

Understanding HIV Program Effects: A Structural Approach to Context Using the Transportability Framework

Megha L. Mehrotra, PhD,^{a,b} Maya L. Petersen, MD, PhD,^c and Elvin H. Geng, MD, MPH^d

Background: Implementation science focuses on evaluating strategies for delivering evidence-based interventions to improve HIV prevention and treatment. The effectiveness of these implementation strategies is often context-dependent and reconciling the desire to produce generalizable knowledge in the face of these contextual interventions is a central challenge for implementation science researchers.

Methods: We provide an overview of the causal transportability theory and conceptualize context under this framework. We review how causal graphs can be used to illustrate the assumptions necessary to apply the results of a study to a new context, and we illustrate this approach using an example of a community adherence group intervention that aims to improve retention in HIV care. Finally, we discuss several key insights highlighted by the transportability theory that are relevant to implementation science researchers.

Results: By adopting causal transportability to consider how context may affect the success of an implementation strategy, researchers can formally diagnose when the results of a study are likely to generalize to a given setting. Moreover, selection diagrams can highlight what additional measurements would be needed in a target population to estimate the effect of an implementation strategy in that target population without having to repeat the initial study.

Conclusions: Transportability translates intuition about context-dependent interventions and external validity into actionable and testable insight.

Key Words: context, transportability, external validity

(*J Acquir Immune Defic Syndr* 2019;82:S199–S205)

Globally, health systems are investing in a range of innovations to improve HIV services, but rigorous scientific evaluation of these strategies must contend with

the fact that context—sociopolitical, organizational, and community characteristics—often affects the effectiveness of these strategies. HIV self-testing, for example, aims to expand testing coverage among hard-to-reach populations, but its success varies across different key populations at risk of HIV infection.¹ Similarly, the effectiveness of differentiated service delivery models, which seek to improve the accessibility and quality of HIV treatment, depends on features of the health systems, clinics, and communities in which they are implemented.^{2,3} For example, community adherence groups (CAGs) are designed to reduce clinic congestion and improve retention in care by allowing members of a group to take turns visiting the clinic to collect medications for the whole group.^{4,5} CAGs are effective in some settings but not others, and these differences are usually attributed to “context.”⁶ Yet, if every implementation strategy is context-dependent, is the scientific task of producing generalizable knowledge through implementation science research impossible?

The broader literature in implementation science expresses an unresolved tension between the notion of context and generalizability. In a review of definitions of implementation science in HIV, Odeny et al⁷ found that 50% of definitions emphasized context, whereas another 50% emphasized the importance of generalizability in implementation research. A later review found 17 distinct frameworks that highlight the role of context in implementation research.⁸ For example, the Consolidated Framework for Implementation Research defines context as the “broad scope of circumstances and characteristics” that affect an implementation effort.⁹ The importance of context for implementation science is widely appreciated, but there is little agreement on what exactly comprises context or how we may produce generalizable knowledge across contexts.⁸ For implementation science research to be useful, a coherent approach that reconciles producing generalizable knowledge in the face of context-dependent effectiveness is crucial.

Recent methodological innovations in formal causal inference offer approaches that could be useful for advancing the implementation research agenda and bringing an analytical framework to context. The causal transportability theory developed by Pearl and Bareinboim¹⁰ is a mathematically-grounded framework used to consider how effects observed in one setting may be applied to another. Rather than describing context as a singular overarching challenge, the framework focuses on the specific contextual factors that are likely to affect a given intervention. In doing so, it lays the groundwork for being able to account for these contextual

From the ^aDepartment of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA; ^bCurrently, Department of Biostatistics, University of California, Berkeley, Berkeley, CA; and ^cDepartment of Biostatistics, University of California, Berkeley, Berkeley, CA.

Supported by grant 5F31 MH111346 to M.L.M. from the National Institute of Mental Health; K24 AI134413 to E.H.G.; and UCSF-Gladstone CFAR SWG Implementation Science Working Group to E.H.G.

The authors have no conflicts of interest to disclose.

Correspondence to: Megha L. Mehrotra, PhD, University of California, San Francisco, 550 16th Street, 2nd Floor #2660, San Francisco, CA 94158 (e-mail: MLMehrotra@gmail.com).

