Better Health Care: How Do We Learn About Improvement?
Salzburg Global Seminar is grateful to our program partners: the USAID ASSIST Project and the New Venture Fund for their generous support of Session 565.

We are also grateful to the following organizations for their additional support: the Austrian Federal Ministry of Science, Research and Economy, the Huffington Family, the McKnight Foundation, the Robert Bosch Stiftung and The Nippon Foundation.

Salzburg Global Seminar would like to thank all the Speakers, Discussion Group Facilitators and Resource Specialists for donating their time and expertise to this session and to all the participants that contributed their intellectual capital and superior ideas.
Better Health Care: How Do We Learn About Improvement?
# Table of Contents

05  Introduction  
07  What are we curious about?  

08  Cracking Open the Black Box of Improvement  
08  Marry practice and research  
10  Improve documentation and develop standardized taxonomy  
11  Highlight the intangibles  
12  Differentiate between generalizable and context-specific elements  

13  FELLOW’S INSIGHT – Nancy Dixon  
What elements of an improvement are transferable, and what adaptations are needed?  

15  Shifting the Paradigm Toward a Marriage Between Improvement and Research  
15  The purpose of evaluation  
16  Understanding causal pathways  
16  Matching evaluation to need  
18  Context as an active player  
19  Improving the rigor, attribution and generalizability of the Ghana Water Quality Case Study  

21  Guidance  
21  Framework for learning about improvement  
22  Theory of change  
23  Context  
25  What  
27  How  
29  Evaluation design  
31  How do we design interventions to make evaluation better?  
33  How do we evaluate better to fully capture the value of interventions?  

35  FELLOW’S INSIGHT – Bejoy Nambiar  
Quality Improvement Interventions: Challenges and Solutions for Evaluation Design  

37  Next Steps  

40  INTERVIEW – Cliff Hughes: Compassion is central to improvement  
41  INTERVIEW – Danika Barry: “Patients are the ultimate stakeholders in Quality Improvement”  

Appendices  
42  I: List of Participants  
44  II: Online Recaps, Videos, Interviews and Fellows’ Blogs
Introduction

The field of health care improvement has come a long way from its origins of tackling routine processes of care, such as waiting times. Today, the field of improvement has turned its attention on the most pressing challenges in health care, from reducing maternal mortality to getting more HIV-positive patients on antiretroviral therapy, and is achieving more significant results at scale. With this shift in the magnitude of the problems that the field is being called on to tackle comes an important, and understandable, shift in the questions that the field is being asked to answer.

To address this shift and the challenges and opportunities it poses, Salzburg Global Seminar, in partnership with the USAID ASSIST Project and the New Ventures Fund, held a week-long session entitled: Better Health Care: How Do We Learn about Improvement? Sixty-one participants from 25 different countries on six continents participated in the session, which was held at the Schloss Leopoldskron in Salzburg, Austria from July 10 to 15, 2016.

The 2016 session followed on in part from the 2012 session Making Health Care Better in Low and Middle Income Economies: What are the next steps and how do we get there? which led to the production of the Salzburg Statement Better Care for All, Every Time: How to Make Health Care Better in Low and Middle Income Countries. Both sessions are part of Salzburg Global Seminar’s long-running multi-year series on Health and Health Care Innovation in the 21st Century.

Returning to Salzburg for the fourth time, chair of the 2016 session, Dr. M. Rashad Massoud, framed several issues at the start of the week by presenting the example of India, where over an 18-month period improvement activities covering 270,000 deliveries led to a reduction in perinatal mortality of 12.7%, and challenging participants to consider the following questions:

- How do we know that these results are attributable to the changes that we are making? Do we know that there were no other factors at play?
- If they are in part attributable to what we are doing, how do we know which part?
- How and why did this intervention achieve these results?
- Do we know that similar results weren’t obtained elsewhere where we weren’t working?
- Can interventions that were successful in one context be replicated in another?
- How reliable is the evidence we are collecting and how much evidence is enough to implement these changes at scale?
Faced with these questions, the organizers of the session engaged participants in a five-day discussion on tangible and practical ways to increase the rigor, attribution, and generalizability of interventions designed to improve the quality of health care without compromising the iterative, adaptive nature of improvement. Central to this discussion was understanding the complex adaptive systems in which health care improvement takes place – in other words, “delving into the black box of improvement” – in order to understand how to best learn from these systems and ultimately bridge the divide between the improvement and research communities.

The five-day program was designed as a practical and participatory working session with the intention of developing a framework for increasing the rigor, attribution, and generalizability of improvement. At the outset, the framework was envisioned as a source of practical guidance for improvers and researchers – laying out different potential evaluation models, the strengths and weaknesses of each, and guidance on the most appropriate situations for using each model. Over the course of the week, it became clear that much more discussion would be required in order to develop a concrete framework and the group shifted focus toward developing broader guidance on key domains that evaluators and improvers need to consider in order to maximize the learning from improvement.

This report provides an overview of the discussions held and the key takeaways, but does not offer a verbatim account nor does it highlight every point that was raised. This report was written under the Chatham House Rule and as such comments, quotations, and viewpoints are not attributed to individual participants except where specific agreement to do so has been obtained.
At the start of the session participants were asked to list the questions that they were most curious about with regards to increasing the rigor, attribution, and generalizability of improvement interventions.

1. What are the best ways to document context to better understand which contextual factors influence the effect of improvement efforts?
2. What is the taxonomy of attribution and generalizability based on “maturity” of the intervention and what are the existing appropriate and feasible methods to measure?
3. What questions do we need to ask to adapt interventions, approaches, and tools to a variety of contexts and settings?
4. What are the characteristics of useful generalizable knowledge?
5. How do we systematically turn stories into a framework that can be replicated or adapted responsibly?
6. Why generalize especially if the intervention was context-specific?
7. How do we describe the context? Which features should we document?
8. When is it more valuable to report generalizable vs. context-specific findings?
9. What is generalizable and who decides? Can there be a theory of generalization as a collaborative effort across levels?
10. How do we capture the invisible and untested changes and what level of evidence is good enough to establish attribution?
11. What aspects are attributable and what is the attributable fraction?
12. How do I take learnings and translate them into messages, methods, and principles that would be useful in other contexts?
Cracking Open the Black Box of Improvement

Health care functions within complex and dynamic systems which are ever evolving. In order to produce meaningful change within a health system, and by doing so effect real change in the health outcomes of the people who engage with these systems, the field of health care improvement has evolved to address challenges through iterative and adaptive methods of testing and implementing changes, using real time data collected and owned by frontline teams to inform these changes. Real time collection of data allows improvement teams to make rapid decisions as to whether changes are producing the desired outcome and what needs to be adapted or even discarded altogether. While improvement interventions have resulted in significant process and patient-level outcomes in both developed and developing country settings, how and why these interventions have led to changes in outcomes is not always well understood.

The program in Salzburg began with participants dividing into five groups to discuss case studies intended to help “crack open the black box of improvement.” The case studies were:

- A project in one region of Uganda aimed at improving outcomes for women with preeclampsia
- A Ministry of Health-led initiative in Mali to reduce the high prevalence of anemia in a single district
- A simulation of a large nurse-run antenatal clinic in a busy hospital in rural India
- A community-based project looking at the relationship between the introduction of a new narrow-mouthed household water storage container and improvement in water quality
- A project aimed at implementing a discharge checklist to improve the quality and standardization of discharge protocol at a small rural UK hospital

Casework and the subsequent discussions uncovered several factors that are important to consider in improving the way we learn from improvement.

Marry practice and research

There is currently a significant dichotomy between the world of health system improvers and that of research. The Salzburg Global session began with a recognition of the need for convergence and a call for a marriage between research and practice. Central to this is an understanding that
evaluating improvement requires a different approach from the traditional, fixed-protocol evaluation designs common in clinical research settings which are inadequate for capturing the iterative, multi-faceted nature of working within complex adaptive systems. A “rubber-scaffold that bends in the wind” was called for, underscoring the need for flexibility and adaptability in evaluation design. A new framework for evaluation design that is responsive to the context within which improvement occurs is required, as is better guidance for improvers on how and when to integrate evaluators into the project lifecycle. The group debated the merits of having internal evaluators embedded throughout the improvement process (versus the more typical approach to academic research where an external evaluator comes in at the end of a project) as a means for deriving generalizable knowledge for broader social good.
**Improve documentation and develop standardized taxonomy**

A recurrent theme throughout the discussion and analysis of the improvement case studies, and indeed the entire five-day seminar, was the importance of adequately describing both the improvement process and the intervention, and developing a standard taxonomy for doing so. The group agreed that in order for the field itself to improve there needs to be a concerted effort to elevate the quality of reporting on improvement activities. Questions arose as to what data is needed to both understand the intervention and to aid researchers in evaluating improvement activities.

Similarly, further clarity and standardization was sought on how to best document improvement processes to make them easily communicable and digestible to others who are hoping to replicate results in different settings. A major emphasis of the discussion was the need to standardize the language used to describe improvement. Another key focus of discussions was also the importance of engaging leadership to ensure the success and sustainability of health care improvement efforts, but there was confusion as to what precisely was meant by “engaging leadership” and how to go about doing so. As one participant noted, an important reason for the continuing gap between the quality improvement community and policymakers is the field’s inability to describe improvement using a common set of principles and taxonomies that are easily understood and relatable to those outside the field and in positions of power.

