

From Reactive to Proactive: How “Stand-by” Evaluators Can Enhance Organizational Learning and Adaptability

AEA November 2025

Maggie Jones and Noon Kuo: Center for Community Health and Evaluation

Maryam Khojasteh: Robert Wood Johnson Foundation

Jess Renger: Kaiser Permanente, National Office of Community & Social Health

About our Organizations

The **Center for Community Health and Evaluation** is a 25-person center within Kaiser Permanente Washington’s Health Research Institute, based in Seattle, WA.

Our mission is to improve the health of communities with collaborative approaches to planning, assessment, and evaluation. We design and evaluate health-related programs and initiatives throughout the United States.

Services

Serve as **evaluation** partner to implement pragmatic, timely, useful evaluations

Facilitate **strategic learning and planning**

Support organizations with **building evaluation capacity**

Design & conduct tailored **assessments** to inform decision making

Focus Areas



Community and public health



Clinical improvement, innovation, and integration



Equity-focused initiatives



Capacity building programs and learning collaboratives



Systems, policy, and cross-sector collaboration

We at RWJF recommit ourselves to creating a future where health is no longer a privilege, but a right ... and we rededicate ourselves to taking bold leaps even in the face of hostility and obstruction.

1

Community Development

We're working to ensure safe, affordable places to live and easy access to healthy food, clean water, and good job opportunities.

2

Economic Inclusion for Families

We envision U.S. tax and caregiving systems that support caregivers, so all families have the resources they need to raise thriving children.

3

Government

We envision a more inclusive, effective, and multiracial democracy where everyone can participate in the political processes that impact health.

4

Health Science Knowledge

We're working to broaden how we collect, share, and use data, including cultural and community knowledge, to prioritize health equity.

5

Healthcare

We're working toward affordable and respectful healthcare, with a diverse workforce that values different cultures' approaches to health.

6

Media

We envision news, TV, movies, and social media with more diversity, less bias, and more complete stories about health.

7

Public Health

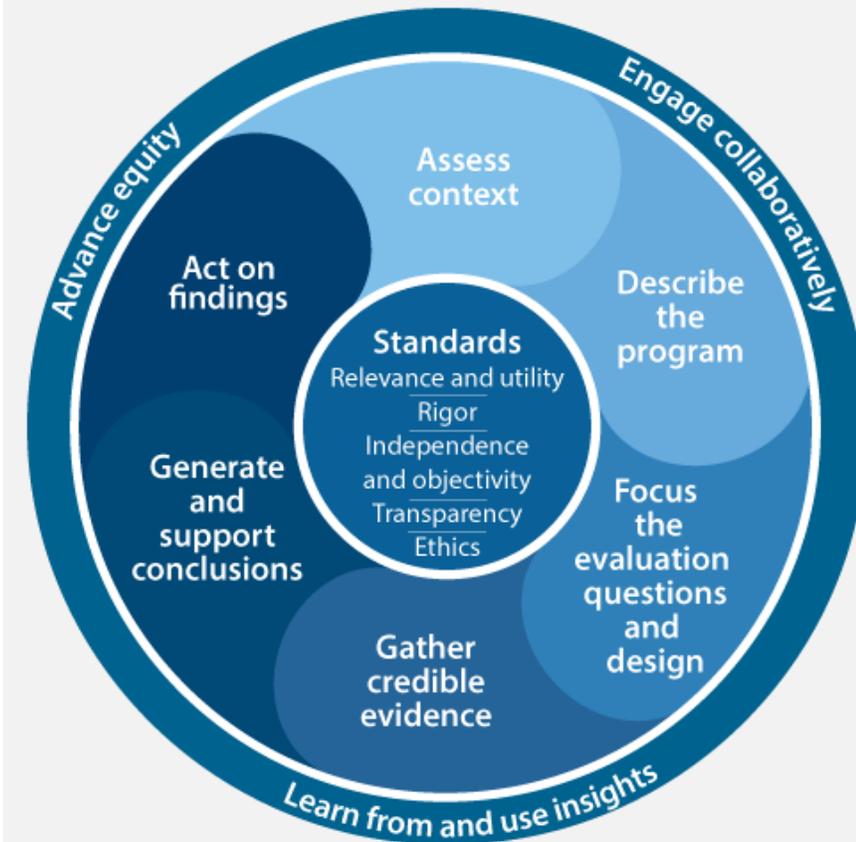
We envision a public health system that is trustworthy and accountable to communities, supported by funding, data, and capacity to advance health equity.

