PART B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS

1. Respondent universe and sampling methods

   Below we describe (1) the selection of staff for staff interviews; (2) the selection of service users for the service user interviews; (3) the selection of staff for the staff activity logs; (4) the selection of enrollees for the enrollee survey; and (5) the selection of providers for the provider survey.

   a. Staff interviews

   We will conduct interviews with staff members at four RETAIN program states during the first site visit and at the same four states during the second. For each visit, the evaluation team will visit two regions of the state.

   SSA and Mathematica will work with each state's RETAIN program administrator to identify project staff to interview. SSA and Mathematica will also identify key staff from partner organizations who can provide a range of perspectives on the implementation of the RETAIN program. The RETAIN program administrators come from many different agencies, some of which focus on the areas of labor and employment. Examples from RETAIN Phase 1 projects include the state employment development department; departments of labor, commerce, or workforce investment; employment and economic development department, job and family services department, and employment security department.[1] Line staff from partner agencies that will deliver RETAIN services could include social service agencies, rehabilitation counselors, employment specialists, and social workers based in health care systems. To better understand the range of perspectives, in states where the program model is such that it is appropriate to do so, SSA and Mathematica will identify up to two different providers that differ on some key dimension (such as area of the state or typical delivery model). Evaluators will return to the same two providers for each visit and seek to interview the same staff whenever possible to avoid repeating information and also to gather observations about changes over time. If staff attrition occurs from one site visit to the next, evaluators will work with RETAIN program leaders to identify new staff for the interviews.

   A precise estimate of the number of staff involved in each state's program is unavailable. In developing the evaluation plan, SSA and Mathematica determined— based on past experience evaluating other demonstrations—that two five-day site visits are enough to collect the information necessary to inform the analyses. We determined the number of staff to be interviewed at each RETAIN program (19) for each visit based on this time frame and the expected length of the interviews. We will select all line staff will be purposefully for inclusion in the interviews so as to reflect a wide range of staff roles, perspectives, and experiences implementing the model.

---

[1] Information about the RETAIN state programs is available at
https://www.dol.gov/odep/topics/SAW-RTW/RETAIN-Phase-1-Recipient-Snapshots.htm (accessed January 29, 2020).

b. RETAIN service user interviews

Mathematica will conduct the RETAIN service user interviews with a convenience sample of 60 enrollees drawn from the universe of the RETAIN treatment group (Exhibit B1). The interviews will take place during August 2022. We will interview each service user separately. To complete 60 interviews, Mathematica will select up to 150 service users per state across the four states and invite each to participate in the interviews. Mathematica will use administrative data to identify RETAIN service users based on their intensity of service use (high, moderate, and low). We plan to interview three subgroups of five service users at each site (i.e., 15 per state). Mathematica will sample from among qualifying enrollees who meet those criteria. Mathematica will send an invitation by mail to all of the potential interviewees in each state. The letter will introduce the evaluation, describe the purpose of the interview, and ask enrollees to call a toll-free number to schedule an appointment to complete an interview.

Exhibit B1-1. Site visit interviewee and service user sample sizes, by RETAIN state

	State 1	State 2	State 3	State 4	Total
Site visit 1					
RETAIN administrators/directors	1	1	1	1	4
RETAIN program line staff	18	18	18	18	72
RETAIN service user interviews					
RETAIN high service users	5	5	5	5	20
RETAIN moderate service users	5	5	5	5	20
RETAIN low/withdrawals	5	5	5	5	20
Staff activity logs					
RETAIN administrators/directors	1	1	1	1	4
RETAIN program line staff	12	12	12	12	48
Site visit 2					
RETAIN administrators/directors	1	1	1	1	4
RETAIN program line staff	18	18	18	18	72

[a] At the time of this submission, specific states have not been selected for Phase 2. We therefore list states generically in this table.

The main intent of the interviews is to provide qualitative information on the experiences of service users. Given the relatively small sample sizes, the resulting sample will not be representative of all RETAIN service users. Resource constraints and the study design prevent SSA from conducting additional interviews. Nonetheless, the qualitative information is important in providing in-depth information about key issues regarding RETAIN service usage that is unavailable in any other planned data source available for RETAIN.

c. Staff activity logs

Manage and line staff who are substantively involved in each RETAIN program represent the population of interest for the activity collection logs. Each program uses a different staffing approach to provide services. We will conduct telephone meetings with each state's RETAIN program administrator to describe data collection for the logs,

discuss the staff positions involved in providing program services, and identify which positions to include in the data collection effort. We will collect activity logs from members of all staff positions (such as administrators and return-to-work coordinators) in which RETAIN activities represent a primary part of their duties. If a program has many people (4 or more) in a staff category, we will consult with the program manager on whether to select a sample of staff (2 to 10 people, depending on the number of staff in that category) to complete the log. If a program has 3 or fewer staff in a position, we will ask everyone in that category to complete the log. We will work with the program manager to identify the staff from each category to include in data collection. We will also ask the program manager to provide additional instructions and activity descriptions to include in the log instructions.

d. Surveys of RETAIN enrollees

The respondent universe for the enrollee survey is all individuals who enrolled in RETAIN during Phase 2. Mathematica will conduct the survey with a sample of 12,000 enrollees, equally distributed across the four state programs (3,000 per state) (Exhibit B1-2). If a state plans to enroll more than 3,000 individuals, Mathematica will select a representative subsample of enrollees.[2] There are two rounds of survey interviews: the first will occur 2 months after enrollment and the second will take place 12 months after enrollment. The same samples will be used for both surveys; Mathematica will recontact all eligible R1 sample members in R2. The RETAIN evaluation design assumes a 24-month enrollment period.