“We need to capture the story of improvement that we will tell our families at home, not just the story we tell at work.”
Another major emphasis of the discussion on improving documentation was the need to differentiate between what elements of an improvement intervention were essential to obtaining results – the core components – and what elements could and should be adapted to fit different contexts. The discussion of documentation was underscored by the recognition that telling the story of what did not work, the unexpected challenges that were encountered, and the difficulties of making improvement happen, can be equally, if not sometimes more, powerful for improvers looking to learn from and adapt improvement approaches in other settings.

**Highlight the intangibles**

The Mali and Uganda case studies highlighted the importance of measuring and studying the intangible results of improvement activities and the impact that these factors have on the replicability of outcomes in different contexts. In Mali, a quality improvement project aimed at reducing the prevalence of anemia in an agriculture rich district resulted in an increased awareness and community ownership of the impact of anemia in children under five and pregnant women. The case discussion centered on the importance of better understanding the changes in mindset among community members and health providers as the government of Mali considers scaling up this initiative to the rest of the country. Participants felt that while documenting and understanding the details of the clinical interventions was important, more critical to the scale-up effort was finding a way to measure and learn from the changes in behavior at both an individual and system-wide level. Tools to measure shifts in mindset were discussed, as was the importance of linkages between the health system and other sectors, such as education and community development, for the long-term sustainability of outcomes. Similarly, in the Uganda case study an improvement intervention aimed at bettering outcomes for pregnant women with preeclampsia resulted in unintended improvements in malaria and anemia outcomes presumably as a result of an attitudinal shift both by leadership and frontline providers. By giving health care workers the ability to improve and demonstrating the power of collecting and analyzing data, health care providers were empowered to tackle problems that were not specifically within the scope of the project. Spreading these changes, both within the local context and beyond, will require further study and analysis of the impact of increasing pride and ownership of work on health outcomes.

Lastly, and arguably most importantly, is the need to fully understand and document the role that context plays in the success or failure of improvement interventions. As one participant put it: “Context isn’t the backdrop…it is an active player in the game.”

“**We can change the world, but not always in conditions of our own choosing.**”

Danielle Charlet

Bruno Bouchet

Mahesh Devnani
Differentiate between generalizable and context-specific elements

Central to the discussion of the case studies and to the larger goals of the Salzburg session was the question of generalizability and the critical role that context plays in improvement. Participants felt that there was a need to differentiate and specify between what can be generalized and what context-specific elements are useful for learning.

The Leopoldskron Antenatal Clinic was set up as a simulation case study based on a real-life example of a large antenatal clinic in a busy rural hospital in India. As part of the simulation, the case group looked at reorganizing systems to improve patient flow and the care experience. Through this process, the group found that while there was a need to define specific work roles as part of the improvement effort (an action that was context-specific and the details of which are not particularly generalizable), the use of a systematic, structured approach to analyzing and reorganizing care, which led to the conclusion that work roles needed to be better defined, is learning that is portable to different settings. The group also felt that evaluating the process improvements that took place could lead to practical knowledge for those looking to replicate outcomes in other settings. While evaluation efforts typically focus on the clinical outcomes of an intervention, focusing on the method of improvement and answering questions such as “How does a team decide which issues to address first?”, “How do you form an improvement team?” and “How do you promote cohesion and communication among team members?” may ultimately be of more use to facilities and health systems looking to reorganize care elsewhere.

The idea that the actual process of improvement is generalizable was further explored through a case study focused on improving the quality and consistency of the discharge process at a small rural UK hospital through the implementation of checklists. The project itself faced a series of setbacks and, while ultimately successful, exposed the true iterative nature of improvement and the need for an adaptive evaluation approach rather than the standard fixed protocols of traditional research design. Despite general consensus that learning about the process of improvement itself was useful and more easily generalizable across a variety of contexts, the point was raised that the field lacked well-documented processes and methods for learning about methods.
What elements of an improvement are transferable, and what adaptations are needed?

Nancy Dixon
Principal, Common Knowledge Associates, Austin, TX, USA

One of the questions discussed during the program was: “What elements of an improvement are transferable, and what adaptations are needed?”

“Transferable” implies moving the knowledge that was gained in one place, to another place. The term “adaptation” acknowledges that the improvement must change to fit into the new context. It is the process of adaptation that I want to focus on in this blog. The issue is: who does the “adapting”? The customary approach to accomplish adaptation is a two-step process: 1) the originator tells the receiver (often in a presentation format) what his/her team did and why, and then engages in a Q&A with the receiver to explain anything further that is needed; and 2) the receiver takes those ideas back home, then adapts them into his or her own context. In this format, the responsibility for adaptation lies with the receiver.

I suggest that a more effective format is to give the responsibility for adaptation to the originator. In this format, the interaction is not a presentation but is instead a group discussion: collective sensemaking. Rather than starting with the originators explaining what they did, the conversation starts by the receivers outlining their context, what they are trying to accomplish, why that is important to them, and what they have already tried. During this explanation, the originators ask questions about the context to be able to put themselves in the shoes of the receivers as much as possible. Then the receivers ask the originators what they have learned from their implementation that could address the receiver’s specific context.

The receivers are not asking the originators to start with the actions they took, but rather to start with what Perry et al. calls, the core concepts. However, the task of the originators is not to recite the core concepts, but to consider how the core concepts would apply to the receiver’s situation. The actions that the originators took may not work in a new context, but the reasoning behind the actions should be of help. This adaptation activity provides deeper understanding not only for the receivers but for the originators as well. By striving to apply what they have learned, the originators come closer to understanding why it worked.

Because this adaptation activity is collective sensemaking, it is thereby most effective when engaged in by a collective. That is, several members of the receiver’s team are present to engage in conversation with several members of the originator’s team, providing the needed diversity of perspectives. Six principles are applicable in the adaptation discussion:

1. Set aside the time required to reflect together
2. Include multiple perspectives
3. Create a psychologically safe space
4. Connect people through circles
5. Connect before content
6. Make ideas visible

The adaptation format can be made even more powerful, if the originators are each from a different
group that has successfully implemented the intervention. If multiple originators are present, both they and the receivers can begin to ferret out the core concepts from across their different experiences. The goal is then twofold: first, to understand the core concepts and then, to adapt the intervention to the receiver’s context.

This format of giving the originator responsibility for the adaptation is effective because of the nature of our tacit knowledge. When working in a complex system, much of what we are learning is tacit. Although we skillfully call upon that tacit knowledge to take action, we are often hard pressed to articulate the reasoning behind those actions. A story told by Gary Klein in his book, Sources of Power, provides a useful example of this paradox. He recounts an incident of a lieutenant of a fire crew acting in a dangerous situation.

It is a simple house fire in a one story house in a residential neighborhood. The fire is in the back in the kitchen area. The lieutenant leads his hose crew into the back of the building to spray water on the fire, but the fire just roars back at them. “Odd,” he thinks, “the water should have more of an impact.” They try dousing it again and get the same results. They retreat a few steps to regroup, then the lieutenant starts to feel as if something is not right. He doesn’t have any clues; he just doesn’t feel right about being in that house, so he orders his men out of the building. As soon as his men leave the building, the floor where they had been standing collapses. Had they still been inside, they would have plunged into the fire below.

Klein reports that in an interview with the lieutenant it was only after persistent questioning that the lieutenant was able to name the thinking (tacit knowledge) that led him to the decision that saved the lives of his men. His best explanation was to offer pieces of the pattern he had unknowingly put together: “For example, the room was hotter than he would have expected from a small fire, and the fire was very quiet. He subconsciously deduced that there must be an unsuspected basement beneath the living room, and that was where the main fire was; he then “intuited” that his men were in danger.”

As with the lieutenant, much of our tacit knowledge is so deeply embedded that we require the stimulus of a real situation for us to surface it. As Dave Snowden has said, “We only know what we know when we need to know it.” It is for this reason that bringing originators of successful interventions to a real situation where they must figure out how to solve the problem, allows them to surface knowledge that they would otherwise not be able to articulate. As Donald Berwick explains: “Experimentalists have pursued too single-mindedly the question of whether a [social] program works at the expense of knowing why it works.” I suggest this adaptation format as a process that focuses on “why it works.”

It would be too much to expect the first originators of a successful intervention to travel to every site that requires the originator’s knowledge to be adapted. But imagine if every group that became successful, perhaps through engaging in this adaptation activity, was willing to “pay it forward” by engaging in an adaptation session with others. Soon there would be many conversations and a wealth of knowledge growing about the core concepts. All would be researchers as well as implementers.

Resources:
- Gary Klein, Sources of Power: How People Make Decisions, MIT Press, 1999
- Dave Snowden, “7 Principles of Knowledge Management”, Cognitive-edge.com
Shifting the Paradigm Toward a Marriage Between Improvement and Research

After unpacking the black box of improvement, the group considered how the rigor, attribution, and generalizability of the five improvement case studies could have been strengthened, and then moved into a discussion of the successes and failures of past efforts to evaluate improvement using three additional case studies. Evaluation experts were called upon to lead a comparative discussion of different evaluation methods and their applicability to improvement.