Kaiser Permanente – Evaluation, Learning, Measurement, and Reporting Team (ELMR)

- Kaiser Permanente – National Office of Community and Social Health
- The ELMR team is an internal evaluation group that supports the full lifecycle of community and social health initiatives within the National Office and promotes a comprehensive evaluation approach across all initiatives.



Evaluation Steps



Why engage a stand-by or emergent evaluator?

Stand-by or emergent evaluator defined:

An open and flexible evaluation contract to support emergent priorities related to evaluation, measurement, planning/design, and organizational learning

- **Informing learning and strategy**
 - Cross-program synthesis
 - Learning strategy/approach
- **Program planning/development**
 - Literature/evidence reviews
 - Environmental scans
 - Analyzing existing data/extracting learnings
- **Evaluation planning**
 - Developing theories of change/logic models
 - Evaluation planning/scoping/identifying questions
 - Evaluation RFP development, Identifying potential partners
- **Small projects**
 - Analyzing data sets
 - Synthesizing existing data
 - Conducting small scale evaluations
 - Writing manuscripts/supporting dissemination



The case for emergent evaluation support at RWJF

- Time and process to plan/implement external evaluation is significant – can take 6-10 months (i.e., develop learning questions, determine qualifications, search/develop RFP, invite/review proposals, grant agreement process)
- Ad-hoc and unanticipated needs emerge that cannot be addressed in a timely way
- Internal staff don't always have the capacity or expertise to meet emergent needs
- External evaluations usually focus on a single investment or portfolio, but do not support cross-team learning or look at connections across investment areas
- Overtime, emergent evaluation partner increases their knowledge of the complex body of work and can more efficiently/effectively plug into new areas – balancing internal / external perspectives

The case for emergent evaluation support at KP

- With the transition to a national Community and Social Health strategy, there was increased demand for a comprehensive measurement strategy to tell a unified story of impact.
- The ELMR team does not have sufficient capacity or all the specialized expertise to respond to every request from leadership.
- CCHE had longstanding relationships with KP through evaluations of long-term initiatives and brought complementary skills and expertise that enhanced ELMR's evaluation and measurement strategies.
- CCHE also provided an external perspective, which strengthened the comprehensiveness and objectivity of the evaluation work.

What does the work entail?

How does it differ from traditional evaluations?

	Traditional Evaluation	Emergent Evaluation
SOW	Specific and concrete	Outline with many TBD, can be based on examples of past activities
Timeline	Interim/final deliverables mostly preset	Clear contract dates, but interim deadlines/timelines TBD or can shift
Flow/pace	Workplan is established at the outset	Uncertain; must monitor, manage ebbs and flows
Ability to Plan	Can plan ahead with reasonable certainty	Varies depending on assignment and capacity
Funding / staffing	Budget and staffing plan set at beginning. Relatively stable during the project.	Often an up-to amount depending on work needed; staffing/FTE will fluctuate to meet demands.
Scope / size of requests	Difficult to support small or time sensitive requests due to contracting	Can be more responsive to time sensitive requests, and can take on smaller projects
Evaluation dynamics	<ul style="list-style-type: none"> • Often viewed as external • SOW dictates work • May not deeply understand internal dynamics 	<ul style="list-style-type: none"> • More partnership • Can be more embedded in strategy/design • Deeper understanding of internal dynamics and decisions

How emergent evaluation support has been used – RWJF

Example	Activities	What emergent evaluation contributes
Evaluation RFP development	<ul style="list-style-type: none">• Development of RFP from idea conception to launch• Working internally with program and grant management teams	<ul style="list-style-type: none">• A process that helped make complex ideas programmable• Assisted with group consensus and working through different PoVs
Water bibliography/ lit review	<ul style="list-style-type: none">• Systematic review of gray and academic literature on the history of structural racism in water infrastructure	<ul style="list-style-type: none">• A useful product that helped RWJF and the grantees with making a better case for why focusing on equity in water infrastructure matters
Learning agenda	<ul style="list-style-type: none">• Creating an evaluation index to centralize the team's learning• Mapping the team's programming against indicators of progress• Synthesizing progress reports	<ul style="list-style-type: none">• Assistance with institutionalizing learning practices• In-time identification of areas for further investigation
Mid-course tobacco strategy assessment	<ul style="list-style-type: none">• Conducted retrospective evaluation of RWJF's tobacco strategy for the past 5 years with a cohort of 40+ grantees	<ul style="list-style-type: none">• Timely support for a complex evaluation of high priority and quick turn-around