We will finalilze the sampling strategy when the states are selected for Phase 2 and we have finalized the enrollment targets. We will stratisfy the enrollee survey sample proportionally by enrollment month, state program, and study assignment group (treatment or control). We may consider other strata, including age and employment status, after we know the Phase 2 programs. The sampling strategy for the enrollee survey seeks to ensure that individuals we select for the survey reflect the state's overall RETAIN enrollee population.

e. Surveys of RETAIN providers

The respondent universe for the provider survey is all RETAIN medical providers offering services. Mathematica will conduct the provider survey with 400 RETAIN providers, equally distributed across the four program states (100 per state) (Exhibit B1-2). We anticipate that these providers will reflect a range of medical fields, including primary care physicians, physical therapists, occupational therapists, nurse practitioners, and mental health providers. If a state has more than 100 providers, Mathematica will select a representative subsample. We will administer the provider survey twice. The first round will take place 4 months after the launch of Phase 2, and the second round will take place 16 months after the launch of Phase 2.

At this time, we do not yet know the universe of providers. We will finalize decisions that affect the sampling strategy for this survey when we select the states for Phase 2, and we know the number of providers in each state. In RETAIN states that are using

---

[2] States will include their enrollment targets in the Phase 2 applications.

clustered random assignment, we anticipate that the full provider list will be available at the start of Phase 2. In the states that are using individual random assignment, we will identify providers as individuals enrolled in the demonstration, so the universe of providers will change between survey rounds. Mathematica plans to draw a sample of providers for any state where the universe of providers is greater than 100. The survey sample will include the universe of all providers if a given state program has 100 or fewer providers. Mathematica's will finalize plans for stratifying the provider survey sample when we select the program states. Potential strata for the provider survey strata include geographic region, practice size, and provider type. Eligibility for the second round of the provider survey is not contingent upon participation in the first round. We will refresh the sample R2 survey to account for providers who have joined RETAIN since the R1 survey. We will consider providers ineligible for the survey if they are either no longer providing services at the practice organization or if they are deceased. We will include providers who report that they have no knowledge of RETAIN or their participation in the demonstration in the survey sample, as this provides important evidence for the impact analysis.

Exhibit B1-2. Sample sizes, by RETAIN state

	State 1	State 2	State 3	State 4	Total
Survey of RETAIN enrollees R1	3,000	3,000	3,000	3,000	12,000
Survey of RETAIN enrollees R2	3,000	3,000	3,000	3,000	12,000
Survey of RETAIN service providers R1	100	100	100	100	400
Survey of RETAIN service providers R2	100	100	100	100	400

[a] At the time of this submission, specific states had not been selected for Phase 2. We therefore list them generically in this table.

2. Procedures for collecting information

a. **Staff interviews**

The first step in this data collection process will be to send an email to each RETAIN state liaison to seek his or her cooperation before each site visit. Jackson Costa, SSA project manager for the RETAIN evaluation will send the email, to lend credibility to the study and to further encourage cooperation. The evaluation team member who will lead the site visit will send a follow-up email to the RETAIN state liaison to schedule a phone call to discuss in greater detail the information to be gathered from RETAIN stakeholders during the site visits and in-person interviews. We will ask RETAIN state liaisons to identify interview respondents who will be able to provide the required information and for information about their general schedule constraints. We will work with the RETAIN state liaison and partner contacts to develop a schedule for the site visits and interviews that meets respondents' needs and causes minimal disruption to operations. Approximately two weeks before a site visit is to take place, Mathematica will email the final interview schedule to the RETAIN state liaison. The email will also contain contact information for the evaluation team member who will be leading the site visit so respondents can reach that person if a need arises to change the schedule or if other issues emerge. Giving the RETAIN stakeholders adequate information early and

in a professional manner will help build rapport and ensure that stakeholders are available and responsive and that interviews go smoothly.

The evaluation team member who will lead the site visit will use an interview guide crafted from the interview topic list in Attachment B to conduct the interviews with the administrator and staff. Before each interview, the team member leading the interview will ask the respondent for verbal consent to participate in the interview and for permission to record the interview. We anticipate that each interview will take approximately 60 minutes. After completing all interviews for a particular RETAIN program, the evaluation team member leading the site visit will develop a summary of the information collected during the site visit and phone interviews (if any). Upon completing the visit, Mathematica will email a thank you letter to the state liaison and the other RETAIN stakeholders involved in organizing the visit.

b. Interviews with RETAIN service users

The first step in this data collection process will be to reach out to the RETAIN service users whom we have identified to participate in the interviews. We anticipate sending an invitation to enrollees to participate. The invitation will explain Mathematica's request and provide a toll-free number to call to schedule an interview or to decline participation. The letter will come from Jackson Costa, SSA project manager for the evaluation, as this will lend credibility to the study and further encourage cooperation. The evaluation team will monitor the toll-free number and will schedule interviews at times that are convenient for RETAIN these individuals. When booking interviews, staff will confirm the best telephone number to reach the individual as well as that person's preferred means for receiving a reminder the day before the interview.

