THE PATHWAYS CLEARINGHOUSE GUIDE FOR RESEARCHERS

People who run programs for job seekers with low incomes need evidence on the interventions and strategies that can help their clients succeed in the labor market. Moreover, researchers want to offer practitioners, policymakers, and other stakeholders evidence to guide decision making.

To provide reliable, accessible information about what works to help job seekers find and keep gainful employment, the Office of Planning, Research & Evaluation at the Administration for Children & Families launched the Pathways Clearinghouse. The Pathways Clearinghouse is built on a foundation of rigor, credibility, and accessibility.

The Pathways Clearinghouse identifies interventions that aim to improve employment outcomes, reduce employment challenges, and support self-sufficiency for people with low incomes. The Pathways Clearinghouse systematically evaluates and summarizes the evidence of interventions' effectiveness. Researchers can also use the Pathways Clearinghouse to understand the extent and quality of existing research on employment and training interventions, as well as the contexts and populations that have been studied.

This guide introduces researchers to the Pathways Clearinghouse by providing a high-level overview of its processes for selecting studies for review, assigning study quality ratings, and assessing the evidence of effectiveness for an intervention. It then describes the Pathways Clearinghouse review standards and the reasoning behind them, and outlines common pitfalls that might cause studies to receive lower quality ratings.

This guide is for researchers conducting rigorous evaluations of employment and training interventions. It introduces the Pathways to Work Evidence Clearinghouse and the standards and procedures it uses for assessing the quality of causal research studies.

How can researchers use this guide?

The information in this guide can support researchers and research funders with understanding the following:

- How to conduct studies that receive the highest possible quality ratings from the Pathways Clearinghouse.
- Why their study was or was not included in the Pathways Clearinghouse, why their study received its study quality rating, and how their study was grouped with other studies to assess intervention effectiveness.
- Common pitfalls for study authors to avoid to ensure their research achieves a high quality rating and therefore reaches the broadest possible audience through Pathways Clearinghouse dissemination efforts.

To find out more, visit https://pathwaystowork.acf.hhs.gov.

The Pathways Clearinghouse screening and review process

The Pathways Clearinghouse assesses existing research through a screening and review process (Exhibit 1). The team first screens existing research for eligibility. Then, reviewers assess the strength of the study's evidence and assign a study rating. Finally, the Pathways Clearinghouse groups studies into interventions and assigns effectiveness ratings to interventions based on the available evidence. This guide describes each step in detail.

Exhibit 1. Overview of Pathways Clearinghouse screening and review process

1 Screen. The Pathways Clearinghouse first identifies research that might be eligible for review by systematically searching electronic databases, examining existing literature reviews, and coordinating with other federal evidence clearinghouses (see the Protocol for the Pathways to Work Evidence Clearinghouse: Methods and Standards for more detail on these search processes). The Pathways Clearinghouse also issues calls for papers to encourage authors and other stakeholders to share studies directly with the Clearinghouse. Next, a team of trained screeners determines the eligibility of these manuscripts.

For research to be eligible for review, it must satisfy eight **eligibility criteria** (Exhibit 2), related to the characteristics of the intervention being studied and the features of the study design, analysis, and reporting.

Exhibit 2. Pathways Clearinghouse Eligibility Criteria

Intervention Characteristics	Study Characteristics
Has the aim of directly or indirectly improving employment or earnings outcomes*	Assesses effectiveness using quantitative methods and eligible study designs
Serves individual job seekers, in a specific context	Analysis conducted in 1990 or later
Serves adults or youth age 16 and older with low incomes	Intervention services are clearly articulated in the research
Implemented in the United States or Canada	Examined the effect of an intervention on employment or earnings outcomes

*Before March 2021, the Pathways Clearinghouse included only research examining employment and training programs. It has since been expanded to include other types of programs that aim to indirectly improve employment-related outcomes.

Eligible research must examine an intervention that aims to directly or indirectly improve employment or earnings. Eligible interventions can include employment and training programs, or interventions that could indirectly improve employment or earnings outcomes. For example, although housing assistance does not directly help people get jobs, it might indirectly increase employment (and, therefore, earnings) by reducing a key employment barrier.

The Pathways Clearinghouse focuses on interventions serving people with low incomes who are age 16 and older. Pathways examines research focused on people who are 16 and older as this captures a full age range of individuals who may participate in the labor market. Groups meeting the low-income criteria include people who are experiencing homelessness, people with prior justice system involvement, means-tested public benefit recipients, and disconnected youth (people age 16 to 24 who are not working or in school), as well as populations characterized as being "low income" or "low skill" by a study's authors. The Pathways Clearinghouse does not classify unemployed workers or recipients of Unemployment Insurance or Social Security Disability Insurance as low income, unless they meet one of the other criteria above.

What study designs are eligible for review?

1. Randomized controlled trials **(RCTs)**
2. Comparison-group quasi-experimental designs **(QEDs)**

Both types of designs compare study participants in an intervention group to those in a comparison group, and compare outcomes for the two groups to assess the effects of the intervention.

The Pathways Clearinghouse only includes research on interventions delivered in a specific context, in which a set of services, conditions, or policies are delivered to or apply to one group and not another comparable group (see Box: What study designs are eligible for review?). It excludes research that solely seeks to understand the effects of federal or state programs, policies, policy adaptations, or funding streams unless those programs are studied in the context of a particular implementation or bundle of services. Although national or statewide evaluations of a policy that affects that entire geographic area are excluded (for example, an evaluation of the national Workforce Innovation Opportunity Act Youth program), the focus on specific context means that a local evaluation of a program supported by a national or statewide funding stream (for example, Workforce Innovation and Opportunity Act-funded local programs) would be eligible for review.

