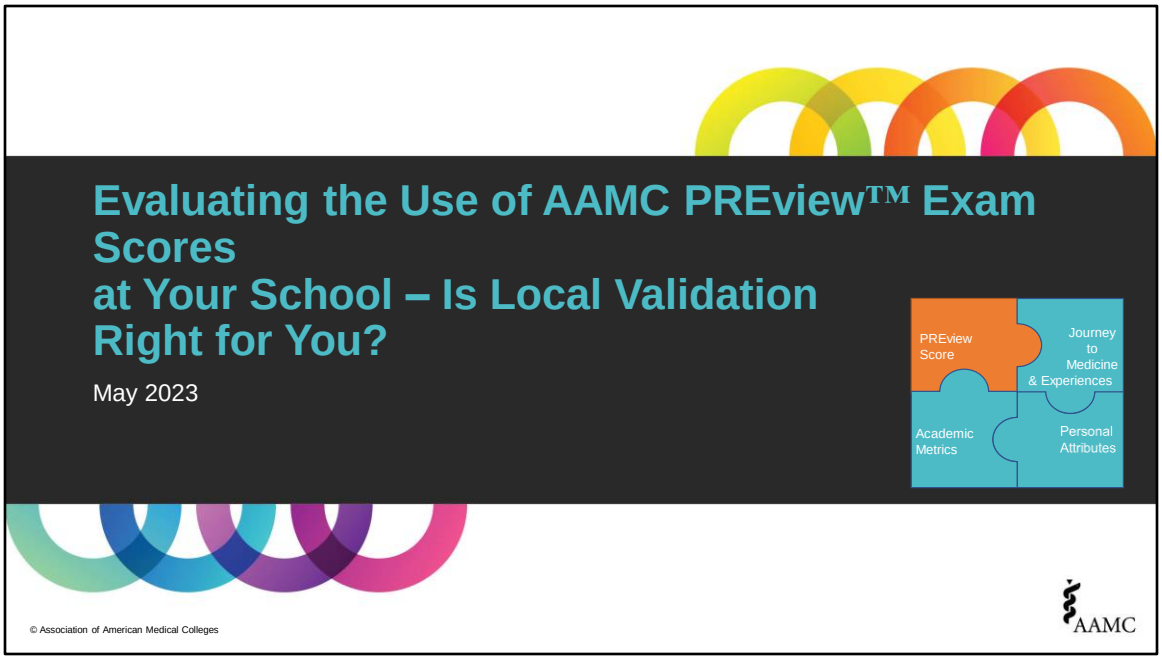
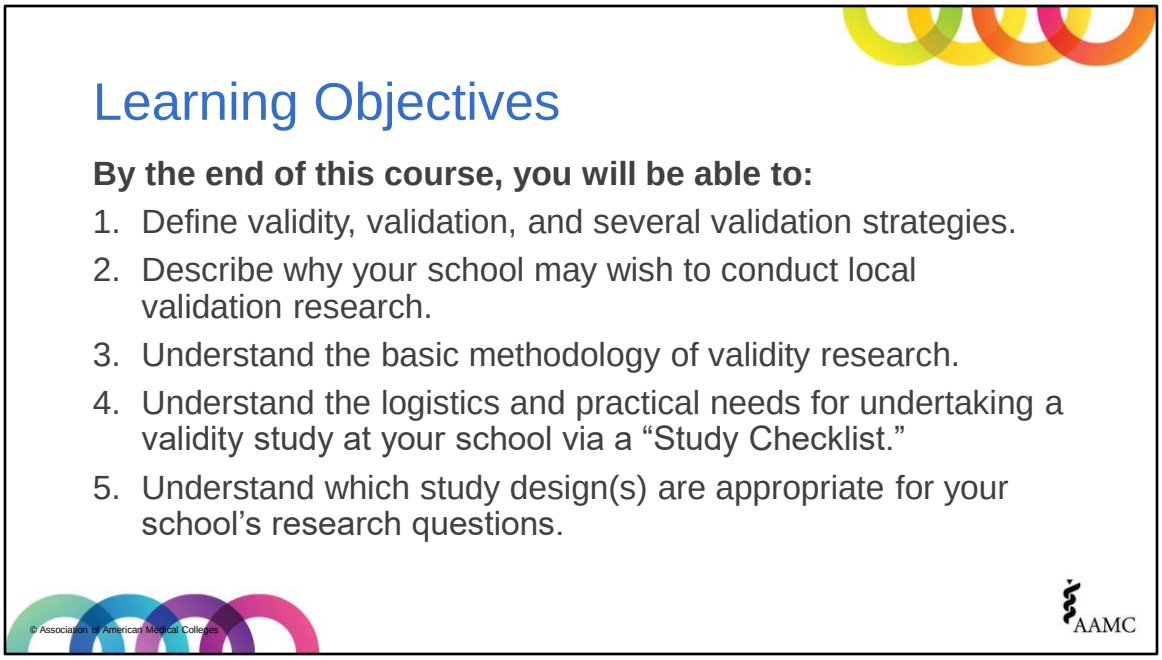


Welcome, and thank you for joining the first course in our series on **Local Validation Research: Evaluating the use of AAMC PREview™ Exam Scores at Your School – Is Local Validation Right for You?**

As you may know, the AAMC PREview exam is designed to assess examinees' understanding of effective and ineffective pre-professional behavior across eight core competencies for entering medical students.