Evaluation staff will reach out to all enrollees on the day before the scheduled appointment using the enrollees' preferred mode of contact. During the interview, the team member leading the interview will obtain the service user's verbal consent to participate and will also request consent to digitally record the interview. The interviewer will explain the benefits and risks associated with participation, confidentiality of the information shared during the interview, and the voluntary nature of participation. The interviewer will inform service users that they may request that the interviewer suspend recording at any time and will assure them that the interviewer will not request personally identifying information during the interview. The interviewer will use an interview guide, based on the topic list in Attachment C, to conduct the interviews. We expect each interview will last up to 30 minutes. Mathematica will mail a thank you letter with a $30 gift card to each service user who completes the interview.

c. Staff activity logs

As a first step, we will schedule telephone interviews with each state's RETAIN program manager to discuss the staff activity logs. During this meeting, we will discuss the logs, review the staff categories and number of staff in each category, consider which categories to include in the data collection activity, and identify staff in each category to complete the logs. We will also ask the RETAIN program manager to identify two one-week periods during which staff could track their time. These should

be periods that represent typical weeks in program service delivery, not times when uncommon activities such as staff conferences or trainings are occurring. The periods will vary for each program, depending on the timing of the site visit.

For each period and program, we will follow a similar schedule to administer the staff activity logs:

- About one to two weeks before data collection begins, we will ask the RETAIN program manager to inform staff that Mathematica will be contacting them about the logs.

- We will send an email to selected staff the Wednesday before the target week of data collection. The email will contain the staff activity log (Attachment I) in Excel and PDF formats, and the body of the email will explain the rationale for the log, provide instructions on how to complete it and where to go for more information, and present options for returning the log to Mathematica. The log is self-administered and respondents can complet it either electronically or with pen and paper.

- On the Monday after staff complete their logs, we will send emails to those who have not yet returned the logs, asking them to do so.

When we receive the logs, we will enter the data into an Excel spreadsheet for analysis.

d. Enrollee surveys

Mathematica will collect the contact information for enrollees during the enrollment process and will use these data to field the enrollee survey. Information about enrollees' preference for English or Spanish will determine the language used for each sample member's initial survey letter and subsequent nonresponse follow-up efforts. The surveys will use a sequential, mixed-mode design, offering sample members the opportunity to complete the questionnaire on the web, by mail, or by telephone. Each round of the enrollee survey will follow the same methodology for sample release, eligibility, and nonresponse follow-up.

We will send all sample members a $5 pre-pay cash incentive with the survey invitation letter. By deploying a prepayment, resources can target sample members who are likely to require intensive efforts to locate, contact, or engage for interviews. All survey respondents will receive a $25 gift card. We estimate the R1 survey will take 12 minutes to complete, and the R2 survey will take about 18 minutes.

Contact with an enrollee survey sample member begins when Mathematica sends the survey invitation letter to all eligible sample members during the first week of the field period (Attachment F). This letter will include a link that enrollees can use to log into the web survey, along with a personalized login code. The letter will also include a toll-free telephone number that enrollees can call to complete the survey over the telephone, or to seek support in completing the survey online. Mathematica will utilize a "push to web" methodology, encouraging self-reporting by web for as many sample members who are willing and able to do so (Dillman 2017). Accordingly, subsequent

reminder postcards will also include the survey link and password information.[3] Mathematica will send a paper version of the questionnaire via first-class mail during week three of the enrollee survey field period, with another copy sent to nonrespondents approximately three weeks later. Each paper questionnaire will be followed by a postcard reminder, encouraging sample members to return the completed survey.

After we offer sample members the opportunity to take part in the survey by web and by mail, Mathematica will begin telephone follow up with nonresponding sample members. Telephone interviewers will receive project-specific training. Further, all of Mathematica's Spanish-speaking interviewers will complete professional certification to ensure that they are qualified to conduct interviews in Spanish. Telephone follow up will begin during week six of the 12-week field period for sample members without a viable mailing address. It will occur later for the remainder of the sample. Mathematica will send reminder mailings during the remaining weeks of the survey period to all outstanding sample cases to: (1) encourage them to participate in the survey, (2) respond to concerns they may have about the study, and (3) notify them the survey will end soon and provide them with a paper survey to complete. Enrollee survey mailings will provide text that stresses the voluntary nature of participation and that the decision to participate in the survey will not affect enrollees benefits, now or in the future.

SSA and Mathematica will design the enrollee instruments to accommodate a wide range of disabilities. We will make the wording of the questions as simple as possible to allow accessibility to those with mild cognitive disabilities. We will train phone interviewers to offer breaks, where needed, to accommodate enrollees with illnesses or injuries that cause stamina limitations. Although we cannot design instruments that will address every disability we may encounter, these basic design characteristics will enable us to interview most enrollees in the study without the use of proxies. We will incorporate alternative question wording for proxies to allow for instances when an enrollee cannot complete an interview independently.

Over the enrollee survey field period, Mathematica's data collection managers will use a range of production reports to monitor the data collection and ensure it aligns with production, cost, and quality goals. Mathematica will carefully monitor response rates overall, and for each enrollment cohort, program, and group assignment. Mathematica will use its sample management system (SMS) to execute several key tasks that support the success of the survey and minimize burden on sample members. For example, we will use this system to: (1) schedule the release of eligible cases; (2) mail invitation and reminder letters and incentive payments; (3) track and store sample cases' updated contact information; (4) cease nonresponse follow up to sample members who have completed the survey (via any mode); and (5) provide production statistics that could help inform outreach strategies to future cohorts.

