Received: 20 October 2016 | Revised: 4 December 2017 | Accepted: 5 December 2017

DOI: 10.1002/hpm.2483

WILEY

PERSPECTIVE

Critical evaluation of international health programs: Reframing global health and evaluation

Chunhuei Chi[1] | Anaïs Tuepker[2] | Rebecca Schoon[1,3] | Alicia Núñez Mondaca[4]

[1] Center for Global Health, College of Public Health and Human Sciences, Oregon State University, Corvallis, Oregon, USA

[2] Division of General Internal Medicine and Geriatrics, Oregon Health and Science University/ Portland Veterans Affairs Health Care System, Portland, Oregon, USA

[3] Public Health Program, Pacific University, Forest Grove, Oregon, USA

[4] Department of Management and Information Systems, School of Economics and Business, University of Chile, Santiago, Chile

Correspondence
Chunhuei Chi, Center for Global Health, College of Public Health and Human Sciences, Oregon State University, 13 Milam Hall, School of Biological and Population Health Sciences, Oregon State University, Corvallis, OR 97331, USA.
Email: chunhuei.chi@oregonstate.edu

Summary

Striking changes in the funding and implementation of international health programs in recent decades have stimulated debate about the role of communities in deciding which health programs to implement. An important yet neglected piece of that discussion is the need to change norms in program evaluation so that analysis of community ownership, beyond various degrees of "participation," is seen as central to strong evaluation practices. This article challenges mainstream evaluation practices and proposes a framework of Critical Evaluation with 3 levels: upstream evaluation assessing the "who" and "how" of programming decisions; midstream evaluation focusing on the "who" and "how" of selecting program objectives; and downstream evaluation, the focus of current mainstream evaluation, which assesses whether the program achieved its stated objectives. A vital tenet of our framework is that a community possesses the right to determine the path of its health development. A prerequisite of success, regardless of technical outcomes, is that programs must address communities' high priority concerns. Current participatory methods still seldom practice community ownership of program selection because they are vulnerable to funding agencies' predetermined priorities. In addition to critiquing evaluation practices and proposing an alternative framework, we acknowledge likely challenges and propose directions for future research.

KEYWORDS

community ownership, evaluation framework, global health, international health programs, upstream and midstream evaluation

1 | INTRODUCTION

Who has the right to determine the direction of a society's health development? Decisions about ***who*** chooses health programs and ***how*** they do it are unavoidably normative and political, yet the positivist mainstream approach to

evaluation tends to treat these decisions as technical, engineering questions. We argue here that the matter of who has the right to choose a program, its objectives, and effectiveness, as well as on what basis this right rests, should be central to an evaluation framework. If people intimately affected by decisions have little or no voice in determining what programs are made available to help reach their (health) goals, then a purely technical evaluation of "success" is problematic and largely irrelevant.

In recent years, several academics and field practitioners in international health (IH) have drawn attention to significant changes in the funding sources and organizational structures defining the agenda of IH programs.[1-5] These changes include the rise of private foundations and international coalitions (whether supra-national or a-national) as funding sources and arbiters of health priorities. The semantic shift, in recent years, away from the term "international health" towards "global health" has been argued to de-emphasize specific national contexts of health issues and the important role of national policies in finding solutions[6]; agreeing with this critique, we retain the older term "international health." Several analysts observe the potentially negative impact of these collective changes on the ability of national and regional communities to focus resources on significant health challenges of their own choosing, unless they are able to align their choices with the priorities of funders.[7,8] Additionally, many of these authors call for greater involvement of (usually national) communities in setting their own health agendas. However, little attention has been paid to the powerful way in which standard evaluation methods work against self-determination in health.

This article challenges mainstream IH program evaluation practices and proposes an alternative framework of Critical Evaluation for IH programs. This framework introduces the non-optional consideration of self-determination as an evaluative criterion. Focusing on the fundamental upstream questions of how and by whom a program is selected reframes the context of the midstream and downstream evaluation phases that follow.

We envision this evaluation approach as applicable to both formative and summative evaluation efforts. By necessity, its emphasis on self-determination has implications for program planning, but we envision the approach as also useful in post-hoc settings where evaluation was not integrated into program planning efforts. Several approaches to program planning, such as Intervention Mapping,[9] emphasize the importance of including communities in the planning process. However, these planning approaches do not explicitly call for a process of community involvement in defining the arena for intervention. In most existing planning efforts, the health problem to be addressed has already been identified by the evidence base, by expert consensus, or a combination of both (the Millennium Development Goals and the Sustainable Development Goals provide an example). Community input is used to find the most effective and acceptable way of addressing the problem, but such approaches do not require communities to be involved in the choice and definition of the problem in the first place.

Various forms of participatory research and planning, such as Community Based Participatory Research (CBPR)[10] and Action Research,[11] do emphasize more explicit community engagement at the earliest stages of problem identification and program development. CBPR appears to be more common in North America, and Action Research on other continents. Everywhere, these participatory approaches remain the exception rather than the rule, especially in the early stage of program planning and the late stage of program evaluation, our focus here. Even when participatory approaches are implemented, major variations in the levels of community participation occur, often without being explicitly noted in published accounts of the program's results. Singh[12] summarizes 5 different levels of community participation in CBPR, from "co-option" to "co-learning," with varying levels of engagement. The first 4 describe scenarios where decision-making power remains with outsiders/researchers. Only the fifth level of "co-learning," we would argue, represents community ownership. In co-learning, people served by a program and outsiders share their knowledge to create new understandings and work together to form action plans, with outsiders' facilitation. At this highest level of "community participation," there is shared ownership of the program in which outsiders are given equal power with the community. With its emphasis on the innately collaborative nature of both knowledge and action, Action Research is consistent in theory with our proposed model; it has not been widely applied to evaluation, however. In practice, participatory approaches are rarely able to operate at the level of "co-learning" because priority-setting is typically already done at the phase of funding allocation prior to community involvement.