The purpose of evaluation

Having a clear understanding of what to evaluate, for whom, and to inform what decisions was repeatedly stressed as one of the most important factors to consider when developing an evaluation design. Defining an evaluation aim, audience, and purpose can set the parameters for questions on attribution, as well as on generalizability. These parameters then allow for teams to select the most appropriate and rigorous evaluation design to accomplish the agreed upon learning objectives. Other factors to consider are the objectives of the funding agency, the time available to complete an evaluation, and the resources required. As part of the evaluation design, it is important to describe the context and intervention with considerable detail, construct adaptive logic models, evaluate whether or not the right indicators and the right monitoring mechanisms have been chosen, and explore the use of existing data and data collection systems.
Understanding causal pathways

A significant portion of the discussion focused on the importance of understanding the causal pathways that underlie the improvement effort. Whereas most clinical interventions have relatively straightforward and linear causal pathways, improvement typically involves complex interventions where the causal pathways are not always clear. This necessitates different mechanisms for linking interventions to outcomes that are able to capture the complex root causes and solutions at the heart of improving health care delivery. Key to this is finding evaluation mechanisms that are responsive to the highly adaptive contexts where health care improvement occurs as opposed to the more consistent settings of clinical drug trials. In clinical trials, interventions tend to be much more consistent and the main effect much more important than the contextual factors. In improvement, however, context carries a lot more weight as the focus centers more on “how” an intervention works rather than “if” it does. Improvement interventions are typically already known to be efficacious. The suggestion was made that the goal of the session be to provide “insights” and “guidance” to improve research and evaluation of quality improvement, rather than a means for generating hard evidence; however, it was also noted that strength of evidence is an important consideration for answering questions on attribution and generalizability.

Subsequent discussions raised the point that the core components of research are also the core components of the improvement process and that while engaging in quality improvement we are essentially trying to answer two questions: 1) What is the best way of improving processes and equally important, 2) What is the best way of learning? The discussion ended with a reminder that human beings are at the center of what we are dealing with when we attempt to improve health systems and as such complexity is inherent in the process.

Matching evaluation to need

It is important to parse out the core elements of evaluation models, determining what elements are essential to the model and what can be adapted to fit the context before committing to an approach. The group discussed three phases at which evaluation can and should occur: the innovation phase where there is a low degree of belief that the change ideas being tested will lead to the desired outcomes; the testing phase in which there is a moderate degree of belief; and the scale-up phase in which degree of belief is high. Each phase requires a different evaluation approach as there are different questions that need to be answered. The group discussed the differences between content theory and execution theory as the basis of a good logic model, wherein content theory helps explain how we predict that

“We need to distinguish not between simple and complex but also between complicated and complex. Sending a spaceship to the moon is complicated, raising a child is complex.”

Karen Wolk Feinstein

Ezequiel Garcia Elorrio
change concepts and improvement drivers will lead to desired outcomes and execution theory explains what improvement teams will do that will lead frontline workers to adopt the changes described by the content theory. Distinguishing between content and execution theory and evaluating both throughout the innovation, testing, and scale-up phases allows us to answer the questions “What are we confident about?” and “What do we need to learn more about?”, and allows for further refinement of the evaluation model.

The group debated the benefits and challenges of embedding evaluators throughout the improvement process. While many felt that having an evaluator embedded as part of the process would help the team not only define their theory of change and expected effect size but also improve the improvement itself (through better use of tools such as the Plan-Do-Study-Act (PDSA) cycle, for example), others contended that opening up the door to having evaluators dictate how improvement is conducted is a slippery slope that could compromise the integrity of the improvement effort. Participants offered their experiences with varied levels of embeddedness, describing situations in which evaluators worked in tandem with improvers, but without actively intervening. Overall, there was consensus around a need for a continuum of evaluation models which can be customized to the requirements of improvers, funders, and context, and which range from randomized-control trials as an example of a method designed to provide strong evidence of effectiveness, to methods that are more adaptable to change but focus less on “if” an intervention works and more on the question of “how” an intervention works. A key consideration in choosing a model is if the evidence at hand is strong enough for generalization and replicability of results or if further evidence is needed, and whether the methods match the needs and resource availability.

“All models are wrong, some are useful.”
Lastly, it was noted that the cost of an embedded evaluation needs to be taken into consideration and that there is an increasing trend from funding agencies to no longer separate funding for implementation from funding for evaluation, necessitating a more efficient and cost-effective way of going forward.

**Context as an active player**

A recurring topic throughout the five-day session was the importance and influence of context. As was repeatedly noted, context is not a passive backdrop for improvement activities but rather an active player in its success or failure. Participants noted that context as it relates to improvement is often hard to evaluate as context is often times modified or artificially created to provide extra resources that enable a change idea to work (especially during the “innovation phase” where change ideas are being piloted on a small scale). As a result, these environments are not indicative of whether a change idea will lead to improvement in the “real world.”

The idea of adapting the context to the innovation was also raised. Where often discussions on adaptation revolve around modifying interventions to fit a given context, it was noted that perhaps we need to be asking whether there are specific and generalizable aspects of a particular context that are favorable to improvement and whether we can adapt context to incorporate these aspects.
Improving the Rigor, Attribution and Generalizability of the Ghana Water Quality Case Study

On the second and third days of the Salzburg Global Seminar session on Better Health Care: How Do We Learn About Improvement? participants reconvened in the case study groups from the first day to discuss how to improve the rigor, attribution, and generalizability of their cases.

The groups underwent a “reverse knowledge café,” receiving visits from five expert faculty in succession who advised them on what to consider in their evaluation approach. Below is summary of the discussion from one case group.

Summary of case
Poor water quality in Ghana is a source of illness, including diarrheal illnesses. Improving water quality is an important public health goal. The goal of the project was to improve the water quality in communities in Ghana. While water quality can be affected by what is happening at the source of water collection as well as how water is transported, the group focused on household water storage. Based on prior evidence and baseline data, the project introduced a change package including a narrow-mouth water container thought to reduce E. coli concentrations and protocols for cleaning, maintenance, and storage. Based on the results of the pilot in Ghana, the project is now planning to implement the intervention in Burkina Faso.

Aim
To improve the quality of water at the community level in Burkina Faso, using the experience from Ghana.

Using a community engagement approach, the group developed an acceptable change package that showed high uptake, a significant reduction in self-reported diarrhea, and a reduction in household water contamination that did not reach statistical significance. Based on this, a decision was made to spread the intervention package to Burkina Faso using the same community engagement approach. The group then focused on finding a suitable evaluation design for the Burkina Faso effort.

Approach
Based on the learning from Ghana, the group suggested having a researcher join the improvement team to help design the effort in Burkina Faso. The group felt that this would allow for better implementation and better evaluation and would generate the knowledge needed for further scale-up – knowledge that was not captured by the Ghana pilot. It was recommended that another pilot be conducted in Burkina Faso before any further scale-up. The group then turned toward choosing an evaluation design based on the following considerations:

1. Who are the customers of the evaluation and what evidence do they need?
2. How confident do we need to be that outcomes were due to the interventions?
3. What is the local causal-chain (logic model) used to identify the outputs/intermediate outcomes used as measures of success?
4. Do we need a control or comparison group, or is a pre-post design enough?
5. What mix of qualitative and quantitative data is needed for measuring adoption/acceptability/reach?
6. When are the results from the evaluation needed?
7. What is the budget for the evaluation?

Continues overleaf
Suggestions for increasing rigor, attribution, and generalizability in this case:

Rigor
- Adopt a mixed methods approach with a pre-post design to answer questions about generalizability and attribution for extension to other countries beyond Burkina Faso
- Because of our knowledge from Ghana, we can get attribution information without a control group through a sequential explanatory mixed method design

Attribution
- The following data is needed to demonstrate attribution:
  - Pre and post E-Coli rates
  - Qualitative data on use of narrow-mouthed container and proper cleaning

Generalizability
- Study the results of the evaluation using contextual data and comparison with Ghana

Conclusion
The group concluded that the following constraints to the evaluation needed to be considered: time, geography, community capacity for E. coli monitoring, and budget for the evaluation. Overall, the case study yielded the following lessons of value to the field at large:
- Involving an evaluation specialist from the beginning will allow for development of a customized design that is rigorous enough to answer stakeholders’ questions
- Involve the community in both the design and evaluation process

Further case studies are available online: [www.SalzburgGlobal.org/go/565](http://www.SalzburgGlobal.org/go/565)
Guidance

Participants spent the last day of the Salzburg session in small groups focused on synthesizing the discussions and lessons learned over the week in order to generate guidance for improvers on how to better evaluate their work in a way that increases the rigor, generalizability, and attribution of results. The main audience for the guidance is improvers themselves with the aim of cementing the “marriage” between improvement and research that began in Salzburg. Importantly, the guidance aims to provide improvers with tools to integrate evaluation into their work in a way that convinces the larger community that improved outcomes are a result of the changes being implemented. The proposed guidance fell largely into the following five domains which constitute a framework for learning about improvement.

