labour force participation rate	OECD average wage data
Transport	Utilisation of transport via questionnaire	Expert opinion, statistical data
Social services (social service utilisation, care provision, counselling)	Service utilisation data, routine data, NHS data, CORE counselling questionnaire, end of therapy questionnaire	Unit costing, assumptions based on published costs

Out-of-pocket expenses (personal expenses	Questionnaire	Estimation on questionnaire responses
Education (therapy, special education, home education, tutoring, counselling, support to teachers/students, financing students)	Questionnaire data, estimates from the literature	Estimation of costs through unit costs and prices of additional support
Housing (living support, usual living situation)	Questionnaire data	Not yet defined
Medication	Routine health service data	Unit costs
Economy and business (impact on local economy, cancelled events, tourism disruption)	Not yet defined	Not yet defined
Justice (reduction in crime)	Crime statistics	Not yet defined
Environment (water and sewage,	Not yet defined	Not yet defined

Measurement and valuation of resource use in different sectors

Of all costs identified, direct healthcare service use was the most common, appearing in 85% of included studies. The second most common cost that was measured was productivity found in 73% of all studies. Typically, this was measured by statistical data, including the Bureau of Labour Statistics, Census data, OECD labour force participation rate, or OECD average wage data. The descending frequency of other costs were transport (25%), social services (17%,), out-of-pocket expenses (13%), education (13%), housing (8%), medication (6%), economy and business (6%), justice (6%) and environment (4%).

5.6 Evaluative frameworks

The most used method of evaluation identified in the T4.3 review of cross-sectoral evaluations was cost-effectiveness analysis (CEA) (Table 15). CEA is a form of economic evaluation that estimates the associated cost per unit of pre-defined outcome achieved and compares the different alternatives for the same condition (Office for Health Improvement and Disparities, 2020a). The next most common evaluative framework identified was social return on investment (SROI). SROI is another form of economic evaluation that can quantify social, environmental as well as economic outcomes of a given intervention. The main outcomes of SROI evaluations are SROI ratios presenting the financial benefit of the quantified outcomes for every unit invested (Olsen & Nicholls, 2005). Cost-benefit analysis (CBA) was also identified in the review. CBA is another form of economic evaluation that is employed to present and compare the costs and

effects of alternate interventions, both in monetary terms (Office for Health Improvement and Disparities, 2020b). The final evaluative framework identified was cost studies. Cost studies reflect a partial economic evaluation (in that they only consider the costs of interventions and not also measuring outcomes).

Table 15. Evaluative framework methods and frequency of use in cross-sectoral evaluations identified in T4.3

Evaluative Framework	Task 4.3 (Scoping review of cross-sectoral evaluations)
Cost-effectiveness analysis (CEA)	(n = 15, 31%)
Social Return on Investment (SROI)	(n = 13, 27%)
Costing study	(n = 8, 17%)
Cost-benefit analysis (CBA)	(n = 12, 25%)

The included papers mostly considered healthcare and workforce sectors combined in the analysis (35%), or an approach considering multi-sectoral analysis of costs and outcomes (i.e., including three of more sectors) (31%). Social care, education, and workforce were covered by 4% of studies each. Within studies that focused on healthcare, the most common additional sector was workforce, using different methods to measure and monetise productivity (82%). Other additional sectors were transport (23%), social care (15%), education (13%), then justice (5%) and environment (5%). A breakdown of lead sectors involved in both SCI examples (T2.1) and cross-sectoral evaluations of healthcare interventions (T4.3) is available in Annex 3.

7 Discussion and next steps

Here we have generated and report on study designs for the future evaluation of the real-world impact of SCI models. We have developed proposals for testbed evaluation (Tables 7, 9, 11, and 13) as well as an overarching guiding checklist (Table 20). These tables can be used in the second tranche of testbeds. We built on previous experience and success in the SELFIE study which presented and defined a range of study designs and evaluative approaches, and the I4H scoping reviews from T2.1 and T4.3.

The presentation of outcomes, study designs and approaches, and evaluative frameworks for evaluations of cross-sectoral impacts provides the groundwork for subsequent formalised study protocols in each testbed to assess the real-world impact of SCIs. This work also informs a formal MCDA task in WP4 that will aggregate the benefits across the sectors identified. Testbed partners will continue to iteratively refine their studies using the work set out in this deliverable and subsequent work in WP4. Testbeds will need to become more specific about which resources from whom will be pooled and what is the repurposing of these resources. In doing so, they will contribute to production of a more representative picture of their respective funding landscapes both now and in the future and how this can be orientated towards prevention.

8 References

Campbell, M., Fitzpatrick, R., Haines, A., Kinmonth, A. L., Sandercock, P., Spiegelhalter, D., & Tyrer, P. (2000). Framework for design and evaluation of complex interventions to improve health. *Bmj*, *321*(7262), 694-696. https://doi.org/10.1136/bmj.321.7262.694

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*, *337*, a1655. https://doi.org/10.1136/bmj.a1655

Craig, P., Katikireddi, S. V., Leyland, A., & Popham, F. (2017). Natural experiments: An overview of methods, approaches, and contributions to public health intervention research*. Annual Review of Public Health, 38*, 39–56. https://doi.org/10.1146/annurev-publhealth-031816-044327

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the economic evaluation of health care programmes* (4th ed.). Oxford University Press.

Lawson, K., Mason, H., McIntosh, E., & Donaldson, C. (2014). Priority setting in public health. *Encyclopedia of Health Economics,* 155-162. https://doi.org/10.1016/B978-0-12-375678-7.00410-7.

Moore, G., Audrey, S., Barker, M., Bond, L., Bonell, C., Cooper, C., Hardeman, W., Moore, L., O'Cathain, A., Tinati, T., Wight, D., & Baird, J. (2014). Process evaluation in complex public health intervention studies: the need for guidance. *Journal of Epidemiology & Community Health*, *68*(2), 101-102. https://doi.org/10.1136/jech-2013-202869

Office for Health Improvement and Disparities. (2020a). *Guidance: Cost-effectiveness analysis: health economic studies.* UK Government. https://www.gov.uk/guidance/cost-effectiveness-analysis-health-economic-studies

Office for Health Improvement and Disparities. (2020b). *Guidance: Cost benefit analysis: health economic studies*. https://www.gov.uk/guidance/cost-benefit-analysis-health-economic-studies

Olsen, S., & Nicholls, J. (2005, May). *A framework for approaches to SROI analysis*. https://apsocialfinance.wordpress.com/wp-content/uploads/2013/01/2005-framework-for-approaches-to-sroi-analysis.pdf

Skivington, K., Matthews, L., Simpson, S. A., Craig, P., Baird, J., Blazeby, J. M., Boyd, K. A., Craig, N., French, D. P., McIntosh, E., Petticrew, M., Rycroft-Malone, J., White, M., & Moore, L. (2021). A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ*, *374*, n2061. https://doi.org/10.1136/bmj.n2061

Walker, S., Griffin, S., Asaria, M., Tsuchiya, A., & Sculpher, M. (2019). Striving for a societal perspective: a framework for economic evaluations when costs and effects fall on multiple sectors and decision makers. *Applied Health Economics and Health Policy*, *17*, 577-590. https://doi.org/10.1007/s40258-019-00481-8

Walker, S., Griffin, S., Asaria, M., Tsuchiya, A., & Sculpher, M. (2019). Striving for a societal perspective: a framework for economic evaluations when costs and effects fall on multiple sectors

and decision makers. *Applied Health Economics and Health Policy*, *17*, 577-590. https://doi.org/10.1007/s40258-019-00481-8

Weatherly, H., Drummond, M., Claxton, K., Cookson, R., Ferguson, B., Godfrey, C., Rice, N., Sculpher, M., & Sowden, A. (2009). Methods for assessing the cost-effectiveness of public health interventions: key challenges and recommendations. *Health Policy, 93*(2-3), 85-92. https://doi.org/10.1016/j.healthpol.2009.07.012

Webber, S., & Prouse, C. (2018). The new gold standard: The rise of randomized control trials and experimental development. *Economic Geography*, *94*(2), 166-187. https://doi.org/10.1080/00130095.2017.1392235

York Health Economics Consortium. (2016). *Perspective*. University of York. https://yhec.co.uk/glossary/perspective/

ANNEX

Annex 1. Overview of available study designs

There are three types of study designs, **experimental designs**, **quasi-experimental designs**, and **observational designs**, and below we include examples of relevant research methods. This list is organised by ease of attributing causality in descending order. This list is repurposed from the "Sustainable intEgrated care modeLs for multi-morbidity: delivery, FInancing and performancE" (SELFIE) study.