2 **Review studies.** After determining a study is eligible for review, the review team assesses the strength of the causal evidence based on the Pathways Clearinghouse review standards, described in detail in the next section. Two reviewers examine each study selected for review. The first reviewer documents all relevant information and assigns a study rating. The second reviewer thoroughly checks the review to ensure the review standards were applied correctly and that the review captures all relevant information. In consultation with senior staff, reviewers reconcile any differences in reviews.

If the information needed to complete the review is not available in the manuscript, the Pathways Clearinghouse will conduct an **author query** to request this information from the study author. If the study author does not provide the necessary information, the review team will complete the review by assigning the highest possible rating based on the information available. The Pathways Clearinghouse also conducts author queries to obtain information necessary to calculate an impact's effect size.

3 **Rate studies.** Reviewers then assign the study a **study quality rating** (Exhibit 3). The goal of the study quality rating is to provide a systematic, consistent, unbiased rating of the quality of the causal evidence in reviewed studies.

Exhibit 3. Pathways Clearinghouse study quality ratings

	Description	Types of studies with this rating
High	A **high rating** indicates a relatively low risk that the study produces biased estimates of an intervention's causal effect. A reader can be confident that the findings are attributable to the intervention examined.	This rating is reserved for findings from high-quality RCTs.
Moderate	A **moderate rating** indicates some risk of biased estimates because of the study's design or analysis.	Well-executed QEDs can receive this rating, as can RCTs that do not meet the criteria for the high rating but satisfy other design criteria.
Low	A **low rating** indicates a relatively high risk that the study produces biased estimates of an intervention's causal effect. A reader should not assume that the findings are attributable to the intervention examined.	This rating is assigned to RCTs and QEDs that, because of flaws in the study design or analysis, do not meet the criteria for either the high or moderate rating.

QED = quasi-experimental design; RCT = randomized controlled trial.

What is a study?

The Pathways Clearinghouse review team defines a study as an analysis of a distinct implementation of an intervention.

One manuscript can include more than one study. In other cases, a study might be reported across multiple manuscripts (a manuscript can be published or unpublished).

- If authors present study impacts for a pooled sample, the review team will consider all individuals in the pooled sample as receiving the same distinct implementation of an intervention (in other words, this would be considered one study).
- If authors do not present pooled results and only present results separately by cohort or site, each subgroup analysis will be considered a distinct study.
- If authors present impacts for both pooled and site- or cohort-specific effects but indicate that the subgroup analyses are the primary estimates of interest and the pooled impact estimates are exploratory or secondary, the Pathways Clearinghouse classifies each subgroup analysis as a separate study.

Rate intervention effectiveness. After reviewing the studies and assigning quality ratings, the Pathways Clearinghouse synthesizes information across studies to rate the evidence of effectiveness for the interventions being studied. The Pathways Clearinghouse groups studies into interventions based on the services examined in the study and the participation requirements for those services (whether participation in services is mandatory or voluntary). That is, two studies are considered to examine the same intervention only if the same services were offered in both studies. In addition, studies in which participation in services is mandatory are classified as examining different interventions than studies with voluntary participation. The Pathways Clearinghouse rates intervention effectiveness for

What are short- and long-term outcomes?

For the purposes of the Pathways Clearinghouse, **short-term outcomes** are those measured within 18 months after an individual is assigned to a study group (the intervention or the comparison group). **Long-term outcomes** are those measured 19 to 60 months after an individual is assigned to a study group. **Very long-term outcomes** are those measured 61 months or more after assignment.

each of 10 **outcome domains** (or groups of related outcomes). The outcome domains are defined by the type of outcome measure (employment, earnings, public benefit receipt, and education and training) and the time span over which they are measured (short-term, long-term, and very long-term—for all outcome types except education and training).[1]

The Pathways Clearinghouse assigns an **intervention effectiveness rating** (Exhibit 4) for each outcome domain. This intervention rating indicates the level of evidence to support the likelihood that the examined intervention will improve outcomes in a given domain if the intervention were replicated. The assigned rating depends on the extent of favorable and statistically significant effects among high and moderately rated studies that examine the given intervention.[2]

See the Protocol for the Pathways to Work Evidence Clearinghouse: Methods and Standards for further detail regarding the requirements an intervention must meet to receive each effectiveness rating.

Exhibit 4. Intervention effectiveness ratings, by outcome domain

Well-supported	There is strong and consistent evidence that the intervention produces favorable results in the specified outcome domain. These interventions have at least two impact studies of moderate or high quality that show evidence of favorable findings within the domain.
Supported	There is some evidence that the intervention produces favorable results in the specified outcome domain. These interventions have at least one study of moderate or high quality that shows evidence of favorable findings in the domain, but the evidence is less conclusive than that for well-supported interventions.
Mixed support	There is some evidence that the intervention produces favorable results for a specific outcome domain, and some evidence that the intervention worsens outcomes in that domain.
Not supported	There is strong evidence that the intervention is unlikely to produce substantial favorable results in the specified outcome domain. Studies with moderate or high quality ratings of these interventions have found patterns of null effects, unfavorable effects, or both in the specified outcome domain.
Insufficient evidence	There is some evidence from moderate or high quality studies of the intervention's effect in the specified outcome domain, but this evidence is not sufficient to assign one of the other effectiveness ratings.
No evidence	There is no high or moderate quality evidence of the intervention's effects in the specified outcome domain.

The Pathways Clearinghouse review standards

To consistently rate study quality across eligible research, the Pathways Clearinghouse review team uses a set of uniform review standards. The review standards provide rules for prioritizing outcomes to review and assigning study ratings.

[1] Education and training outcomes are considered at just one time point (in other words, the Pathways Clearinghouse does not define this outcome domain over multiple time spans).

[2] The Pathways Clearinghouse aimed to explore interventions that help people become more economically self-sufficient; therefore, it considers decreases in public benefit receipt to be favorable. In contrast, increases in all other outcomes cataloged in the Pathways Clearinghouse are considered favorable.