“...We need to conduct research that is of more use to improvers in developing and developed countries. What can we learn from developing countries to take to developed countries where frugal improvement is quickly becoming the order of the day?”
**Theory of change**

Guidance on developing a strong theory of change was thought to be the core principle of whatever eventual framework for evaluating improvement projects came out of the Salzburg session and indeed for all aspects of an improvement initiative and evaluation approach. While some elements of the guidance in this domain are unique to theory of change many cross across the other four domains. The Theory of Change group set out the following parameters for the guidance it would develop:

1. That it would provide guiding questions or statements accompanied with a range of examples segmented by improvers and investigators;
2. That it would provide links to resources, training materials and further reading;
3. That it would clarify terminology by providing operational definitions for key terms and concepts; and
4. That it would provide a communication plan and recommendations for various stakeholders to engage them in the design and ownership of the Theory of Change guidance.

While the group did not emerge with draft guidance statements, several important features were discussed that could be useful in the production of guidance following the seminar. Firstly, the group envisioned a product that could be used to engage stakeholders in addition to guiding improvers in designing their project and evaluation. Secondly, the group felt it was important to shift from using theory of change as a model for evaluation to using it as part of the intervention itself, thus marrying what has typically been a tool of researchers into the field of implementation, recognizing that a theory of change, like improvement, is not a static concept but rather dynamic and adaptable as learning is generated by the improvement effort. The theory of change should include both the core elements of the intervention that have strong evidence of effectiveness and other elements that might require adaption based on the context and organizational culture. Furthermore, the theory of change should also include specific details on other domains such as the “What” (the intervention), the “How” (the implementation method), and the “Context.” The importance of defining these terms and being clear on terminology was reiterated.

Lastly, the group stressed the necessity of understanding the social side of the theory of change, involving diverse stakeholders such as policymakers and frontline workers in documenting a theory of change in order to build trust and a shared understanding of the different perspectives and drivers that make up the human side of improvement.
Context

Context was acknowledged as one of the most challenging factors to adequately document and capture into generalizable guidance for improvers and evaluators. The Context group specified four key principles of context to be included in a guidance document.

Context is dynamic

Context is an active player and requires constant evaluation. Improvement can change context, necessitating exploration of key contextual factors before, during, and after the intervention. The contextual issues provide the framework within which the improvement and evaluation design operate. Adaptive design of the improvement intervention is required while recognizing the possible complications of constant adaptation. Context is part of the improvement process as well as part of the analysis and evaluation – a flexible scaffold that changes and develops as improvers and evaluators gain deeper understanding of its nuances.

Understand the underlying assumptions and beliefs

Understanding context requires investigation of the underlying assumptions, beliefs, and attitudes of improvers, health care providers, patients and their families, as well as other key stakeholders such as policymakers and donors. The group stressed the importance of analyzing these assumptions to understand why they exist and constantly monitoring them for changes. As part of this, improvers should consider hierarchies and how they can act as enablers or barriers to changing assumptions and beliefs. In addition, the group suggested that studying resistance to change within a population would provide valuable knowledge to inform the change effort and the
evaluation. Given the need to ensure the sustainability of improvements, it is also critical to understand which environments are more or less conducive for long-term continuity. Factors to consider include the policy landscape, donor expectations, organizational priorities, individual provider priorities, the attitudes and expectations of patients and families, and the infrastructure and resources required to sustain the intervention.

Study the key elements of context
Context is a rich and broad domain but there are common elements across contexts that are known to exert significant influence, and often barriers, to the success of improvement efforts. These include stakeholders; culture and underlying belief systems; regulatory influences from government, professional bodies, or internal guidelines; past and present experiences of change, both positive and negative; resources for improvement and evaluation, including time and space for improvement, human resources, and funding; political and policy pressures; and the critical need for change as a result of external pressures, internal events, or other factors. In addition, incorporating a strategic assessment of both team and management buy-in and readiness for change into an analysis of the key elements of context could provide valuable insight.

Develop an understanding of the experiences of providers, patients and their families
Understanding past and present experiences of change and the relationships between the key players at the frontline is the critical foundation of understanding context in relation to improvement. The group discussed several factors to consider such as whether there has been previous top-down directives for change that had been unsuccessful, whether there are existing relationships between the key actors that could either positively or negatively impact a change effort, whether the team possesses the technical capacity for improvement, whether there exist local experts who can enhance and facilitate the project, and whether the context rewards or sanctions change.
What

Four categories for defining and understanding the “what” as it relates to evaluating improvement were discussed, each of which set out questions that teams should ask when designing their improvement interventions (see below). A key takeaway from the discussion centered on the interdependence of each of the domains of guidance – that the “what” could not be defined without simultaneously addressing issues of context and implementation – and the need to define and be consistent with terms between each domain. The group discussed the need for concrete examples throughout that could serve as a further guide for frontline teams considering the guidance and help to clarify any confusion that might arise from differing terminology. The notion of maintaining fidelity to the intervention versus fidelity to means of implementation was discussed in relation to defining a set of core components of the intervention. It was noted that it is often easier to maintain and measure fidelity to an intervention than to the way it is implemented.

Background for designing improvement

- What are we trying to address? What have stakeholders explained is the problem?
- What is the general evidence base regarding the problem and what interventions have been used in the past to address it?
- Has a driver diagram and process map been done for this particular problem?
Planning

- Of the drivers identified, which are the most relevant for this particular context?
- What are the specific improvement outcomes that are being targeted and why?
- Who should be involved in designing the “what”?
- What should be tested and why?
- Who should be involved in the testing and what kinds of tests are needed?

Creating the “what” (the change package or theory of change)

- What was tested?
- What data was used?
- What conclusions were drawn?
- Are there a set of core elements that are essential for the success of the intervention and if so, what are they? Are the components that are essential differentiated from elements that can be adapted in another context?
- How are those core elements defined and why are they the core elements? Is the team able to explain to frontline staff why these are core elements? Are the core elements defined well enough to test fidelity to the model?

Implementation phase

- What resources are needed for the implementation of the “what”? In essence, how do we get from the “what” to the “how”?
- Who should be implementing the intervention? What are the roles of each actor/stakeholder in the implementation phase?
- What are the potential implementation challenges?
How

The “How” breakout group began with the overall guidance that a detailed taxonomy for improvement is needed, a theme that was repeatedly stressed throughout the week in Salzburg. The group urged the field to more consistently define and use specified terms, building on existing knowledge. The “How” guidance was then broken into several sub-categories:

Adaptation

Improvers should identify the core components of the intervention that should not be adapted, can be adapted, and should be adapted. Secondly, there is a need to better understand when and how components can be adapted to context, where the context itself can and should be adapted, and how to identify when the intervention is not appropriate for the context. The group stressed the importance of using tools to capture adaptation and fidelity to better evaluate if the adaptation worked.

Improving interventions

The group suggested three priorities for improving the rigor of interventions:

1. Better define and document methods and activities and ensure there is a shared understanding of what each entails;
2. Better define and document what actually happened during the testing and implementation of interventions and find more effective ways of capturing this knowledge; and
3. Pause at appropriate intervals to jointly reflect on learnings and decide what to apply internally and what to adapt.
“[The] question isn’t if the work is practical or academic but rather if the academic work is useful. Is there salience? If there isn’t, there isn’t any point.

Likewise if you have improvement activities that don’t produce generalizable knowledge, then that is a missed opportunity. The onus is on both sides.”

**Improving improvement**

In order for the field to progress there is a need to learn from and develop upon the existing knowledge generated by other fields dealing with complex adaptive systems such as anthropology, sociology, and organizational science and apply those learnings to improvement.

**Sustainability**

Sustainability should be planned for and integrated into improvement design from inception.

**Translating practice into knowledge**

Discover what actual practice has been and develop additional advice and evidence on how best to apply that practice or suggest workable alternatives.

**Patient perspectives**

Use qualitative methods to better capture and incorporate patient perspectives into improvement efforts.

**Cost**

Do a cost analysis on the degree to which improvement interventions are using the appropriate level of resources.
Evaluation design

The Evaluation Design breakout group began by precisely stating what was in the scope of the Evaluation Design guidance and what was not. The scope was narrowed into two mutually-informing domains, particularly targeted for the audience of implementers but also relevant for evaluators – as it underscores the dynamic, mutually-informing relationship between the two.

• How do we design interventions to make evaluation better?
• How do we evaluate better to fully capture the value of interventions?

Doing both of these well – in a mutually-informing, iterative, and adaptive way – will likely lead to interventions with more highly embedded evaluations, in a way that could allow implementers to learn more from evaluation and vice versa. The group was solely focused on how to improve interventions inasmuch as they help improve evaluations, and similarly, how to improve evaluation inasmuch as they help to improve interventions, and saw efforts outside of these aims as outside of the scope of their guidance. The group felt it important to clearly state the value of their guidance at the outset and considered areas such as minimizing bias, more robustly understanding and describing casual pathways, gaining a better understanding of interventions, and making improvement more effective, efficient, and equitable as where they could add the most value.
The group realized that implementers and evaluators must have a shared understanding of the theory of change behind interventions in order to understand their potential impact. Discussion on how to achieve this shared understanding led the group to note that there is a spectrum of potential models that vary by how and when implementers and evaluators interact, in other words, by the degree to which evaluation is embedded within implementation. For example, on one extreme of this spectrum are purist, external evaluation models in which a blinded evaluator comes in at the end of the intervention to perform the evaluation. On the other end of this spectrum are highly embedded evaluation models in which the evaluator works with the implementer during the initial intervention planning phase and at every step along the implementation process in an iterative and adaptive fashion, with feedback mechanisms to continuously inform and improve the intervention.

