Healthy Cities Phase V evaluation: further synthesizing realism

Evelyne de Leeuw1,*, Geoff Green2, Agis Tsouros3, Mariana Dyakova4, Jill Farrington5, Johan Faskunger6, Marcus Grant7, Erica Ison8, Josephine Jackisch9, Leah Janss Lafond10, Helen Lease11, Karolina Mackiewicz12, Per-Olof Östergren13, Nicola Palmer14, Anna Ritsatakis15, Jean Simos16, Lucy Spanswick17, Premila Webster18, Gianna Zamaro19, June Crown20, Ilona Kickbusch21, Niels Rasmussen22, Gabriel Scally23, Marian Biddle24, Suzanne Earl25, Connie Petersen26, and Joan Devlin27, on behalf of the World Health Organization European Healthy Cities Network

1Glocal Health Consultants, Geelong, Australia, 2Sheffield Hallam University, Sheffield, UK, 3World Health Organization Regional Office for Europe, Copenhagen Ø, Denmark, 4Division of Health Sciences, Warwick Medical School, Coventry, UK, 5Leeds University, Leeds, UK, 6Proactivity, Stockholm, Sweden, 7Healthy Urban Planning, Bristol, UK, 8Health Impact Assessment, Oxford, UK, 9Lund, Sweden Copenhagen Ø, Denmark, 10The Women’s Sports Network, London, UK, 11Planning and Development, Bristol, UK, 12Baltic Region Healthy Cities Association, Turku, Finland, 13Lunds Universitet, Lund, Sweden, 14La Trobe University, Melbourne, Australia, 15Independent Researcher, Melissa, Greece, 16University of Geneva, Geneva, Switzerland, 17Barwon Medicare Local, Melbourne, Australia, 18Public Health Specialist Training Programme, Oxford, UK, 19Udine Healthy City, Udine, Italy, 20Independent Researcher, London, UK, 21Global Health Programme at the Graduate Institute of International and Development Studies, Geneva, Switzerland, 22Independent Researcher, Copenhagen, Denmark, 23WHO Collaborating Centre, University of the West of England, Bristol, UK, 24Melbourne, Australia, 25Survey Monkey Programmer, USA, 26Healthy Cities Support, Copenhagen, Denmark, and 27WHO Healthy Cities Network Belfast Secretariat, Belfast, UK

*Corresponding author. E-mail: evelyne@glocalhealthconsultants.com; e.deleeuw@latrobe.edu.au

Summary

In this article we reflect on the quality of a realist synthesis paradigm applied to the evaluation of Phase V of the WHO European Healthy Cities Network. The programmatic application of this approach has led to very high response rates and a wealth of important data. All articles in this Supplement report that cities in the network move from small-scale, time-limited projects predominantly focused on health lifestyles to the significant inclusion of policies and programmes on systems and values for good health governance. The evaluation team felt that, due to time and resource limitations, it was unable to fully exploit the potential of realist synthesis. In particular, the synthetic integration of different strategic foci of Phase V designation areas did not come to full fruition. We recommend better and more sustained
REALIST SYNTHESIS: A DOSE OF REALITY

Evaluations of the earlier Phases of the WHO European Healthy Cities Network were found to have drawbacks. As discussed in this Supplement (de Leeuw et al., 2015), the first Phase harvested commonalities between a small group of highly committed cities and mapped the conditions for becoming a Healthy City (Ashton et al., 1986; Draper et al., 1993). This research was neither evidence or theory based, but produced some of the most seminal texts used globally in Healthy Cities development (e.g. Tsouros, 1991, World Health organization, 1992). The research collaboration funded by the European Commission for Phase II limited the geographical scope mainly to Western Europe (Berkeley and Humphreys, 1998; De Leeuw et al., 1998; Capello, 2000) and Price and Tsouros (Price and Tsouros, 1996) compiled a collection of case studies that seems to have had more impact than the scholarly enterprise. A more integrated, mixed-method approach was advocated from Phase III onwards (De Leeuw, 2009), but the proposed MARI (Monitoring, Accountability, Reporting and Impact) research framework stretched local research commitments beyond their limits and resourcing the comprehensive inquiry into a multinational socio-ecological health paradigm was virtually impossible. With a highly condensed approach to MARI through Annual Reporting Templates (Tsouros and Green, 2013), a methodology started to crystallize in which research questions and approaches were defined and refined in collaboration with all relevant stakeholders, reflecting the naturalistic fourth-generation evaluation approach advocated by Lincoln and Guba (Lincoln and Guba, 1986).