Experimental Designs

Experimental methods involve the process of conducting controlled research to reach specific conclusions regarding a hypothesised statement. They involve random allocation of participants to different groups in an experiment, and groups are followed parallel to one another. Experimental designs are chosen because they offer a more robust design to infer causality. It is important to note the issues related to transferability of findings into real-life practice.

Examples of experimental designs include a randomised controlled trial (RCT), which controls factors not under direct experimental control, and a Cluster RCT in which groups of subjects are randomised.

Table 16. Features of experimental designs.

Randomised Controlled Trials (RCT)	• Patient-level randomisation to exposed (e.g. integrated care) and unexposed (e.g. usual care) groups.
	• Prospective parallel-group comparison between exposed and unexposed groups.
	• RCTs are not possible or desirable:
	  ○ in complex interventions when controlling all factors/context is impossible and contamination between groups is likely;
	  ○ if randomisation (withholding an intervention in the control group) is considered unethical;
	  ○ if population in RCT not representative for real-world population.
	• A trade-off needs to be made between internal and external validity.

Cluster RCT	• Group-level randomisation to exposed (e.g. integrated care) and unexposed (e.g. usual care) groups. Clusters of patients in a group (e.g., neighbourhood, department, GP Practice) are randomised. • Randomising a group (instead of an individual) prevents contamination (how persons undergoing the intervention influence those not undergoing the intervention) and increases the logistic feasibility of implementing the intervention. • The power of the design is determined by the number of clusters. If too few clusters are included, controlling by chance for all factors that might differ between groups is impeded and internal validity is compromised.

Quasi-experimental designs

Empirical research designs that aim to establish a cause-and-effect relationship between an independent and dependent variable. Quasi-experiments do not rely on random assignments. Examples include a **stepped-wedge design** which is a type of RCT; **propensity score matching** which is a statistical matching technique to estimate the impact of an intervention; **instrumental variables** which estimate causal relationships when controlled experiments are not feasible; **regression discontinuity** which is a pretest–post-test design to evaluate programmes that have a cutoff point for participation; and **difference in difference** which estimate the effect of a treatment or an outcome using observational data.

Table 17. Features of quasi-experimental designs.

Stepped-Wedge Design	• The stepped wedge randomised trial is a modification of the individual or cluster RCT in which the exposure (e.g., intervention) is sequentially rolled-out to all subjects over consecutive time periods. • Selection bias is prevented by randomising the order by which subjects receive the exposure (e.g., intervention). • Because all subjects receive the intervention ethical issues of withholding the intervention are solved. • Changes in intervention based on lessons learned in previous step are possible before the next step. • Possible measured effects: both short and long-term effects, fade out effects, and the natural course of the condition under study. • Larger sample sizes are needed for this method to ensure enough statistical power. • There may be a higher burden on participants and researchers due to necessity to repeatedly collect data. Thus, the design is most feasible if data can be (partly) routinely collected at the appropriate time intervals in a reliable and valid way. • Statistical analysis is more complex since it is necessary to include a random coefficient for cluster and a fixed effect coefficient for time.

Propensity Score Matching (PSM)	• Propensity scores can be estimated using logistic regression analyses modelling the exposure as the dependent variable and the potential determinants of being exposed as independent variables.
	• In PSM the average characteristics of the exposed versus unexposed group are similar;
	  o Nearest neighbour matching: matches an unexposed individual with the closest propensity score (PS) of an exposed individual.
	  o Nearest neighbour matching within a specified caliper distance.
	  o Kernel weighting: each individual in the exposed group gets a weight of 1 and each unexposed patient gets a weight that depends on the distance of their PS from the PS of the exposed patients. Higher weights represent better matches.
	• Inverse probability weighting using propensity score:
	  o Uses weights based on the propensity score to create a synthetic sample in which the distribution of measured baseline covariates is independent of treatment assignment.
	  o A subject's weight is equal to the inverse of the probability of receiving the treatment that the subject received.

Instrumental Variable	• In this method, it is hypothesised that an instrumental variable (IV) can 'randomise' by adjusting for measured and unmeasured confounding. • This IV should be a proxy for the exposure (e.g., intervention), but cannot have a direct association to the outcome (only via the exposure), and also cannot be associated with unmeasured confounders. • The ratio of the effect of the IV on the outcome on the one hand, and of the instrumental variable on the exposure on the other hand, shows the true effect of the exposure on the outcome. • Choosing the correct instrumental variable is a challenge! • This method is often used in trials when there is non-compliance and/or for intention to treat analyses.
Regression Discontinuity Design (RDD)	• This method can be used when there is some kind of criterion that must be met before subjects can be exposed (i.e. included in an intervention). • In RDD, the exposure (e.g., intervention group) is assigned to subjects that score [just] above a certain 'cut-off point' (on a continuous variable); and the unexposure (e.g., control group) is made up of individuals that score [just] below that cut-off point. • The main assumption in RDD is that individuals that just score on either side of the 'cut-off point' or threshold, belong to the same population. Allocation of these subjects to either exposed or unexposed is therefore considered 'random' – assuming that subjects cannot manipulate the threshold value. • Intervention effects are estimated by comparing the outcomes of the group that meets the criterion with the group that just not meets the criterion using regression techniques: non-parametric (local linear regression) or parametric (polynomial regression). • Assumption that individuals around threshold are similar is often debatable. • The design requires larger sample sizes than an RCT to achieve sufficient statistical power. • Data collection required for all individuals considered for intervention (also those who are not included in the intervention).

Difference in Differences	• Also known as 'double difference' method, compares changes in the outcome over time between two groups (i.e., exposed and unexposed) to estimate an effect. This way the difference at baseline between the groups is 'removed'. • It assumes that the outcomes of interest in the unexposed group follow the same trend over time as the outcomes would do in the exposed group in the absence of the intervention; thus that there is no confounder that influences one group and not the other. • The assumption of parallel trends between intervention and control groups can be relaxed in a differential trend model.