While the first model, which uses a totally independent evaluator, may decrease the risk of bias and thus have greater internal validity, it may generate less useful or relevant knowledge than the latter model. The spectrum of models in between these two extremes could include features like an evaluator who is present with implementers at the beginning, but which involves implementers in the continuous collection and analysis of data to inform interventions, with the evaluator checking in with implementers periodically.

The group stressed that the proposed guidance should not advocate for a monolithic mode for evaluation, but instead, detail a spectrum of models with varying degrees of intervention-evaluation embeddedness, including the pros and cons of each, as well as the associated costs. This would lead to guidance on how to choose a model, for example on the basis of how well understood the causal pathway is, the purpose of the evaluation, the timeframe for the intervention-evaluation, the trade-offs between the pros and cons, as well as the costs.
How do we design interventions to make evaluation better?

Make interventions more explicit – specify the theory of change: this needs to happen a priori, and must be continuously adapted. The components of the theory of change include:

a) **Specify the content theory**: the relationship between the clinical intervention and the outcomes.
   i) Identify proximate impacts that specify the proper interim measures we need to collect in our evaluation.
   ii) Identify distal impacts based on the level of certainty we currently have in our causal model.

b) **Specify the implementation theory**: detail the activities that will lead to behavior change.

c) **Specify the likely impact size of the entire causal chain**: this is important for appropriately powering evaluations, as well as for other evaluation design factors, such as which indicators we need to collect data for. These would include:
   i) The confounders, mediators, moderators and mechanisms, including relevant external factors related to context, and other co-interventions either implemented by the project itself or in projects running in parallel within the same catchment area.

d) **Specify potential spillovers**: that is, the emergent outcomes beyond the intended outcomes. These should include both the predicted positive and negative spillovers.

e) **Specify the timeline of likely effects and impact**.

f) **Specify factors that are likely to impede intervention success**, and develop risk mitigation strategies, which inform how both the intervention and evaluation design will need to be adapted.
Overall there was consensus among the seminar participants that the purely external evaluator model is not particularly helpful for the field of improvement at this current moment, and instead, the outcome of the seminar should be to encourage the field to move toward more highly embedded models which can tell us not just whether an intervention “worked” but why and how it worked and what we learned.

The group felt that the guidance being developed had the potential to greatly advance the field of improvement by transforming the way we think about the roles of implementers and evaluators, shifting the paradigm towards a marriage between the two, and evolving emerging realms of “research practice” or “empowerment evaluation.” It is imagined that building shared understanding and capacity among implementers and evaluators on the models included in the spectrum of intervention-evaluation embeddedness would allow them to better negotiate with funders and policymakers for the use of a design which optimizes learning and improvement impact.

Finally, in addition to the guidance generated, the group stressed the importance of outlining the areas where the field is missing knowledge, and prioritizing these areas for innovation. The following preliminary areas were discussed:

1. Methods for explicitly studying the confounders, mediators, moderators, and mechanisms of effect using quantitative and qualitative data. In other words, developing a toolkit for looking inside the “black box of improvement” as traditional randomized controlled trials and other existing methods do not do this well.

2. A better understanding of how we learn from adaptation and how to design adaptable evaluations.
How do we evaluate better to fully capture the value of interventions?

Evaluators need to understand the theory of change via dialogue with implementers, and document changes in the areas included in the theory of change. This will inform the following areas of guidance:

**Measure what matters**
Measure as close to the end outcome/impact in the theory of change as possible, understanding that the closer we get to measuring the end outcome/impact, the more time, money or difficulty this may involve, and understanding that how close we need to get to measuring the end outcome/impact depends on the level of confidence we have in our causal pathway.

To illustrate this with an example, we can imagine a basic causal pathway in which we predict that our intervention (with specified inputs) will lead to an increase in the percentage of nurses trained in counseling patients on blood pressure medication adherence by 60%, which in turn leads to increased patient adherence to blood pressure medications by 30%, which leads to a 10% reduction in hypertension, ultimately decreasing the incidence of stroke by 2%.

We may not need to measure stroke incidence because we have high confidence (based on existing evidence) of the causal link between decreased hypertension and stroke. We may similarly not need to measure reductions in hypertension if we have high confidence in the strength of the causal link between blood pressure medication adherence and hypertension reduction. However, stopping our measurement at the percent of nurses trained in counseling would inspire less confidence over whether our intervention achieved the intended impact, as the existing evidence for the causal link between nurse training and increased blood pressure medication adherence is not as strong.

Furthermore, even when measurement on the end outcomes/impacts is possible, measurement of the process and input indicators should still be undertaken. Otherwise, we have not added to our understanding on whether the intervention was appropriate or effective, nor have we added to our understanding of the causal model linking the inputs, activities, outcomes, and impacts. The ability to make conclusions on these areas is key for attribution as well as for generalizability.

Finally, while in an ideal world one would measure all of the relevant effects between inputs, process indicators, outcomes, and impacts, and also be able to estimate all of the parameters in our theory of change—in the real world, we must be strategic, and we must also understand that the parameters are often very context-specific.

For example: The evaluation of a study in rural Gujarat, India may find that the effect of increasing nurse education on blood pressure medication adherence counseling was a 10% increase in patient adherence. One of our effect moderating parameters, like supportive hospital leadership, may be highly context-specific. Measuring this leadership parameter allows us to better understand differences in effect size when the intervention is implemented in other contexts with different degrees of leadership support, and thus allow us to adapt our interventions in data-informed ways to achieve the desired results.

*Continues overleaf*
How do we evaluate better to fully capture the value of interventions?

Continued

**Have a plan to capture and learn from heterogeneity.**

**Have plan for data validation and data reliability to ensure data quality.**

**Use methods which account for time and strengthen causal inference, including:**

i) Display data in time-ordered manner.

ii) Use a comparison group that is similar to the intervention group pre-intervention.

iii) Specify the counterfactual especially regarding secular trends.

iv) Understand or explicitly study the confounders, mediators, moderators, and mechanisms of effect using quantitative and qualitative data.

**Use economic evaluation that captures cost implications.** This could include for example, the cost of the intervention, opportunity costs, and positive and negative externalities.

**Consider evaluating long-term effects.**

**Consider the sustainability of the intervention.**
FELLOW’S INSIGHT

Quality Improvement Interventions: Challenges and Solutions for Evaluation Design

Bejoy Nambiar
Researcher, Institute for Global Health, University College London, Malawi

The choice of comparison groups in the evaluation of complex QI interventions remains a challenge both at a conceptual level as well as at operational level.

At a conceptual level, quality improvement models are based on the Diffusion of Innovation concept where the rapid spread of new ideas or practices happens largely by imitation through individuals categorized as innovators, early adopters, early and late majority, and laggards. The “innovators” form a small proportion of this category but depending on their network linkages and interpersonal relationships, they can engage the “early adopters” who then go on to influence the “early majority” and so on. The spread is thus organic and dictated largely by interpersonal network, which is not limited by the boundaries of intervention groups. Thus, comparison groups can be considered as a hindrance to the organic spread of QI intervention.

In quality improvement interventions, the selection of groups or clusters (typically these are facilities within a health system) are influenced by their level of commitment or “readiness” to be part of the intervention. Such facilities are bound to possess individual and organizational characteristics that are inherently different from their comparison groups. Thirdly, evaluation designs using a comparison group requires them to be selected at the beginning of the intervention. The diffusion of innovation approach might be a more organic and perhaps pragmatic approach from an implementation perspective, but this could be in conflict with most conventional evaluation designs where the selection of a comparison group happens a priori.

From an evaluation perspective, focusing on only the low-hanging fruits does not tell us if the intervention has the same effect on the “not-so-enthusiastic” facilities. A comparison group gives the opportunity to analyse and understand the mechanisms in play with “innovators” and “early adopters” as well as the “laggards.” However, from an improvement science perspective, psychology plays a role in the organic spread of QI interventions. Focusing on the low-hanging fruit and managing to get a critical mass of facilities that will adhere to QI principles is a strategy that is more likely to influence the intervention acceptance by “later adopters” and “laggards.” This aspect of psychology of influencing a larger stakeholder group by building a critical mass to eventually influence the outcome is difficult to measure using a comparison group. It gets further complicated if the comparison groups have to be randomly allocated.

At an operational level, there are limitations to the identification and function of comparison clusters. The constitution of a comparison group is determined largely by the focus of the evaluation design. Evaluations usually focus on probability design to tell if an intervention works or not, while plausibility designs attempts to answer the “how” question. Thus probability designs focus on the measurement of intervention outcomes while plausibility designs focus on the intervention mechanisms. Comparison groups are usually a
feature of probability designs and helps to measure attribution of the QI intervention to the outcome. The choice of comparison group is therefore very important and is determined by the level at which the intervention takes place as well as the level at which outcomes are measured. QI interventions can focus at different levels, such as improving the processes of care or service delivery mechanism or re-organization of the healthcare systems. In some cases it is a combination of these different levels. Furthermore, these different levels are inter-dependent, thus making it difficult to isolate intervention effects even in the presence of a comparison group. Having a comparison group at a higher level, such as district or organizational level, raises a challenge for adequate sample size for a quantitative analysis, while having a comparator cluster at a much lower level, such as individual service provider level, raises challenges for appropriate comparative sample for outcome measurement.