With advances in methodological approaches to ‘real world’ policy research and development called ‘realist evaluation’ (Pawson and Tilley, 1997) and the increased computing power of software managing qualitative data (Bazeley and Jackson, 2013), it was felt that responding to the needs, requirements and assets of local governments in health development in evaluations for Phase V and onwards became feasible and appropriate. Adding to this conviction was the emergence of the ‘realist synthesis’ approach, pioneered in Britain (Pawson et al., 2004) and implemented as the European Union funded DECIPHER (2015, Developing an Evidence-Based Approach to City Level Public Health Planning and Investment in Europe) project. In such a vision, it was found that it is not necessary to investigate the full logical causal and final sequence of proximal and distal (social, political and commercial) determinants of health to demonstrate the evidence of effectiveness of systems interventions on health, well-being and equity. Rather, the available evidence from other sources could be inserted in a programme logic to approximate the likelihood of effectiveness under certain (complex) conditions that such system interventions work.

From the start of Phase V we negotiated this approach with the stakeholders in the WHO European Healthy Cities Network. More than any other group of stakeholders, the designated cities and their coordinators were quickly convinced of the legitimacy and feasibility of the novel approach. As a programme essentially driven by local government, WHO adopted the programme logic required for the implementation of a realist synthesis approach. It should be noted, however, that one of the persistent drawbacks of earlier Phase evaluations remained: full resourcing of a comprehensive socio-ecological research framework extending across dozens of highly diverse Healthy Cities would require a long-term investment in a broad and (virtually permanently available) group of institutional research actors—a situation that is not within the governance prerogative and funding capabilities of the World Health Organization (e.g. Bernier and Clavier, 2011).

REALIST SYNTHESIS: WHAT WAS ACCOMPLISHED

Reflecting on the analyses and reporting in this Supplement, the realist synthesis approach has yielded a number of insights that go above and beyond what has been found earlier and elsewhere.

First of all, there is a complete absence of any research (either published in the peer-reviewed scholarly literature or outside it in the ‘grey literature’) that collects, compiles and analyses a diversity of qualitative and quantitative data from a set of 99 local government areas across a WHO region that is as heterogeneous as the European one. Figure 2 in De Leeuw et al. (De Leeuw et al., 2015) shows how we have classified four different geographical regions. Figure 1 further shows that governance determinants within the European Region can also be seen as reflections of varying commitments between WHO, Council of Europe and European Union.
The analyses carried out on, in and with local governments in this enormously heterogeneous part of the world in itself is a testament to the power of the methodology deployed. These cities, and their actions, constitute a unique living laboratory of health innovation at the local level, and the research reported here builds on that exceptional accomplishment. At the same time—this diversity creates methodological and analytical issues that we will address below.

Across the European Healthy Cities, independent of their geographical, demographic or governance context, we show in our analysis that

- Local governments increasingly transcend traditional lifestyle-only interventions and generate policies and programmes that include and endorse issues of sustainability, urban planning and equity. Cities move away from short-term demonstration projects and make sustained efforts to develop longer-term policies to address these complex (or ‘wicked’) issues. They attribute the potential and feasibility to move into that realm to the inspiration and connection with other cities across Europe, and to the facilitative role of WHO to access information and models of good practice.

- Cities themselves start to ‘connect the dots’ of the evidence generated around social determinants of health, governance and equity. Again, making those connections may have been facilitated by external factors (such as WHO and international city networking), but at the local level momentum is clearly building to involve other sectors and engage in multi-stakeholder consultation and action to develop visions and strategies that—often without explicit acknowledgement by cities—build towards lasting Health in All Policies.

