

Session 3 – Webinar Summary

Process measures are measures that examine how your program activities are delivered. They focus on different aspects such as participation, readiness to change, mastery of skills, and program delivery.

Core Process Measures

- Selected because they are applicable across programs and are easy to collect
 - Program enrollment
 - Program attendance
 - Program activities: in-session & between session
- Goal is to engage the target audience, meet enrollment goals, and that participants complete activities in and out of the session, but the reality is that this does not always happen. Need to use process data proactively to enhance engagement and participation in your programs.

NET-WORKS (Now Everybody Together for Amazing and Healthy Kids) Trial

- NET-WORKS was a part of the Childhood Obesity Prevention and Treatment Research (COPTTR) Consortium. It was funded for seven years across four field centers—two focused on obesity prevention and two on obesity treatment. Each center conducted different interventions.
- NET-WORKS MN site included the following components:
 - A brief primary care component
 - Home visits (main component of the intervention)
 - Community-based parenting classes
- Idea was to use these three synergistic components to influence parents or primary caregivers to ultimately influence child BMI by changing home environments and parenting practices.
- Two key process measures were documented in the study database across all the interventions:
 - Dose delivered – amount and length of intervention sessions offered
 - Dose received – participant attendance and completion of the offered sessions
- MN site put many systems in place to track what was happening:
 - Tracked which families had no participation after certain time periods and implemented reengagement protocols to figure out how to ensure the intervention works for them
 - Developed packets to give families some content when they were not participating to bridge active participation in the program
- Despite families disengaging, they were able to reengage families and work with them for enough time over the course of the program.
 - Being flexible and responsive to the unique needs of families was key
 - Tracking systems and reengagement strategies must be tailored to specific contexts

Stakeholders

- It is critical to measure and discuss process measures with stakeholders, because they can inform the sustainability of your program
- Stakeholders include municipalities, hospital systems, faith communities, community organizations, etc.

- There can be many stakeholders for community-based programs; it is important to know each stakeholder is different and understand what those differences will bring the table.
- Ask stakeholders what they want out of the program, doing so in different ways to ensure you understand their expectations.
- You need different messages for different stakeholders.
- Stakeholders have many expectations, including:
 - Return on investment
 - Financial
 - Workforce productivity
 - Community health
 - Improve quality of life
 - Reasonable deliverables
 - Patient satisfaction
- Help your stakeholders understand that changing BMI is difficult, and that your goals stem around process measures is critical

Community

- First, you must define your community.
- Ask your community what they want out of the program, who are the key players, and which partners are there for the long haul.
- Before asking your community what they want, consider the following:
 - Time expectations
 - The big picture
 - Clarity on roles
 - Commitments that need to be kept on both ends
 - Be prepared to lead the work on process measures
- Make the cost-benefit case for your program to community
 - Difficult to make on a monetary basis alone.
 - Health care systems present special challenges, but you may be able to make a case with certain health care systems.
 - Think globally about the value of your program to the community.
 - You goal may be to be as revenue neutral as possible.
 - Think about your program as infrastructure that benefits the community overall and value it brings to other entities.
 - It may be easier to make your case in communities where obesity is recognized as a crisis.
 - Individual participant stories are important for promoting your program, but they need to be superseded by outcomes that demonstrate a community benefit.
 - Keep in mind that our work still suffers from “weight bias”—treating a personality flaw versus a public health crisis.

Treatment Fidelity

- Treatment fidelity has two components:
 - How well was it done?
 - How much was done?

Measures of Provider Skill and Communication Style

These measures are necessary because they can help you monitor quality control, ensure fidelity, and conduct dose-response analyses. There are two primary approaches for measuring interventionist skill and style. One approach uses an expert who rates real-time or transcribed intervention sessions. Another approach relies on patient report of what occurred.