For complex QI interventions that have long implementation periods, comparison sites are prone to be influenced by the intervention (referred to as “spill-over effect”) and contamination by other interventions cannot be ruled out.

Given the challenges of intervention and systems complexity and intervention fidelity, a general agreement in evaluation is for plausibility designs to be considered alongside probability designs. A little explored area in this regard is the role of comparison groups in plausibility designs where there is a need to understand differences in context and mechanisms between intervention and comparison sites. Unlike control groups in clinical trials, which are methodologically considered to be inert, comparison groups in mixed methods evaluation have a different set of context and mechanisms operating, even in the absence of QI intervention and whose dynamic nature is influenced by time and internal context. Such a comparison group can be part of a research strategy, which combines a theory based evaluation approach along with impact evaluation designs such as step-wedge design and simulation models or Bayesian methods. Recently, realist RCTs have been proposed as an approach to evaluate complex interventions such as QI. In such trials, the choice of comparison groups is likely to be related to the key intervention processes and functions, rather than the precise activities itself.

References:
Next Steps

At the beginning of the five-day Salzburg Global Seminar session on *Better Health Care: How Do We Learn About Improvement?* the organizers described the week as a “wedding ceremony” bringing together the practice of improvement with traditional research and evaluation. At the end of the weeklong seminar, many felt that the marriage had indeed started to come to fruition. Overall, the session ended with a general sense that progress had been made in laying out a path for the future of the field, starting with the refinement of the initial guidance that was developed in Salzburg.

The closing session included heated debate on the role of evaluation in improving improvement itself. Some participants argued against opening up the door to having evaluators teach improvers how to do improvement, while others maintained that by not integrating evaluation into the actual improvement process we are missing an important tool to help refine aims and strengthen the proper and rigorous use of key improvement methods. It was noted that in an ideal world we would not need evaluators as part of the improvement team as everyone would have all of the necessary competencies to make improvement successful; however in reality, frontline teams have been struggling with key aspects of improvement like picking an appropriate
aim and having the courage and skill to measure data correctly – knowledge that currently sits more with the evaluator community than with improvers.

In the end, it was agreed that there is no one right model for the relationship between the evaluator and implementer, but that what the field needs is to have the advantages and disadvantages of different levels of interaction clearly laid out, recognizing that the relationship will ultimately depend on the given context, stakeholder priorities, and funding. As one participant noted, “[The old] model of evaluators coming in at the end and saying thumbs up and thumbs down isn’t really helpful. We all agree on this. Our job is not to say: ‘This is the right model.’ Our job is to lay out the benefits, the costs, the alternative models and push the field in this direction because it generates a lot more value and gets us to a model that doesn’t ask ‘Did this work?’ but ‘What do we need to do to make this happen?’”

The session ended with a discussion of next steps which included the refinement of the guidance begun at Salzburg and the distillation of the concepts discussed during the week into papers to be published in a special supplement of the *International Society for Quality in Health Care (ISQua) Journal*. In addition, the group committed to creating a community of practice that would continue to share their learning and reaffirmed the importance of widespread dissemination of the outcomes of the seminar at upcoming conferences and meetings and, with the leadership of the World Health Organization, at the global policy level.

As the session closed, the group was reminded by one participant of the ultimate goal of the Salzburg session and indeed improvement work more broadly:

“*Our patients – people – that’s who we are here for, to improve their quality of care.*”
Cliff Hughes: Compassion is central to improvement

“Pushing a trolley down the corridor, feeding a patient, administering a pill, or doing an operation - all of these activities relate to the patient,” explains Cliff Hughes, President of the International Society for Quality and Health Care (ISQua) and a senior advisor in Patient Safety at Guy’s & St Thomas Trust in London UK.

“We need to recognize that everyone in the health care system ultimately interacts with the patient,” Hughes elaborated.

Hughes has played two important roles in improvement: direct patient care provider and evaluator. As a cardiothoracic surgeon in Sydney for 35 years, Hughes believes that quality of care is built upon direct engagement with patients in order to fully understand the challenges they face. He says, “The whole reason I was at work was for the benefit of the patient and their family. I think that can get lost in the translation, particularly for people that are coming to what is essentially a bureaucracy and might not get a chance to talk to the people who are delivering the care.

“I think there are two solutions to that problem. One is that the people delivering the care have to talk to the managers and secondly, the managers have to walk around the wards and see what really happens. Or walk around villages and see what problems emerge from lack of safe water, or proper sewage or any natal care.

“If we don’t actually interact with the patients, we miss all of that, wherever we happen to practice.”

In providing care and later evaluating areas in which it can be improved, Hughes says that it is important to first recognize that quality is often defined differently by caregivers and patients; “Interestingly, if you ask doctors and nurses about what is happening in their hospitals, they focus on safety when things go wrong and patients get harmed. However, if you ask patients, it is about quality, which is the experience they went through even though they might have been sick or even facing death.”

There are multitude of other factors besides the patient-clinician interaction that impact care. Hughes explains that care is also affected by the patient’s family and surrounding staff, as well as the pressures placed on the health care system itself, from the perceived needs of the ministries of health to the funding obstacles of treasury departments.

“Compassion is central, because it means having a respect for and with the patients, the people that we are trying to serve, and it also means having a respect for and with the staff that we work with,” he explains, which he says translates into recognizing these external pressures but compassionately working to provide the best care possible.

Noting that the quality improvement movement has grown dramatically since the inception of ISQua over 30 years ago, Hughes sees value in the growth of understanding around defining quality improvement.

“One of the roles of our organization is to act as a broker for some of that information and education while doing implementation work. I think that this seminar is starting to build a framework around the science of quality improvement; there has been a science for many years, but the rigor around the scientific endeavor in an academic sense or around products like drugs or devices is very different than the science around quality improvement.

“Quality is actually defined best by the patient experience, so it is a personal and subjective approach to doing things in better way.”
INTERVIEW

Danika Barry: “Patients are the ultimate stakeholders in Quality Improvement”

Danika Barry, a student at Harvard Medical School and a former health care improvement fellow with the USAID ASSIST Project and the Quality & Performance Institute at University Research Company, LLC., spoke to Salzburg Global during the session Better Health Care: How Do We Learn About Improvement? underscoring the value of direct patient interactions in the scope of improving and evaluating health care systems.

Barry spent several years working with community-based primary health care programs in India, Uganda and later Ethiopia where she became involved in the field of quality improvement in particular, whilst working on a project titled: Maternal and Newborn Health in Ethiopia Partnership.

“It was a pivotal time in my career trajectory, because it was a way of seeing how communities themselves can really lead changes that were important to them and also have ownership over the data collection, evaluation, and analysis processes, and I think that shift in power is what was really so key to that project’s success. The direct focus on systems which is so key for quality improvement was also very appealing to me.”

Through an increased perspective of the changes occurring in communities due to work to improve access to quality care and scale-up complex interventions, Barry saw the critical role that frontline clinicians play in developing and developed nations alike.

“It lead me to want to become deeply immersed within health systems as a care provider and then I think for other reasons, I wanted to be directly engaged in care provision and working with people and patients as it led me to understand what the key challenges were.”

As she transitions to a new role as a direct care physician, her participation in the session comes at a unique inflection point: the completion of her first year of medical school. Barry says, “some of the reflections that I have so far: a much deeper appreciation of what it might be like to be a frontline worker, so someone who is on the frontlines implementing changes and someone who is a key stakeholder immersed in a complex system. I have also been able to learn a lot more about the patient perspective by being able to spend so much time directly with patients, talking with patients about what their concerns are, what their hopes are and what they value in their care and I think that is an absolutely essential perspective.”

While many in the health care improvement and evaluation field were first and foremost direct care providers and then later evolved into roles addressing intervention and methods of quality evaluation, Barry’s perhaps reversed experience has led to her interest in working in the intersections of clinical practice, research and policy.