- Local governments suffer from the fall-out of the Global Financial Crisis and often operate under stringent conditions of austerity. However, this is not to the detriment of the above commitment to broader intersectoral systems strategies for health. A key reason for this lasting vision is a strong local recognition of the assets that communities and other local stakeholders can bring to bear on health development.

- Following the programme logic that drives both Healthy City designation and their actions, as well as our realist synthesis methodology, it can be asserted at an aggregate level that European Healthy Cities are bound to reduce health inequity through investment in evidence-based policies and programmes that are grounded in our accounts of best practice in policy development and a recognition of good governance for health. These include an increase in the adoption of Health Impact Assessments, community-wide investments in Healthy Urban Environments and Designs and the willingness to integrate the locally and internationally produced evidence in organizational and strategic learning. We do recognize, of course, that addressing health inequities does also involve increased sophistication in proportionate universality (Marmot and Bell, 2012) as a strong evidence base is building that some well-intended interventions in fact increase such inequity (Lorenc et al., 2013).

A key question that drove our inquiry, and the preceding negotiations over research approaches and action priorities, was ‘Do Healthy Cities make a difference?’ Cities believe they do. They report changes, over time, in their ways of dealing with complex health issues. They manage to embrace abstract notions like ‘equity’ and ‘governance’ and integrate these in concrete and operational ways in policies, strategies and interventions that contribute to healthier choices for healthier lifestyles, and better conditions to allow for those healthier choices.

If one overarching pattern has become clear in this evaluation it is that European Healthy Cities have indeed moved to put health high on social and political agendas without sacrificing tried-and-tested public health action. Davies et al. (Davies et al., 2014) and Kickbusch (Kickbusch, 2007) argue that public health evolution has come in four (Davies) or five (Kickbusch) waves, with the most recent one a cultural paradigm shift to

Fig. 1: Membership of international organizations of WHO European Region member states.
health (rather than disease) and health promotion (rather than disease prevention). In arguing shifts from a biomedical model towards a social model of health many seem to have contended that one development superseded a prior one—but in order to reach a health culture with appropriate governance for health it is important to appreciate and operationalize the importance of each ‘wave’ in an integral vision of global health (De Leeuw, 2001). Healthy Cities do precisely that (Figure 2).

GROWING AND MATURING A REALIST SYNTHESIS APPROACH TO HEALTHY CITIES EVALUATION

Reviewing and reflecting on the work carried out that led to this collection of papers and further documentation aimed at providing specific guidance for individual cities and their networks we can also identify a number of challenges and missed opportunities.

We recognized a broad range of confounding factors and biases that may have skewed our analyses and tainted findings [De Leeuw et al., 2014], p. 21]. The dynamic nature of these confounders is further exacerbated by the fact that entry into a Phase is continuous, and that cities may seek designation up to the very last moment a Phase runs (Figure 3).

This means that a significant number of cities may not have been exposed to particular preconditions and requirements that designation entails. With more detailed analyses, we might have been able to control for these differences, and qualified some of our findings.

More importantly, though, is the obstinate—and partly justified—critique that answering the question ‘Do Healthy Cities make a difference?’ can only happen when applying a methodology that would allow for a case–control or pseudo-experimental design. In doing so we would have had to recruit a matched set of 99 non-Healthy Cities (or, within designated Healthy Cities, communities and/or neighbourhoods that were explicitly excluded from Healthy City-like strategies and actions). Apart from logistical and possible moral–ethical barriers, it would have been operationally impossible, within the resource opportunities that were available to us, to do this. But proponents of the realist evaluation paradigm would claim that it is indeed the very purpose of this methodological approach to demonstrate that, under varying contexts, different policy expressions can yield different yet effective outcomes.