This series of courses was developed to provide information and guidance that may help you and your school decide if and how to conduct your own research to evaluate the validity and usefulness of PREview scores in your admissions process.

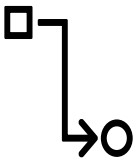


This is the first courses in a multi-part series of short recorded courses. At the end of this introductory 101 course, you will be able to:

1. Define validity, validation, and several validation strategies.
2. Describe why your school may wish to conduct local validation research.
3. Understand the basic methodology of validity research.
4. Understand the logistics and practical needs for undertaking a validity study at your school.
5. Understand which study design(s) are appropriate for your school's research questions.

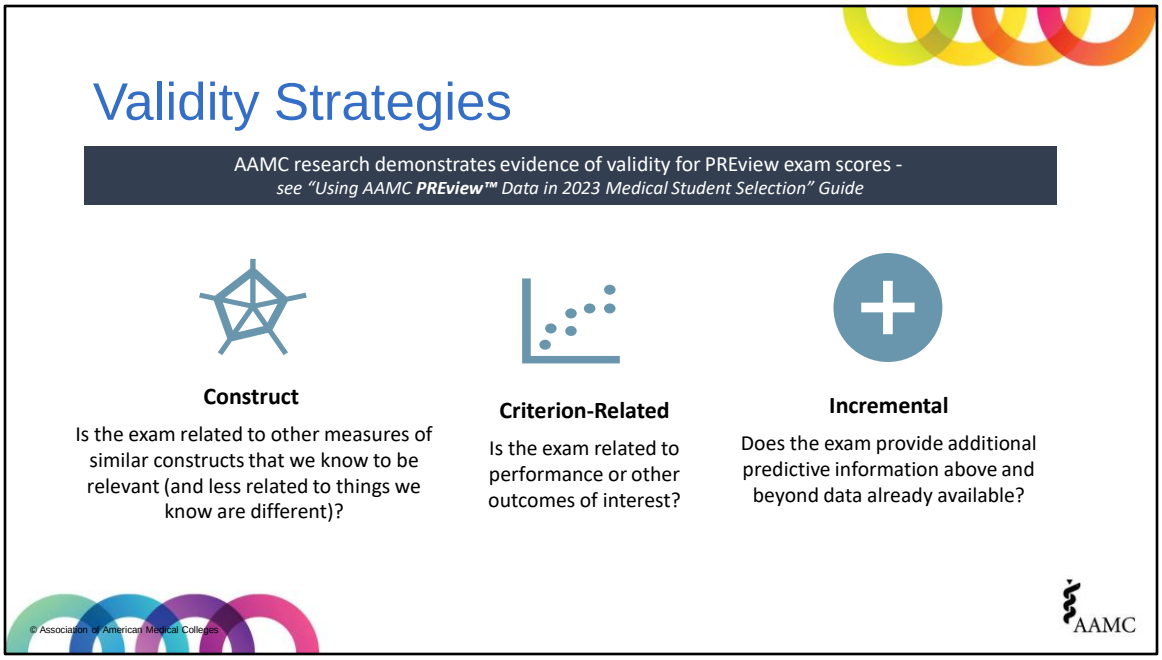
What Do We Mean by "Validity" and "Validation"?

Validity = "The degree to which evidence and theory support the interpretations of test scores for proposed uses of tests."



Validation = "The process by which evidence of validity is gathered, analyzed, and summarized." [1]

AAMC



Here, we'll briefly review the different validation strategies.

Construct Validation helps answer the question of does the exam relate to other assessments in ways we would expect it to? Constructs in this case represent traits, states, attitudes, and abilities like leadership, teamwork, professionalism, personality, or academic ability. These kinds of variables are difficult to observe and measure directly, and construct validation is a way to link the constructs measured by the exam (understanding of pre-professional competencies) with other admissions data.

So, in this case: *What is the relationship between PREview scores and other admissions data?*

Criterion Validation helps answer the question – does the exam statistically predict (or correlate with) important performances outcomes?
So, in this case: *How well do PREview scores predict students' non-academic, pre-professional performance?*

And finally, **Incremental Validation** helps answer the question – does the exam provide additional predictive information above and beyond data already available?