Prioritizing outcomes for review

Many studies of employment and training interventions report findings from numerous outcome measures across multiple time spans. Including all of these outcomes in the review process might increase the likelihood of finding false positive effects of the intervention simply by chance. To avoid this, and to better compare findings across studies, the Pathways Clearinghouse review team selects a limited set of priority outcomes on which to focus the study review. These priority outcomes fall into one of the 10 Pathways Clearinghouse outcome domains.

Exhibit 5 provides examples of priority outcomes and commonly excluded outcome measures within each broad outcome domain, in no particular order. See Appendix Exhibit A.1 for the standardized approach Pathways Clearinghouse reviewers use to select outcomes for review.

Exhibit 5. Examples of outcomes prioritized for review by the Pathways Clearinghouse

Outcome area	Time spans selected	Examples of prioritized outcomes	Examples of excluded outcomes
Employment 	Short-term Long-term Very long-term	• Employment status at the time of follow-up, or the most recent time period • Cumulative measures of employment status (for example, number of quarters employed)	• Employment rates for each year after random assignment
Earnings 	Short-term Long-term Very long-term	• Annual earnings for the latest elapsed year of the follow-up period	• Earnings measures for only people who are employed • Measures of income combining earnings and unearned income • Measures of earnings above a specific threshold (such as the poverty line)
Public benefit receipt 	Short-term Long-term Very long-term	• Indicators of benefit receipt and the (dollar) amount of annual benefits received for the latest elapsed year of the follow-up period	• Benefit amounts for only people who receive benefits • Supportive services payments received through the intervention
Education and training 	Attainment over the follow-up period	• Individual measures of educational milestones (such as training completion or degree attainment)	• Credits earned

The Pathways Clearinghouse tries to select outcome measures for review that are easy to compare across studies. For example, the percentage of people employed at the time of follow-up is an outcome that is easily interpretable across studies. In contrast, the number of academic credits attained might be difficult to compare because of differences in how credits are measured in different contexts. Some outcomes are comparable given certain conditions. For example, outcomes measured in dollars can be compared over time if the year of measurement, or the year that authors used to index dollars, is taken into account.

Some outcome measures are excluded, such as measures of employment at points in time other than the end of the follow-up period; these findings might raise concerns of false positives because of multiple comparisons. Although study authors might have certain outcomes in mind when beginning an analysis (for example, the annual employment rate), they might have ultimately examined more finely grained outcomes (such as quarterly employment) to identify effects. The Pathways Clearinghouse specifies preferred points in time for priority outcome measures to prevent artificially inflating the chances of finding a statistically significant effect. For example, if study authors report quarterly employment effects over a five-year period after random assignment, the Pathways Clearinghouse would select employment for only the latest quarter in each time span (short-term, long-term, and very long-term) as outcomes to review.

When study authors use multiple data sources to assess outcomes in a domain (for example, long-term earnings are measured in survey data and administrative data), the Pathways Clearinghouse review team uses the guidelines for prioritizing outcomes (Appendix Exhibit A.1) to select one set of outcomes from each source.

Causes of bias in randomized controlled trials (RCTs)

The Pathways Clearinghouse examines studies for evidence that the study design and analysis can credibly estimate that the intervention being studied caused the effects that researchers observed. Studies receive the highest possible rating if the design and analysis choices can reasonably rule out other potential causes of the estimated effects. In an RCT, researchers randomly assign study participants to either an intervention group or a comparison group. Random assignment can produce the most reliable evidence of causal effects because, if done correctly, it creates intervention and comparison groups that have no systematic differences in observable and unobservable characteristics before the intervention.

Two factors could jeopardize the validity of an RCT and lead to a low study quality rating.

1. **Confounding factors** are components of a study whose effects combine with the effect of the intervention in ways that are difficult to disentangle. These confounding factors lead to differences between the intervention and comparison groups that are impossible to distinguish from the effects of the intervention.[3]

Examples of studies with confounding factors include the following:

- An intervention is implemented in Chicago, where all intervention group members live, but all comparison group members live in Detroit.
- A study examines the effects of a job search assistance program on employment for people with low incomes who have certificates. The comparison group consists of people with low incomes who have high school diplomas.
- A study examines the effects of supportive services for Temporary Assistance for Needy Families (TANF) recipients. Study authors measured outcomes for intervention group members in 2006, and outcomes for comparison group members in 2008.

What does the Pathways Clearinghouse consider a confounding factor?

A study component is considered a **confounding factor** in the Pathways Clearinghouse if it is:

- Observed
- Completely aligned with only one study group
- Not an intended component of the intervention being evaluated

Confounding factors that satisfy these conditions can be systematically identified from the study design and context.

[3] Note that aspects of group formation (such as people choosing to participate in the intervention or being deemed ineligible) are not typically considered to be confounding factors for the purposes of the Pathways Clearinghouse.

2. **Mishandling missing data** can also cause bias that might lead to incorrect conclusions regarding the effectiveness of an intervention. Study authors might use a variety of approaches to handle data that are missing at baseline (before the intervention) or at follow-up.

If a study uses a method of handling missing data that is acceptable according to the Pathways Clearinghouse review standards, it is eligible to receive a high study quality rating. If the study does not use one of these methods, it will receive a low study quality rating.[4] The methods for handling missing data that are not acceptable, such as certain forms of hot deck imputation, might pose risks to the validity of study findings because they might create artificial differences between the intervention and comparison group that are not caused by the intervention (Andridge and Little 2010; Allison 2001; Little and Rubin 2019). These requirements for handling missing data are based on those used by the What Works Clearinghouse Version 4.1 (WWC 2020), which are in turn based on recommendations and evidence from the existing literature, such as Little and Rubin (2019) and Wooldridge (2002).