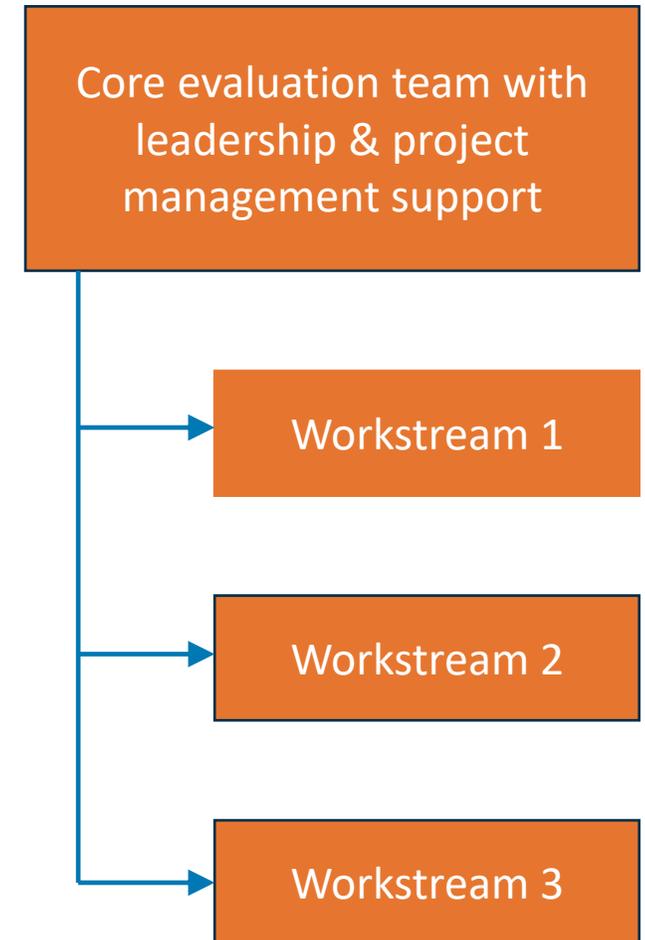
How emergent evaluation support has been used – KP

Example	Activities	What emergent evaluation contributes
SNAP Evaluation	<ul style="list-style-type: none">• Analyze data from a texting intervention to help focus on utility and impact• Quick pragmatic studies to test language, modalities, etc.• Longer-term ongoing evaluation support	<ul style="list-style-type: none">• Subject matter expertise• Expertise designing complex evaluations and analyzing messy data• Research and analytic skills for statistical inference
Impact Modeling	<ul style="list-style-type: none">• Projecting longer term outcomes from social health initiatives• Sharpen focus on intermediate and longer-term outcomes• Look for commonalities across initiatives	<ul style="list-style-type: none">• Long-standing relationships with people in different parts of KP• Time and expertise for in-depth literature review and meta-analysis• Support communication and storytelling by highlighting areas of impact
Literature reviews	<ul style="list-style-type: none">• Critique/analysis of articles to determine level of evidence• Rapid evidence review• Comprehensive literature review to inform strategy/planning• Draft background section of a manuscript	<ul style="list-style-type: none">• Quick turnaround support for tight deadlines• Content or methods expertise• Based on the evidence, consult on initiative design, strategic planning, intervention selection

How is the work structured?

Evaluation team structure

- Small, **core evaluation team** lead portfolio to maximize knowledge, collaboration, efficiency
 - Size/structure will depend on the scope of the budget/work
 - Number of additional staff needed will vary with different workstream needs
 - Subject or methods experts may need to be brought in
- Requires strong **internal communication** between core evaluation team and workstreams
- **Managing staff capacity** given ebbs and flows of work
 - Balance between holding capacity of staff to be responsive with being good stewards of funding
 - Have ongoing projects during slower times for core staff



Funder Structure

Key features include:

1. Having some level of autonomy in making decisions
2. Main point of contact with some evaluation expertise, who:
 - Coordinates and navigates SOW
 - Prioritizes work with evaluation team
 - Engages with institution around leadership buy in, funding, etc
3. Engage other people at their organization/team to lead workstreams
4. Regular communication and coordination



How do you set up these relationships for success?