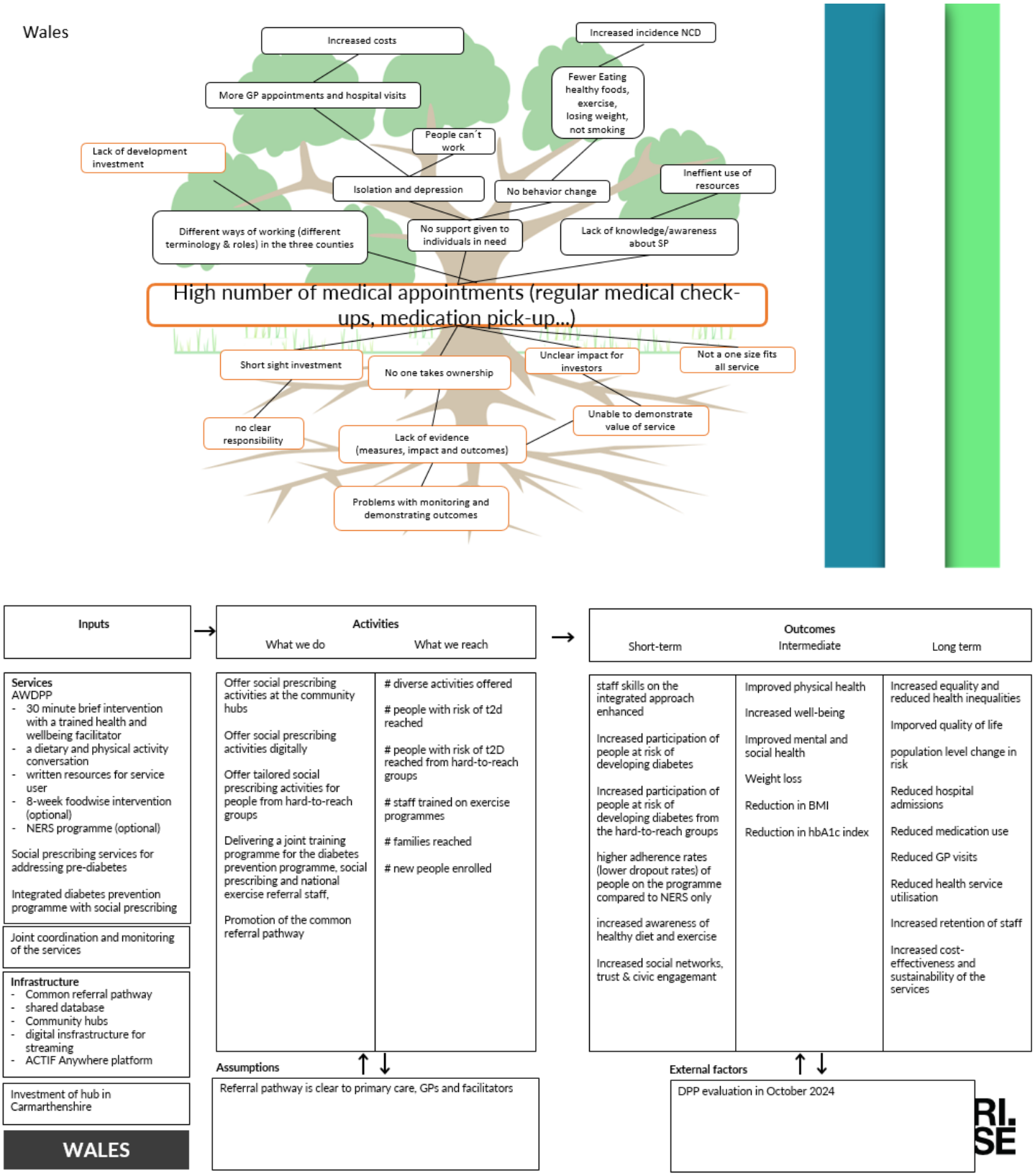
Observational Designs

Table 18. Features of observational design

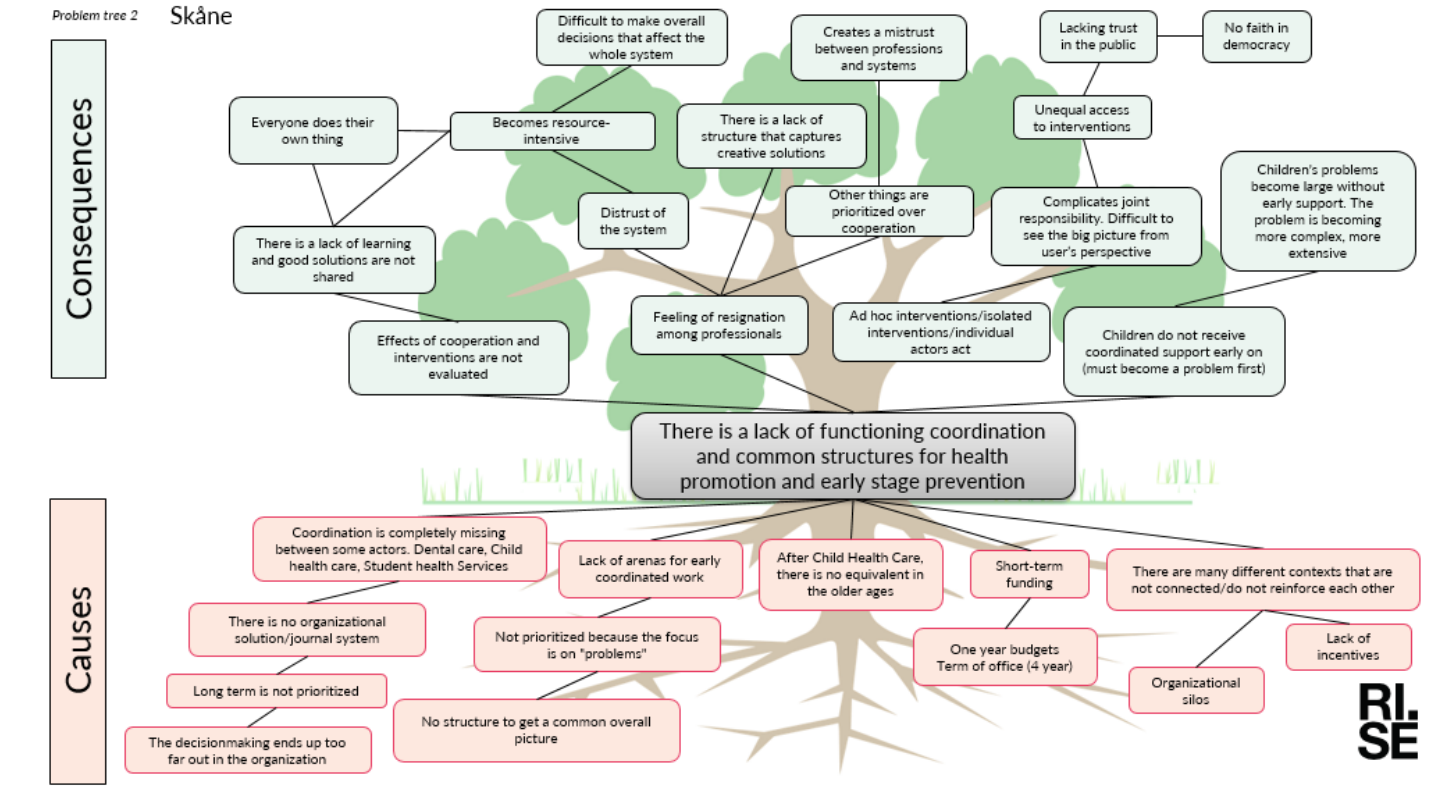
Observational studies look at the effect of an intervention without manipulating who is or is not exposed to it. It answers the research question on what the researcher observes and there is no interference. Observational designs include **a cohort study** which samples a group of people who share a defining characteristic or a common event; **a case control study** in which groups of participants with and without a condition of interest are identified and compared; and **interrupted time series analysis** which is a method of statistical analysis involving tracking a long-term period before and after a point of intervention to assess the intervention's effectiveness.