What methods of handling missing data does the Pathways Clearinghouse accept?

Prior research has determined that these methods do not impose unnecessary bias on the estimated effects.

- **Complete case analysis,** or including only people with non-missing data in the analysis.
- **Regression imputation,** in which study authors use a regression model to predict values for the missing data. Acceptable forms of imputation include single and multiple imputation.
- **Maximum likelihood,** in which study authors estimate findings using a process that accounts for patterns in the data that is missing.
- **Nonresponse weights,** in which study authors determine weights based on an estimated probability that an observation would be missing an outcome. This method is acceptable for missing outcome data only.
- **Dummy variable imputation,** in which study authors replace all missing values with a constant, and include a dummy variable in the model for observations with missing data. This method is acceptable for missing baseline data only and only for RCTs.

For more information, see the What Works Clearinghouse Standards Handbook, Version 4.1 (WWC 2020).

Two factors could cause a moderate risk of bias in an RCT, leading the RCT to be reviewed as if it is a quasi-experimental design (QED) study. These studies can receive a rating of moderate or low (but not high).

1. **Compromised random assignment** can cause differences between the intervention and comparison groups at baseline that could contribute to the estimated effects of the intervention.

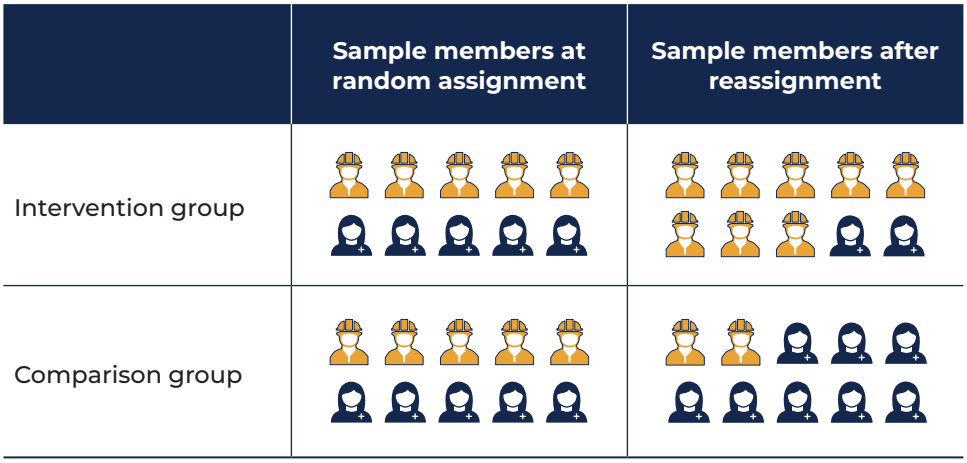
Compromised random assignment can occur if study authors use a random process to assign participants to study groups but analyze participants in a group to which they were not randomly assigned. A common example of this occurs when not all members of the intervention group received the intervention (for example, because some individuals offered a training program never showed up to receive training), and the authors use statistical methods or sample selection to estimate the effect of the intervention on only the people who received it, rather than on all people offered the intervention. Researchers often refer to this as estimating the treatment-on-the-treated effect rather than the intent-to-treat effect.[5]

[4] Note that the Pathways Clearinghouse calculates attrition based on the share of the sample with observed outcomes; study authors' imputation cannot be used to lower the attrition rate.

[5] An intent-to-treat analysis can be complemented by additional exploratory analyses estimating treatment-on-the-treated effects.

As another example, suppose a study wants to estimate the effect of a job search assistance program. After the researchers perform the random assignment, half of participants in each of the intervention and comparison groups were pursuing jobs in the construction sector, and half in the health care sector. However, after seeing workers in the construction sector struggling to find jobs, researchers decide to replace 30 percent of the participants in the intervention group with construction participants from the comparison group. As seen in Exhibit 6, this type of reassignment changes the composition of the two research groups, making it difficult to disentangle any estimated impacts of the program from differences in the characteristics of the intervention and comparison group members. If the study authors analyzed the data based on the original random assignment status (an "intent-to-treat" analysis), this would not cause a risk of bias. Reassignment is only considered to compromise random assignment in an RCT if study authors analyze participants in a group to which they are not randomly assigned.

Exhibit 6. Reassignment after random assignment



Compromised random assignment can also occur if study authors **vary the probability of random assignment** to the intervention or comparison group and do not account for this difference in their analysis (Exhibit 7). Varying the probability of assignment to the intervention or comparison group creates differences in the characteristics of the research groups at baseline. Not accounting for these varying probabilities can cause a risk of bias due to differences in outcomes that are not a result of the intervention, but rather a result of differences in the composition of the two groups. For example, imagine that the study authors in the previous example decided before random assignment that participants interested in the health care sector would have a 50 percent probability of assignment to the intervention group, and those interested in the construction sector would have a 70 percent probability. As shown in Exhibit 7, the share of each type of participant in the intervention and comparison groups then becomes unequal.

To ensure a causal interpretation of the results from an RCT with varying probability of assignment, authors can use one of several acceptable techniques:

- Using weights to account for the different probabilities of random assignment.
- Including an indicator variable for all individuals with the same probability of assignment.
- Separately estimating effects for all individuals with the same probability of assignment, and pooling the results to provide the overall estimate.

Studies using an acceptable technique to account for varying probabilities of random assignment in an RCT are eligible to receive a high study quality rating

Lastly, compromised random assignment can occur if the analysis sample is selected based on a variable or factor that could be influenced by the intervention (for example, examining outcomes only for those who are employed). In this case, it is difficult to disentangle whether the estimated impacts are caused by the intervention itself, or by differences between the intervention and comparison group that are created as a result of this sample restriction.