Our model may be thought of as an extension of Action Research's resolutely open and diverse framework into a particular setting (IH) and particular practice (program evaluation). In short, Critical IH Program Evaluation is novel in that it seeks to influence research and program planning methods in the long term, by evaluating and judging success with a more explicit emphasis on self-determination and community ownership and a more explicit focus on the dynamics of the field of IH. Progress in the development of more participatory planning methods will not translate into the funding and proliferation of programs that meet community priorities until evaluators of those programs change their metrics of success. We believe that challenging and offering alternatives to current evaluation norms will allow and, ideally, over time transform the way program funders and planners work with communities to define problems as well as solutions.

2 | THE NEED FOR CRITICAL EVALUATION

A short parable can illustrate why evaluation practice needs to change. We use a parable rather than a specific program example to point out systemic rather than individual program failings; we combine elements which we feel will be familiar to many colleagues working in IH. Imagine a woman in Lagos wants to install solar panels on her home, so she can have electricity at home for a small sewing business, which she hopes will let her save money to move to another part of town that does not have as many mosquitoes and thus has a lower malaria rate. Having heard that an NGO is working on malaria prevention, she and several other neighborhood women with the same idea visit the NGO and ask for small business loans. The NGO staff think this is a great idea, but their current funding only supports the distribution of insecticide treated bednets—would the women like those, for now? The women take them (adding to the NGO's tally of nets distributed), leaving with some benefit but not with their long-term goals met. Some of them try to sell the nets to raise the money for their solar panels, but because the NGO gives the nets away for free in their area, the women end up traveling for many hours to rural areas where the nets can be sold for a little money. The NGO staff meanwhile write a request to their major donor for additional funding to support small business loans, but it is unsuccessful. The donor insists this would be a risky distraction from their very successful ITN distribution program—looks at how many nets have been distributed! The program has been declared a success by the NGO's evaluators, who is employed by the donor to make sure the money is being spent "effectively."

Some readers will find this story absurd, but the real world landscape of IH program selection and evaluation exhibits significant parallels.[7] Many donors and evaluators employ metrics of success that they prefer; many middle-and-low-income nations, like the women, are at the mercy of NGOs and Official Development Assistance to determine what programs they can access; many well-meaning NGOs and other programs know better than to accept pre-determined definitions of success, but are caught in the middle. Donors and evaluators frequently argue that programs are effectively providing benefits that would not exist otherwise. The inconvenience, hidden cost, and (possible) insult to those served, whose preferences were ignored, are rarely acknowledged. If program planners are working to change this state of affairs—and some are—they need the attendant efforts of an evaluation approach designed to recognize the problem of interest: not whether the program did what it was supposed to do (according to the program designers), but whether it did what was desired by those it served.

2.1 | Shortcomings of current IH program evaluation

As overall IH funding has risen in recent decades, evaluation of IH program efforts has achieved greater prominence.[13-15] Mainstream methods generally approach evaluation as an engineering process, wherein evaluators take objectives as given and focus on whether the program in question achieved its stated objectives.[13] Current evaluation norms emphasize scientific criteria for ascertaining evidence of causal relationships between programs and health outcomes. This dominant technocratic approach aligns well with the biomedical paradigm.[16] It also fits well with the stated preference of some funding agencies, including the Bill & Melinda Gates Foundation, for technology-

driven solutions that can be decontextualized from social, political, and economic determinants.[16] Further, the biomedical paradigm and engineering approaches to evaluation tend to treat health programs as interventions that can be isolated from any community and applied universally. Thus, the goal of evaluation is to identify generalizable "best practices" beyond the specific context in which they were implemented. Health programs, however, are integrated into health systems, which in turn are social institutions reflecting a particular society's norms and values.[17] Viewed this way, the whole horizon of evaluation changes. When analyzed through a political economy lens, the technocratic approach avoids some important questions:

1. Who determines what is a "successful" IH program?
2. Who decides which program(s) a nation/community should implement?
3. Who decides the objectives a program should achieve?
4. Who decides the direction and priorities for health development in a nation/community?

Calls for better evaluation have not, for the most part, asked whether programs being evaluated are aligned with the priorities of the people and communities that they impact. Instead, external funding agencies frequently unilaterally determine which programs and countries they fund, what the program objectives will be, and what evaluation criteria will be applied.[18] Ample evidence suggest that countries and/or regional communities often have little or no voice in selecting program objectives, let alone power to prioritize programs to be funded.[8,18-21] Particularly stark misalignments between funding directives and community priorities have been revealed by MacKeller[22] and Schiffman[23], whose respective studies both concluded that levels of IH funding correlate neither to the priorities expressed by recipient countries identified in Poverty Reduction Strategy Papers nor to the burden of disease. The World Health Organization (WHO)'s Maximizing Positive Synergies Collaborative Group has published similar conclusions including, for example, how external funders chose to direct more than 40% of the total external health funding given to Cambodia towards sexually transmitted infections such as HIV/AIDS, despite a national priority to improve primary health care overall.[1] Other researchers have qualitatively examined how international funding agencies shape policy and programs, sometimes against the priorities of researchers, providers, and officials implementing health programs.[24]