“There is this sense of urgency or even impatience that I feel among the Fellows and I think that is actually really positive because it forces us to make sure the questions that we are asking are relevant and that the solutions that we try to come up with will be relevant to the ultimate stakeholders, and those are our patients and our communities.”
Appendix I

Chair

M. Rashad Massoud
Director, USAID ASSIST Project; Senior Vice President, Quality & Performance Institute, University Research Co. LLC, Bethesda, MD, USA

Participants (positions correct at time of session – July 2016)

Bruce Agins
Medical Director, NYSDOH AIDS Institute; Director, HEALTHQUAL International, New York, NY, USA

Maina Boucar
Director, Regional Office for Francophone Africa USAID ASSIST Project, Niamey, Niger

Joshua Bardfield
Knowledge Management Director, HEALTHQUAL International, New York, NY, USA

Danika Barry
Healthcare Improvement Fellow, USAID ASSIST Project, University Research Company, LLC, Boston, MA, USA

Viktor Boguslavskiy
Regional Director for East Africa, Europe & Eurasia, USAID ASSIST Project, University Research Company, LLC, Bethesda, MD, USA (Ukraine)

Bruno Bouchet
Director, Health Systems Strengthening, FHI 360, Washington, DC, USA (France)

Rhea Bright
Quality Improvement Advisor, USAID, Arlington, VA, USA

Danielle Charlet
Technical Advisor, University Research Company, LLC, Bethesda, MD, USA (Canada)

Vijay Chattu
Senior Lecturer – Public Health, The University of the West Indies, St. Augustine, Trinidad and Tobago (India)

Astou Coly
Senior Improvement Advisor, University Research Company, LLC, Bethesda, MD, USA (Senegal)

Mahesh Devnani
Assistant Professor of Hospital Administration, Postgraduate Institute of Medical Education and Research, Chandigarh, India

Houleymata Diarra
Chief of Party, USAID ASSIST Mali, University Research Company, LLC, Bamako, Mali

Nancy Dixon
President, Common Knowledge Associates, Austin, TX, USA

Lisa Dolan-Branton
Senior Improvement Advisor, University Research Company, LLC, Bethesda, MD, USA

Amel Farrag Hammad
Quality Improvement Consultant, Ministry of Health and Population, Cairo, Egypt

Kate Fatta
Improvement Advisor for Knowledge Management, University Research Company, LLC, Bethesda, MD, USA

Karen Wolk Feinstein
President and CEO, Jewish Healthcare Foundation, Pittsburgh, PA, USA

Ezequiel Garcia Elorrio
Director, Institute for Clinical Effectiveness & Health Policy (IECS), Buenos Aires, Argentina

Donald Goldmann,
Chief Medical and Scientific Officer, Institute for Healthcare Improvement, Cambridge, MA, USA

Cecil Haverkamp
Country Director, University Research Company, LLC, Gaborone, Botswana (Germany)

James Heiby
Medical Officer, USAID, USAID, Arlington, VA, USA

Jorge Hermida
Director for HCI Programs in Latin America, University Research Co. LLC, Quito, Ecuador

Lisa Hirschhorn
Director, Implementation and Improvement Science, Ariadne Labs, Boston, MA, USA

Clifford Hughes AO
President, International Society for Quality in Health Care (ISQua), Dublin, Ireland (Australia)

Usman Iqbal
Health Informatics Researcher, Taipei Medical University, Taipei, Taiwan – China (Pakistan)

Donna Jacobs
Regional Director, University Research Company, LLC, Pretoria, South Africa

Rachel Jean-Baptiste
President, Oxford Epidemiology Services, LLC, Washington, DC, USA (Haiti)

Ashish K. Jha
Director, Harvard Global Health Institute, Harvard University, Cambridge, MA, USA

Mary Kawonga
Senior Lecturer, School of Public Health, University of the Witwatersrand, Johannesburg, South Africa (Zambia)

Edward Kelley
Director, Service Delivery and Safety World Health Organization, Geneva, Switzerland (USA)

Peter Lachman
CEO, International Society for Quality in Healthcare (ISQua), Dublin, Ireland
## Participants (continued)

<table>
<thead>
<tr>
<th>Name</th>
<th>Title and Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teresa Li</td>
<td>Assistant Director of Family and Elderly Health Services, Department of Health Hong Kong China, Hong Kong SAR</td>
</tr>
<tr>
<td>Nigel Livesley</td>
<td>Project Director, ASSIST India, University Research Co., LLC New Delhi, India (UK)</td>
</tr>
<tr>
<td>Deneil Logiudice</td>
<td>Quality Improvement Consultant, The Billions Institute, Boston, MA, USA</td>
</tr>
<tr>
<td>Charles Kimani</td>
<td>Monitoring &amp; Evaluation Advisor, University Research Company, LLC, Nairobi, Kenya</td>
</tr>
<tr>
<td>Michael Marx</td>
<td>Professor for International Health, University of Heidelberg, Heidelberg, Germany</td>
</tr>
<tr>
<td>Nana Mensah-Abrampah</td>
<td>Technical Officer, Service Delivery and Safety Department, World Health Organization, Geneva, Switzerland (USA)</td>
</tr>
<tr>
<td>Brian Mittman</td>
<td>Senior Scientist, Kaiser Permanente Department of Research and Evaluation, Pasadena, CA, USA</td>
</tr>
<tr>
<td>Mohamed Mohamed</td>
<td>Director – Health Quality Assurance, Ministry of Health, Dar es Salaam, Tanzania</td>
</tr>
<tr>
<td>Andrew Muhire</td>
<td>Sector Monitoring &amp; Evaluation and Report Lead Specialist, Ministry of Health, Kigali, Rwanda</td>
</tr>
<tr>
<td>Bejoy Nambiar</td>
<td>Research Associate, Institute for Global Health, University College London, Lilongwe, Malawi (India)</td>
</tr>
<tr>
<td>James Nganga Ndirangu</td>
<td>Southern Africa Regional Advisor, University Research Company, LLC, Pretoria, South Africa (Kenya)</td>
</tr>
<tr>
<td>Katie O’Connor</td>
<td>Public Health Analyst, Health Resources and Service Administration (HRSA), Washington, DC, USA</td>
</tr>
<tr>
<td>John Øvretveit</td>
<td>Director of Research and Professor of Health Care Innovation Implementation and Evaluation, Medical Management Centre, Karolinska Institutet, Stockholm, Sweden (UK)</td>
</tr>
<tr>
<td>Gareth Parry</td>
<td>Senior Scientist, Institute for Healthcare Improvement, Cambridge, MA, USA (UK)</td>
</tr>
<tr>
<td>Mihajlo Rabrenovic</td>
<td>Associate Professor, School of Business Studies, Belgrade, Serbia</td>
</tr>
<tr>
<td>Rohit Ramaswamy</td>
<td>Associate Professor, The University of North Carolina, Chapel Hill, NC, USA</td>
</tr>
<tr>
<td>Julie Reed</td>
<td>Deputy Director and Academic Lead, National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) for North West London (NWL), London, UK</td>
</tr>
<tr>
<td>Amy Reid</td>
<td>Director, Institute for Healthcare Improvement, Cambridge, MA, USA</td>
</tr>
<tr>
<td>Alexander Rowe</td>
<td>Medical Officer, Malaria Branch, US Center for Disease Control &amp; Prevention, Atlanta, GA, USA</td>
</tr>
<tr>
<td>Sylvia Sax</td>
<td>International Public Health Consultant; Research &amp; Lecturer, Institute of Public Health, University of Heidelberg, Heidelberg, Germany (Canada)</td>
</tr>
<tr>
<td>Lisa Simpson</td>
<td>President &amp; Chief Executive Officer, AcademyHealth, Washington, DC, USA</td>
</tr>
<tr>
<td>Kelly Suzan</td>
<td>Adjunct Health Instructor, University of Minnesota, Minneapolis, MN, USA</td>
</tr>
<tr>
<td>Abraham Wandersman</td>
<td>Professor, University of South Carolina, Columbia, SC, USA</td>
</tr>
<tr>
<td>Aaron Yao</td>
<td>Assistant Professor of Public Health Sciences, University of Virginia, Charlottesville, VA, USA (China)</td>
</tr>
<tr>
<td>Roman Yorick</td>
<td>Deputy Chief of Party, USAID RESPOND Project, Pact Inc., Kiev, Ukraine, (Russian Federation)</td>
</tr>
<tr>
<td>Nancy Zionts</td>
<td>Chief Operating and Chief Program Officer, Jewish Healthcare Foundation, Pittsburgh, PA, USA</td>
</tr>
</tbody>
</table>

## Staff

<table>
<thead>
<tr>
<th>Name</th>
<th>Title and Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paul Jansen</td>
<td>( Former) Program Director</td>
</tr>
<tr>
<td>Antonio Riolino</td>
<td>Program Associate</td>
</tr>
<tr>
<td>Clare Shine</td>
<td>Vice President &amp; Chief Program Officer</td>
</tr>
<tr>
<td>Louise Hallman</td>
<td>Editor</td>
</tr>
<tr>
<td>Jan Heinicke</td>
<td>Fellowship Manager</td>
</tr>
<tr>
<td>Beth Pertiller</td>
<td>Director of Operations</td>
</tr>
<tr>
<td>Jessica Franzetti</td>
<td>Communications Intern</td>
</tr>
<tr>
<td>Ha Ram Hwang</td>
<td>Communications Intern</td>
</tr>
<tr>
<td>In Ae Joo</td>
<td>Program Intern</td>
</tr>
<tr>
<td>Yeri Kwak</td>
<td>Program Intern</td>
</tr>
</tbody>
</table>

Appendix II

Recaps

Day 1
*Improving Improvement*
[www.salzburgglobal.org/topics/article/day-1-improving-improvement](http://www.salzburgglobal.org/topics/article/day-1-improving-improvement)

Day 2
*Case Study Learnings*
Facilitators summarize the key learnings from the morning’s case studies
[www.salzburgglobal.org/topics/article/day-2-case-study-learnings](http://www.salzburgglobal.org/topics/article/day-2-case-study-learnings)