---

[3] To ensure privacy of this sensitive information, Mathematica will use postcards that feature a seal, which recipients will need to open to view the information provided.

e. Provider surveys

Each participating state program selected for Phase 2 will provide Mathematica with the contact information, including the preferred email address, for all of their participating providers. This information will also provide data on sample geographics and practice organization.

Each round of the provider survey will follow the same methodology for sample release, eligibility, and nonresponse follow up, and each round will have a 14-week field period. The first round of the provider survey will launch four months after the start of Phase 2 and the second round will releas 16 months after the start of Phase 2. The provider survey will use a sequential, mixed-mode design, offering sample members the opportunity to complete the questionnaire on the web, by mail, or by telephone. We will send all sample members a $5 pre-paid incentive with the survey invitation letter and all provider survey respondents will receive a $45 check. The R1 and R2 provider surveys are each expected to take 14 minutes to complete. We anticipate that all providers will complete the provider survey in English. However, Mathematica will offer a professionally translated version of the paper questionnaire, upon request, for providers who prefer Spanish.

Mathematica will send an invitation letter to all presumed eligible providers selected for the survey sample in the first week of the field period (Attachment H), asking them to complete the survey online. This letter will include a link they can use to log into the web survey and a personalized login code. The letter will also include a toll-free telephone number that providers can call to complete the interview over the telephone, or to seek support in completing online, as needed. We will follow the letter with an invitation email, which will provide a customized link through which sample members can log into the provider survey. Mathematica will utilize a "push to web" methodology, encouraging self-reporting using the web (Dillman 2017). Accordingly, subsequent reminder postcards will also include the survey link and password information.[4]

Mathematica will send a paper version of the questionnaire via priority mail approximately three weeks into the provider survey field period, with another copy sent approximately four weeks later. Priority mail will help support messages about the survey legitimacy and importance and will also help set the survey mailing apart from other correspondence sample members may receive. Mathematica will follow each paper questionnaire with a postcard reminder, encouraging sample members to return the completed survey. After offering sample members the opportunity to take part in the survey by web and by mail, Mathematica will begin telephone follow-up to nonresponding sample members. Telephone interviewers will receive project-specific training that includes training in working with gatekeepers, who often screen calls at practice organizations. Telephone follow-up will begin in week 9 of the 14-week field period. Across the 14-week field period, Mathematica will send reminders to all eligible nonresponding provider sample members to (1) encourage them to participate,

---

[4] To ensure the privacy of this sensitive information, Mathematica will use postcards that feature a seal, which recipients will need to open in order to view the information provided.

(2) respond to concerns they may have about the study, and (3) notify them that the survey will end soon and their unique input is critical to the success of the evaluation. All contacts and outreach will emphasize the voluntary nature of participation.

During each round of the provider survey, Mathematica's data collection managers will use a range of production reports to monitor the data collection and ensure that it aligns with production, cost, and quality goals. During the field periods, Mathematica will carefully monitor the provider survey response rate overall, as well as for each state program and stratum (if we use strata in sampling) to minimize nonresponse bias among any one program or subgroup of providers. As with the enrollee survey, Mathematica will use its sample management system to execute several key tasks to support the successful deployment of the provider survey.

Estimation procedures

The objective of the impact analysis is to provide statistically valid and reliable estimates of the effects of each Phase 2 RETAIN program on the outcomes of enrollees. Mathematica is currently working with Phase 1 states to develop potential rigorous impact study designs. At this time, we expect all Phase 2 states to rely on experimental designs to estimate the causal impacts of RETAIN. Random assignment will enable estimation of the net impact of RETAIN by comparing average outcomes across the treatment and control groups. Mathematica's analysis will focus on intent-to-treat estimates, which measure how the *offer* of RETAIN services shaped enrollees' behavior after they enrolled in the demonstration.

States are pursuing two different approaches to random assignment. Some states are proposing individual-level random assignment, whereby eligible individuals would be randomly assigned to a treatment group offered to RETAIN services or to a control group that is not able to access RETAIN services. Other states are proposing clustered random assignment, whereby the state will randomly assign a group of providers or a geographic area to the RETAIN treatment group and the other providers and areas continue to offer business-as-usual services. Both of these random assignment designs provide causal impact estimates, but the designs have implications for the estimation procedures and the precision of the estimates.

To avoid concerns about data mining and reduce the extent of false positives, Mathematica will pre-specify a parsimonious set of primary outcomes in important study domains. A preliminary list of these outcomes includes employment, earnings, and applications for Social Security disability benefits. The evaluation team will refine this list and define specific outcomes prior to conducting the analysis. Evaluation reports will also include results for secondary outcomes related to services received, job characteristics, income, and health and well-being. However, the evaluation team will consider these results exploratory. This approach strikes a balance between addressing the multiple comparisons problem while maintaining the evaluation's ability to detect policy-relevant impacts.

Precision of estimates

Even with an experimental design, we need sample sizes large enough to provide sufficient statistical power for statistically significant impact estimates in cases where the program produces large enough estimates which policymakers or practitioners find meaningful. Each state plans to enroll at least 3,000 individuals in the study and may have an enrollment target as large as 10,000 enrollees. Based on these sample sizes, we present in Exhibit B2-1 the minimum impacts we expect to detect using administrative and survey data for our key outcome of not being employed the year after random assignment. We consider MDIs under two different study designs: individual random assignment and clustered random assignment.

