PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Process evaluation of the scale-up of integrated diabetes and hypertension care in Belgium, Cambodia, and Slovenia (the SCUBY project): A study protocol
AUTHORS	Martens, Monika; Wouters, Edwin; van Olmen, Josefien; Klemenc Ketiš, Zalika; Chhim, Srean; Chham, Savina; Buffel, Veerle; Danhieux, Katrien; Stojnić, Nataša; Zavrnik, Črt; Poplas Susič, Antonija; Van Damme, Wim; Por, Ir; Remmen, Roy; Ku, Grace Marie; Klipstein-Grobusch, Kerstin; Boateng, Daniel

VERSION 1 – REVIEW

REVIEWER	E. Mannucci Azienda Osped Univ Careggi, Diabetes Agency
REVIEW RETURNED	22-Mar-2022

GENERAL COMMENTS	The study protocol presented is culturally interesting for the scientific community and potentially useful for policymakers. I have two main points which require clarification: 1) The details of the integrated care package to be implemented should be provided more clearly. The authors should also state if the package is applied "as it is" in the three contexts, or modified and adapted to local conditions (i.e., adapted to local cultural, environmental, social, and organizational context). The acceptable extent of such adaptation, which remains compatible with the preservation of the nature of the package, should be pre-defined. 2) The authors state that, in this paper, they will report only methods for assessing some outcomes, excluding the "impact" - which, I guess, also includes observed health outcomes. However, the whole project appears to include the assessment of long-term health outcomes. I am aware that the results of such a complex program can be conveniently communicated in more than one paper; I do not understand the need to divide the study protocol in multiple paper. I strongly suggest to include in the present manuscript the part of the SCUBY prohject related to the assessment of impact.

REVIEWER	Anne Katahoire Makerere University, Child Health and Development Centre
REVIEW RETURNED	11-Apr-2022

GENERAL COMMENTS	This is a well-written protocol, it is detailed and all the procedures are clearly outlined

REVIEWER	Rachel Nugent RTI International, Center for Global NCDs
REVIEW RETURNED	27-Apr-2022

GENERAL COMMENTS	The study protocol misses some key pieces. I cannot evaluate the following specific aspects of protocol as they are not described: 1. Quantitative analysis is not described, nor are the outcomes to be measured 2. The quasi-experimental aspect is mentioned but not described 3. Not clear who has defined and named the three scale-up options, and on what basis 4. Why were the three countries chosen? 5. No checks described for internal or external validity or redundancy of the described qualitative instruments 6. There are no documents to show how the data will be compiled, represented for the analysis, and results of the various tools weighed to draw conclusions.

VERSION 1 – AUTHOR RESPONSE

Comments to the Author: **(reviewer 1)**	**Responses from the authors to the reviewer**
The study protocol presented is culturally interesting for the scientific community and potentially useful for policymakers. I have two main points which require clarification: 1) The details of the integrated care package to be implemented should be provided more clearly. The authors should also state if the package is applied "as it is" in the three contexts, or modified and adapted to local conditions (i.e., adapted to local cultural, environmental, social, and organizational context). The acceptable extent of such adaptation, which remains compatible with the preservation of the nature of the package, should be pre-defined.	We thank the reviewer for this comment, as we realise this was not well clarified. The Integrated Care Package (ICP) in the project definition has five components: (a) early detection and diagnosis of people with HT and/or T2D and subsequent (b) treatment in primary care services, (c) health education and (d) self-management support to patients and caregivers, and (e) collaboration between caregivers. The five components are implemented via country-specific delivery models. In all three countries, we identified the ICP to be implemented to a certain degree and with variation in delivery models (see overall protocol paper). The degree to which the five components were implemented at the start of the project is part of the baseline evaluation (currently in process of publication). The ICP grid has been developed as a measurement frame, evaluating to what extent the 5 components have been implemented. The scale-up interventions (as part of the roadmap) aim to increase coverage of the ICP, which can include: increasing the implementation of one or more components; widening the coverage of the current package towards a larger part of the population; optimising the conditions for implementation of the ICP. We do, however, realize that the role of the ICP in our study was not explained with sufficient detail in the previous version of our manuscript – potentially causing the confusion/interpretation signalled by the reviewer. We have rewritten the section and added a sentence referring to the SCUBY protocol paper which explains this