In her assessment of funders' influence in setting global health research agenda, Sridhar concluded that the priorities of funding bodies largely dictate which health issues and diseases are studied.[7] One particular financing mechanism Sridhar noted is *multi-bi financing*, in which donors choose to route non-core funding—earmarked for specific sectors, themes, countries, or regions—through multilateral agencies and to the emergence of new multi-stakeholder initiatives such as the Global Fund to Fight AIDS, Tuberculosis and Malaria, and the GAVI Alliance. This financing mechanism has accounted for most of the increase in global health funding in the last decade. In the example of WHO funding, between 1998 and 2009, its total biennial budget more than doubled from US$1.647 billion in 1998 to 1999 to US$4.227 billion in 2008 to 2009, yet the core budget during this period remained relatively flat, and by 2006 WHO's extra-budgetary funding was nearly 3 times of its core budget. The importance of this, as Sridhar points out, is that extra-budgetary funds do not need to be allocated in accordance with the wishes of member states, or using the same transparent mechanisms for determining program funding. Extra-budgetary funding decisions are at the discretion of donors, while core budget funding allocation is decided by the World Health Assembly. In 2008 to 2009 WHO allocated 25% of its regular budget for infectious diseases, while 60% of extra-budgetary funding of the same biennial was allocated to infectious disease, reflecting the priorities of donors.[25] Similar phenomenon in funding allocations has also been demonstrated with the World Bank.[5,7] The impact of this shift in financing on research priorities was summed up by a former health minister in sub-Saharan Africa nation in this way: "*Everyone is chasing the money—reputable universities, the UN agencies, partnerships, civil society groups, so who is actually doing what developing countries really need, rather than what donors want*?"[26] We would argue the same impact occurs in shaping evaluation: that programs feel pressure to demonstrate success in terms of donor priorities.

Further evidence of the discrepancy between community priorities and donor priorities may be found by evaluating changes in donor funding patterns related to the implementation of WHO's Sector Wide Approaches (SWAps). SWAps were developed and promoted in the 1990s to allow low-income nations greater control over how aid receipts were managed and collocated.[27-29] In their review of SWAps implementation in 46 low-income countries, Sweeney et al[29] found that donors of development aid for health did reduce funding after SWAps were introduced. This shift could be considered an indication that donors do not like losing that control. The authors point out that interpreting this shift is complicated by donors' concurrent perception among greater corruption in the lower income countries; one could argue the lack of support for SWAps was justified by concerns over corruption. Assuming this is the case, there remains an underlying assumption that "donors know best." While it is understandable that some governments—especially authoritarian or corrupt ones—are poor proxies for the wishes of the people they govern, corruption also exists in high-income nations, funding agencies, and NGOs. It may be convenient for donors to pin decisions to move funding away from nationally led SWAps on concerns about efficiency or corruption.

Despite this evidence of systematic practices allowing donors to set the health agenda, the literature on how to conduct program evaluation is largely unconcerned with the fundamentals of program selection and instead is preoccupied with technical expertise in assessing downstream outcomes. Important reference guides like the United States' Agency for Health Research and Quality's Global Health Evidence Evaluation Framework year 2013 focus exclusively on concepts and methods for downstream evaluation. Several reviews of evaluation studies provide no discussion of how the projects evaluated were selected, illustrating that such issues are generally accepted as outside the scope of evaluation.[30,31] This lack of attention to underlying values and priorities as *legitimate* and *necessary* objects of evaluation leads to what we view as a "*golden rule*" with a double meaning in the political economy of IH. In the first meaning, "*he who has the gold, makes the rules*;" in other words, influential funders are able to use their funding to market preferred messages about the relative importance of different health issues.[32] This influence supports the second meaning, where "*those who have the gold, rule,"* through their ability to set the overall agenda of IH.[33,34] We propose Critical IH Program Evaluation as a counterforce to this dynamic. Through reframing the program evaluation framework, it may facilitate IH communities, funding agencies, and countries/communities on the receiving end of aid to reimagine and insist upon community ownership as a new norm to be evaluated in IGH programs.

2.2 | The ethical foundation of critical evaluation

A fundamental concern of upstream evaluation is *who decides what program(s) to implement*. Because such decisions are normative and political, methodologically their evaluation requires specified ethical criteria for assessment. For our proposed framework, we contend that the principles of *human dignity* and *self-determination* are central to the field of IH and hence can provide an ethical foundation for Critical IH Program Evaluation. Self-determination and human dignity are closely linked, and it is our position that the former can only be trumped by a challenging claim from the latter. Human dignity grounds modern conceptions (and legal expressions) of human rights [35] and is a value that both justifies and obligates international collaboration to develop health. The related concept of self-determination was largely developed and codified into a right by the United Nations beginning in the 1940s.[35] Its significance to the field is demonstrated by its inclusion in the WHO's 1978 Declaration of Alma Ata, the text of which clearly articulates values of self-reliance and self-determination.[36] These past articulations of the necessity of human dignity and self-determination to the possibility for health and well-being motivated us to develop an evaluation method capable of considering these values as criteria for program success. Further, we consider health programs as part of a social institution rather than a technical device or market product.[37,38] As a social institution, health programs are inseparable from the cultural, social, and economic context in which they are being formulated. It is in this light that any IH program needs to be conceived and owned by the community where it is to be implemented, which is the essential foundation of this Critical IH Evaluation framework.

2.3 | The practice of critical evaluation

Under our proposed framework, evaluation consists of 3 phases (Figure 1). The first upstream phase evaluates program selection. Program choice is not only a legitimate and necessary focus for evaluation but should be the primary one, because choosing 1 program over another informs whether interventions *can* be effective in delivering the health improvements specified by beneficiaries. The second midstream phase evaluates the selection of objectives and indicators for measuring. The third downstream phase gathers relevant evidence and assesses whether the program achieved its objectives. A comprehensive program evaluation assesses whether a program addressed a community's self-identified priorities *and* whether it achieved its stated objectives.