There is a very robust literature of motivational interviewing (MI) measures:

- Gold standard MI measures that include both count and subjective ratings (both elements of fidelity) include Motivational Interviewing Treatment (MITI) and Motivational Interviewing Skills Code (MISC).
 - Downside of both is cost and training
- One Pass, MI Coach, and Behaviour Change Counseling Index (BECCI) are MI measures that are based on global subjective ratings.
 - Require transcripts which add costs, but have less training and calibration costs
- Medical Objective Structured Clinical Examination (OSCE) and Roter Interaction Analysis System (RIAS) are universal measures.
 - Both are objective measures where experts rate various components of individual encounters.
 - OSCE is used primarily for training medical residents and medical students.
 - RIAS is used primarily as a measure of physician communication style.

MITI

- Has four specific global scores
 - Change talk and sustain talk are specific to MI
 - Partnership and empathy are universal elements that most schools of counseling agree are important predictors of rapport and outcome
- Includes various dimensions of behavioral counts where the rater goes through a transcript to count each utterance (an utterance can fall into only one category)
- Has very rigorous criteria for what counts as an utterance
- Achieving inter-rater reliability is difficult, but after a 2—4-day training people are able to reliably code with the counts

MISC

- Includes most of the same codes as the MITI
- Codes both practitioner and patient speech
 - Coding allows for both sequential and discourse coding, which shows how important it is for a practitioner to briefly respond to specific patient statements
 - How a practitioner responds does predict outcomes
 - Example of a coded patient utterance: “*I tried cooking without butter*”; this is coded as TS+ (taking steps, positive)

One Pass

- Divided MI skill into 23 domains: 19 specific individual items and 4 global measures, all of which are subjective ratings

- Provides one score for the entire session instead of utterance counts
- Slightly more reliable than MITI
- Training can be provided by developers over several days or can be self-taught

BECCI

- Began to be used in the profession about 17 years ago
- Not as up-to-date as other methods
- Items included are similar to other methods
- Free, but requires some training

MI Coach

- More recent self-report measure
- Includes constructs in common with the One Pass

RIAS

- Most widely used, universal measure of physician communication
- Divides communication into four functions
- Relatively similar to the MITI

ComOn Check

- Universal measure that appears to be a nice compromise of the MI measures and the RIAS
- Can be adopted for different types of providers

CBT Measures

- Rather than focus on how well a provider communicated, it looks at coverage and completeness, e.g., *How much of the content was delivered? How much of the topic was covered?*

Health Care Climate Questionnaire

- Universal measure of patient self-reporting of their providers
- Started as an instrument to measure self-determination theory, but has morphed in many ways
- Anywhere from 6–18 items
- Very low cost
- Takes patients about 3–5 minutes to complete
- No training required
- The instrument can be changed to work with any type of provider e.g., nurse, community health worker, etc.
 - Psychometrics remain consistently superb

OSCE

- Can be used if you don't want to focus your fidelity measures on a school of thought such as MI or CBT
- Similar to RIAS, measures generic components of a health professional interview
- Good generic measure of counseling skills

Discussion

N. Klinkhamer indicated she would like to learn more about the coding described by K. Resnicow on a separate call. NCCOR will work to facilitate a call specific to topic for those who are interested.

T. Earle was not sure she heard anything about the “operational flexibility of programs” to operate into a specific organization’s needs. She wondered if this is a measure that is not important to most programs or if there is some measure that programs should use to look at their program overall to be delivered in different settings? K. Resnicow indicated that implementation scientists have both a framework and set of standard measures that are adaptable to different interventions and suggested it may be good to hear from them. B. Belay noted there is not an implementation science session planned, but in thinking about what the toolkit might look like, thought it would be helpful to include this implementation science perspective. Something such as One Pass could be adapted as a tool for community-based programs. He added that some of the CDC grantees may also benefit from this. During her presentation, N. Sherwood indicated she tried to think about common measures that can be adapted, such as enrollment, engagement, and participation. She noted that we also have to think about the quality of the interventions delivered.