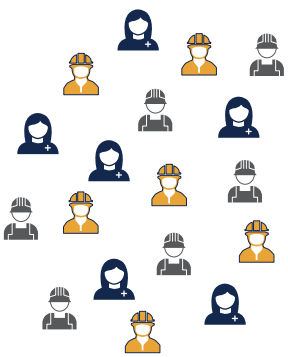
Exhibit 7. Varying the probability of assignment

Group A: 50% assigned to intervention	Group B: 70% assigned to intervention	Sample at random assignment	
		Intervention group	
		Comparison group	

2. **Attrition** refers to loss of sample members over time and can cause bias in estimates. Attrition occurs when some randomly assigned people are not included when calculating the effects of the intervention. Attrition can occur for many reasons—for example, because researchers cannot locate a participant at follow-up, administrative data are missing for a participant, a participant decided to drop out of an intervention, or a participant chose not to respond to a survey.

Attrition can cause bias because it might lead to differences in the composition of the intervention and comparison groups at follow-up that are not a result of the intervention. Exhibit 8 illustrates this. At random assignment, the intervention and comparison group each have equal shares of the three types of participants. However, because of sample loss over time, the intervention and comparison group members observed at follow-up are no longer composed of equal shares of each type of participant. Attrition can also cause bias because sample members observed at follow-up might have different characteristics than those who were not observed.

Exhibit 8. Attrition

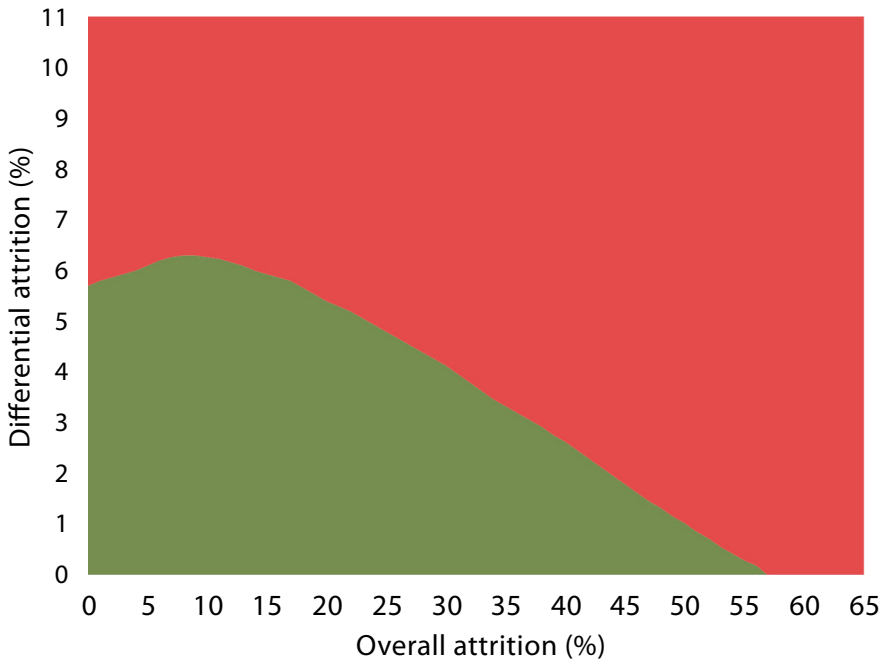
Eligible sample		Sample members at random assignment	Sample members observed at follow-up
	Intervention group		
	Comparison group		

The Pathways Clearinghouse considers two types of attrition as potential contributors to bias: **overall attrition,** or the share of study participants who were randomly assigned but are missing outcome data, and **differential attrition,** or the difference in the share of missing cases between the intervention and comparison groups.

To determine whether attrition poses a risk of bias, the Pathways Clearinghouse applies an attrition threshold that accounts for levels of overall and differential attrition (Exhibit 9).[6] The threshold was constructed based on a model that estimates the level of potential bias associated with various combinations of overall and differential attrition. For example, if overall attrition is high, the differential attrition between the intervention and comparison groups must be relatively low to meet the attrition threshold boundary. This tradeoff between overall and differential attrition has been calculated based on the restriction that attrition should lead to a bias of less than 0.05 standard deviations of the outcome (WWC 2013; WWC 2014).

If the level of attrition in a study is below the Pathways Clearinghouse attrition threshold (shaded in green in Exhibit 9), the study has **low attrition,** implying that the potential level of bias from attrition is sufficiently low. If there are no other identified causes of bias, then the study can receive the highest study quality rating.

Exhibit 9. Pathways Clearinghouse attrition standard



[6] This is equivalent to the What Works Clearinghouse cautious attrition threshold.

Causes of bias in QEDs and RCTs that cannot receive a high study quality rating

In a comparison-group QED, study participants belong to either an intervention group or a comparison group. However, unlike an RCT, these groups are not randomly determined. Instead, researchers typically identify intervention group members who receive services and construct a comparison group using people who did not receive services.

Confounding factors and mishandling missing data, as described earlier, can lead to bias in comparison-group QEDs. QEDs might also suffer from bias because of the way groups are formed. Because participants are not randomly assigned to study groups, intervention group members might differ from comparison group members, based on observed and unobserved dimensions. These differences at baseline might lead to differences in outcomes measured at follow-up that are not a result of the intervention. Because it is difficult to identify all unobserved differences between study groups (for example, differences in motivation), even if there are no observed baseline differences in a comparison-group QED, the causal evidence from a comparison-group QED is weaker than that from an RCT.

RCTs with compromised random assignment, or high attrition, are subject to the same sources of bias as a comparison-group QED. In both types of designs, regardless of how study authors initially assigned people to the intervention and comparison groups, the people remaining in the groups within the analytic sample cannot be considered to have been randomly allocated. For an RCT with high attrition, this lack of randomness stems from nonrandom attrition, rather than nonrandom assignment (as Exhibit 8 demonstrates).