	research strategy in an even more elaborated manner (p.6, third paragraph).
2) The authors state that, in this paper, they will report only methods for assessing some outcomes, excluding the "impact" - which, I guess, also includes observed health outcomes. However, the whole project appears to include the assessment of long-term health outcomes. I am aware that the results of such a complex program can be conveniently communicated in more than one paper; I do not understand the need to divide the study protocol in multiple paper. I *strongly* suggest to include in the present manuscript the part of the SCUBY project related to the assessment of impact.	We agree that the previous version of manuscript – which focused on the process evaluation, but also briefly mentioned the impact evaluation – might have caused some confusion. The goal of this manuscript is to describe an extensive process evaluation framework for a complex, multi-country and multi-dimensional project. The complexity of the topic (scaling-up chronic disease care in different policy contexts) requires an according multifaceted approach rooted in the available (theoretical) literature and frameworks. To render this focus clearly, we have adapted Figure 2 (the evaluation framework) omitting the impact evaluation elements. The reasons for this choice are: evaluating chronic illness care is a complex task in itself as chronic care does not have a clear endpoint (e.g. in comparison to many infectious diseases where being cured is a clear end point). In addition, it requires coordinated action by several stakeholders (doctors, dieticians, nurses, patients, families of patients) along the entire continuum of care (from screening, over diagnosis, follow-up, treatment to illness control). Such a comprehensive impact evaluation thus requires a wide range of data sources (health system data (measuring ICP), health activities data (measuring activities by doctors/nurses), health insurance data (measuring clinical actions, medication use, etc.), laboratory data (measuring health outcomes), etc.) which all need to be linked (and rendered comparable across the three countries of this study). Consequently, the impact evaluation falls outside of the scope of the current paper, which explicitly aims to provide an evaluation framework for evaluating the process of scaling-up integrated care in the three countries. The current version of this manuscript addresses this shortcoming and clearly states its goals (p.7) and limitations (p.18-19). In short, deepening the impact evaluation is beyond scope of this paper – *inter alia* in view of the maximum word count. Moreover, currently, such a paper on impact evaluation (and challenges we have encountered whilst

	collecting data) is being drafted.
Comments to the Author: **(reviewer 2)**	
This is a well-written protocol, it is detailed and all the procedures are clearly outlined.	We would like to thank the reviewer for their support in publishing this protocol paper.
Comments to the Author: **(reviewer 3)**	**Responses from the authors to the reviewer**
The study protocol misses some key pieces. I cannot evaluate the following specific aspects of protocol as they are not described: 1. Quantitative analysis is not described, nor are the outcomes to be measured	We thank the reviewer for this observation. We would like to stress that this process evaluation protocol paper mainly aims to collect and analyse qualitative data – whereas we mostly address implementation outcomes and processes with limited procedural data. The impact evaluation (previously briefly described in the evaluation framework, figure 2) would contain mostly quantitative data. As chronic illness care requires coordinated action by several stakeholders (doctors, dieticians, nurses, patients, families of patients) along the entire continuum of care (from screening, over diagnosis, follow-up, treatment to illness control), a comprehensive impact evaluation requires a wide range of data sources (health system data (measuring ICP), health activities data (measuring activities by doctors/nurses), health insurance data (measuring clinical actions, medication use, etc.), laboratory data (measuring health outcomes), etc.) which all need to be linked (and rendered comparable across the three countries of this study). Consequently, the impact evaluation falls outside of the scope of the current paper, which, in view of the broad presentation of implementation outcomes as part of the process evaluation, explicitly aims to provide a framework for evaluating the process of scaling-up integrated care in the three countries. We have clarified this in the new version, omitting the part on the impact evaluation. Currently, a separate