Each of the presenter’s talks appeared to be at a different level, and there are different levels that a program can invest in at getting data at these levels. C. Bolling agreed the organizational piece alluded to would be next part of the macro piece. B. Saelens added that his programs have done some important D&I work. He finds the fidelity assessments that K. Resnicow presented on to be very important to understanding what we’re doing with our interventions and then we figure out how we work with and assist interventions. He added he has assessed his intervention in a broader system.

K. Rhee was struck by the intensity of this work and how much effort it takes. While some measures such as attendance are easier to collect, evaluating what our partners want evaluated or view important, with fidelity, is difficult unless you have budgeted for this. She thinks this is important, but asked how do you do this and make it available, especially since D&I work is very qualitative? T. Earle felt that gets to the heart of her initial question. It is important, and programs are asked this frequently, but the reality is less people believe in using these measures because of the time and money it takes to implement them. She fears we’re too stuck in the weeds to understand they are all equally important.

Voula, Ken, Nancy, Bill, Brian, Hannah, Renee, Sarah, Sue, Bob, Matt, Teresa, Brook, Chris

Feedback on Session 3 Format

Session Format

- K. Resnicow felt that participation on the webinar was overwhelmingly an academic audience.
 - H. Zaganjor clarified the goal is to target **select** community-based healthy weight program participants who are the key audience for this session; programs will be targeted more widely when materials from this project are disseminated. Several participants from academia who are considered SMEs offer valuable input during the sessions.
 - B. Belay added that when one or two healthy weight programs are unable to attend a webinar or call, the audience can seem overwhelming academic, but we want to keep the final products geared toward the healthy weight program audience.

- S. Armstrong asked for clarification on how topics for each session were determined. She wondered if the community-level programs were not attending as much as we hoped because we need to better understand what they are looking for and what they see as the gaps.
 - B. Belay noted the community programs invited to be a part of the Collaborative Learning Project (CLP) attended the kickoff meeting and were able to provide input then. Due to constraints, NCCOR did not survey programs beyond those who were originally invited to this project.
 - S. Yanovski added that NCCOR held focus groups prior to the launch of this project to help inform it. NCCOR invited a select group of programs to participate in the CLP. The project was not open to all programs because this project is considered more formative to start targeting and then products would be adapted for wider distribution. It was clarified that NCCOR is not opening up the learning sessions to the broader community, but that the number of people attending include only those people who were invited.
- B. Saelens was wondering what is the next step after the formative presentations? Will there be a process to ensure some consistency in the language of the presentations?
 - H. Zaganjor stated we have not discussed this a group but indicated that another in-person meeting to work through next steps with the entire group is being considered.
 - B. Belay added that AAP may be an ideal group to help disseminate products from this project more broadly. They have done a lot of work around virtual collaboratives and connecting primary care with YMCAs.
 - S. Hassink, who is a part of the CLP, has expressed interest, but this project will need to take some next steps to make this more concrete.
 - T. Earle felt that between sessions, participants lose track of what the end product is going to be like, which makes it difficult to assess and answer some questions during the actual sessions. If she were to think about where this project is going in a vacuum, she agrees that you would want sessions to look similar so that others would know it is part of the process. She understands why we would not have done so to date but feels it may be time to start thinking about this. Having the end product in mind makes it easier to digest what is being presented.
 - B. Saelens added we need to have the community-based provider perspective to ensure there is logic and consistency in the information provided.
 - S. Yanovski indicated we have discussed having a “Chinese menu approach” which recognizes that consumers of this information may range from programs with little knowledge of program evaluation to more academic or experienced programs. This is the reason for targeting sessions at a mid-level: as the project develops materials, some could be targeted to programs with more evaluation background and some for those programs doing less-intensive types of evaluations.
- M. Levy felt that instead of having a discussion board, it may be helpful to have a very short survey after each session that captures the essence of that session, e.g., What target audience do you think this is for? Did you find this helpful? Can this information be easily disseminated? etc. Include more targeted questions and then an area for comments. Real-time feedback may be more helpful than the discussion board. Participants on the call agreed that this idea would be much more helpful.
 - S. Armstrong suggested you can even copy and paste what was shared on the surveys to the discussion board. Facilitators will plan to incorporate surveys into the sessions moving forward.
 - M. Levy added to keep in mind the audience and the time for this audience; it may be worth breaking up the sessions that will be disseminated more widely into 10- to 15-minute modules they can progress through.