Copyright © 2019 Wolters Kluwer Health, Inc. All rights reserved.

characteristics and apply study results across contexts. Here, we provide a brief introduction to transportability and conceptualize context under this framework. We discuss analytical and research implications of using transportability for advancing implementation science research on the response to HIV/AIDS.

TRANSPORTABILITY

Why interventions (whether “evidence-based interventions” or “implementation strategies”) may vary in effectiveness across different settings is intuitive: if there are certain characteristics that modify the effect of an intervention, and the distribution of those characteristics varies from setting to setting, then the intervention’s effectiveness would similarly vary. Furthermore, if we can measure and account for all the characteristics that both modify the effect of the intervention and differ between 2 settings, then we should be able to predict how effective an intervention would be if it were to be implemented in the new setting.

Transportability formalizes this intuition by building on the theoretical foundations of observational causal inference. In doing so, simple modifications of existing tools and statistical estimators that are widely used in the causal inference literature can be applied to predict how well an intervention may work when implemented in a new setting where it was not formally tested. That is, transportability provides tools to (1) formally evaluate whether findings in one setting could be used to generate valid estimates in another, and (2) if so, estimate what the effect would have been had the study been conducted in the new setting.

Much in the same way that observational causal inference uses directed acyclic graphs (DAGs)⁹ to identify the variables needed to control for confounding, transportability uses similar causal graphs—called selection diagrams—to assist in isolating the important characteristics that determine whether and how the effectiveness of an intervention may differ between the study population and the population to which we wish to apply the results (the target population). Selection diagrams encode formal assumptions