Lessons learned

1

Ensure strong communication, coordination, and alignment

- Establish roles and identify primary points of contact for both the evaluation team and funding partner
- Identify a core evaluation team that will be accountable for all workstreams
 - Core evaluation team should have mechanisms to support workstream staff to ensure they have needed context to be effective
- Build in time for relationship development
- Check in regularly (e.g., monthly meetings) for feedback, course corrections, approvals, real-time alignment, and (re)prioritization as needed
- Prioritize open communication and discuss challenges as they emerge
- Structured reflection helps identify challenges and tension points in the work & lessons across workstreams

Lessons learned continued

2

Build structure for transparency around priorities and progress

- Establish a mechanism for identifying, vetting, and prioritizing potential projects – a menu of options can be a helpful starting point
- Maintain a dashboard of all ongoing and recently completed projects to ensure shared understanding of priorities and progress
- Create accessible project folders with shared access
- Develop consistent documentation of core materials (e.g., scoping documents, timeline, notes, deliverables) to ensure all workstreams to have a parallel structure
- Track and be transparent about resource use to help inform priority settings and the pace/intensity of the work
- Focus on scaffolding – don't overbuild infrastructure because things will change

Lessons learned continued

3

Anticipate the need for flexibility in managing emergent requests

- Core evaluation team must be flexible, willing to pivot, and have a breadth of expertise to be able to plug into different types of asks
- Plan to bring in and shift additional staff and expertise for specific workstreams, when necessary
- Plan for the pace of various workstreams to align with funder priorities/timelines and evaluation team capacity:
 - Plan for larger projects/workstreams to ensure sufficient capacity
 - Maintain a list of “nice to have” or non-urgent priorities to pick up during slower times

Lessons learned continued

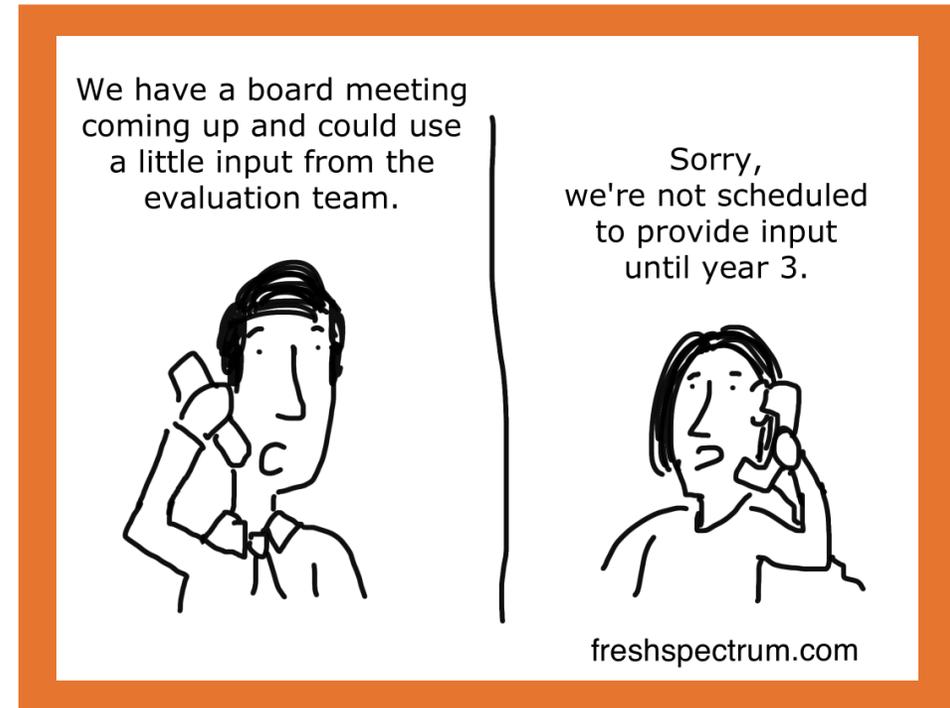
4

Allocate time for co-development of work/priorities

- Work often starts with an idea or a question. Evaluation partner is coming in much earlier. It takes time to translate ideas into actionable workstreams.
- Evaluation team will require support from funder to navigate and prioritize work. Need to have willingness to share internal processes and tensions.
- Evaluation team can provide “external” insights to help inform strategic decisions/initiatives.
- The longer the partnership, the more independent the evaluation team will be because they will have a deeper understanding of context and internal processes.
- Work is sometimes experimental and won't always be linear