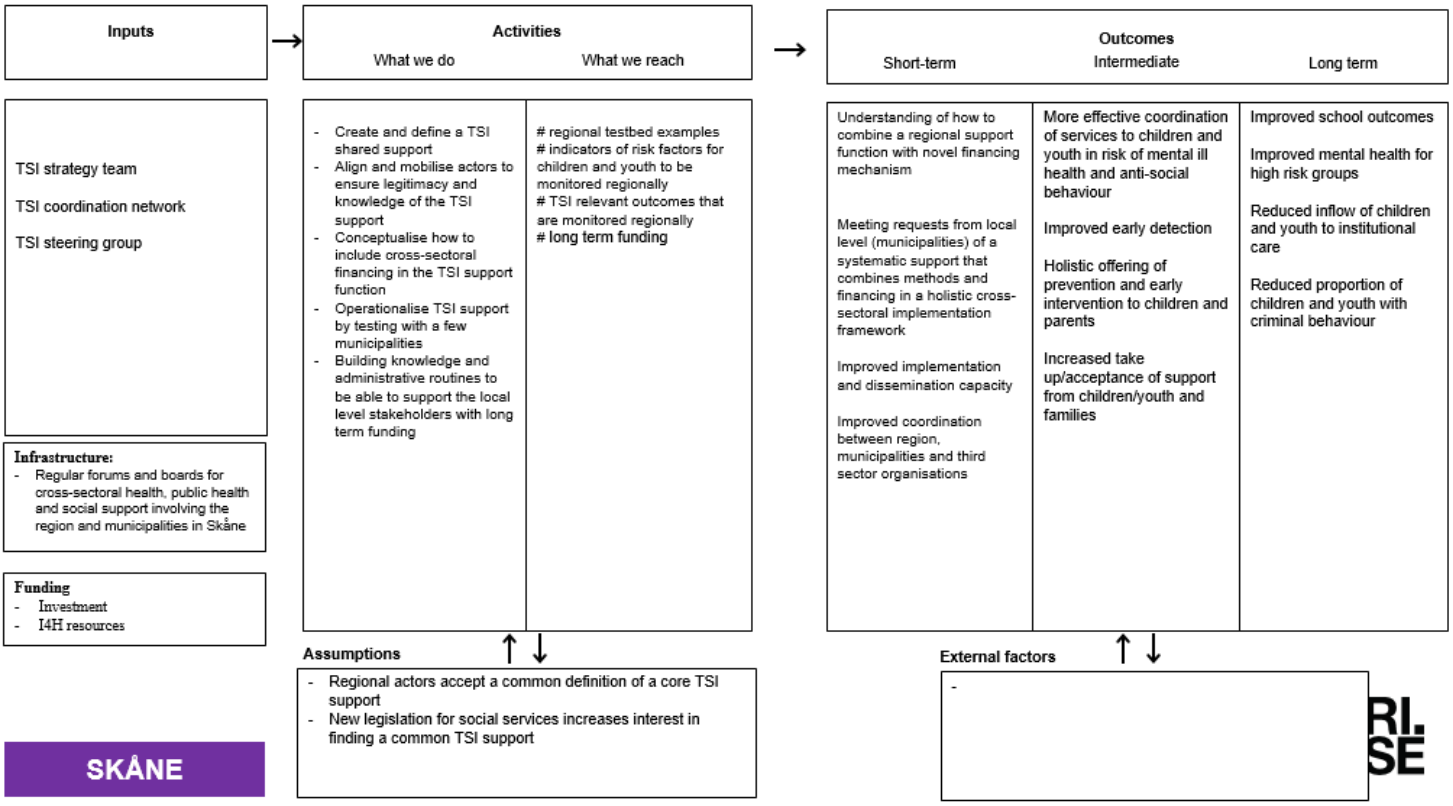
Cohort Study	• The starting point in such studies is the selection of a study population, or a cohort study. • Information is obtained to determine which persons in this cohort are exposed to a specific factor, and which are not (e.g., which persons participated in an integrated care programme, and which did not). • Allocation of participants to either group (exposed vs. unexposed) is not influenced by the investigator. • Outcomes can be collected prospectively or retrospectively.
Case Control Study	• The starting point here is the identification of 'cases' and 'controls' (outcome present or not present). • Then go back and look at which cases and which controls were exposed vs. unexposed (i.e., participated in an integrated care programme versus usual care). • There is also the possibility to do a nested case-control study, in which a subset of controls are matched to cases. X controls are matched to a single case based on certain characteristics, making this a more efficient model
Interrupted Time Series	• Repeated measurements are performed before and after exposure at population level in order to detect whether the intervention has a greater effect than the underlying secular trend (e.g., economic, market, or demographic trend) . • Determining whether the intervention had a larger effect than any underlying trend is estimated by comparing the trend in the outcome after the intervention to the trend in the pre-intervention period. • Relevant design when using routinely collected data (e.g., insurance data) and when looking at macro interventions. • Pros: allows for control over secular trends, ability to evaluate outcomes at population-level, clear graphical presentation of results (see next slide), ability to do stratified analyses and look at intended and unintended consequences of interventions. • Cons: multiple measurement waves are needed (circa 8 before and after exposure (i.e., implementation), had to look at specific elements of exposure.