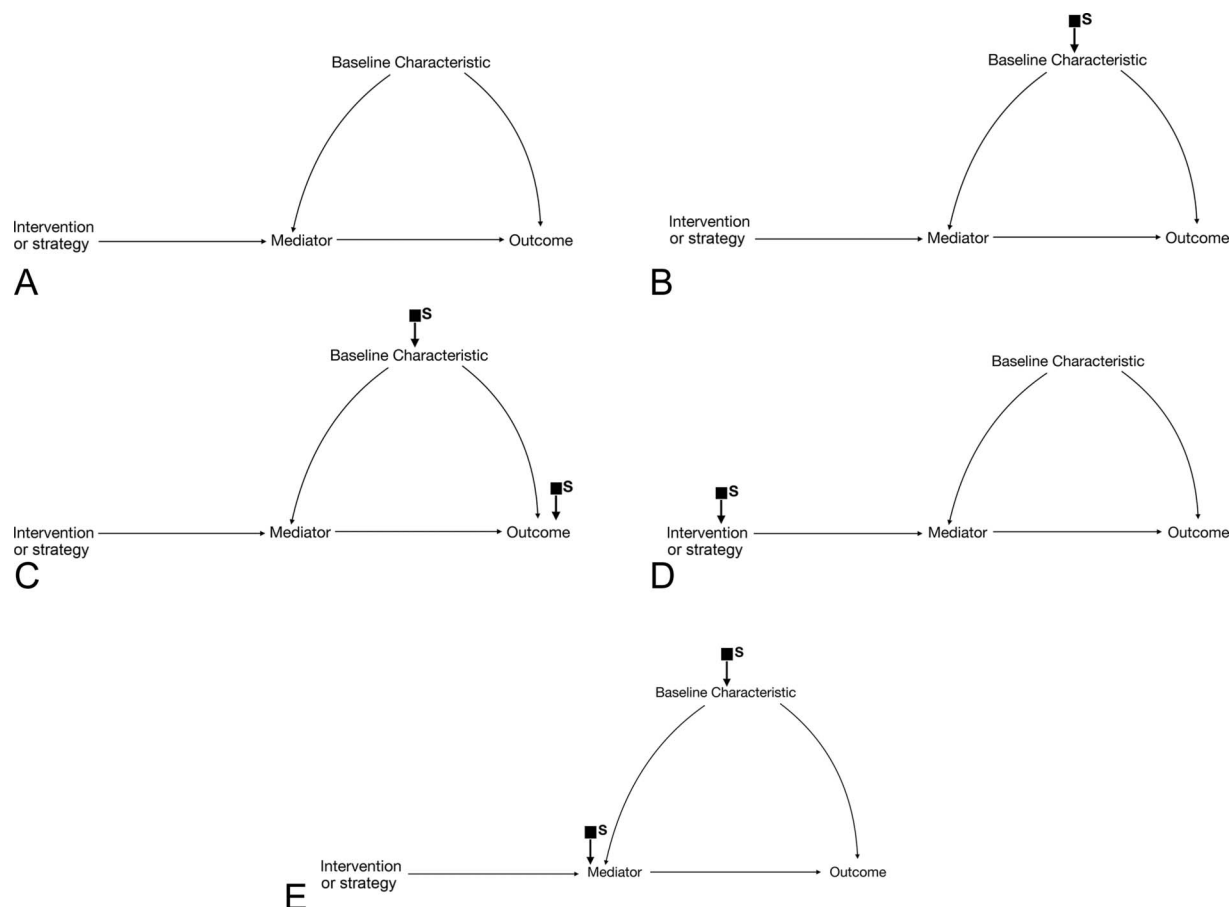


FIGURE 1. A, DAG drawn for the study population. Simplified for pedagogical purposes. B, Selection diagram with a selection node on baseline characteristics. This is transportable. C, Selection diagram with selection nodes on baseline characteristics and on the outcome. This is not transportable (meaning that an effect could be estimated in an external target population). D, Selection diagram with a selection node only on the intervention. This is directly transportable (ie, No additional measurements or adjustments needed). E, Selection diagram that includes selection on the mediator. This is transportable, but would require that the intervention is offered in the target population. The full study would not need to be repeated as the outcome would not need to be measured in the target population.

about the underlying causal relationships and mechanisms through which an intervention is believed to operate in the study population and assumptions about how the target and study populations may differ from one another.

Selection diagrams begin with a traditional DAG representing the study population—paying special attention to the mechanisms through which the intervention is hypothesized to affect the outcome and to any characteristics that may affect the outcome or modify the effectiveness of the intervention (Fig. 1A). Selection diagrams explicitly consider how a target population differs from the study population by including selection nodes indicating these potential differences (Fig. 1B). Unlike standard random variables that are usually included in DAGs, selection nodes do not have probability distributions and cannot be influenced by other variables in the graph. Instead, they function as indicators that point to where the data generating processes may differ between the 2 settings. In other words, they indicate where, if we were to draw a separate DAG for the target population, we may expect the processes that gave rise to the data may differ between contexts. Importantly, the absence of a selection node on a variable indicates that we assume there are no differences in that variable's distribution between the 2 populations given its parents (ie, the most proximal causes of the variable explicitly represented on the DAG).

Selection diagrams reveal first whether the effect of the intervention which was estimated in the study population can be transported to a specific alternative target population given the data available from both the study population and the target population. One approach to this is to measure (and thereby adjust) for enough preintervention variables to render selection nodes independent of the outcome (in a graph in which arrows into the intervention have been removed). In this case, the observed estimate could be used to produce a valid estimate in the target population, and we would deem that the observed estimate is transportable to the target population.¹⁰ Transportability goes beyond this, however, by also helping us to understand which specific postintervention variables can also be used to facilitate transport, and how to use them to derive a specific transport formula to predict the effect in the target population.

In short, encoding knowledge about possible differences in populations and settings through selection diagrams enables a visual assessment of whether the data currently available from both the study and target populations are sufficient to allow for a valid estimate to be generated for the target population, and if not, what, if any, additional measurements would in either the study or target population suffice. We give several simple examples illustrating basic principles here; throughout, we assume that the intervention has not yet been implemented in the target population (observationally or experimentally), and that all variables included as nodes in the diagram have been measured in the study population:

- Selection nodes on the outcome variable itself (Fig. 1C) indicate potential differences in the effect of the intervention on the outcome across populations that cannot be explained by the measured variables. Options for obtaining a valid estimate of the causal effect of the intervention or

strategy in the target population, short of repeating the experimental study, include understanding and measuring additional variables that would explain the differences, and measuring the mediator, outcome, and baseline covariates in the target population.¹¹

- If, on the other hand, no selection nodes point into the outcome or all of the selection nodes only influence the outcome through the intervention itself (Fig. 1D), then the effect of the intervention on the outcome in the study population is identical to what we would expect in the target population, and no new measures or experiments are needed.¹¹ This is known as direct transportability.
- If pretreatment variables alone are sufficient to make all the selection nodes independent of the outcome (Fig. 1B), we need only to measure these variables in the target population. Importantly, this means that we do not need to implement the intervention in the target population to estimate its effect there, and the effect in the target population can be estimated using classic approaches to standardization.¹²
- Finally, when selection nodes point directly at a mediator (Fig. 1E), classic standardization would not be sufficient to transport the effect estimate to the target population. One alternative in this case is to measure the effect of the intervention on the mediator (together with baseline covariates) in the target population. In this scenario, we may not be able to predict the effect of the intervention on the outcome in the target population before implementing the intervention, but we can do so without having to measure the outcome itself in the target population.