This brings us to some of the drawbacks to the realist synthesis methodology that we applied. Although following the ‘negotiated reality’ stage of the development at an early point in Phase V with city representatives demonstrated a deep commitment of all Healthy city coordinators and

Fig. 2: Five evolutionary developments of public health
their political representatives, the actual deployment of the range of research tools caused substantial ‘research fatigue’ and responses that may have displayed considerable social desirability-bias. Through processes of triangulation between different data sources, researchers and theories we have attempted to control for such biases (Creswell and Clark, 2007) and we have furthermore contrasted our response findings with non-response characteristics to ascertain that response fell within credible levels of reliability and validity. In some questions of the General Evaluation Questionnaire this issue of biased self-reporting will have impacted on the degree to which cities gave a positive spin on their performance. They were asked to rate their own performance at three points in time (the start of Phase V (2009), at the end of it (2013), and at the end of Phase VI (2019) in case they sought redesignation) on issues such as health equity, the ability to deal with non-communicable disease and sustainability. Not surprisingly all cities found that they would improve due to membership of the network. Further sophistication should be going into designing such questions to control for this pattern.

There are some features of ‘Healthy Cities realist synthesis in action’ that must be described as substandard. The development of the research tools, based on the programme logic and processes of pre-testing (testing format, language, internal logic, etc.), was slow but effective. Their implementation was delayed because of personnel changes, hitches in decision-making processes and formal approval procedures. As we have described in de Leeuw et al. (de Leeuw et al., 2015) the response rates to each of the instruments were more than satisfactory, and the quantity of raw data was at the limits of the manageable. If double-sided printed versions of all documents would be piled up the stack would tower above UN-City, the office of WHO Europe. In a first rough analysis of case studies and General Evaluation Questionnaires into ‘Mother Reports’ we compiled over a 1000 pages of analyses only—and these excluded designation documents, case studies submitted to WHO throughout the Phase, Annual Report Template responses, and statistics derived from OECD and EuroStat.

The overwhelming wealth of responses, combined with extremely limited resourcing to the research team (For comparison: in a recent grant application for a European urban health investigation over 5 years with one principal investigator, two post-docs and two research assistants, the projected budget was 3 500 000 euros. The available budget for the Phase V evaluation with 15 senior researchers and 3 research assistants was about USD 200 000), has had as a consequence that a full-scale realist synthesis across the programme logic has not been possible. As is clear from the articles in this Supplement, the interface between the different domains that are reported on remain largely unexplored; for instance, we know from our cursory analysis of data across equity, health urban planning, governance and policy-making that this challenging area of work does receive significant attention in a number of cities and from a synthetic point of view such action should make for radically different outcomes for age-friendliness and healthy living. Unfortunately, for the time being it appears that these data remain in the data morgue if not the data graveyard (Custer, 2015). Putting a more optimistic spin on this, one could say that they are a sleeping beauty waiting for the charming prince(ss) to be kissed to lead a vibrant life.

Data collection supervision, data management and initial data processing were carried out in less-than-optimal ways. The Healthy Cities Research Director (EdL) developed a data management strategy and a case study coding strategy using the NVivo software package in consultation with staff at WHO. The Research Director then supervised two research assistants during the data processing stages, with a time difference between the two locations where action was taken of 8–10 h (Geelong/Australia–Copenhagen/Denmark). When the two research assistants returned from their WHO internships to Australia, the same time difference made real-time communications for data interrogation with the evaluation teams challenging. In spite of this situation meaning that the evaluation could effectively happen ‘around the clock’ and pronouncements that ‘working in the cloud’ is the future (Hays et al., 2015) there was a general consensus that being in the same (approximate) time zone with an opportunity for more face-to-face meetings would have enhanced the research process.

Members of the evaluation team were selected because of their research credentials in areas of the programme...
logic for the inquiry; because of their earlier work on European Healthy Cities evaluation and/or because of their association with World Health Organization Collaborating Centres that connect with the Healthy Cities programme. Not all members of the evaluation team were familiar or comfortable, however, with the realist synthesis methodology, its strengths and potential or limitations. In many ways this turned out to be an advantage, as it created a need to be fully unequivocal on research purpose, idiom, scope and tools that are appropriate within this perspective. In a final review exercise we asked the members of the team to reflect on the methodology, the degree to which they felt it met their research needs and capabilities and the extent to which they felt research outcomes aligned with the research outcome requirements expressed by Healthy City representatives early in the research negotiations in Phase V. The feedback was unequivocal.