The minimum detectable impacts (MDIs) in Exhibit B2-1 suggest that the planned study samples will support the detection of meaningful impacts consistent with estimates in the literature. For example, under an individual random assignment design, we expect to detect program impacts of 3.6 percentage points or larger using administrative records, and 4.3 percentage points or larger using survey data. These MDIs represent a 14–17 percent reduction in the share of individuals not employed the year after random assignment. A rigorous evaluation of the Centers of Occupational Health and Education pilot found that the program reduced joblessness at 12 months by 21 percent (Wickizeret et al. 2011), a larger impact than the MDIs presented here.

The MDIs are larger for the clustered random assignment design, but still consistent with findings from the previous literature. For example, under a clustered random assignment design, we expect to detect program impacts of 3.7 percentage points or larger using administrative records, and 4.9 percentage points or larger using survey data. These MDIs represent a 15–20 percent reduction in the share of individuals not employed the year after random assignment.

Exhibit B2-1. MDIs, by study design and data source

	Individual random assignment			Clustered random assignment		
	Sample size	MDI	Relative MDI %	Sample size	MDI	Relative MDI %
Follow-up data from administrative records						
Whether not-employed the calendar year after random assignment	3,500	0.036	14.2	6,000 (175)	0.037	14.7
Follow-up data from surveys						
Whether not-employed one year after random assignment	2,400	0.043	17.1	2,400 (175)	0.049	19.8

The mean value of the outcome for control group members is assumed to be 0.25. MDI calculations assume (1) an equal number of treatment and control members, (2) a 95 percent confidence level with an 80 percent level of power, (3) a one-tailed test, (4) a reduction in variance of 5 percent owing to the use of regression models with the individual random assignment, (4) an intraclass correlation coefficient of 0.025 for the clustered design, (6) a reduction of variance of 5 percent owing to the use of regression models in the clustered design, (7) administrative data obtained on 100 percent of the sample, and (8) survey response rates of 80 percent.

3. Methods to maximize response rates and deal with nonresponse

a. Staff interviews

Mathematica will work with the RETAIN program leaders in each state to determine the best times and formats (group versus individual; phone versus in person) to convene the interviews with RETAIN staff. Mathematica will also limit the interviews to approximately one hour so that data collection imposes only a modest burden on respondents. Mathematica will use separate discussion guides for each potential respondent type so we don't ask respondents about activities or issues that do not apply to them. To further support respondent convenience, data collectors will meet interview respondents in their own offices or at a location of their choice when in-person interviews are the chosen approach.

b. Interviews with RETAIN service users

Because we will draw RETAIN service user interviews from a convenience sample of volunteers, target response rates to ensure a representative population are not at issue. To mitigate interview nonresponse, evaluation team members will contact each individual who has an appointment on the day before the interview, using the person's preferred mode of communication. A team member will confirm the best telephone number to use to reach each enrollee so the appointment takes place when the enrollee is not distracted by other responsibilities. Because we will conduct the interviews by telephone, participants will not face barriers related to transportation to an interview location. Finally, the $30 honorarium will encourage interview participation.

c. Staff activity logs

Our discussions with the RETAIN program managers intend to identify the right people to ask about completing the staff activity logs, to identify the best times for this data collection, and to facilitate staff's completion of the activity. We will send up to three reminder emails to staff after the data collection period to ask for their completed logs. The evaluator will also limit data collection to no more than 5 minutes per day (or 35 minutes over all seven days) to minimize the burden on respondents.

d. Surveys of RETAIN enrollees and providers

The evaluation team designed the survey fielding methods to maximize response rates. The mixed mode of administration will offer potential respondents the flexibility to complete the survey in a manner that is most convenient for them. The evaluation team designed the length of the surveys to balance the evaluation's need for information with the need to minimize burden to encourage response. Survey materials and procedures will assure respondents of the privacy of their responses to questions, which should

address privacy-related reasons for nonresponse. Below, we describe specific methods for minimizing nonresponse in each of the proposed survey data collections.

e. **Enrollee surveys**

The surveys of RETAIN enrollees will target a response rate of at least 80 percent at each round. Mathematica will follow industry best practices for deploying nonresponse follow-up efforts in accordance with the mixed-mode design in a way that minimizes the burden of sample members and does not suppress responses to the survey. Features of our approach include the following:

- **Incentives to motivate survey response.** The survey features both a prepaid and a post-pay incentive, both of which will motivate enrollee sample members to complete the survey. Moreover, respondents from the first round of the survey will likely remember receiving the incentive, further motivating them to participate in the second round.
- **A mixed-mode survey, featuring two modes for self-reporting.** Respondents will have the option to complete the enrollee survey in the mode that they prefer, by web, paper, or over the telephone with a professional interviewer. The web survey will be accessible using any device (tablet, smartphone, laptop, or desktop computer) and will be compatible with assistive technologies respondents may need to participate online. Both the web and paper questionnaires can be completed at the sample member's convenience.
- **Sequential release of the survey modes.** Mathematica will release the web survey before sending the paper questionnaire or conducting outbound telephone calls to nonresponders.[5] The sequential release of modes will avoid potential confusion among sample members or the perception of complexity (Dillman 2014; Medway 2012).
- **Multiple languages for survey administration.** By offering the survey in both English and Spanish, we minimize nonresponse bias for sample members who would not be able to complete the survey in English.
- **Outreach to sample members in different formats.** Varying outreach formats can improve survey response rates because not all modes of follow-up will resonate with all sample members. For example, some may not open their postal mail, but they may call back when we leave a voicemail. Others may find the mailings a reassuring source of legitimacy and mistrust contact by telephone.
- **Informational webpage and toll-free telephone number**. SSA will host a webpage describing the survey to provide a way for sample members to verify study legitimacy and gather more information. Mathematics interviewers will staff a survey toll-free number to answer questions and address sample members' concerns on a range of issues.