- T. Earle added we need to think about who the ultimate end users of the toolkit are. She added it could be equally useful to people commissioning programs (e.g. local funder) not just the providers or creators of these healthy weight programs.

Webinar Follow-up Discussion

- H. Zaganjor reminded the participants that during the Session 3 webinar, participants felt that all the measures described are important to collect, but that many measures—such as those around fidelity or what partners want—may not be collected due to barriers such as budget constraints. It appears that some measures fall in the nice-to-have bucket of the Chinese menu for the toolkit. H. Zaganjor asked participants one of the discussion board questions: What process measures that you collect and think are most important to collect can also be collected across community-based healthy weight programs?
- B. Saelens indicated his program collects a lot of information about what they do to ensure consistency in the interventions they deliver, and when they compare different treatment models they can determine if they are delivering different treatment models. He felt K. Resnicow laid out a very detailed way to evaluate MI, but felt that may be too intense for a community-based program. He stressed the importance of understanding what is actually being delivered. He added that his program also assesses what they think is the dose that people receive—in terms of attendance, families' reactions to what was delivered to them, or other measures of dose. His programs find these measures very important and they always collect attendance and engagement data.
- B. Saelens added he was also struck by the presentation on what stakeholders want. He assumes different stakeholders want different things in terms of process. They may be less interested in fidelity because they assume you are delivering what you say you are delivering, but there may be additional things they'd want you to collect from families about dose and what they feel they are receiving.
- S. Armstrong agreed. In terms of the process measures that could be collected across programs, she felt that Nancy laid it out nicely in her presentation which measures are necessary to collect. She feels measures related to enrollment, attendance, and engagement are important to understand. It's likely that each individual program does not have the sample size to understand predictors of enrollment, attendance, and engagement, but if we are able to compare across the programs, then we may begin to understand which measures are working for which people and under what circumstances. She feels these measures are low-hanging fruit for most programs, although she recognizes they can still be a challenge for some.
- S. Armstrong also built on the comment about fidelity. She thinks it is important to capture across programs the adaptations that were successful. She thinks this is the space where we can learn a lot about what works in different communities—trying to understand where things deviated from protocol, why, and whether a change was positive for the program. She believes collecting both fidelity and engagement measures across programs would be important.
- H. Zaganjor reiterated the point made that by looking at programs as a whole that are evaluating similar measures, we may be able to see what is working.
 - N. Armstrong added it would be important to think about what we would want to know about each program such as setting and dose. We need a basis for comparison to link different types of programs or different intensities of programs with different levels of engagement and so forth.
 - T. Earle agreed: when you compare, it is difficult to ensure you are comparing like for like. This is important because when you get to other process measures which need to include some form of cost-effectiveness or cost-per-person, you want to ensure you are comparing like for like. For example, MEND and the Boys and Girls Club Triple Play

program, which is a primary prevention, cannot be compared because you cannot get near the \$10/person. People do not understand this difference, and this is a way to help them grasp the difference between these programs.