There are two additional factors that the Pathways Clearinghouse considers in reviewing QEDs and compromised or high-attrition RCTs.

1. **Baseline controls** are important features of comparison-group QEDs and compromised or high-attrition RCTs because of the potential for differences across study groups. All studies reviewed as comparison-group QEDs or compromised or high-attrition RCTs must use an analytic technique that adjusts estimates for any differences in the outcome at baseline. Adjusting for these baseline measures of the outcome in the analysis reduces the risk that the estimated effect of the intervention is the result of preexisting differences in the outcome between groups. Studies typically control for baseline outcome measures by including control variables in their regression analysis. Controlling for a propensity score summarizing the probability of group assignment (rather than directly controlling for the baseline or lagged measures used to construct the propensity score) is not an acceptable method of controlling for pre-intervention outcomes.

   Analyses of employment- or earnings-related outcomes should adjust for a measure of employment (for employment outcomes only) or earnings (for both employment and earnings outcomes) measured at least 1 year before the intervention began. In other words, the measure must be from 366 days or more before random assignment or the start of the intervention. This lagged measurement requirement exists because people participating in an employment or training program have been shown to experience a pre-enrollment dip in earnings (Ashenfelter 1978; Heckman and Smith 1999). Thus, measuring employment or earnings immediately before the intervention might not capture the potential differences in outcomes between the intervention and comparison groups. For all other outcomes, the baseline control should be measured shortly before the intervention began.

2. **Baseline equivalence** is also necessary, even if a QED or compromised or high-attrition RCT study controls for baseline differences in the outcome measure. Study authors must demonstrate baseline equivalence in key characteristics of the intervention and comparison groups at the start of the study. Demonstrating baseline equivalence provides reasonable confidence that any differences in the measured outcomes between the study groups are a result of the effects of the intervention, rather than differences that existed at the start of the study.

The Pathways Clearinghouse has identified key measures for which study authors must demonstrate equivalence (Exhibit 10). These key measures are likely to influence the outcomes of interest in the research being assessed. The Pathways Clearinghouse review team considers equivalence to be established between the intervention and comparison groups if the difference in means across the two groups is not statistically significant at the 5 percent level (using a chi-squared test for categorical variables and a two-tailed t-test otherwise).

Exhibit 10. Pathways Clearinghouse requirements for baseline equivalence between intervention and comparison group in QED studies

Key characteristics	Time frame in which to demonstrate equivalence
1. Earnings (or employment for employment outcomes)	At least 12 months before baseline (366 days or more)
2. Socioeconomic status (such as educational attainment or receipt of some means-tested public benefit)	At or near baseline
3. Race and ethnicity	At any point in time
4. Gender	At any point in time
5. Age	At any point in time

A study is eligible to receive a moderate rating if it is a comparison-group QED study or compromised or high-attrition RCT that includes appropriate baseline controls, demonstrates equivalence on key characteristics, and does not suffer from bias because of confounding factors or mishandling of missing data. If any of those conditions do not hold, the study will receive a low rating.

Common pitfalls for study authors to avoid

Often, a study design has the potential to receive a high or moderate evidence rating but does not do so because of features of the analysis or reporting. Exhibit 11 describes common research pitfalls that cause studies to receive a lower evidence rating than might be possible given their design, reporting pitfalls that might lead to an author query, and recommended solutions to avoid the pitfalls. Note that there might still be other issues that preclude a study from receiving a high or moderate rating.

Exhibit 11. Common pitfalls for Pathways Clearinghouse study authors and recommended solutions

Research issue	Common pitfall	Recommended solutions to avoid pitfall resulting in a lower rating
Some individuals assigned to the intervention group did not receive the intervention, whereas some in the comparison group did receive the intervention.	Comparing outcomes in an RCT for people based on the services they actually received, or excluding noncompliers (people in the intervention group who did not receive the intervention, or people in the comparison group who did).	Conduct an intent-to-treat analysis, analyzing differences between the intervention and comparison groups, regardless of the services actually received. (This can be complemented by additional analyses based on actual service receipt.)
Random assignment probabilities differed across participants (for example, based on site).	Not accounting for differing probabilities of random assignment to study groups.	Use weights; include indicator variables for all people with the same probability of assignment; or, separately estimate effects for all people with the same probability of assignment and pool the results to provide the overall estimate.
Accounting for differences in study groups at baseline.	Not controlling for baseline or lagged measures of the outcome, or controlling for a propensity score summarizing the probability of group assignment (rather than directly controlling for the baseline or lagged measures used to construct the propensity score).	Include controls for pre-intervention measures of the outcome, measured at least 366 days before baseline for employment or earnings outcomes.
Assessing differences in preprogram earnings.	Demonstrating baseline equivalence in immediate preprogram earnings but not a measure of earnings or employment representative of participants' typical earnings.	Demonstrate equivalence of earnings measured at least 366 days before baseline.
Reporting issue	**Common pitfall**	**Recommended solutions to avoid pitfall leading to an author query**
Reporting sample sizes.	Not reporting all relevant sample sizes in an RCT.	Report sample sizes in RCTs for the intervention and comparison group both at the time of random assignment and at the time of analysis (for the analytic sample).
Reporting on baseline equivalence.	Not reporting information on baseline equivalence.	Report the means, standard deviations, and sample sizes for key characteristics (listed in Exhibit 9).

Note: These pitfalls are presented in no particular order.
RCT = randomized controlled trial.

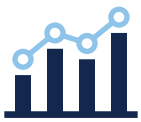
The Pathways Clearinghouse considers information about a study's design, context, analytic methods, and findings when assigning quality ratings. In some cases, study authors do not report the necessary information required to accurately assess the quality of a study. Exhibit 12 outlines the core information the Pathways Clearinghouse looks for to describe a study and related interventions. When study authors provide more complete information, the Pathways Clearinghouse team can more easily synthesize study findings so a range of interested stakeholders, such as TANF administrators or policymakers, can easily digest these findings and translate that evidence into policy.