*Case Studies*
- Water Quality in Ghana
  [www.salzburgglobal.org/topics/article/day-2-case-studies-water-quality-in-ghana](http://www.salzburgglobal.org/topics/article/day-2-case-studies-water-quality-in-ghana)
- Anaemia in Mali
  [www.salzburgglobal.org/topics/article/day-2-case-studies-anaemia-in-mali](http://www.salzburgglobal.org/topics/article/day-2-case-studies-anaemia-in-mali)
- Mystery Improvements in Uganda
  [www.salzburgglobal.org/topics/article/day-2-case-studies-mystery-improvements-in-uganda](http://www.salzburgglobal.org/topics/article/day-2-case-studies-mystery-improvements-in-uganda)
- Antenatal Care in India
- Patient Discharge Checklists in the UK
  [www.salzburgglobal.org/topics/article/day-2-case-studies-patient-discharge-checklists-in-the-uk](http://www.salzburgglobal.org/topics/article/day-2-case-studies-patient-discharge-checklists-in-the-uk)

*Reverse Knowledge Café*
Participants further their understanding of their case studies through input from expert faculty
[www.salzburgglobal.org/topics/article/day-2-reverse-knowledge-cafe](http://www.salzburgglobal.org/topics/article/day-2-reverse-knowledge-cafe)

Day 3
*Improving Rigor, Attribution and Generalizability*
- Water Quality in Ghana
- Anaemia in Mali
- Mystery Improvements in Uganda
  [www.salzburgglobal.org/topics/article/day-3-improving-rigor-attrition-and-generalizability-pre-eclampsia-in-uganda](http://www.salzburgglobal.org/topics/article/day-3-improving-rigor-attrition-and-generalizability-pre-eclampsia-in-uganda)
- Antenatal Care in India

Day 4
*Synthesis of Approaches for Evaluation*
[www.salzburgglobal.org/topics/article/day-4-synthesis-of-approaches-for-evaluation](http://www.salzburgglobal.org/topics/article/day-4-synthesis-of-approaches-for-evaluation)

Videos

Pre-session series: How Do We Learn About Improvement?
Participants discuss what quality improvement means for them and their organizations and why they are taking part in the Salzburg Global Seminar program

**Don Goldmann**
Chief Medical and Scientific Officer at the Institute for Healthcare Improvement (IHI)
[youtu.be/-Hbt7huo94s](https://youtu.be/-Hbt7huo94s)

**Ed Kelley**
Director of Service Delivery and Safety at the WHO

**M. Rashad Massoud**
Senior Vice President of the Quality & Performance Institute, and Director of the USAID ASSIST Project
[youtu.be/Zb7mdOHT11Y](https://youtu.be/Zb7mdOHT11Y)

**Sylvia Sax & Michael Marx**
Institute of Public Health at Heidelberg University, Germany
[youtu.be/kp7fHfuxQ](https://youtu.be/kp7fHfuxQ)
Interviews

Danika Barry
“Patients are the ultimate stakeholders in quality improvement”
Harvard Medical School student discusses the value of improvement for the most important stakeholders—patients and communities.
www.salzburgglobal.org/topics/article/danika-barry-patients-are-the-ultimate-stakeholders-in-quality-improvement

Rhea Bright
Making care better
Technical adviser and Global Health Fellow in the Office of Health Systems at USAID explains her original interest in quality improvement and her subsequent work in the field.
youtu.be/6uo7zZzzhr8

Nancy Dixon
How do we spread knowledge?
Health care researcher and consultant considers the value of tacit knowledge in improvement.
youtu.be/98BnBlww3kE

Cliff Hughes
Compassion is central to improvement
Long-time cardiothoracic surgeon and current President of ISQua discusses clinician-patient interactions and defining quality of care.
www.salzburgglobal.org/topics/article/cliff-hughes-compassion-is-central-to-improvement

Ed Kelley
Achieving global health goals through capacity building
Director of the Department of Service Delivery and Safety at the WHO explains how health improvement is intertwined in the building of domestic capacities.
www.salzburgglobal.org/topics/article/ed-kelley-achieving-global-health-goals-through-capacity-building

Rashad Massoud
Making lives healthier through quality improvement
Session Chair speaks about the importance of convening to improve health care.
youtu.be/JYjxYeQ3Kaw

Andrew Muhire
Using data to improve health care
M&E and report lead specialist in Rwanda’s Ministry of Health discusses the importance of data in quality improvement.
youtu.be/WZRYmJJ9Zbw

Fellows’ Blogs

Bejoy Nambiar
Quality Improvement Interventions: Challenges and Solutions for Evaluation Design
Researcher at the Institute for Global Health, University College London, shares his opinion on challenges and solutions in evaluating designs in quality improvement interventions.

Nancy Dixon
What elements of an improvement are transferable, and what adaptations are needed?

Andrew Muhire
When measuring health care improvement, how do we know that no other factors are influencing the results?
www.salzburgglobal.org/topics/article/nancy-dixon-when-measuring-health-care-improvement-how-do-we-know-that-no-other-factors-are-influ

Academic turned consultant shares her opinion ahead of July session Better Health Care: How do we learn about improvement?
Salzburg Global Seminar Staff

Senior Management
Stephen L. SALLYER, President & Chief Executive Officer
Benjamin W. GLAHN, Vice President – Development & Operations
Clare SHINE, Vice President & Chief Program Officer
Daniel SZELENTI, General Manager – Hotel Schloss Leopoldskron
Pia VALDIVIA, Vice President & Chief Finance Officer

Program and Administrative Staff
Chanel Bell, Program Associate – Mellon-Global Citizenship Program (M-GCP)
Thomas Biebl, Director, Marketing and Communications
Ian Brown, European Development Director
Jemma Clerk, Program Associate (on leave)
Michelle Dai Zotti, Development Associate
Kristina Dortschy, Program Development Assistant
Charles E. Ehrlich, Program Instructor
Marty Gecek, Chair – Salzburg Seminar American Studies Association (SSASA)
David Goldman, Program Consultant – (M-GCP)
Michaela Goldman, Internship Program Manager
Barbara Grodecka-Poprawska, Program Associate (on leave)
Louise Hallman, Editor
Jan Heinecke, Fellowship Manager
Andrew Ho, US Development Director
Lisa Karl, Assistant Director Finance, Salzburg
Danielle Karnoff, Development Associate
Allison Kingery, Senior Manager, Individual Giving & Campaign Planning

Astrid Koblmüller, Program Manager
Kevin Kolesnikoff, Program Associate Trainee
Brigitte Kraibacher, Assistant, Admissions Office
Tatsiana Lintouskaya, Program Director
John Lotherington, Program Director
Sharon Marcoux, Senior Finance Manager, US
Paul Mihailidis, Program Director – Salzburg Academy for Media and Global Change
Klaus Mueller, Program Consultant – Salzburg Global LGBT Forum
Beth Pertiller, Director of Operations
Bernadette Prasser, Program and Admissions Officer
Michaela Radanovic, Finance Assistant, Salzburg
Ursula Reichl, Assistant Director Finance, Salzburg
Manuela Resch-Trampitsch, Director Finance, Salzburg
Antonio Riolino, Program Associate
Susanna Seidl-Fox, Program Director, Culture and the Arts
Sarah Sexton, Special Assistant to the President
Nancy Smith, Program Consultant – (M-GCP)
Molly Walker, Davidson Impact Fellow

Hotel Schloss Leopoldskron Staff
Richard Aigner, Hotel Operations Manager
Jürgen Chum, Executive Chef
Niklas Geelhaar, Front Office Supervisor
Karin Maurer, Reservations and Revenue Supervisor

Sebastian Rechberger, Banquets Manager
Matthias Rinnerthal, Maintenance Supervisor
Karin Schiller, Sales and Marketing Manager
Marisa Todorovic, Executive Housekeeper

Interns (at time of program)
Jessica Franzetti, Communications
Ha Ram Hwang, Communications
In Ae Joo, Program
Emma Kehilbeck, Program

Yeri Kwak, Program
Allia Veaz, Program – Global Citizenship Alliance
Galen Voohees, Development
Report Author:

Anjali Chowfla is an Improvement Specialist working on the USAID Applying Science to Strengthen and Improve Systems project being implemented by University Research Co., LLC and supports project activities in six countries in Southern Africa. She previously worked for Columbia University where she helped to establish a regional research center in Mumbai, India to leverage the university's diverse intellectual capacity around issues of importance to the South Asia region. Anjali holds a MPA with a focus on global health and international development from Columbia University’s School of International and Public Affairs and a B.A. in Development Studies from the University of California, Berkeley.
Health and Health Care Innovation in the 21st Century series

Salzburg Global Seminar has long been a leading forum for the exchange of ideas on issues in health and health care affecting countries throughout the world. At these meetings agendas have been re-set affecting policy and practice in crucial areas, such as patient safety and the engagement of patients in medical decision making.

In 2010, Salzburg Global Seminar launched a multi-year series – **Health and Health Care Innovation in the 21st Century** – to crystallize new approaches to global health and health care in the face of emerging challenges affecting us now and set to continue on through the coming generation.

FOR MORE info. on the series PLEASE VISIT:
health.SalzburgGlobal.org