Upstream IH program evaluation (Phase 1) focuses on 3 key questions: *WHO* decides to implement a program, *HOW* the program was chosen, and *WHY* the program chosen (making the rationale explicit, not implicit). For the question of *who* decides on a program, our ethical foundation argues that, in the absence of strong countervailing claims regarding human dignity, community members should make the final decision (directly or through representatives).[39] This claim of *community ownership*, unlike *community participation*, leaves little room for ambiguity or variation. The focus of upstream IH program evaluation is on *process* rather than outcome, echoing emerging frameworks for priority setting such as Accountability for Reasonableness.[40,41] Thus, this framework departs from the mainstream evaluation framework, which is heavily consequentialist. Alternatively, we may view Critical Evaluation as redefining the "outcome" we are assessing in upstream evaluation: the extent to which the community (in which the program is to be implemented) was able to choose, shape, and thus own the program. In this Critical Evaluation framework, the "process" becomes a crucial "outcome." As more programs introduce community ownership of choosing the program area, Upstream Evaluation will progress as a field by focusing on assessing the methods through which community preferences were obtained and the extent to which they can be judged to reflect a community's will and values. As already referenced, different participatory methods exist—the challenge is not inventing them, but normalizing their inclusion in program evaluation.

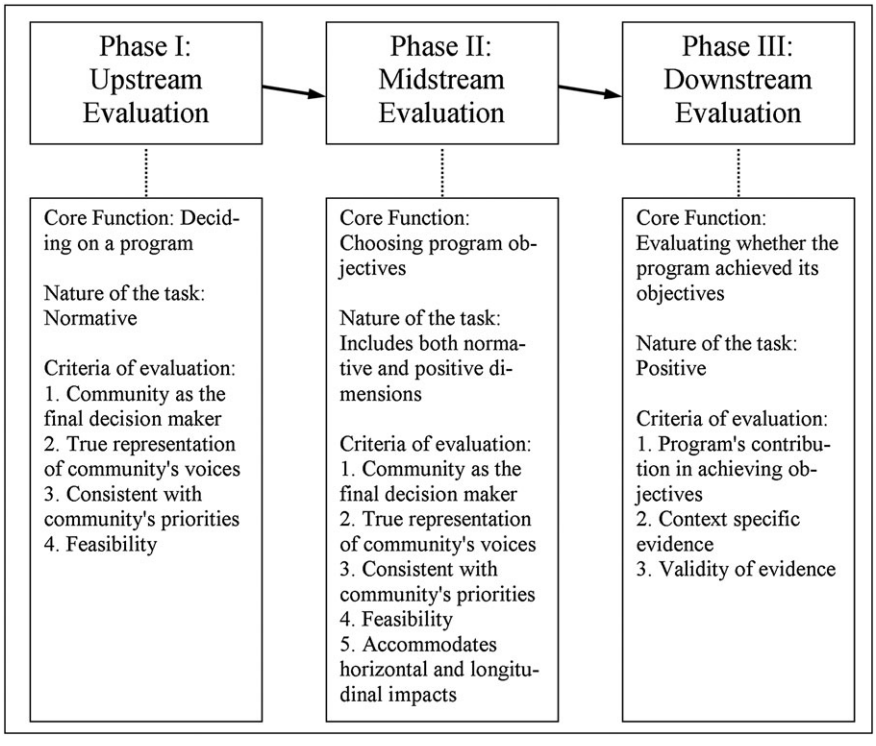

FIGURE 1 The framework of critical international health program evaluation

Upstream evaluation must be conducted prior to midstream evaluation (Phase 2), and the first informs the conclusions of the second. If upstream evaluation concludes that a program was entirely determined by outsiders, rather than giving the community people and organizations most directly affected by it the power of final decision, then it is appropriate to conclude that, from the community's perspective, the program is a failure, regardless of how well it achieves its stated objectives. If the upstream evaluation concludes that the program is a failure, it is irrelevant within this critical framework whether the Midstream or Downstream levels are successful. In practice, evaluators can decide whether they want to continue assessing these subsequent levels.

Midstream IH program evaluation is developed in a manner similar to upstream evaluation. This stage asks the same 3 critical questions, but in relation to the *WHO*, *HOW*, and *WHY* of deciding on program objectives. Given that a program—for example, health infrastructure development—has been chosen by a community, that community should have the final word on which components among the vast array of health infrastructure should be given priority. Public health professionals will still play a large role in the selection of program objectives and indicators, especially in regard to whether the objectives selected are appropriate in terms of feasibility. To do this, public health professionals will need to have a thorough understanding and acceptance of overall community priorities in health, and apply their technical knowledge to provide information on which objectives will be feasible and are likely to best serve those priorities. The focus of midstream IH program evaluation is on both process and outcome. Given that few interventions are shown by evidence to work equally well for all populations or without any socio-environmental unintended consequences, even relatively straightforward scientific objectives, such as reduction of a specific disease's prevalence, require normative choices to be made about priorities, such as whose prevalence reduction to target (those at highest risk, or the general population, or groups for whom the illness would have the most serious social/economic consequences). Public health professionals do not have expertise on value or normative choices and should honor community's decisions.

Only once a program's upstream and midstream evaluations are conducted and passed a minimum criterion of "success" in terms of reflecting community priorities, does a comprehensive evaluation of the critical framework consider downstream performance (Phase 3). Two critical questions remain for this phase of evaluation: whether the program achieved its objectives, and what factors contributed to the program's success or failure to achieve its objectives. Because this phase of evaluation is the dominant form of mainstream evaluation and its methods are well developed in the literature, we will not elaborate upon this phase. This is not to suggest that improvements to more technical evaluation cannot be made; recent work has fruitfully focused attention on the need to better identify, report, and evaluate the unintended harms of health interventions.[42] Rather, our point is that for many evaluators this phase is so well understood and studied that it has mistakenly come to constitute the whole of evaluation. We seek for evaluation to be able to take an upstream approach and a broader view.