Although some information listed in Exhibit 12 is helpful for the Pathways Clearinghouse to fully synthesize findings across studies, other reported information is necessary for the team to assign a study quality rating. In RCT studies, authors commonly do not report the sample sizes for the intervention and comparison groups at the time of random assignment and at the time of analysis (for the analytic sample). Without this information, it is impossible to assess the levels of attrition in the RCT to evaluate the quality of the study. In comparison-group QED studies, authors commonly do not report the means, standard deviations, and sample sizes for the key characteristics (listed in Exhibit 10) at the required point in time. Even if equivalence is satisfied in the study context, authors must report this information in the manuscript or provide it in response to an author query from the Pathways Clearinghouse in order to meet the baseline equivalence condition.

Equity and reporting

It's important to know whether an intervention improves outcomes not just on average but for diverse sets of clients, including across genders, races, and ethnicities. The Pathways Clearinghouse encourages study authors to report subgroup analyses by race, ethnicity, and gender when possible. This information is currently cataloged during Pathways Clearinghouse reviews. The Pathways Clearinghouse team is currently investigating how to incorporate this information into the website and future analyses.

Exhibit 12. What Pathways Clearinghouse looks for in a study

Study characteristics 	★ What services are offered to the intervention group? To the comparison group? • Where did the intervention take place? • What are the characteristics of the study participants? • How was the intervention implemented (such as dosage, costs of implementation, and any deviations between actual and intended implementation)? • What was the implementing organization, and who funded the program and study?
Study design and analysis 	★ How were the intervention and comparison groups formed? What is the unit of assignment? ★ If the study team attempted to collect outcome data for only a subset of the people in the intervention and comparison groups, how was this subset selected? ★ What outcome measures were used to assess the impacts of the intervention? When were these outcomes measured? ★ How did the analysis account for missing data?
Study data 	★ What are the sample sizes, unadjusted means, and standard deviations for each key characteristic (reported in Exhibit 10) of the intervention and comparison groups at baseline?[a] ★ What are the sample sizes of the intervention and comparison group for each analysis performed in the study? • What are the estimated findings (including standard deviations, standard errors, or *p*-values) for each outcome measure in the analysis?

[a] This information is required for QEDs and compromised or high-attrition RCTs, to establish baseline equivalence

Note: ★ (stars) indicate information that is necessary for the team to assign a study quality rating. Information that is not starred is used by the Pathways Clearinghouse team to fully synthesize findings across studies, and may be needed to ensure findings from a study are included in Pathways Clearinghouse research synthesis products.

Making sure the Pathways Clearinghouse enhances the reach of your research

Researchers and the Pathways Clearinghouse seek to advance our collective understanding of the programs and policies that are supported by evidence and improve outcomes for people with low incomes. When assessing the evidence of interventions' effectiveness, synthesizing existing evidence, and creating other Pathways Clearinghouse products, the Pathways Clearinghouse includes only studies that receive a high or moderate study quality rating. These products contribute to stakeholders' and other decision makers' understanding of the effectiveness of these employment and training interventions. Research studies that follow the guidelines outlined here are more likely to receive the highest quality rating and, as a result, be included in the Pathway Clearinghouse resource.

What should researchers do to receive the highest quality rating possible?

When developing reports and manuscripts, researchers should do the following:

1. Review the common pitfalls described in this guide when designing studies.
2. Be mindful of the reporting details that the Pathways Clearinghouse attends to, as described in this guide.

Researchers can review the Protocol for the Pathways Clearinghouse (Rotz et al. 2020) for additional details regarding the process, methods, and standards for the Pathways Clearinghouse review effort.

References

Allison, P.D. (2001). *Missing data*, Thousand Oaks, CA: Sage Publications.

Andridge, R.R., and R.J. Little (2010). A review of hot deck imputation for survey non-response, *International Statistical Review* 78(1): 40-64.

Ashenfelter, O. (1978). Estimating the effect of training programs on earnings, *The Review of Economics and Statistics* 60(1): 47-57.

Heckman, J.J., and J.A. Smith (1999). The pre-programme earnings dip and the determinants of participation in a social programme. Implications for simple programme evaluation strategies, *The Economic Journal* 109(457): 313-348.

Little, R.J., and D.B. Rubin (2019). *Statistical analysis with missing data*, Hoboken, NJ: John Wiley & Sons.

Rotz, Dana, Emily Sama-Miller, and Paul Burkander (2020). *Protocol for the Pathways to Work Evidence Clearinghouse: Methods and standards*, OPRE Report #2020-44, Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

What Works Clearinghouse (2013). *Assessing attrition bias, version 2.1*, Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Available at https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_attrition_v2.1.pdf.

What Works Clearinghouse (2014). *Assessing attrition bias-addendum*, Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Available at https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_attrition_v3.0.pdf.

What Works Clearinghouse (2020). *What Works Clearinghouse standards handbook, version 4.1*, Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Available at https://ies.ed.gov/ncee/wwc/handbooks.

Wooldridge, J.M. (2002). *Econometric analysis of cross section and panel data*, Cambridge, MA: MIT Press.

Appendix A

Exhibit A.1 presents the standardized approach that the Pathways Clearinghouse review team uses to select findings for review. Studies often report findings from numerous outcome measures and time points. The Pathways Clearinghouse uses this standardized approach to limit the chance of false positives and better facilitate the comparison of findings across studies. This process occurs independently for each data source in a study.