Annex 2: Testbed problem trees and logic models

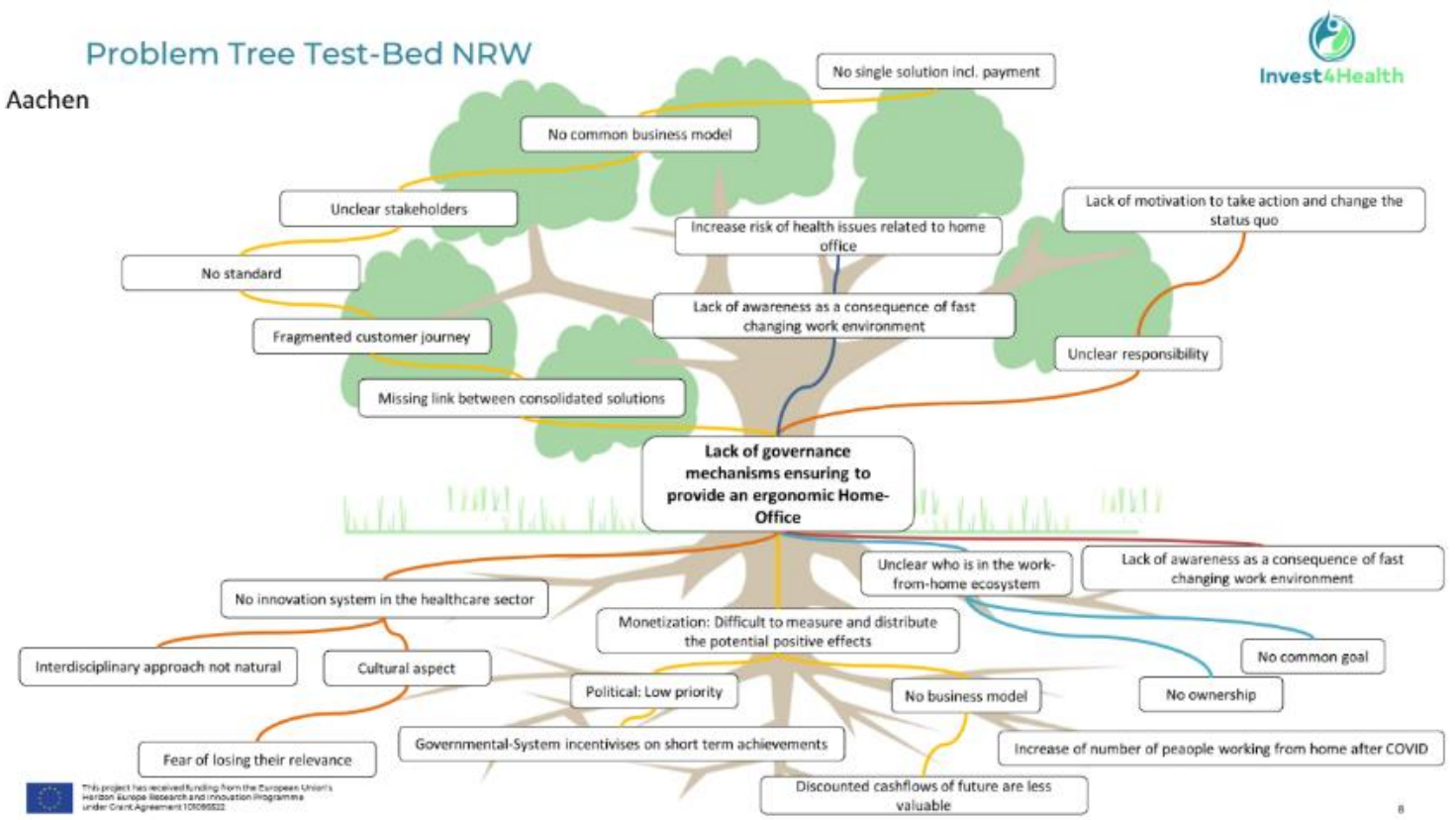


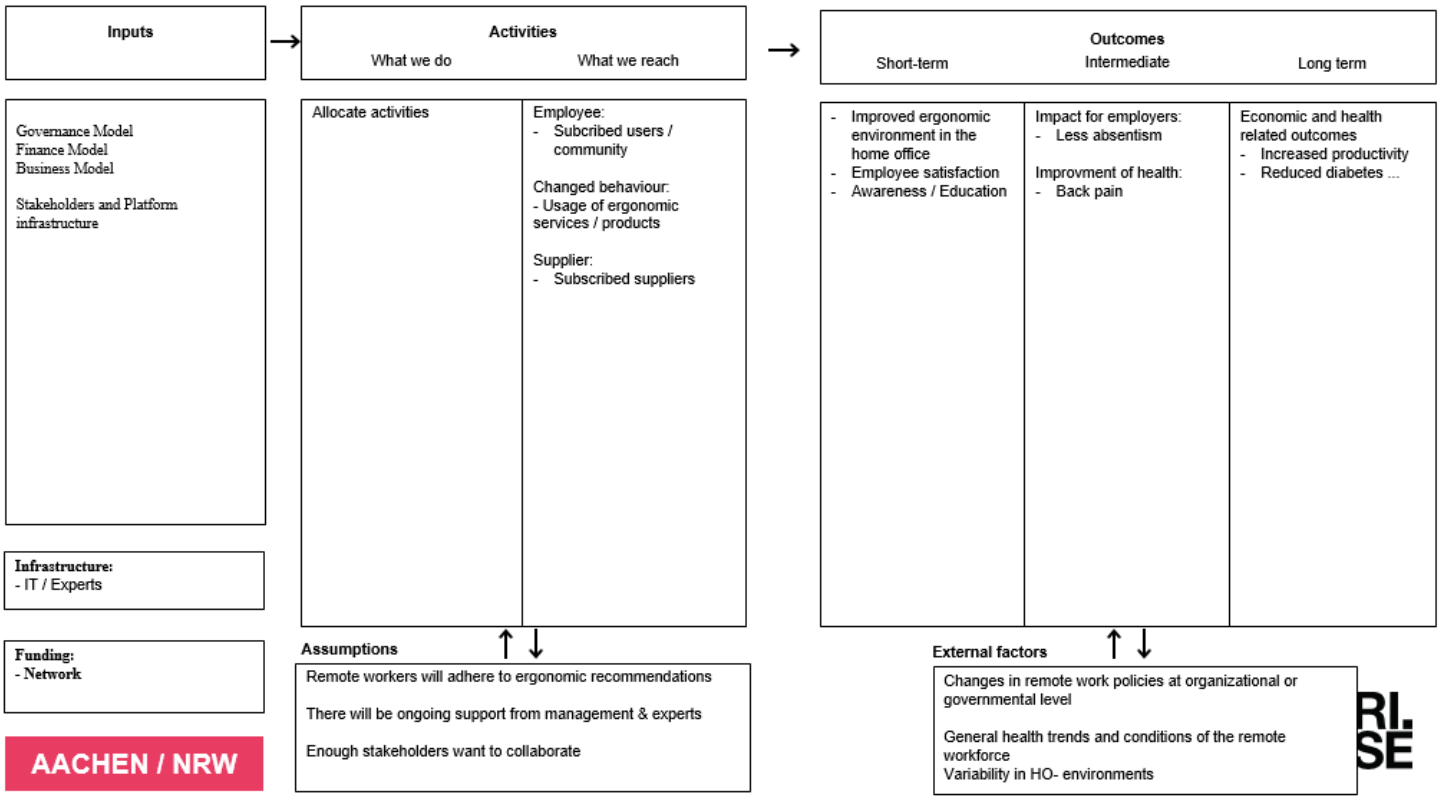
Figures 4, 5: Hywel Dda (Wales) testbed problem tree and Logic model