A limitation is still the lack of time/capacity to cohere better as a strong trans-disciplinary team. The writing retreat was good—we need one every year. We need to better access the heart of city and national network thinking to help give input to research strategy as we develop it. One could envisage a process akin to best practice PPI (patient public involvement) in research, where the ‘subjects’ are more central to research committee itself.

Realist evaluation and synthesis is a useful and potentially powerful methodology for assessing the influence of Healthy Cities. But it can be difficult to comprehend if you are not an expert.

Using it (experimenting/piloting it) for the first time in the Phase V evaluation has provided essential background understanding, know-how and practical experience. It has also brought the evaluation team(s) together—to look, think over, analyse and synthesise across the themes and silos—something, I think happens for the first time in the HCs evaluation history. It (...) should be refined, channeled, supported, facilitated and taken forward to Phase VI where it would be essential in the evaluation of the key Health 2020 multidisciplinary/cross-sectoral (incl. horizontal & vertical) approaches.

I did not have sufficient time to follow through what individual cities said they would do in their application and then what they actually did—and why and how it differed. I agree with comments (of another team member) that pity we did not have time to explore more fully the abstracts and submissions to Conferences during the Phase.

In the same reflective exercise, without exception, the members of the evaluation team expressed frustration with the fact that (we did not pay... ) special attention and further development to the links (‘arrows’) between the prerequisites, activities, results, etc., i.e.—how do we realise in practice this transition in terms of cross-thematic analysis/synthesis and linkage between ‘input’, ‘process’, ‘structure’ and ‘output’/outcomes.

Indeed, the key of the rationale to realist synthesis is that it purports to bring more sense to the logical linkages in realist cause and effect in diverse contexts, and while we do have the data to more comprehensively describe and explain what happens in the arrows in our programme logic (Figure 4) the evaluation team neither had the time nor the resources to bring all the material together.

This is all the more unsettling as Healthy City coordinators, throughout the evaluation exercise, seemed more interested in learning about processes of good practice, rather than the outcomes these hypothetically would lead to. This is of course no surprise (Hancock and Duhl, 1986) at the very kick-off of this movement outlined the core stances of the experiment to develop health at the local level as a process: ‘A healthy city is one that is continually creating and improving those physical and social environments and expanding those community resources which enable people to mutually support each other in performing all the functions of life and in developing to their maximum potential’.

A WAY FORWARD

The substantive papers in this manuscript show great progress towards the goals of the WHO European Healthy City Networks. Health did acquire a significant position on local social and political agendas. Designated cities have moved from small-scale and time-limited projects on lifestyle change to larger, long-term policies and programmes that explicitly deal with social determinants of health and systems parameters for good governance for health. Inserting existing evidence of pathways to close the health inequity gap would suggest that this shift will lead to better local health and decreases in health inequities. Clearly, this will require an increased sophistication in proportionate policies and programmes and a continuing recognition of the importance of political determinants of health (Clavier and de Leeuw, 2013; Hunter, 2015; Kickbusch, 2015).

This continues to make the selection of a realist synthesis methodology the most appropriate for evaluations of diverse Healthy Cities across Europe and elsewhere. The methodology was welcomed by all respondents and other stakeholders in the research process, but upon reflection it has become clear that maintaining strong support structures and explanatory narratives to sustain the
focus of the evaluation needs to be in place throughout the process. This would include—as has happened with a selection of Healthy City coordinators after the conclusion of Phase V—consultations on the types of feedback, deliverables and messages that would be derived from the available rich data. Ultimately, deliberate and conscientious integration of a realist evaluation paradigm in every stage and element of the Healthy Cities programme would be a prerequisite for the fuller and even more successful implementation of the research endeavour.

The realist synthesis paradigm and the associated programme logic that we developed have led to very high response rates, particularly for this type of mixed-methods research. This has led to enormous quantities of data, and the limited resources put in place did not anticipate either the wealth of information, or the colossal level of sophistication that could be added to our analyses. Integrating realist synthesis evaluation throughout Phase VI and allocating adequate research support and resources, in collaborating with all stakeholders, would allow for a significant further increase in the capacity of the network to understand, explain and adapt its approaches to the dynamics of systems change for health. This Supplement provides the foundations and arguments to take this forward.

REFERENCES