By leveraging a formal mathematical framework for causal inference, transportability ensures that as long as the knowledge about the data generating processes and any differences in these processes between the study and target population are accurately represented, the statistical target parameter provided by the transport formula will be equal to the effect of the intervention in the target population. This means that familiar statistical estimation procedures can be readily adapted and used to estimate transported effects in new populations.^{13,14}

CONTEXT THROUGH THE LENS OF TRANSPORTABILITY

Although there is broad appreciation for the importance of the general idea of context for implementation science research, there are few definitions of what specific factors comprise context or how to address contextual differences once we've identified them. By characterizing context in the transportability framework to focus on the specific contextual factors that matter for a given intervention, we can use this information to solve many previously intractable challenges arising from contextual differences.

Defining Context

Proctor's conceptual model of implementation research posits that context can influence the effects of an implementation strategy by affecting several different steps along the

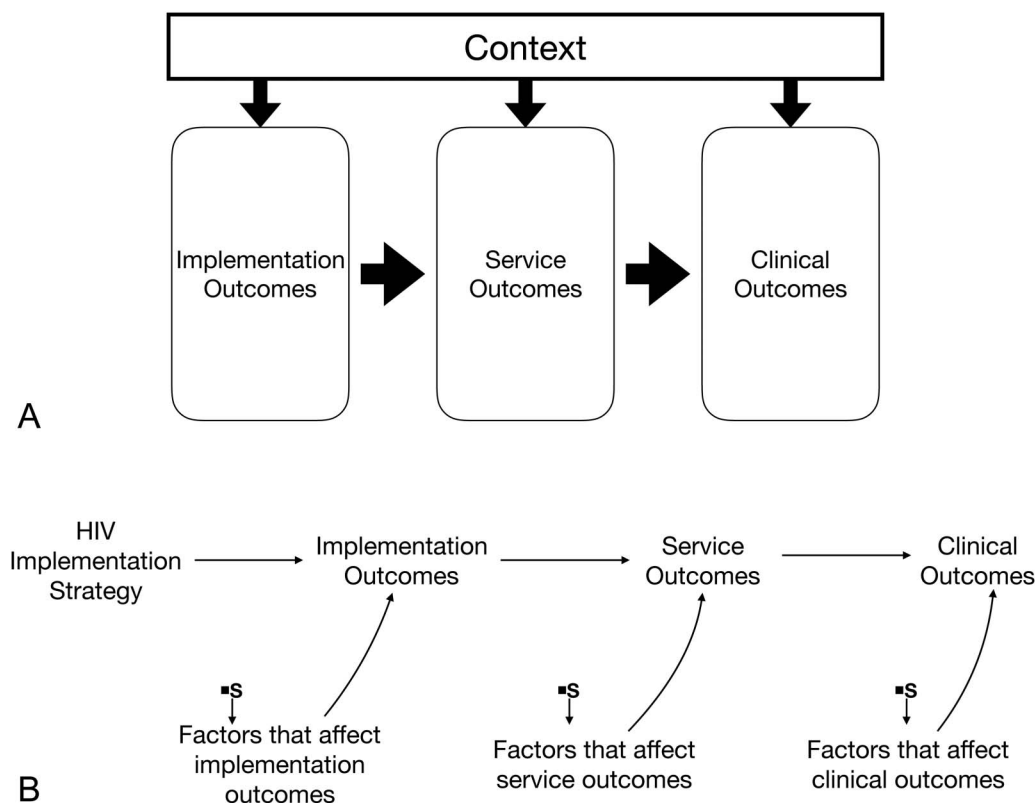


FIGURE 2. A, Proctor's conceptual model of implementation science outcomes. B, Proctor's conceptual model of implementation science outcomes drawn as a selection diagram.

way from design to health outcome: implementation outcomes, service outcomes, or clinical outcomes (Fig. 2A). We represent Proctor's conceptual model as a selection diagram, where implementation outcomes and service outcomes act as mediators of the effect of the implementation strategy on the clinical outcomes. Selection nodes represent context and depict where and how the effects of an implementation strategy may vary between the 2 populations (Fig. 2B).