2.4 | Implications and next steps

> "The process of defining a problem is identical to the process of finding its solution."[43]

If mainstream IH agencies and academics see the challenges of IH program evaluation as primarily technical puzzles, then future efforts will be concentrated in refining downstream techniques. Continuing along this path, the field of IH development runs the very real risk of providing support to an existing paternalistic world order where the *golden rule* reinforces power structures and relationships that undermine globally shared and articulated principles of self-determination and human dignity.

Evaluation is far from a politically neutral and positive enterprise. In the absence of an upstream evaluation of the appropriateness of program selection, evaluators run the risk of becoming unintentional accomplices to funding agency values, which are not necessarily aligned with community values and preferences.[17] Put more provocatively, evaluators implicitly choose sides and support the *status quo* political economy when they do not explicitly investigate

the upstream aspects of program decision making. While asking *who* makes decisions about program selection is normative, disregarding this question and focusing only on downstream evaluation is also normative.

The Critical Evaluation framework radically departs from mainstream program evaluation and makes an argument in support of "strong" implementation of existing participatory research and evaluation practices. As mentioned, new methods for eliciting community preferences do not need to be invented; rather, they need to be normalized as part of our collective expectations of measuring whether a program was successful. By locating the consideration of community ownership in the evaluation process, there is greater opportunity to focus the field towards fundamental questions of power imbalances in the provision of health services, and to develop public awareness needed to address these imbalances.

We recognize several potential challenges, both conceptual and practical, which must be resolved through further research and practice. One is the extent to which external funders can legitimately claim to be part of a shared global community and thus entitled to a voice in the decision-making process. The lines between "global," "international," "national," and "local" community membership are not always easy to draw, because individuals and organizations may have legitimate claims as stakeholders in each of these categories (for an ethnographic example of the difficulty in defining a "local" non-governmental agency, see Berry [44]). Our framework will benefit from deeper consideration of models of shared agency and governance like that proposed by Ruger [45]. Work needs to be done to develop a new understanding of the right balance between international collaboration and national health sovereignty.[46]

A related challenge is that the Critical Evaluation framework requires a well-defined but adaptable definition of "community," as well as sophisticated methods of soliciting, aggregating, and operationalizing community preferences. Capabilities approaches, originally developed in the work of Amartya Sen [47], engage with the challenges of community definition and align well with our model, because they develop frameworks for valuing the ability of individuals and communities to pursue and achieve objectives of their choosing. Although there have been applied applications of a capabilities approach in policy development, we see opportunities for greater use of this framework in direct elicitation of community values in the context of IH programs. As mentioned, the literature of CBPR, Action Research, and other approaches such as empowerment evaluation also offer practical guidance on how to evaluate the presence or absence of genuine community voices in policy setting.[48,49] The emerging literature around the use of citizens' juries in formulating health policy also is relevant to this challenge.[50,51] Although to date most citizens' juries have been used to inform policy around pre-set health issues, we see potential for such mechanisms as an effective tool for identifying broad health priorities within a national/regional community. These priorities could guide the development of multiple health programs, instead of having to seek community priorities for each newly proposed program. Any such processes used to identify priorities will need to be context specific to accommodate a community's preferences, while also proving able to reconcile different voices within a community, especially those coming from politically weaker groups. We recognize this is no easy task.

We identified 3 major structural barriers for implementing this Critical Evaluation framework. They are the lack of intellectual infrastructure to implement it, the prevailing norms and paradigm of downstream evaluation that implicitly support the golden rule, and the political economy and power dynamic that facilitated and continue to sustain it. To implement the Critical Evaluation framework and change the landscape of program evaluation, underdeveloped infrastructure in our evaluation capacities for community engagement will have to be strengthened. Besides defining who is a member of a *community* when one invokes community ownership, a related term that requires more research and understanding is *representation*. A direct participatory approach of soliciting community preferences and values in program choice might be feasible when the "community" in question is relatively small in terms of population size. When the community is large, such as a province or a nation, direct participatory approaches become less feasible. In these contexts, researchers/evaluators often have to adopt an indirect representative approach to gather community values and preferences. The issue of representation goes far beyond statistical sampling. How a community may identify a legitimate framework of representation requires further research and development. To begin a more intensive infrastructural development on defining "community" and identifying legitimate method of

"representation," both academic and professional worlds of program evaluation will need to shift from the current evaluation paradigm of predominantly outcome legitimacy to process legitimacy.

The lack of development in these infrastructures, however, lies in the current dominant frame of mind in program financing decision and program evaluation. That is, when the main paradigm continues to consider program evaluation as a technical enterprise, what will be funded to research and improve are technical methods of gathering evidence of effectiveness, while the intellectual infrastructures that are essential for Critical Evaluation framework will continue to be ignored and underdeveloped. What is essential for changing the norm and paradigm of downstream evaluation is to shake the foundation of golden rule.