Exhibit A.1. Selecting findings for review, by outcome domain and measure

Rules for selecting findings based on outcome measures	Notes
Employment	
Select the finding examining the outcome measure that is first in this list: 1. Employment status at the time of follow-up 2. Employment status during the latest available month 3. Employment status during the latest available quarter 4. Employment status over the entire follow-up period (for example, employed since random assignment) Also select findings examining cumulative measures of employment status (such as duration of employment, quarters employed, employment over consecutive quarters, or number of consecutive time periods of employment) for the longest elapsed period (for example, 18 or 12 months for short-term, 3 years for long-term).	**Include** • One set of findings for short-term outcomes and one set for long-term outcomes (and an additional set for very long-term outcomes, if applicable) • Findings for outcomes capturing both overall and unsubsidized employment if both are considered (or unsubsidized and subsidized employment if presented in this way) • Findings for outcomes for full- and part-time employment separately if a combined measure is not available **Exclude findings related to the following outcome measures** • Employment by job characteristics (for example, percentage employed in a job offering benefits) • Point-in-time measures of employment other than those at follow-up (for example, exclude a measure such as "Employed in Quarter 1")
Earnings	
Select the findings examining the outcome measure that is first in this list: 1. Annual earnings for the latest elapsed year of the follow-up period 2. Average annual earnings over the follow-up period 3. Total earnings over the follow-up period 4. Quarterly earnings for the latest elapsed quarter of the follow-up period 5. Monthly earnings for the latest elapsed month of the follow-up period 6. Average hourly wage rate at follow-up 7. Median hourly wage rate at follow-up	**Include** • One set of findings for short-term outcomes and one set for long-term outcomes (and an additional set for very long-term outcomes, if applicable) • Findings for outcomes capturing both overall and unsubsidized earnings if both are considered (or earnings from unsubsidized and subsidized employment if presented in this way) **Exclude findings related to the following outcome measures** • Earnings measures for only people who are employed

Rules for selecting findings based on outcome measures	Notes
Public benefit receipt	
Select the findings examining the outcome measures that are first in this list: 1. Indicators of benefit receipt both overall and by specific benefit type and amount (dollars) of annual benefit receipt, for the longest elapsed follow-up year (for example, receipt in Year 4 of a four-year follow-up) 2. Indicators of benefit receipt both overall and by specific benefit type and average amount of annual benefits over the follow-up period (for example, average benefits Years 1–8) 3. Indicators of benefit receipt both overall and by specific benefit type and amount of total benefits received over the follow-up period (for example, total benefits collected Years 1–3) 4. Indicators of benefit receipt both overall and by specific benefit type and amount of benefits received for the latest elapsed follow-up quarter 5. Indicators of benefit receipt both overall and by specific benefit type and amount of benefits received for the latest elapsed follow-up month	**Include** • One set of findings for short-term outcomes and one set for long-term outcomes (and an additional set for very long-term outcomes, if applicable) • Findings for decompositions of benefit receipt if they are presented by study authors (for example, measures of receiving Temporary Assistance for Needy Families, Supplemental Nutrition Assistance Program, or Unemployment Insurance benefits) • Measures of months of benefit receipt if indicators of overall benefit receipt are not available **Exclude findings related to the following outcome measures** • Benefit amounts for only people who receive benefits • Measures of credit attainment • Measures of educational attainment at a specific institution or group of institutions, unless the group of institutions for which data is available is (1) similar to the set of institutions covered in an established data source (for example, the National Student Clearinghouse), or (2) similar to the set of institutions attended by all study participants[a]
Education and training	
Select the findings examining measures of educational attainment over the follow-up period (for example, acquisition of a GED, associate's degree, bachelor's degree, attaining a certificate or credential)	**Exclude findings related to the following outcome measures** • Decompositions of the measures over time (for example, obtained GED within one year) • Measures combining different educational milestones (such as training completion or degree attainment) if individual measures are available • Measures of credit attainment • Measures of educational attainment at a specific institution or group of institutions, unless the group of institutions for which data is available is (1) similar to the set of institutions covered in an established data source (for example, the National Student Clearinghouse), or (2) similar to the set of institutions attended by all study participants[a]

[a] Similarity is judged using the attrition threshold. That is, the authors should demonstrate that the overall and differential differences in the rate at which study participants enroll in the two sets of institutions is below the thresholds for overall and differential attrition, as Exhibit 9 shows.

Goals of the Pathways Clearinghouse

The Pathways Clearinghouse systematically evaluates and summarizes the evidence on the effectiveness of interventions that aim to improve employment outcomes, reduce employment challenges, and support self-sufficiency for populations with low incomes. It has several goals:

- Conduct a transparent, comprehensive search to identify studies of employment and training interventions designed to improve employment, increase earnings, support self-sufficiency, or advance education and training for populations who are low income.
- Rate the quality of those studies to assess the strength of the evidence they provide on the different interventions.
- Determine the evidence of effectiveness for those interventions.
- Share the results, as well as other Clearinghouse products, on a user-friendly website to help state and local TANF administrators, policymakers, researchers and the general public make sense of the results and better understand how this evidence might apply to questions and contexts that matter to them.
- Synthesize the overall state of evidence in the field by creating and disseminating a variety of reports, briefs, and other products.

For more information, see https://pathwaystowork.acf.hhs.gov.

January 2022

OPRE Brief # 2022-08

Project Officer: Kimberly Clum

Project Monitor: Clare DiSalvo, Amelie Hecht, Amelia Popham
Office of Planning, Research, and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services

Project Director: Diana McCallum
Mathematica
1100 First Street, NE, 12th Floor
Washington, DC 20002-4221

Suggested citation: Shiferaw, Leah, Diana McCallum, and Dana Rotz. (2022). Pathways Clearinghouse Guide for Researchers. OPRE Brief # 2022-08, Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

This brief was funded by the Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, under contract number HHSP233201500035I/HHSP23337034T.