---

[5] Mathematica's Survey Operations Center will field inbound calls across the survey field period, as the toll-free telephone number for the survey will be on all survey mailings.

- **Accommodations and proxy respondents.** Mathematica's telephone interviewing team will be well-prepared to address accommodation needs of respondents with disabilities, whether the accommodations are stamina related, or with the use of assistive technology. Moreover, if a sample member's disability or impairment is too severe to allow them to participate in the survey via self-report, we have the necessary protocols in place to complete the interview with a designated proxy, such as another adult in the household.
- **Paradata to inform optimal days of week and times of day for telephone outreach.** Because of the rolling release of cohorts for the enrollee surveys, we will be able to leverage survey paradata from early cohorts to inform follow-up efforts for later cohorts. For example, we will be able to to align the contact attempts with the time periods that have generated the highest completion rates.

These strategies, combined with Mathematica's intensive oversight of production statistics across the field period, will help maximize response rates for the enrollee survey.

f. Provider surveys

The surveys of RETAIN providers has a target response rate of at least 80 percent in each round. As with the enrollee survey, Mathematica will follow industry best practices for deploying nonresponse follow-up efforts in accordance with the mixed-mode design in a way that minimizes sample member burden and does not suppress survey response. We will use each of the approaches described for the enrollee survey in section B3.1, including (1) monetary incentives to motivate survey response; (2) a mixed-mode survey with two modes offered in self-reporting format; (3) a sequential release of the survey modes; (4) multiple languages available for survey administration; (5) a variety of outreach formats (email, postal mail, and telephone calls); (6) a toll-free number and informational webpage; and (7) paradata to inform optimal times to contact providers. In addition, trained Mathematica telephone interviewers will work with providers' frontline office staff to ensure that the providers are aware of the survey request and to help facilitate completion of the questionnaire. These strategies, combined with Mathematica's intensive oversight of production statistics across the field period, will help maximize response rates for the provider survey and minimize nonresponse.

Addressing nonresponse in the enrollee and provider surveys

Nonresponse is a key source of survey error that impacts the quality of the data collected in both the enrollee and provider surveys. In sections B3.1 and B3.2, we discussed our approach to minimizing unit nonresponse bias and achieving target response rates for both surveys. Below we describe the proposed approaches for addressing both item and unit nonresponse as the survey data for use in the impact analysis.

Item nonresponse

Although the RETAIN evaluation team's past experience conducting surveys for similar evaluations suggests that rates of item nonresponse on the enrollee and provider surveys will be very low, some item nonresponse is inevitable, especially when using self-administered modes for survey data collection. Both the enrollee and provider surveys collect data on outcome measures for the evaluation team to use in the impact analysis. Imputation of outcome data could lead to biased estimates due to imperfect matches on observables when using a hot-deck procedure (Bollinger and Hirsch 2006). We will exclude observations with missing outcome data, unless we knew the outcome to have a specific value for some cases that was conditional on the value of another variable. To minimize the risk of bias from this source, we will use multivariate techniques to input missing data.

Unit (individual-level) nonresponse

As with almost any survey, some nonresponse in each round of data collection is inevitable, especially during later rounds of follow up. For example, the evaluation team will not be able to locate some sample members and others will not be able or willing to respond to the survey. SSA expects to attain a response rate of at least 80 percent for each round of the enrollee and provider surveys. In the event that response rates are lower, Mathematica will conduct a nonresponse analysis using various data items from enrollment. The nonresponse bias analysis will consist of the following steps:

- *Compute response rates for key subgroups.* The evaluation team will compute the response rate for the subgroups using the American Association for Public Opinion Research definition of the participation rate for a nonprobability sample: the number of respondents who provided a usable response divided by the total number of individuals from whom they requested participation in the survey (AAPOR 2016). The evaluation team will conduct comparisons of the response rate across key subgroups, including most notably the treatment group and the control group, as well as any subgroups used to stratify sampling for the survey. The goal of this analysis is to determine whether response rates in specific subgroups differ systematically from that of other subgroups or from the overall response rate. This could inform the evaluation team's development of nonresponse weights for use in the analysis.

- *Compare the distributions of respondents' and nonrespondents' characteristics.* Again, using data from RETAIN enrollment, the evaluation team will compare the characteristics of respondents and nonrespondents. They will assess the statistical significance of the difference between these groups using *t*-tests. This type of analysis can be useful in identifying patterns of differences in observable characteristics that might suggest nonresponse bias. However, this approach has low power to detect substantive differences when sample sizes are small, and the large number of statistical tests conducted can also result in high rates of Type I error. Consequently, the evaluation team will interpret the results of this item-by-item analysis cautiously.