Perhaps, the greatest challenge to implementing a Critical Evaluation framework will be convincing major IH development funding agencies to change course, which may also require the change of power dynamics in IH development financing. If funding agencies prove immune to the effects of a more critically energized evaluation field, an alternate effective strategy to reduce the power of the *golden rule* is to minimize a community's dependence on external gold. The health cooperation within the Global South pioneered in Latin America offers an example of how to reduce dependency on high-income nations and to promote self-determination and regional self-sufficiency.[46,52,53] Arguably, such collaboration, grounded in a regional form of self-determination, might result in more programs that would successfully meet the first and second phase evaluation criteria for success. In the last decade, the political economic dynamics and landscape of IH financing, program selection, and evaluation have moved the opposite direction from what we perceive as conducive to Critical Evaluation framework. The dramatic increase in multi-bi financing, as demonstrated by the growth of donor influence in the H8 (WHO, the World Bank, GAVI, the Global Fund, UNICEF, UNFPA, UNAIDS, and the Bill and Melinda Gates Foundation, see Clinton and Sridhar[5]), has made the implementation of this Critical Evaluation framework even harder. This multi-bi financing model was established to accommodate greater donor control and influence in choosing nations, programs, and their objectives. Parallel to this development is the growing influence of *philanthrocapitalism* that further strengthens the golden rule's second more vicious meaning, "*he who has the gold, rule*." [33,34] Those major financing agencies gained extensive influence without explicit mandates or democratic process, nor accountability.[54]

Although the tide in IH development financing and programming seems to go against the Critical Evaluation framework, there is a strategy to swim against that tide. The simplicity of Critical Evaluation's primary question at the upstream and midstream levels—at its core, does the program seek to deliver the priorities that communities have chosen?—may be difficult to answer exhaustively for all community voices, but it is relatively easy to answer and investigate partially. It can also be done independently from program implementation; although, in our ultimate vision, it will be integrated into program design, during this transition phase its independence is an asset. For those with existing access to the community members who will need to be consulted, it requires relatively limited financial resources to conduct an independent upstream and mid-stream evaluation, making the approach less vulnerable to the golden rule. This article provides researchers and professionals an alternative framework and a strategy to promote Ruger's strategy of norm internalization. If more researchers adopt this Critical Evaluation framework and implement it, there is a strong likelihood that over time, we may put the gold-holders on defensive and shift the norms of program evaluation and perhaps even the financing structures that underpin them.

3 | CONCLUDING REMARKS

The concerns raised in this article connect importantly to central themes in recent IH literature. Our discussion of improving both practice and evaluation mirrors the field's predominant concern with how better governance and inclusion of stakeholders can make aid more effective. [55-58] They stop short, however, of considering national/community priorities as an essential factor in the definition of either "effectiveness" or "success." Further, the concept of *community ownership* is largely absent. We argue that, by turning the framework and notion of "*effectiveness*" on its head, critiques of current practices would be powerfully aided by a Critical Evaluation practice methodologically

enabled to identify a lack of community-identified priority alignment as a mark of failure. While donors might continue to fund such failed programs, this methodological shift would serve to inculcate a new practice and mode of thought among program evaluators, shifting norms that we argue are in dire need of reconstruction. Further, by transforming the framework of IH program evaluation, we may gradually change the academic institutions that educate and train the new generation of IH professionals in their future research and practices. As more IH professionals adopt practices of the Critical Evaluation framework, they may in turn influence the practices of gold-holders.

Advocates for participatory approaches often advance consequentialist arguments that community/stakeholder buy-in achieves better outcomes [59,60] and minimizes unintended harms. [42] While highly sympathetic to this perspective, we make a different argument: it is morally imperative to support communities in deciding their own futures. This is the same argument that the international community adopted in ending colonialism. Our critical framework compels the pursuit of a world consistent with the principles of human dignity and self-determination for all, especially those whose wellbeing is most ignored and challenged in the world today. In this endeavor, rather than conforming to the norm set by gold-holders, we will put them on the defensive. Ruger [45] has proposed approaches that may be helpful to promote upstream IH program evaluation, through persuasion, value change, and norm internalization. These approaches challenge us to reconsider the political economy of the default framework that has shaped the landscape of IH development in general, and program evaluation in particular, constraining our collective intellectual imagination. This article is a first step to persuade, open the way for change, and make norms visible and open to contestation. In the long run, if we are successful, Critical IH Program Evaluation may provide a building block towards the reconstruction of IH governance upon a transformed professional consensus around how programs are funded, planned, implemented, and evaluated. This same framework can also be applied to domestic program evaluation where the "gold holders" might be a national government. Only when this Critical IH Evaluation framework and related critical approaches in planning and research have become the norm in the IH field, can we then expect the gold holders to change their practices, thus liberating low-income nations' health development, together with researchers and professionals, from the yoke of golden rule. Ultimately, it may transform the IH development landscape.

ETHICS STATEMENT

This article does not contain any studies with human participants or animals performed by any of the authors.

ACKNOWLEDGEMENTS

This research was not supported by any funding or grant.

ORCID

Anaïs Tuepker http://orcid.org/0000-0002-9138-5043
Alicia Núñez Mondaca http://orcid.org/0000-0001-5407-5583

REFERENCES

1. WHO. An assessment of interactions between global health initiatives and country health systems. *The Lancet*. 2009;373(9681):2137-2169. https://doi.org/10.1016/S0140-6736(09)60919-3
2. Szlezak NA, Bloom BR, Jamison DT, et al. The Global Health System: actors, norms, and expectations in transition. *PLoS Med*. 2010;7(1):e1000183. https://doi.org/10.1371/journal.pmed.1000183
3. Sridhar D. Seven challenges in international development assistance for health and ways forward. *J Law Med Ethics*. 2010;38(3):459-469. https://doi.org/10.1111/j.1748-720X.2010.00505.x
4. Ottersen OP, Dasgupta J, Blouin C, et al. Global governance for health. *The Lancet*. 2014;383(9917):630-667.
5. Clinton C, Sridhar D. *Governing Global Health*. New York: Oxford University Press; 2017.
6. Kuhlmann AS, Ianonotti L. Resurrecting "international" and "public" in global health: has the pendulum swung too far? *Am J Public Health*. 2014;104(4):583-584.