Example

Consider an example of a large cluster-randomized controlled trial that found that implementing CAGs in clinics improved HIV viral suppression and mortality. We would expect that numerous factors in the outer context, inner context, health care workers, and populations may affect the success of CAG implementation if introduced in new settings. For example, health care worker training and motivation could affect whether the model is offered to patients; acceptance and participation in CAGs may be influenced by perceived HIV stigma, clinic wait times, or how far patients need to travel to the clinic (ie, if they live nearby, then the benefits of reduced visits are less obvious); retention in CAGs may be influenced by patient demographics (such as income or education) or availability of health workers to supervise the CAGs. Finally, comorbidities such as tuberculosis (TB) prevalence may affect clinical outcomes such as mortality (Fig. 3A). In reality, many of these factors are likely correlated

with one another, but we use this simplified example for pedagogical purposes to highlight the role of selection diagrams.

In a simplified example, suppose a study of CAG's was done in an urban environment, and a more rural clinic is interested in implementing CAGs. To inform their decision, they would like to know what the results of the study would have been had it been conducted in their rural clinic. In the target population (the rural clinic), patients tend to live further away than patients attending the urban clinics that were included in the study, but clinic wait times in the rural clinic are typically not as long as in the wait times in the urban clinics. In addition, in the rural clinic, there are fewer available community health workers, the patients often experience higher HIV stigma, and the patient population is generally lower socioeconomic status. The trial and target populations are otherwise similar.

To assess whether we could transport the results of the trial to the target clinic, we begin by drawing a selection diagram illustrating the hypothesized causal relationships between CAGs, the various types of outcomes, and the factors that may influence these outcomes (Fig. 3B). The selection diagram includes selection nodes directed at distance to clinic; clinic wait times; number of community health workers; stigma; and socioeconomic status. By excluding any additional selection nodes, we are assuming that there are no other relevant differences between the 2 populations. Based on this selection diagram, we can estimate the anticipated effect of introducing CAGs into the target clinic just by measuring distance to the clinic; clinic wait times; number of community health workers; stigma; and socioeconomic status