Impact of engaging with a “stand-by” evaluator

- ✓ Contributes to more robust cross-program/initiative learning to inform strategy
- ✓ Expands capacity for small data/evaluation projects:
 - Broader methodological and/or content expertise
 - Ability to meet needs/answer questions as they emerge
- ✓ Brings external perspective to internal processes/decisions
- ✓ Allows for exploration and experimentation to test ideas/approaches



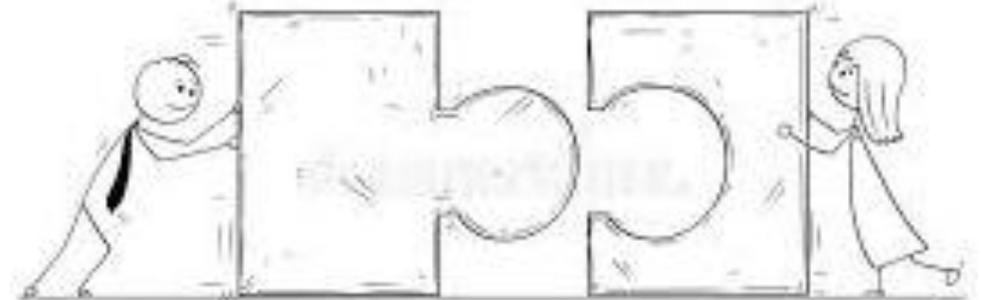
When emergent evaluation works well....

Conditions

- Strong relationships already exist
- Funder can make a long-term investment with flexible deliverables
- Funder point person has some evaluation background
- Evaluation team has multiple areas of expertise
- Evaluation team can manage ebbs and flows in the work
- Shared commitment to learning & experimentation

Types of work

- Time sensitive requests
- Evaluations that require strong understanding of internal context/strategy
- Cross-program or strategy-level evaluations
- Efforts to support internal learning and/or strategy
- Exploratory projects that aren't fully "baked"



Participant questions & responses

Participant questions

Foundation Board/leadership support: *With the varying levels of commitment to eval and learning in philanthropy, how did you gain the support of your board support for this approach?*

In both cases, the amount for this contract did not require Board approval. However, it was important to build a case to executive leadership. It may be helpful to start with a smaller investment to experiment with this approach, show value, and then expand. Much of our emergent work is focused on providing leaders with information that they need, so when they see the value, they are more likely to continue to support the investment in subsequent years.

Soliciting partners for this work: *It sounds like these partnerships came out of existing relationships, not an RFP process. If you (KP or RWJF) WERE to release an RFP for an emergent evaluator, what are a few items that would be key expectations? For CCHE, what do you think should be included?*

RWJF did issue an RFP for this initially, but it was limited invitation based on evaluators that they had worked with in the past and who knew them/their work. If you were developing an RFP, it would be important to assess the evaluators' approach to evaluation, comfort with ambiguity, willingness to be flexible, broad subject matter expertise, and learning orientation. The response should focus more on approach and less on a concrete workplan and deliverables, as those will shift. We'd also recommend including an interview as part of the process because this type of work is highly collaborative and interpersonal relationships matter – so you need to select someone you can easily work with. If it's a new partnership, you will need to build in time for relationship building and it will likely require more upfront investment from your internal team before the evaluator will be able to work independently.

After the initial year, the renewals have included examples of the work we have completed as part of the contract and the impact of those efforts and specifies that we expect to do similar work during the next contract period.

Participant questions continued

Using internal vs. external teams for emergent requests: *For internal evaluation teams, how do you define the work you take on versus the standby evaluator and justify internally the need for investment in your own team and the external partner?*

This differs by organization:

At RWJF, the Research, Evaluation, and Learning (REL) team does not conduct evaluations, so that's a clear distinction. Much of the learning strategy work could be done internally if staff had capacity, but this partnership has been helpful to extend internal capacity.

At KP Community Health, the internal Evaluation, Learning, Measurement, and Reporting (ELMR) team does do some internal evaluation work, but the emergent evaluator is brought in for specific methodological or subject matter expertise that the team doesn't have internally and also used as capacity extenders for the internal team.