- *Identify the characteristics that best predict nonresponse and use this information to generate nonresponse weights.* This is a multivariate generalization of the subgroup analysis described previously. The evaluation team will use logistic regression models to assess the partial associations between each characteristic and response status; propensity scores obtained from such models provide a concise way to summarize and correct for initial imbalances (Särndal et al. 1992). Examples of automated procedures they could use to produce these weights efficiently include: (1) using prespecified decision rules, such as those described by Imbens and Rubin (2015) and Biggs, de Ville, and Suen (1991) to select covariates and interactions between them; and (2) identifying and addressing outliers by, for example, trimming weights in a way that minimizes the mean-square error of the estimates (Potter 1990).

- *Compare the nonresponse-weighted distribution of respondent characteristics with the distribution for the full random assignment sample.* In this last step, the evaluation team will compare the weighted distribution of respondent baseline characteristics to the unweighted distribution of the full set of RETAIN enrollee that went through random assignment. They will make these comparisons for the whole sample and for key subgroups, as described earlier in this subsection. This step will include validation of the nonresponse weights using outcomes measured in the program data for the full sample (but not used in the construction of the weights). This analysis can highlight measures in which the potential for nonresponse bias is greatest, even after weighting, in which case they should exercise greater caution in the interpretation of the observed findings.

4. Tests of procedures or methods to be undertaken

Interviews with RETAIN staff and service users

We will conduct no pre-tests of the interview protocols. Mathematica will make minor modifications to the data collection procedures and discussion guides, if necessary, based on the experiences of the early interviews.

Surveys of RETAIN enrollees and providers

Mathematica conducted a pretest of the enrollee and provider questionnaires. The pretest provided an accurate estimate of respondent burden as required by OMB, and the evaluation team also assessed flow and respondent comprehension.

We completed the R1 and R2 enrollee survey pretests in English and Spanish with nine adults who had experienced an illness or injury that could preclude their participation in work (so as to mirror the RETAIN enrollee population). Mathematica conducted the R1 provider survey pretest with nine providers from a range of specialties. Because the R2 provider instrument mirrors the R1 instrument, we did not conduct additional testing of the R2 provider instrument. Following the pretests, Mathematica revised the survey instruments based on their findings.

5. Individuals consulted on statistical aspects of the design and on collection and/or analyzing data

   In addition to their collaboration with staff from DOL, SSA has also organized a Technical Working Group (TWG) to provide input on key research questions, evaluability considerations, feasible experimental and nonexperimental methods, survey designs, analysis strategies, and interpretation and presentation of results. The TWG consists of researchers and clinicians with expertise in the areas of disability, early intervention, and evaluation design. The external experts include:

   - Thomas Wickizer, Ph.D., Ohio State University College of Public Health

   - Glenn Pransky, former director at Center for Disability Research at the Liberty Mutual Research Institute

   - Carolyn Heinrich, Ph.D., Vanderbilt University

   - Jack Smalligan, M.A., Urban Institute

   - Frank Neuhauser, Ph.D., University of California at Berkeley's Institute for the Study of Societal Issues

   - Douglas Martin M.D., Medical Director, UnityPoint Health – St. Luke's Occupational Medicine

   - Marianne Cloeren, M.D., M.P.H., University of Maryland School of Medicine

   - Benjamin Doornink, M.B.A., Kootenai Health

   An interdisciplinary team of economists, disability policy researchers, and survey researchers on staff at Mathematica or at the evaluation subcontractor (Tree House Economics, LLC) contributed to the design of the overall evaluation. The team consisted of:

   - Jillian Berk, Ph.D., Mathematica

   - Kenneth Fortson, Ph.D., Mathematica

   - Rosalind Keith, Ph.D., Mathematica

   - Gina Livermore, Ph.D., Mathematica

   - Holly Matulewicz, M.A., Mathematica

   - David Wittenburg, Ph.D., Mathematica

   - David Stapleton, Ph.D., Tree House Economics, LLC

REFERENCES

American Association for Public Opinion Research (AAPOR). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. Ninth edition. Oakbrook Terrace, IL: AAPOR, 2016.

Anand, Priyanka, and Yonatan Ben-Shalom. "The Promise of Better Economic Outcomes for Workers with Musculoskeletal Conditions." Roosevelt House Faculty Forum, Hunter College, May 2017.

Ben-Shalom, Yonatan, and Hannah Burak. "The Case for Public Investment in Stay-at-Work/Return-to-Work Programs." Issue brief. Washington, DC: Mathematica Policy Research, March 2016.

Berlin, M., L. Mohadjer, J. Waksberg, A. Kolstad, I. Kirsch, D. Rock, and K. Yamamoto. "An Experiment in Monetary Incentives." *Proceedings of the Section on Survey Research Methods*. Alexandria, VA: American Statistical Association, 1992.

Biggs, David, Barry de Ville, and Ed Suen. "A Method of Choosing Multiway Partitions for Classification and Decision Trees." *Journal of Applied Statistics*, vol. 18, no. 1, January 1991, pp. 49–62.

Bollinger, Christopher R., and Barry T. Hirsch. "Match Bias in the Earnings Imputations in Current Population Survey: The Case of Imperfect Matching." *Journal of Labor Economics*, vol. 24, no. 3, July 2006, pp. 483–520.

Bureau of Labor Statistics. "Nonfatal Occupational Injuries and Illnesses Requiring Days Away From Work, 2015." November 2016. Available at https://www.bls.gov/news.release/osh2.nr0.htm. Accessed May 25, 2018.