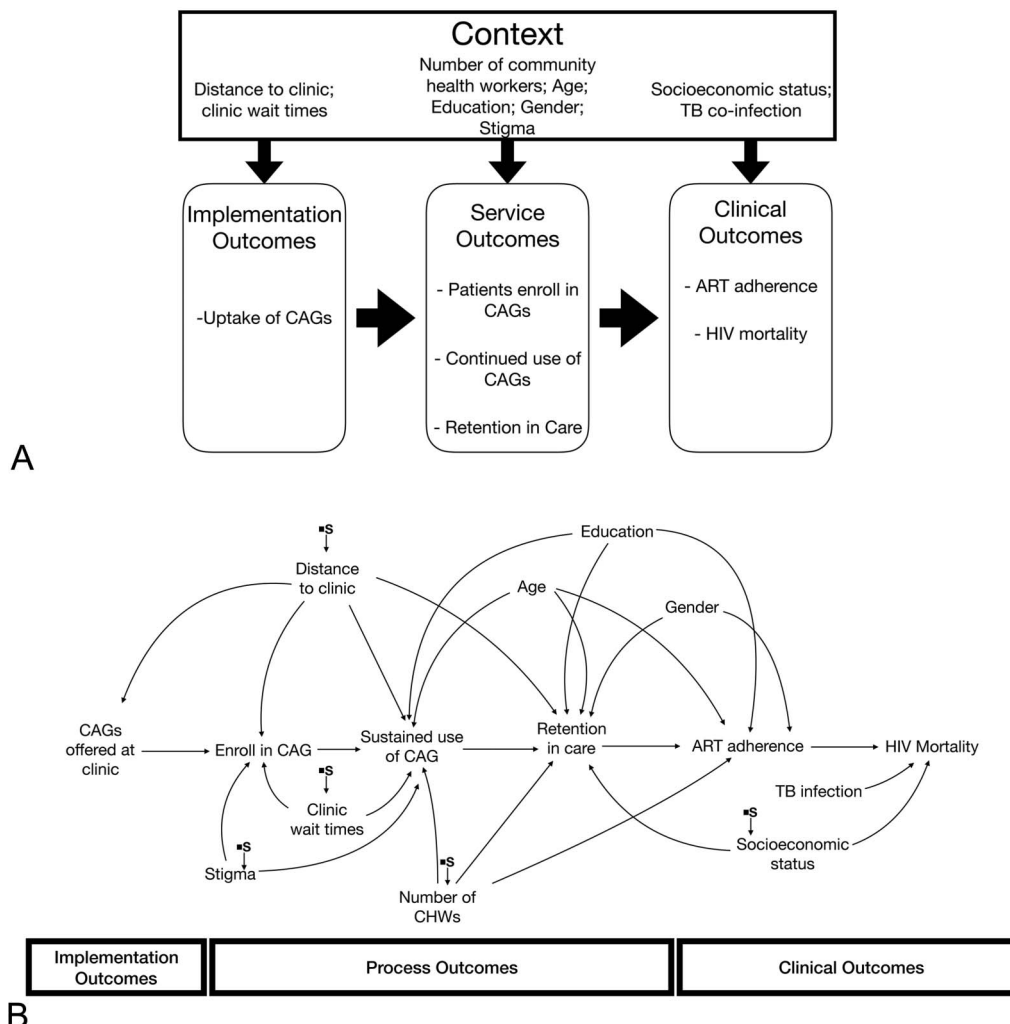


FIGURE 3. A, Proctor's conceptual model of implementation science outcomes for CAGs. B, Selection diagram. Relevant contextual differences are: clinic wait times; distance to clinic; number of community health workers; stigma; and socioeconomic status.

in the target population. Because none of these variables are mediators between CAGs and HIV mortality, we can apply standardization using these variables to estimate the effectiveness of CAGs in the target population.¹⁵ Framing the problem of context through a selection diagram allows us to identify which specific contextual factors are important for generalizing study results to the new population and how we can estimate the anticipated effect of an implementation strategy in that population without having to repeat the study.

IMPLICATIONS FOR IMPLEMENTATION SCIENCE

A Single Finding That Applies Everywhere: A White Whale in Implementation Research?

Whether the results of a trial are transportable to a target population depends on how that specific target population differs from the trial population; a trial may be transportable to some populations, but not others. This demands a more

nuanced view of the traditional idea of “external validity” in which a trial could be considered generally externally valid or not irrespective of who the target population may be. Instead, transportability does not solely focus on the trial, but rather highlights the relationship between the trial population, context, and a specific target population.

In the pursuit of maximizing external validity, some researchers have championed pragmatic trials as a means of evaluating the effectiveness of implementation strategies in usual care settings.¹⁶ However, in the face of heterogeneous effects of implementation strategies, the quest for a single effect that applies universally is hopeless even if the trial design conforms to all the established features of a pragmatic trial (eg, no recruitment restrictions, flexible interventions). So, unless the trial population is a random sample of all future target populations (an implausible concept given that even in the same location, the data generating process may evolve over time), even results of pragmatic trials will need to be formally transported to produce evidence that is relevant for different settings (or the same setting in the future). To avoid

altering standard care practices, pragmatic trials often minimize the number of measurements taken over the course of the study. However, formal transport requires individual-level measurements of variables that modify the effectiveness of the implementation strategy. To the extent that pragmatic trials minimize the number of characteristics measured, they may reduce the chance that their results can be validly transported to external settings, and thus undermine the goal of generating study results that can be applied to a range of target populations. Instead, if the objective of pragmatic trials is to produce more generalizable knowledge, it is essential that these studies understand and measure variables that mediate and modify the implementation strategies being evaluated.

Context Is Contextual

Another key insight that emerges from transportability is that the contextual factors that need to be measured to transport a given implementation strategy will vary depending on the specific trial and target populations. For example, returning to the hypothetical example described in Figure 3, suppose that we are interested in transporting the results of the trial to a different target clinic specializing in TB care. The clinic is located in the same general region as many of the clinics that were included in the trial, but has a much higher number of patients who have TB. To transport the trial results to this target population, we would only need to measure and account for TB prevalence in the trial and target populations (Fig. 4).