- B. Belay asked when it comes to collecting information on fidelity specifically around MI, are there ways to look at some of the tools K. Resnicow presented on to choose a few items from those tools that could become low-hanging fruit? He added when it comes to the community-based programs that may be struggling with evaluation of fidelity, what can we offer them?
 - K. Resnicow indicated there is the potential to create a short form on global measures; it is not as easy to do with the count measures. He is willing to develop a short form for his instrument, One Pass, and would attempt to contact other developers to see if they could create a short form for their instrument. B. Belay and H. Zaganjor will set-up a follow-up call with Ken and any interested participants to discuss this further.
 - H. Zaganjor wondered if the Health Care Climate Questionnaire, which has published a 6-item version, would be feasible for community-based programs to use. Participants felt that would be reasonable. K. Resnicow indicated he published a 6-item version on cancer care and found it worked very well. He believes the toolkit can include abbreviated versions of tools for this audience. He added another thing to consider is using EHRs to track fidelity. He has spent time during the past year tracking billing codes and CPT and ICD-10 combinations as a way to compute if obesity counseling took place. There has been a fair bit of work done to infer counseling sessions. Participants indicated they would be interested in discussing this further.
 - M. Levy indicated trying to catalogue this without spending a lot of money is impossible. Therefore, his colleagues use data about claims made s proxies. How you capture the things you are doing in an electronic format is an interesting discussion.
 - N. Sherwood mentioned that the Health Care Systems Research Network, previously referred to as HMO Research Network, published a paper called *Automating Assessments of Lifestyle Counseling in Electronic Health Records*. She added using EHRs to track counseling is not easy. B. Belay indicated he will talk to CDC colleagues, as they are considering how to create the electronic phenotype to capture what providers are doing.
- S. Armstrong noted C. Bolling's presentation highlighted the cost-benefit case. She indicated that, due to the Medicaid transformation, information on how much programs cost to deliver is becoming more important and wondered if this discussion should be included in this session.
 - T. Earle agreed that this is a very important discussion to have if we want to have long-standing value from the programs. The sustainability of the programs depends on this. Participants agreed this discussion would fit nicely into Session 6 on sustainability.

In summary, participants had great feedback for how to improve collecting input for the sessions. Future sessions will include a brief survey to garner feedback from participants immediately after calls.

Participants also shared key process measures that they feel should be included across the board for healthy weight program evaluations. It will also be critical to consider what information is collected on the types of programs to be able to use evaluation data for comparison across programs. Although the measures discussed are important, as the project evolves, it will be critical to consider what is manageable to include in final products for this project based on the audience. While EHRs are the future for data collection, if included in this project they need to be within scope. Given the audience, it will be important to have something in the toolkit about working with providers to have some sufficient way to connect to their EHRs.

In-depth Discussion on Fidelity Measures

A few participants had requested an additional call with K. Resnicow to have a deeper discussion on fidelity measures. The small group began by asking about the 6-item version of the HCCQ that was developed for cancer care.

- K. Resnicow indicated that questionnaire is one that patients or parents fill out to rate their clinicians. Core items on the HCCQ include, “My doctor listens to me, my doctor respects my decisions, and my doctor elicits my opinion.” The questionnaire allows for substitution of any professional within the stem such as “my dietician” or “my exercise counselor. The original questionnaire was used by self-determination theory researchers to measure autonomy support. The HCCQ makes it clear that the emphasis is on patients rating to what extent their providers supports their choices and feelings. The tool does very well at predicting outcome: when patients report high levels of autonomy from their providers, they report more positive outcomes such as losing more weight, quitting smoking, and taking medications. The original questionnaire is about 15 items, but it does not mention anything about nutrition or physical activity. K. Resnicow’s team took the original questionnaire and added items on diet and physical activity such as “the dietician supported my choices around screen time, supported my choices around nutrition.” This is the “obesity-specific” version. His team has a short form (6-items) that is not obesity-specific. They have used this form for cancer and heart disease, and it does not specifically mention screen time, nutrition, or physical activity.
- K. Resnicow indicated he would run the psychometrics on the longer form to determine what is the shortest version he can develop for the CLP toolkit.
- T. Earle noted the most helpful aspect is the application of understanding how people felt about their providers. The additional questions that contextualize nutrition and physical activity are picked up in the many questionnaires MEND requires. Parents and implementers have criticized MEND for the number of questionnaires to fill out. The area where MEND has less feedback is the “how did the providers do in the actual delivery” section.
 - T. Earle added that everything delivered in MEND has been evaluated extensively, but as they begin to adapt MEND to be deliverable by one health coach, that is where they get to the unknown because they do not have as much contact with these coaches. T. Earle finds that it is necessary to get to this point—that you know that your curriculum can be delivered by those coaches to the true fidelity of the program.
 - K. Resnicow noted that, while we think it’s about the content of the programs we create, the delivery of programs by providers ends up being as big a predictor of outcomes.
- K. Resnicow indicated they are working on a machine learning program (MPL) version of their tool. They have trained a computer to pick up empathy from providers; it is 95% accurate, but still a few years away from being functional. Ultimately, if you record sessions, the program will be able to detect the following: does the practitioner express empathy, does the practitioner reflect back what the patient says, etc.
- B. Belay asked about sharing the adapted tools publicly in the toolkits and whether there will be any proprietary issues. K. Resnicow indicated the original HCCQ tool and their tool are public domain, but because their tool was developed with the AAP, K. Resnicow will get their permission to share.
- S. Yanovski added that when K. Resnicow shortens his tool to get a 5–6 item tool that captures most of the variance, NCCOR can encourage people to see how this briefer tool works on prediction.
 - K. Resnicow added there was meta-analysis done in all the health care studies that use the HCCQ. Although it does not include many studies, the effect/magnitude was