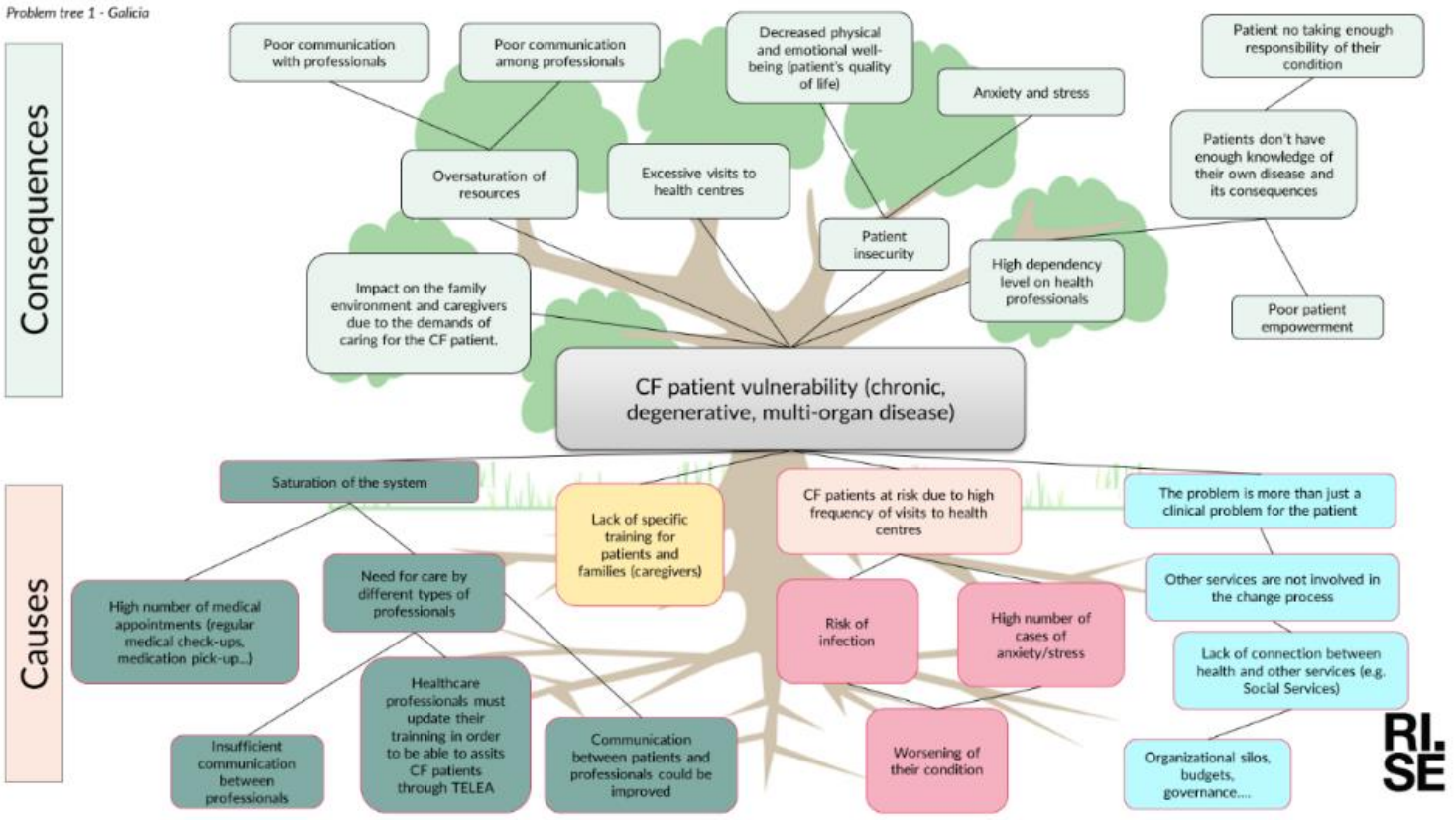


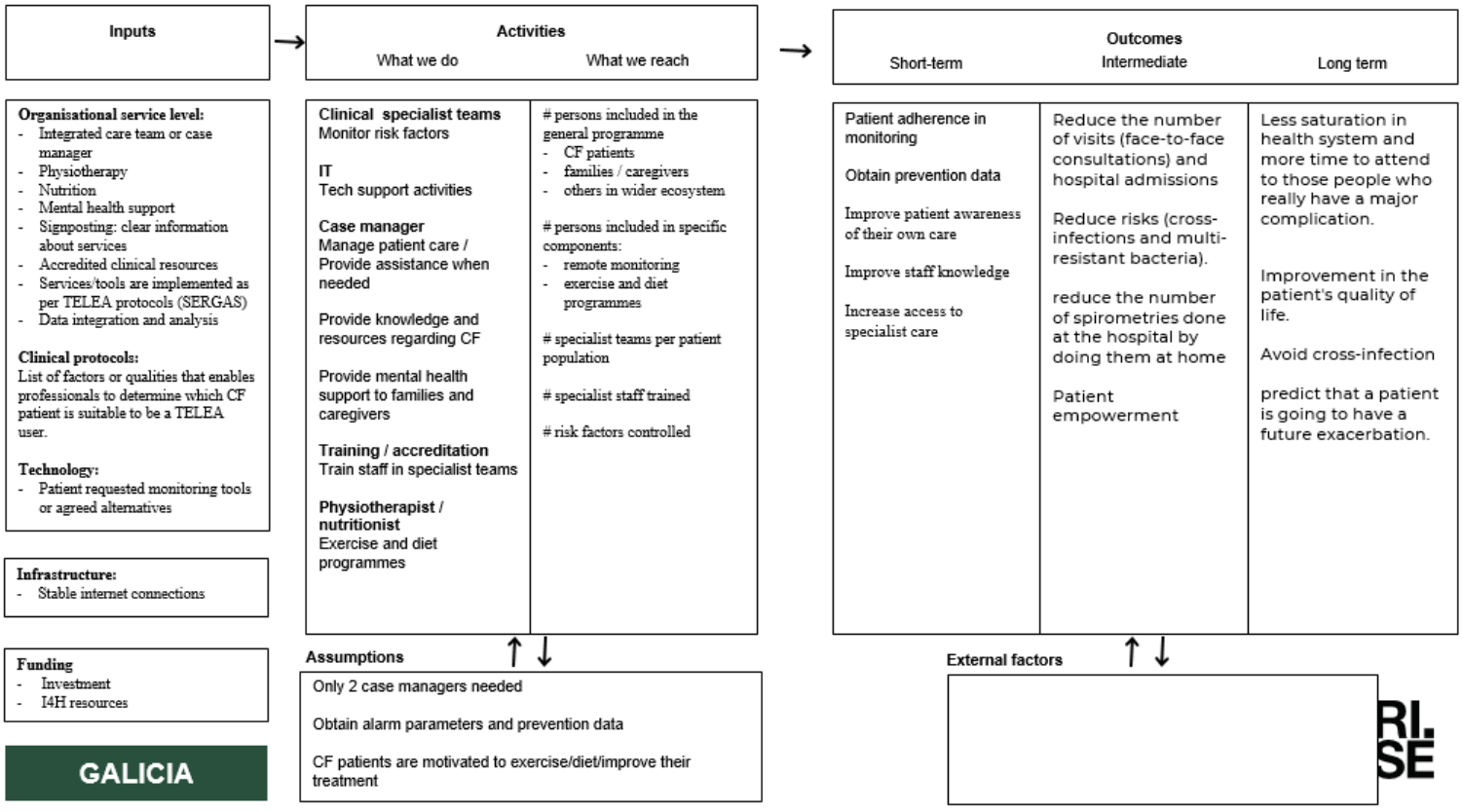
Figures 6, 7: Region Skåne testbed problem tree and Logic model





Figures 8, 9: North Rhine Westphalia (Aachen) testbed problem tree and logic model





Figures 10, 11: Galicia testbed problem tree and logic model

Annex 3. Sectors involved in cross-sectoral evaluation

Table 19. Breakdown of scoping reviews of SCI and cross-sectoral evaluations by lead sector.

Lead Sector	Task 2.1 (Scoping review of SCI examples) (N=78)	Task 4.3 (Scoping review of cross-sectoral evaluations) (N=48)
Health	(n = 26, 34%)	(n = 39, 81%)
Education	(n = 5, 6%)	(n = 2, 4%)
Employment/Workforce (including Occupational Health)	(n = 10, 13%)	(n = 4, 8%)
Social Care	(n = 10, 13%)	(n = 2, 4%)
For profit organisations Business	-	(n = 1, 2%)
Academia	(n = 1, 1%)	-
Environmental	(n = 4, 5%)	-
Not for profit (including responsible finance)	(n = 6, 8%)	
Housing	(n = 5, 6%)	-
Justice	(n = 1, 1%)	
Sports and recreation	(n = 3, 4%)	-
Social enterprise	(n = 1, 1%)	-

Annex 4. An evaluation checklist for the testbeds.

This evaluation checklist is developed in two parts.

- Questions 1–6: Evaluation of the prevention activity itself.
- Questions 7–12: Evaluation of the SCI mechanism.

We also provide a 'traffic light system' approach for grading the results of the evaluation checklist. Each tick is one point, each cross is zero points. Each section is scored out of six points. Totals are to be calculated for each section at the end of the form. Evaluate each of the two components in turn. Users can also total the overall score for their prevention activity however should be mindful to ensure adequate quality of both components.

4+ Points: Green. This evaluation has returned a good or excellent score indicating that this component has worked well. Users can be confident in the approach.

Next steps: none.

2–3 Points: Amber. This evaluation has returned an average score indicating that this component has worked moderately well. Users need to revisit the questions that have returned a negative score to reconsider any required adjustments in order to make the prevention activity, SCI mechanism, or cross-sectoral impact more effective.

Next steps: Reevaluate when concerns have been addressed.

0–1 Points: Red. This evaluation has returned a poor score indicating that the component works inadequately or not at all. Users should exercise caution or potentially discontinue the prevention activity, SCI mechanism, or cross-sectoral impact until improvement measures are in place.