Furthermore, there is no guarantee that the specific contextual factors that are important for transporting one implementation strategy will be pertinent for another. This suggests that thinking about context broadly is less useful than focusing on identifying and measuring all characteristics that may modify any portion of the causal path between a specific implementation strategy and the clinical outcome of interest in our trial populations and may differ across contexts; although this does not ensure our ability to accurately estimate how effects may differ in a new setting, it helps move closer to that goal. This helps narrow the scope of what variables we need to understand about a setting to predict how well a strategy may perform.

Learning From Heterogeneity

The most straightforward application of the transportability framework helps us answer the question of external validity: “if I had conducted this study in a different population, what would my results have been?” More broadly, however, selection diagrams provide a flexible language in which assumptions about mechanisms of action and differences between populations are clearly articulated. These diagrams unlock the ability to test our hypotheses and better understand why implementation strategies vary in effectiveness between populations.

For example, transportability can be used to understand heterogeneous results within a single study.^{17,18} If we observe heterogeneous effects across sites or sub-populations of a study, we can use transportability to test our understanding of the sources of discrepant results. For instance, imagine our trial of CAGs included both public and private clinics and that the intervention appeared most effective in private clinics. Suppose patients that attend private clinics are different from those patients that attend public clinics across several demographic factors. In interpreting the results of the trial, we may be unsure whether the differences in effectiveness of CAGs between private and public clinics were driven by these differences in patient characteristics or whether there are other clinic-level differences that impact how well CAGs work.

To address this question through the lens of transportability, we can estimate what the effect of CAGs would have been in the public clinics had those clinics’ patient populations resembled those of the private clinics. We can then compare this transported estimate to the observed effectiveness in private clinics; if the transported estimate is similar to what was observed, then we can conclude that differences in patient characteristics alone are sufficient to explain the discrepant results. If the transported estimate is not similar to what was observed in the private clinics, then this suggests that other factors may have been responsible for the incongruous results.¹⁷

A similar approach can also be used to understand discrepant results across trials of similar delivery strategies. For example, suppose we have 2 studies that both evaluated the use of CAGs for improving HIV care, but the results differed between the studies. If, after drawing

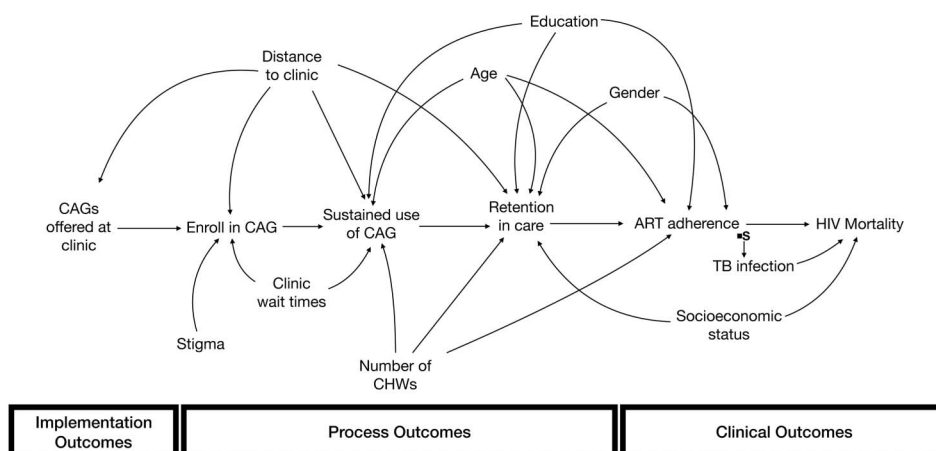


FIGURE 4. Selection diagram. Relevant contextual differences are: TB infection.

a selection diagram, we believe we have measured all the important contextual factors that should explain the differences in effectiveness between the 2 studies, we can transport the results from the first to the second study and compare the transported results to what was actually observed in the second study. If the transported estimate equals the observed, then we can conclude that our measured set of characteristics was sufficient to explain the differences in the results. Otherwise, if the transported estimates do not match the observed results, then we may have omitted an important contextual factor that affects the effectiveness of CAGs.