7. Sridhar D. Who sets the global health research agenda? The challenge of multi-bi financing. *PLoS Med*. 2012;9(9): e1001312. https://doi.org/10.1371/journal.pmed.1001312
8. Sridhar D, Batniji R. Misfinancing global health: a case for transparency in disbursements and decision making. *The Lancet*. 2008;372(9644):1185-1191. https://doi.org/10.1016/S0140-6736(08)61485-3
9. Bartholomew LK, Parcel GS, Kok G. Intervention mapping: a process for developing theory and evidence-based health education programs. *Health Educ Behav*. 1998;25(5):545-563.
10. Schulz AJ, Israel BA, Coombe CM, et al. A community-based participatory planning process and multilevel intervention design: toward eliminating cardiovascular health inequities. *Health Promot Pract*. 2011;12(6):900-911.
11. Reason P, Bradbury-Huang H (Eds). *Handbook of Action Research: Participatory Inquiry and Practice*. London: Sage Publications; 2001.
12. Singh N. Community participation. http://www.slideshare.net/chauhan25/community-participation-and-techniques-ofr-it-oct-2014. Accessed March 1, 2015.
13. Bryce J, Gilroy K, Jones G, Hazel E, Black RE, Victora CG. The accelerated child survival and development programme in West Africa: a retrospective evaluation. *The Lancet*. 2010;375(9714):572-582. https://doi.org/10.1016/S0140-6736(09)62060-2
14. Deghati M. Evaluation: the top priority for global health. *The Lancet*. 2010;375(9714):526. https://doi.org/10.1016/S0140-6736(10)60056-6
15. Murray CJ, Frenk J. Health metrics and evaluation: strengthening the science. *The Lancet*. 2008;371(9619):1191-1199. https://doi.org/10.1016/S0140-6736(08)60526-7
16. Birn A-E, Pillay Y, Holtz TH (Eds). Introduction. In: *Textbook of International Health*. 5th ed. New York: Oxford University Press; 2009:5-11.
17. Mooney G. *The Health of Nations*. London: Zed Books; 2012 17-19, 29-31, 149-150, 190 p.
18. McCoy D, Kembhavi G, Patel J, Luintel A. The Bill & Melinda Gates Foundation's grant-making programme for global health. *The Lancet*. 2009;373(9675):1645-1653. https://doi.org/10.1016/S0140-6736(09)60571-7
19. Brown TM, Cueto M, Fee E. The `and the transition from—international to—global public health. *Am J Public Health*. 2006;96(1):62-72. https://doi.org/10.2105/AJPH.2004.050831
20. Organization for Economic Cooperation and Development. Evaluating development cooperation. Summary of key norms and standards. OECD, Paris; 2008.
21. Vaughan JP, Mogedal S, Kruse S-E, Lee K, Walt G, de Wilde K. Financing the World Health Organisation: global importance of extra budgetary funds. *Health Policy*. 1996;35(3):229-245. https://doi.org/10.1016/0168-8510(95)00786-5
22. MacKellar L. Priorities in global assistance for health, AIDS, and population. *Population and Development Review*. 2005;31(2):293-312. https://doi.org/10.2307/3401362
23. Shiffman J. Donor funding priorities for communicable disease control in the developing world. *Health Policy Plan*. 2006;21(6):411-420. https://doi.org/10.1093/heapol/czl028
24. Behague D, Tawiah C, Rosato M, Some T, Morrison J. Evidence-based policymaking: the implications of globally-applicable research for context-specific problem solving in developing countries. *Soc Sci Med*. 2009;69(10):1539-1546. https://doi.org/10.1016/j.socscimed.2009.08.006
25. Stuckler D, King L, Robinson H, McKee M. WHO's budgetary allocations and burden of disease: a comparative analysis. *The Lancet*. 2008;372(9649):1563-1569.
26. Global Economic Governance Programme, 2008. High-level working group of developing country health ministers, meeting report. Available: http://www.globaleconomicgovernance.org/wp-content/uploads/Working%20Group%20Report%20May%202008.pdf. Accessed January 31, 2011.
27. Cassels A. *A Guide to Sector-Wide Approaches for Health Development*. Geneva: World Health Organization; 1997.
28. Chansa C, Sundewall J, Mcintyre D, Tomson G, Forsberg BC. Exploring SWAp's contribution to the efficient allocation and use of resources in the health sector in Zambia. *Health Policy Plan*. 2008;23(4):244-251.
29. Sweeney R, Mortimer D, Johnston DW. Do sector wide approaches for health aid delivery lead to 'donor-flight'? A comparison of 46 low-income countries. *Soc Sci Med*. 2014;105:38-46.
30. Institute for Health Metrics and Evaluation. Impact evaluations. http://www.healthmetricsandevaluation.org/research/team/impact-evaluations%23/publications-presentations/publications/all Accessed September 7, 2014.
31. Poverty Action Lab. Evaluations. http://www.povertyactionlab.org/search/apachesolr_search?filters=type:evaluation Accessed August 16, 2014.
32. Birn A-E. Gates' grandest challenge: transcending technology as public health ideology. *The Lancet*. 2009;366(9484):514-519. https://doi.org/10.1016/S0140-6736(05)66479-3