	paper on impact evaluation is being drafted. For the current paper on process evaluation, the part where we mention quantitative outcomes are the scale-up dimensions of 'coverage' and 'expansion'. We have added more details on both dimensions (p.11).
2. The quasi-experimental aspect is mentioned but not described	As indicated above, the current paper aims to provide the readership with an evaluation framework to assess the process of scaling up integrated care for chronic diseases across different contexts. Given the complexity of this topic, the impact evaluation (also complex as the impact evaluation of chronic care requires assessing outcomes along the entire continuum of care) falls outside of the scope of this article. The previous version of this manuscript contained some references to the impact evaluation (which will be presented in another publication) containing the quasi-experimental design. The revised version of the manuscript aims to address this and solely focuses on the process evaluation framework – thus leaving the description of the quasi-experimental design to the impact evaluation paper.
3. Not clear who has defined and named the three scale-up options, and on what basis	Here, we would like to refer to the source we have cited (reference 21, p.6), the SCUBY protocol paper. We also importantly note that these scale-up dimensions have been previously conceptualised in literature. To portray this better, we have added two additional sources (idem, p 6): • World Health Organization & ExpandNet. (2010). Nine steps for developing a scaling-up strategy. World Health Organization. https://apps.who.int/iris/handle/10665/44432; and • Greenhalgh, T, G Robert, F Macfarlane, P Bate, and O Kyriakidou. (2004). Diffusion of Innovations in Service Organizations: Systematic Review and Recommendations. Milbank Quarterly 82, no. 4 (4AD): 581–629. doi: 10.1111/j.0887-378X.2004.00325.x). The three-dimensional framework (shown in appendix 1) however, has first been put forward by the SCUBY consortium (ref 21). Please note that there is no conceptual uniformity when it comes to scale-up dimensions/options across sources, but we mention this in the manuscript (p10, below). Within the SCUBY project, we have aimed to make it into a visual

	framework, which can facilitate discussion on scale-up in other settings (see e.g. https://bmjopen.bmj.com/content/12/4/e053122).
4. Why were the three countries chosen?	On page 6, we mention: *"The selection of the three cases was based upon their health system characteristics and current focus on scale-up strategies"*. Additionally, we describe the characteristics of these scale-up strategies (horizontal, vertical or diversification) and what these different strategies entail across countries, which is further elaborated on in the SCUBY study protocol (reference 21). We have added one sentence in which we highlight how we aim to draw lessons from these different health systems and we are also interested in their therefore different needs for scale-up: *"These three countries were chosen in view of the lessons that can be drawn from these diverse health system contexts: a developing health system in a lower middle-income country (Cambodia); a centrally steered health system in a high-income country (Slovenia); and a publicly funded highly privatised health-care health system in a high-income country (Belgium)."* More information can be found in the main protocol paper in Global Health Action (reference 21).
5. No checks described for internal or external validity or redundancy of the described qualitative instruments	We thank the reviewer of this remark; we had indeed previously not mentioned how we aim to ensure the quality (validity and reliability) of our process evaluation. The credibility (internal validity) and transferability (external validity) of findings is largely ensured by means of data, methods and investigator triangulation. We have added on p.17: *"By means of employing multiple methods, data sources and a larger analysis team (independent researchers conducting the analysis and feeding back to country research teams for discussion on the findings), we wish to cross-check information and conclusions drawn from the data via triangulation and data saturation and thereby ensuring the credibility of the data."* Hence, we argue rather than evaluating redundancy of the described qualitative instruments, we aim to corroborate findings via these instruments. In terms of transferability or external validity, having clear, rich and detailed descriptions makes it possible to provide others with the context and a certain transferability of the situation, e.g. a successful policy

	dialogue. Data saturation is exemplified by the number of interviews that are conducted until there is no additional useful information due to saturation.
6. There are no documents to show how the data will be compiled, represented for the analysis, and results of the various tools weighed to draw conclusions.	We have provided templates for the data collection tools in the appendices and given an overview of the outcomes we are interested in (cf. figure 2). We have now added a flowchart in appendix 8, visualising how and which data collection methods contribute to specific analyses. Almost in all cases, the data collection tools serve multiple purposes (e.g. the policy dialogue reporting form contains information regarding the context as well as policy dialogue implementation outcomes, and roadmap implementation outcomes) and are triangulated to corroborate our findings. We also argue that the weighing of the tools is not appropriate in this instance, as we have used the tools for different parts of the process evaluation and our aim is to saturate and triangulate the data in order to improve the trustworthiness of our research.