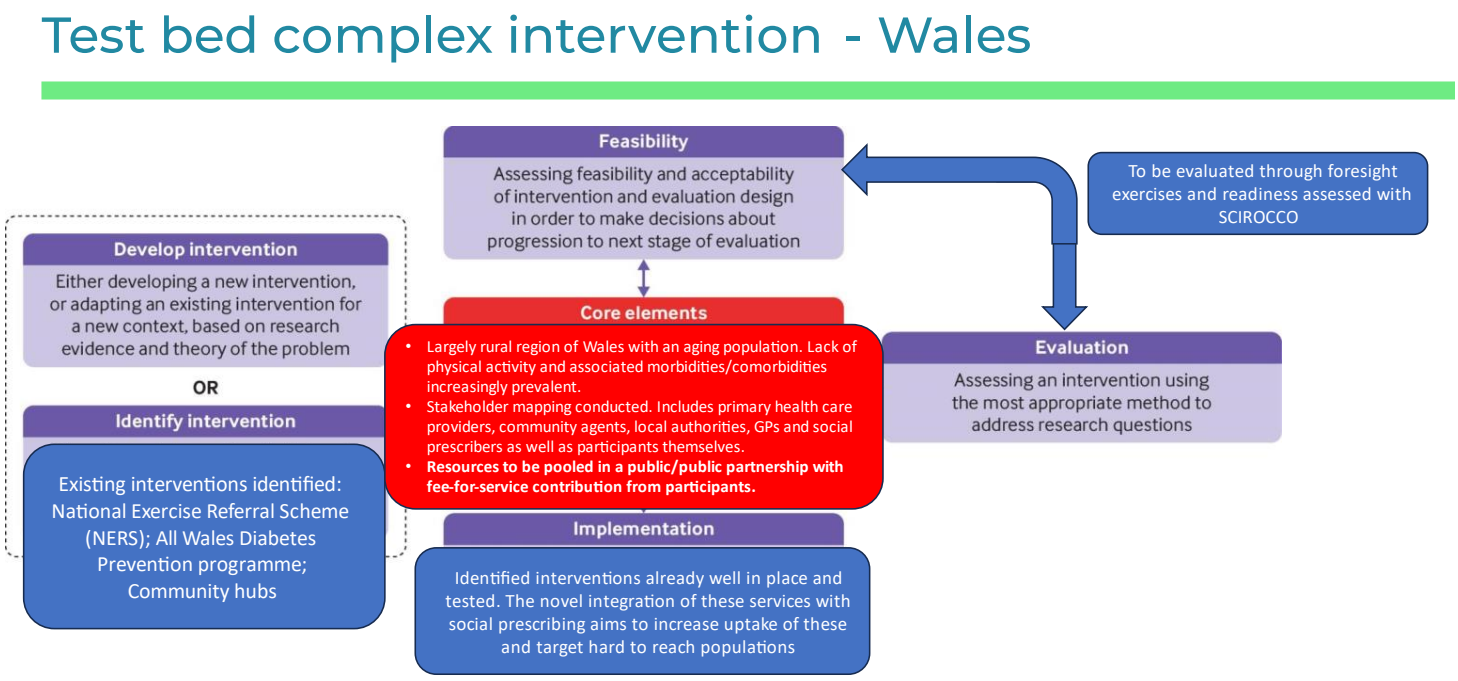
Next steps: Reevaluate when concerns have been addressed.

Table 20. Testbed evaluation checklist.

Part 1: Evaluation of the prevention activity.		
1	Measured against a counterfactual population e.g. matched population with similar health needs in another part of the local health economy, was there a global improvement in life expectancy and/or health -related quality of life of the stratified population concerned e.g. frail elderly (to prevent the need to admission or re-admission to hospital)?	✓ ✘
2	Matched against a counterfactual population with similar health needs in the local health economy, was there a change or reduction in global demand on health and social care systems for this stratified population group?	✓ ✘
3	Is there evidence that inevitable demands on these systems relating to disability, chronic disease and health risks, is being pushed into the future,	✓ ✘

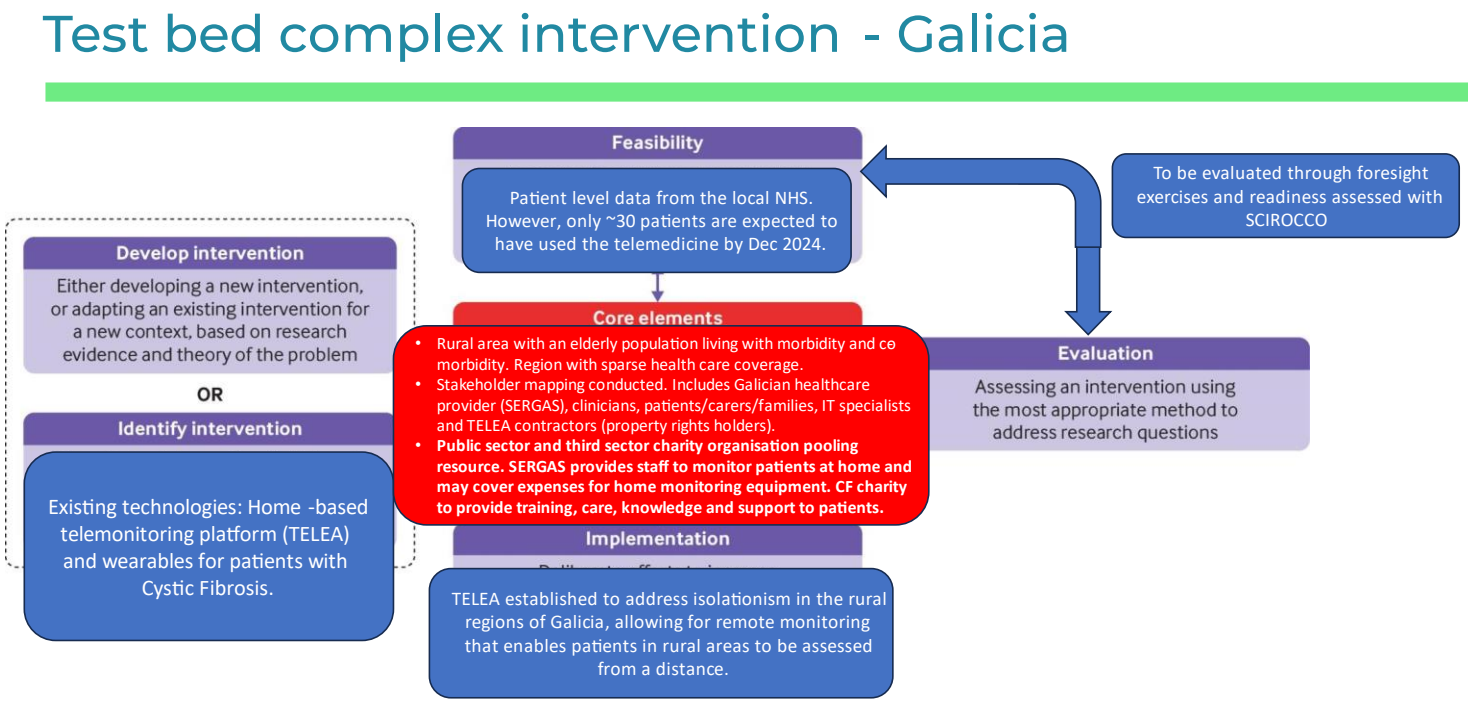
	relieving pressures on services in the immediate future, facilitating better health-related quality of life now, whilst not leaving patients without care?	
4	Do patients and public representatives report positively on the local, place based, impact of the prevention measure funded through a SCI?	✓ ✗
5	Do health professionals report positively on the local, place-based prevention measure funded through the SCI?	✓ ✗
6	Has the prevention measure funded by a SCI reduced or improved any variation of population health status with the rest of the relevant region?	✓ ✗
	Total score for Part 1	
Part 2: Evaluation of the SCI mechanism.		
7	Over a suitable period of e.g. five years, has the SCI investment e.g. through a social enterprise been recouped and paid back into that fund as an evergreen resource for prevention, either through cashable savings or some other means?	✓ ✗
8	Would the SCI investor be willing to provide more funding in this way and are there other investors in the local health economy interested in coming on board?	✓ ✗
9	Has there been appropriate governance with sufficient distance between the SCI provider and oversight board for the organisation and delivery of the prevention measure?	✓ ✗
10	Is there evidence of a feedback loop in the governance and business model during the investment period that has fine-tuned the process?	✓ ✗
11	How innovative or novel is the SCI in the local health economy, and is there interest in wider roll-out or franchising in the relevant local region?	✓ ✗
12	What do stakeholders in the process of initiating, running and evaluating the SCI think relative to their experience of both public and alternative sector finance models? Is this thought of positively?	✓ ✗
	Total score for Part 2	
	Total score for the activity (total parts 1 & 2)	

Annex 5. Amended frameworks for developing and evaluating complex interventions (Figure 1) representative of each testbed's complex intervention.