33. Birn A-E. Philanthrocapitalism, past and present: the Rockefeller Foundation, the Gates Foundation, and the setting(s) of the international/ global health agenda. *Hypothesis*. 2014;12(1):e8. https://doi.org/10.5779/hypothesis.v12i1.229
34. Birn A-E, Richter J. U.S. Philanthrocapitalism and the global health agenda: the Rockefeller and Gates Foundations, past and present. In: Waitzkin H, the Working Group for Health Beyond Capitalism, eds. *Health Care under the Knife: Moving Beyond Capitalism for Our Health*. Monthly Review Press. Posted in May on http://www.peah.it/2017/05/4019/ Accessed on October 31, 2017.
35. United Nations, 1948. Universal declaration of human rights. http://www.un.org/en/documents/udhr/index.shtml Accessed September 10, 2013.
36. World Health Organization (WHO). 1978. Declaration of Alma Ata. http://www.euro.who.int/en/who-we-are/policy-documents/declaration-of-alma-ata,-1978 Accessed September 2, 2014.
37. Mooney G. *Challenging Health Economics*. New York: Oxford University Press; 2009:171-173.
38. Mackintosh M. Do health care systems contribute to inequalities? In: Leon D, Walt G, eds. *Poverty Inequality and Health*. New York: Oxford University Press; 2001:175-193.
39. Mooney G. The need to build community autonomy in public health. *Aust N Z J Public Health*. 2000;24(2):111. https://doi.org/10.1111/j.1467-842X.2000.tb00121.x
40. Daniels N, Sabin J. *Setting limits fairly: can we learn to share medical resources?* NewYork: Oxford University Press; 2002.
41. Maluka S, Kamuzora P, Sebastián MS, Byskov J, Ndawi B, Hurtig A-K. Improving district level health planning and priority setting in Tanzania through implementing accountability for reasonableness framework: perceptions of stakeholders. *BMC Health Serv Res*. 2010;10(1):322. https://doi.org/10.1186/1472-6963-10-322
42. Allen-Scott LK, Hatfield JM, McIntyre L. A scoping review of unintended harm associated with public health interventions: towards a typology and an understanding of underlying factors. *Int J Public Health*. 2014;59(1):3-14. https://doi.org/10.1007/s00038-013-0526-6
43. HWJ R, Webber MM. Dilemmas in a general theory of planning. *Pol Sci*. 1973;4(2):155-169. https://doi.org/10.1007/BF01405730
44. Berry NS. Did we do good? NGOs, conflicts of interest and the evaluation of short-term medical missions in Solola, Guatemala. *Soc Sci Med*. 2014;120:344-351. https://doi.org/10.1016/j.socscimed.2014.05.006
45. Ruger JP. Global health justice and governance. *Am J Bioeth*. 2012;12(12):35-54. https://doi.org/10.1080/15265161.2012.733060
46. Birn A-E. Remaking international health: refreshing perspectives from Latin America. *Rev Panam Salud Publica*. 2011;30(2):106
47. Sen A. *The Idea of Justice*. Cambridge, MA: The Belknap Press of Harvard University Press; 2009.
48. Fetterman D, Wandersman A. Empowerment evaluation yesterday, today, and tomorrow. *Am J Eval*. 2007;28(2):179-198. https://doi.org/10.1177/1098214007301350
49. Minkler M, Wallerstein N (Eds). Introduction to community-based participatory research. In: *Community-Based Participatory Research for Health: From Process to Outcomes*. 2nd ed. San Francisco, CA: Jossey-Bass; 2008:5-23.
50. Wiseman V, Mooney G, Berry G, Tang KC. Involving the general public in priority setting: experiences from Australia. *Soc Sci Med*. 2003;56(5):1001-1012.
51. Street J, Duszynski K, Krawczyk S, Braunack-Mayer A. The use of citizens' juries in health policy decision-making: a systematic review. *Soc Sci Med*. 2014;109:1-9.
52. Amorim C, Douste-Blazy P, Wirayuda H, et al. Oslo ministerial declaration—global health: a pressing foreign policy issue of our time. *The Lancet*. 2007;369(9570):1373-1378. https://doi.org/10.1016/S0140-6736(07)60498-X
53. Feinsilver JM. Cuba's medical diplomacy. In: Font M, ed. *A Changing Cuba in a Changing World*. New York: Bildner Publication; 2009:273-285.
54. Okma KGH, Kay A, Hockenberry S, Liu J, Watkins S. The changing role of health-orientated international organizations and nongovernmental organizations. *Int J Health Plann Manage*. 2016;4:488-510.
55. Dussault G. A United Nations Global Health Panel for Global Health Governance: a commentary on Mackey. *Soc Sci Med*. 2013;76:16-17. https://doi.org/10.1016/j.socscimed.2012.09.039
56. Hein W. A United Nations Global Health Panel for Global Health Governance: a commentary on Mackey. *Soc Sci Med*. 2013;76:18-20. https://doi.org/10.1016/j.socscimed.2012.09.041
57. Mackey TK, Liang BA. A United Nations Global Health Panel for Global Health Governance. *Soc Sci Med*. 2013a;76(1):12-15. https://doi.org/10.1016/j.socscimed.2012.09.038
58. Mackey TK, Liang BA. Response to comments on—a United Nations Global Health Panel for Global Health Governance. *Soc Sci Med*. 2013b;76:24-27. https://doi.org/10.1016/j.socscimed.2012.09.036

59. De Silva-Sanigorski AM, Bolton K, Haby M, et al. Scaling up community-based obesity prevention in Australia: background and evaluation design of the health promoting communities: being active eating well initiative. *BMC Public Health*. 2010;10(1):65. https://doi.org/10.1186/1471-2458-10-65

60. Mooney G, 2010. A Handbook on Citizen's Juries. http://www.newdemocracy.com.au/item/194-a-handbook-on-citizens-juries. Accessed May 1, 2015.

How to cite this article: Chi C, Tuepker A, Schoon R, Núñez Mondaca A. Critical evaluation of international health programs: Reframing global health and evaluation. *Int J Health Plann Mgmt*. 2018;33:511–523. https://doi.org/10.1002/hpm.2483