Staffing “standby” team: *Are all your “standby” evaluators already full-time members? We tend to bring in short-term support for larger projects or when we are at capacity, but they tend to be external.*

At CCHE, almost all of our emergent project work is managed by full-time team members, this has been important because typically these requests require rapid response (which doesn't allow time for contracting) and benefit from our knowledge of the organization and the context in which the work is occurring. For some of the larger emergent projects, when the capacity of our evaluation team was limited, we have brought in consultants in a more supportive role to ensure the project was implemented in a timely way.

Participant questions continued

Infrastructure: *Can panelists share example(s) of resources/infrastructure they've created that hits the mark of supporting evaluation without overbuilding?*

The main things we've developed to manage these portfolios include a one-page Scope of Work template, so as new requests are made, we can ensure we're on the same page with the goals, timeline, and resource allocation – and have documentation of those agreements. We also have a high-level dashboard to track priorities across the portfolio, which we review monthly. This just has a quarterly timeline, key deadlines, and status update for each workstream in the portfolio.

Types of projects: *Can you share more about some of the smaller, but more impactful projects that have been carried out?*

Slides 12-13 provide examples of projects that were completed.

For RWJF, the water bibliography is a good example of a small project that wouldn't have otherwise happened but has had a large impact. The project was a systematic review of the history of structural racism in water infrastructure. The initial plan was for this to inform internal efforts, but the product ended up being so useful it was shared externally and with grantees to help build the case for focusing on equity in water infrastructure matters.

For KP, the impact modeling work has gotten a lot of attention internally. In this work, CCHE is projecting longer term outcomes of KP's social health initiatives to help communicate the potential value of these investments. These projections have helped with communication efforts to make a stronger case for investing upstream.

Participant questions continued

Budget: *Is there a minimal contract size (budget) that makes a relationship like this doable, given the investment in structure/processes involved?*

This model can be scaled to different levels. If you are testing out this emergent model, it may be helpful to start small and pair it with a more concrete evaluation project – i.e., build in extra resources for some emergent support in addition to a specific evaluation project. For CCHE, the contracts are big enough that we can have 1-2 larger projects and maintain responsiveness to smaller requests. It is helpful to have a mixed portfolio so you have some larger projects, with more defined scopes that can flex to ensure capacity for managing the smaller and time sensitive requests efficiently.

Time for visioning: *How much time should we budget (more or less) for the co-development of vision?*

We had established relationships so didn't need to have as much time for building trust and vision. But it is important to build in time for building the relationship and ensuring alignment on the overall vision. We would recommend meeting at least monthly to coordinate, review, and reflect on how things are going.

You should also expect to have initial projects require more collaboration as you are beginning the working relationship. As time goes on, the emergent evaluation partner should be able to work more independently as they have a better understanding of your organization's needs.

For individual projects, the timeline for reaching agreement on goals and vision can vary significantly depending on how concrete it is and the urgency of the request. We've had some projects that have had to move quickly, while we've had other "ideas" that we've been experimenting with for years.

Participant questions continued

***Use of AI in literature reviews:** You mentioned how often literature reviews are needed. How are you utilizing AI to do those literature reviews? Do you use Notebook to speed up the work? Is there still have heavy human involvement, but a sort of iterative process with AI? Or are you not using AI at all?*

We do use AI to support literature reviews. As with many AI-supported tasks, we've learned some tips and tricks to efficiently get accurate results. Perplexity.ai is our preference. We set the universe to be Pubmed articles or specific websites only, and and ask for references to be listed. We specify language, date ranges, article type, etc.; all the different types of filters you could do in a more traditional Pubmed search. While we do ask AI for syntheses in addition to searches, we still have a fairly large human component to read each article. We find there are nuances that AI can't (yet) capture, like how to differentially weigh information from different study designs, which study populations are most applicable to the question at hand, assessing limitations, weighing overall quality and generalizability of the results.

If you have additional questions or would like to discuss further, please reach out!

Maggie Jones, Center for Community Health and Evaluation – maggie.e.jones@kp.org

Elena "Noon" Kuo, Center for Community Health and Evaluation – elena.s.kuo@kp.org

Maryam Khojasteh, Robert Wood Johnson Foundation – mkhojasteh@rwjf.org

Jess Renger, Kaiser Permanente, National Office of Community & Social Health – jessica.a.renger@kp.org