Skivington K, Matthews L, Simpson S A, Craig P, Baird J, Blazeby J M et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance BMJ 2021; 374 :n2061 doi:10.1136/bmj.n2061

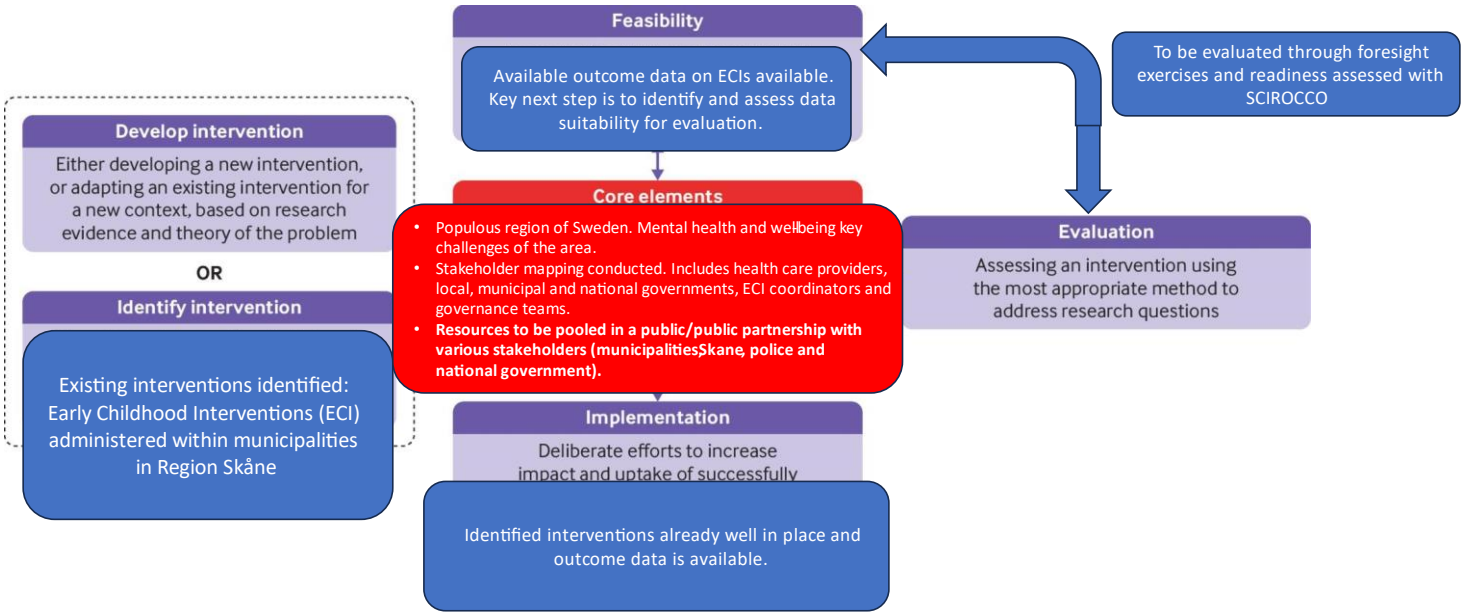
Figure 12. Framework for developing and evaluating complex interventions: Wales.



Skivington K, Matthews L, Simpson S A, Craig P, Baird J, Blazeby J M et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance BMJ 2021; 374 :n2061 doi:10.1136/bmj.n2061

Figure 13. Framework for developing and evaluating complex interventions: Galicia.

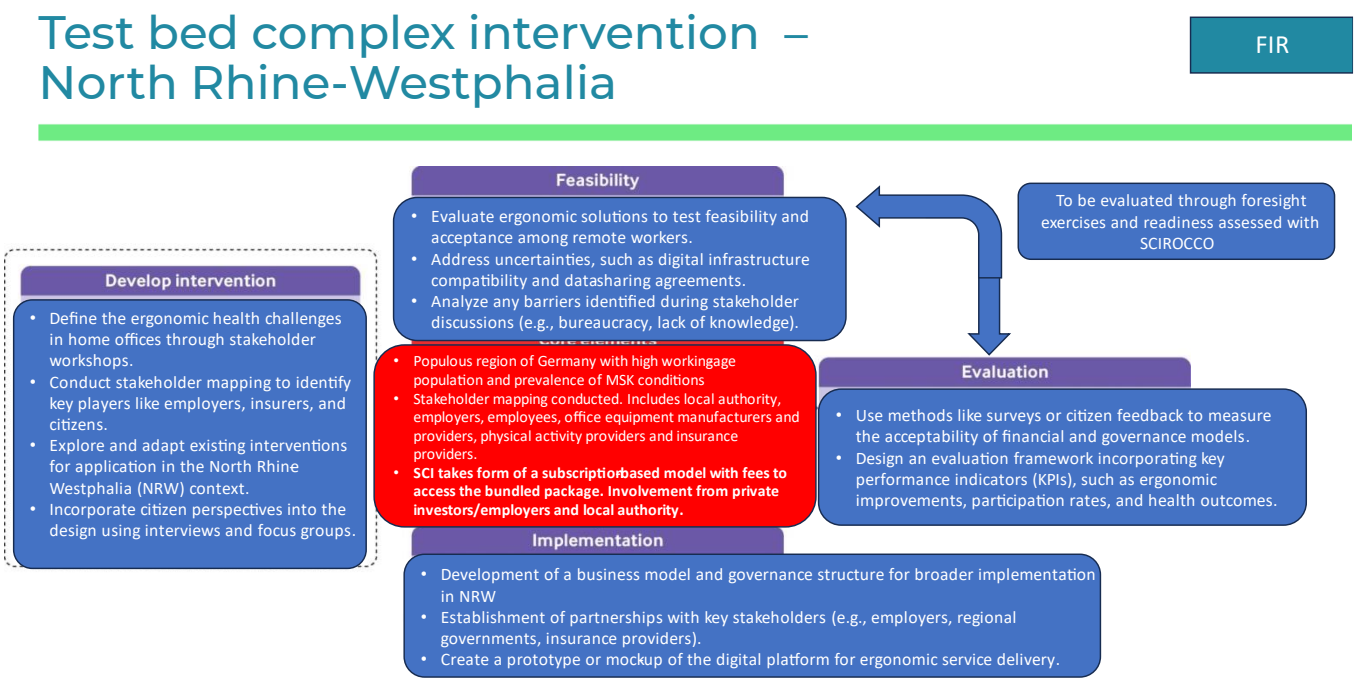
Test bed complex intervention - Skåne



Skivington K, Matthews L, Simpson S A, Craig P, Baird J, Blazeby J M et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance BMJ 2021; 374 :n2061 doi:10.1136/bmj.n2061

Figure 14. Framework for developing and evaluating complex interventions: Skåne.

Figure 15. Framework for developing and evaluating complex interventions: Germany.



Skivington K, Matthews L, Simpson S A, Craig P, Baird J, Blazeby J M et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance BMJ 2021; 374 :n2061 doi:10.1136/bmj.n2061

This project has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement 101095522