The information learned from these sorts of exercises is useful for identifying which specific contextual factors are most important for the success of a given implementation strategy, and this knowledge allows future researchers or policy-makers to ensure that the relevant characteristics are measured and accounted for when planning implementation in new target populations. In doing these exercises, if we learn that we have omitted an important contextual factor that affects the effectiveness of the intervention, researchers should prioritize identifying what these other factors may be (through qualitative or other exploratory analyses).

CONCLUSIONS

Overall, transportability theory empowers implementation science researchers to translate intuition about context into actionable and testable science. Selection diagrams are powerful tools that facilitate a deeper understanding of the mechanisms that matter most to the success of an implementation strategy. This approach is not at odds with previous work describing how and why context matters for implementation science. Rather, it allows us to move beyond describing these challenges and applying powerful and emerging insights to carry out analyses, determine what measurements to make and how to conduct science that is externally valid and maximally relevant.

REFERENCES

1. Sabapathy K, Van den Bergh R, Fidler S, et al. Uptake of home-based voluntary HIV testing in sub-saharan Africa: a systematic review and meta-analysis. *PLoS Med*. 2012;9:e1001351.
2. Kaplan HC, Brady PW, Dritz MC, et al. The influence of context on quality improvement success in health care: a systematic review of the literature: quality improvement success in health care. *Milbank Q*. 2010; 88:500–559.
3. Ehrenkranz P, Grimsrud A, Rabkin M. Differentiated service delivery: navigating the path to scale. *Curr Opin HIV AIDS*. 2019;14:60–65.
4. Mukherjee JS, Barry D, Weatherford RD, et al. Community-based ART programs: sustaining adherence and follow-up. *Curr HIV/AIDS Rep*. 2016;13:359–366.
5. HIV/AIDS: community models of care explained | médecins sans frontières (MSF) international. Available at: <https://www.msf.org/hivaids-community-models-care-explained>. Accessed February 12, 2019.
6. Nachega JB, Adetokunboh O, Uthman OA, et al. Community-based interventions to improve and sustain Antiretroviral therapy adherence, retention in HIV care and clinical outcomes in low- and middle-income countries for achieving the UNAIDS 90-90-90 targets. *Curr HIV/AIDS Rep*. 2016;13:241–255.
7. Odeny TA, Padian N, Doherty MC, et al. Definitions of implementation science in HIV/AIDS. *Lancet HIV*. 2015;2:e178–180.
8. Nilsen P, Bernhardsson S. Context matters in implementation science: a scoping review of determinant frameworks that describe contextual determinants for implementation outcomes. *BMC Health Serv Res*. 2019; 19:189.
9. Damschroder LJ, Aron DC, Keith RE, et al. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci*. 2009;4:50.
10. Pearl J, Bareinboim E. Transportability of causal and statistical relations: a formal approach. 2011 IEEE 11th International Conference on Data Mining Workshops, Vancouver, BC, Canada: IEEE; 2011, 540–547. doi: 10.1109/ICDMW.2011.169.
11. Pearl J, Bareinboim E. External validity: from do-calculus to transportability across populations. *Stat Sci*. 2014;29:579–595.
12. Rosenbaum PR. Model-based direct adjustment. *J Am Stat Assoc*. 1987; 82:387–394.
13. Westreich D, Edwards JK, Lesko CR, et al. Transportability of trial results using inverse odds of sampling weights. *Am J Epidemiol*. 2017; 186:1010–1014.
14. Rudolph KE, van der Laan MJ. Robust estimation of encouragement-design intervention effects transported across sites. *J R Stat Soc Ser B Stat Methodol*. 2017;79:1509–1525.
15. Lesko CR, Buchanan AL, Westreich D, et al. Generalizing study results: a potential outcomes perspective. *Epidemiology*. 2017;28:553–561.
16. Ware JH, Hamel MB. Pragmatic trials—guides to better patient care? *N Engl J Med*. 2011;364:1685–1687.
17. Rudolph KE, Schmidt NM, Glymour MM, et al. Composition or context: using transportability to understand drivers of site differences in a large-scale housing experiment. *Epidemiology*. 2018;29:199–206.
18. Mehrotra ML, Westreich D, McMahan VM, et al. Baseline characteristics explain differences in effectiveness of randomization to daily oral TDF/FTC PrEP between transgender women and cisgender men who have sex with men in the iPrEx trial. *J Acquir Immune Defic Syndr*. 2019; 81:e94–e98.