- impressive; it always predicted better outcomes and is consistent across health behaviors. K. Resnicow indicated he has some data from his BMI2 study and could correlate that with outcomes too, but that would take a little more time.
- K. Resnicow indicated they also have the ONE-Pass, which is an expert rating where someone rates a clinician's audio or video encounter. It is more labor intensive and robust because it allows you to get very specific clinical feedback.
 - R. Kuczarski indicated that it would be interesting to do this. It seems like the results would be proportional to the time the practitioners have with patients. There may be situations with specialized lifestyle interventionists (e.g., trained RDs with these skills) who are trained to do bonding strategies. The social bonding is what kept programs like DPP and Look Ahead going and facilitated the sustained involvement of the participants if they had that relationship. You would expect to see a difference if you have resources such as trained RDs.
 - K. Resnicow added that, interestingly, diabetes educators do not score very well naturally on their measure because they are very proactive "information dumps." K. Resnicow has always wanted to compare various professions on these skills.
 - K. Resnicow closed by indicating that one thing this group can eventually discuss is data pooling. For example, if we had a few community-based programs that have used these tools and if we could find a way to share their data, we could conduct more robust analyses.

Discussion Questions

1. What process measures does your program evaluate and why?

B. Saelens indicated, "Our process measures are generally dictated by the research question we are addressing. However, across our various treatment trials, we have routinely assessed parental and child behavioral skills use (such as frequency of food and activity monitoring/tracking and change in the home food environment) throughout treatment as these are critical components to our conceptual model.

Our fidelity measures are focused on evaluating what our interventionists do and don't do in their delivery of the treatment. We capture this by audiotaping the delivery of the intervention. We then code a portion of the treatment sessions to evaluate on 1) behaviors we would want interventionists to engage in that are consistent with our conceptual model, 2) behaviors we would want interventionists not to engage in because they are not consistent with our conceptual model, and 3) general engagement behaviors (e.g., empathy, motivational approaches) that are not specific to our conceptual model but are important for engagement."
2. What processes do you follow to evaluate your program's process measures outlined in Q1?
3. What tools and/or methods do you use to evaluate your program's process measures outlined in Q1?
4. What barriers have you experienced evaluating process measures for your program?
5. What processes do you follow to evaluate the use of readiness assessments to identify children and families most likely to benefit from a HWP?
6. Are there process measures that your program does not evaluate, but which you believe are critical for HWPs to evaluate? If so, what are they and why do you believe they are critical to evaluate?
7. What process measures do you believe should be prioritized for HWP evaluation?
8. How does your program use process evaluation data to enhance engagement and outcomes?