Bureau of Labor Statistics, U.S. Department of Labor. "Occupational Employment and Wages, May 2018: 11-9151 Social and Community Service Managers." March 29, 2019a. Available at https://www.bls.gov/oes/current/oes119151.htm. Accessed November 13, 2019.

Bureau of Labor Statistics, U.S. Department of Labor. "Occupational Employment and Wages, May 2018: 21-1015 Rehabilitation Counselors." March 29, 2019b. Available at https://www.bls.gov/oes/current/oes211015.htm. Accessed November 13, 2019.

Bureau of Labor Statistics, U.S. Department of Labor. "May 2018 National Occupational Employment and Wage Estimates, United States." April 2, 2019c. Available at https://www.bls.gov/oes/current/oes_nat.htm. Accessed November 13, 2019.Cho, Y. I., T. P. Johnson, and J. B. VanGeest. "Enhancing Surveys of Health Care Professionals: A Meta-Analysis of Techniques to Improve Response." *Evaluation & the Health Professions*, vol. 36, no. 3, September 2013, pp. 382–407.

Dillman, Don A. "The Promise and Challenges of Pushing Respondents to the Web in Mixed-Mode Surveys." *Survey Methodology,* vol. 42, no. 1, June 2017, pp. 3–30. Available at https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2017001/article/14836-eng.pdf?st=UIIot_RI]. Accessed January 25, 2020.

Dillman, Don A., Jolene D. Smyth, and Leah Melani Christian. *Internet, Mail and Mixed-Mode Surveys: The Tailored Design Method.* Hoboken, NJ: John Wiley & Sons, 2014.

Hollenbeck, Kevin. "Promoting Retention or Reemployment of Workers after a Significant Injury or Illness." Washington, DC: Mathematica Policy Research, October 2015.

Imbens, Guido, and Donald Rubin. *Causal Inference in Statistics, Social, and Biomedical Sciences.* New York: Cambridge University Press, 2015.

Jäckle, Annette, and Peter Lynn. "Respondent Incentives in a Multi-Mode Panel Survey: Cumulative Effects on Nonresponse and Bias." Working paper presented to the Institute for Social and Economic Research, University of Essex, Colchester, United Kingdom, 2007.

James, Jeannine M., and Richard Bolstein. "The Effect of Monetary Incentives and Follow-Up Mailings on the Response Rate and Response Quality in Mail Surveys." *Public Opinion Quarterly,* vol. 54, no. 3, Autumn 1990, pp. 346–361.

Jurisic, Maja, Melissa Bean, John Harbaugh, Marianne Cloeren, Scott Hardy, Hanlin Liu, Cameron Nelson, and Jennifer Christian. "The Personal Physician's Role in Helping Patients with Medical Conditions Stay at Work or Return to Work." *Journal of Occupational and Environmental Medicine,* vol. 59, no. 6, June 2017, pp. e125–e131.

Kay, Ward R. "The Use of Targeted Incentives to Reluctant Respondents on Response Rates and Data Quality." *Proceedings of the American Association for Public Research.* Montreal, Canada: American Association for Public Opinion Research, 2001.

McLeod, C. C., C. N. Klabunde, G. B. Willis, and D. Stark. "Health Care Provider Surveys in the United States, 2000–2010: A Review." *Evaluation & the Health Professions,* vol. 36, no. 1, March 2013, pp. 106–126.

Medway, Rebecca L., and Jenna Fulton. "When More Gets You Less: A Meta-Analysis of the Effect of Concurrent Web Options on Mail Survey Response Rates." *Public Opinion Quarterly,* vol. 76, no. 4, Winter 2012, pp. 733–746.

Potter, Francis J. "A Study of Procedures to Identify and Trim Extreme Sampling Weights." In *Proceedings of the American Statistical Association, Section on Survey Research Methods.* Alexandria, VA: American Statistical Association, 1990, pp. 225–230.

Särndal, Carl-Erik, Bengt Swensson, and Jan Wretman. *Model-Assisted Survey Sampling.* New York: Springer-Verlag, 1992.

Schwartz, Lisa K., Lisbeth Goble, and Edward M. English. "Counterbalancing Topic Interest with Cell Quotas and Incentives: Examining Leverage-Salience Theory in the Context of the Poverty in America Survey." *Proceedings of the American Association for Public Research.* Montreal, Canada: American Association for Public Opinion Research, 2006.

Singer, Eleanor, and Richard A. Kulka. "Paying Respondents for Survey Participation." In *Studies of Welfare Populations: Data Collection and Research Issues,* edited by Michele Ver Ploeg, Robert A. Moffitt, and Constance F. Citro. Washington, DC: National Academies Press, 2001. Available at www.nap.edu/read/10206/chapter/6. Accessed April 22, 2019.

Social Security Administration. "Annual Statistical Report on the Social Security Disability Insurance Program, 2016." SSA Publication No. 13-11826. Washington, DC: Social Security Administration, October 2017.

Wickizer, T.M., G. Franklin, D. Fulton-Kehoe, J. Gluck, R. Mootz, T. Smith-Weller, and R. Plaeger-Brockway. "Improving Quality, Preventing Disability and Reducing Costs in Workers' Compensation Healthcare: A Population-based Intervention Study." *Medical Care,* vol. 49, no. 12 (December 2011I), pp. 1105–1111.