So, in this case, *does the PREview test provide unique predictive information as compared to admissions data such as UGPA and MCAT® scores*?

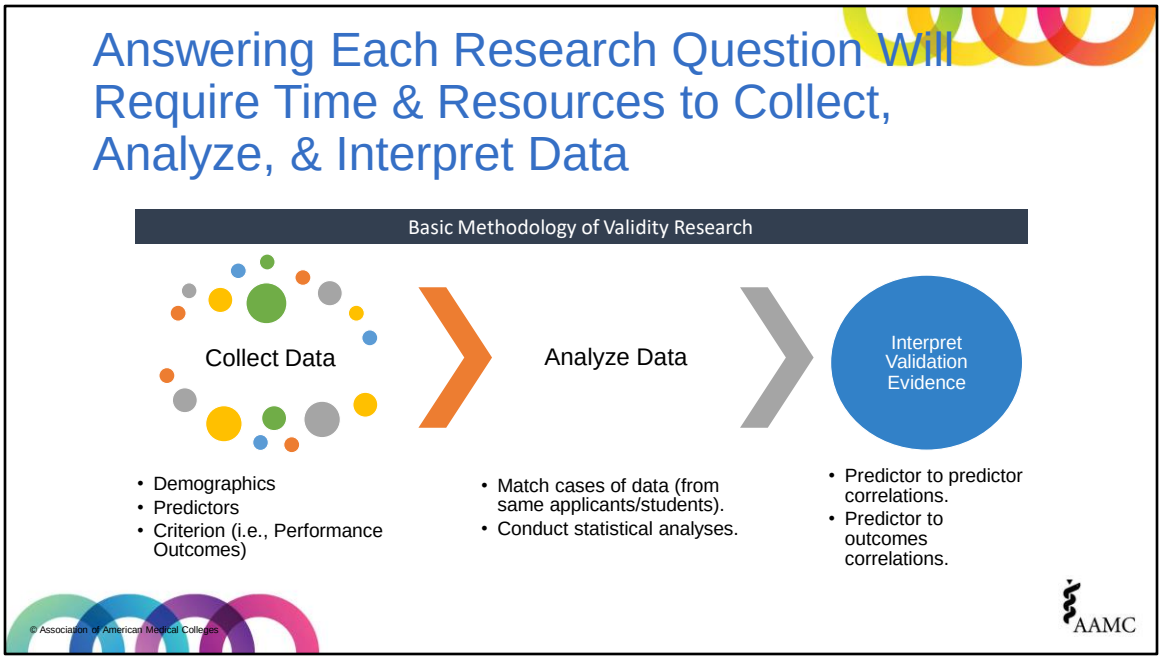
As noted at the top of the screen, the AAMC has conducted research that demonstrates each type of validity using data from a consortium of schools currently using the PREview assessment.

From here, let's move to discussing why and how you might approach your own local validation study.

Given that AAMC researchers have examined and documented results regarding its validity, why are we suggesting that you consider conducting local research at your school?

There are several reasons.

- Stakeholders at your school may find local research more compelling in support of decisions regarding the use of PREview scores.
- Your school may use sources of admissions data that were not included in the AAMC study, or the content of some of these data may differ across schools (e.g., applicant interviews). This could lead to different findings regarding both the relationship of PREview scores with other admission data and the added predictive value (incremental validity) of PREview scores at your school.
- Your school may have available or be able to collect – and may value - different performance and other outcome measures than those used in the AAMC study.
- Your school may have an interest in exploring how the PREview test should fit in to your school's mission and holistic review process.



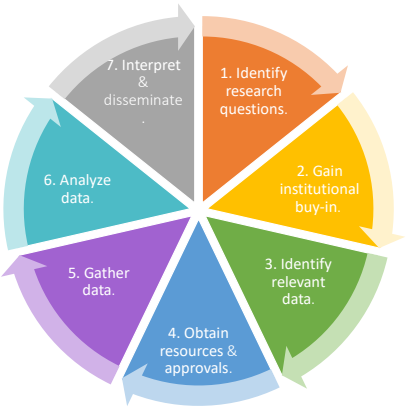
At the outset, it is important to understand upfront that this research will require time and resources to collect, analyze, and interpret data. Unless you contract with an external consultant to conduct the research and/or analyses, most costs should be internal (employee time). Nevertheless, most studies will require considerable time and resources to do properly.

At its most basic, this research will involve conducting quantitative, correlational research, which will involve:

- Collecting data,
- Matching cases and conducting statistical analyses (mainly correlations), and
- Interpreting the validation evidence to look for construct, criterion, or incremental validity.

Study Checklist

- ❑ **1. Identify your research questions.**
- ❑ **2. Gain institutional buy-in.**
- ❑ **3. Identify relevant data.**
  - ❑ *Conceptual considerations*
  - ❑ *Methodological considerations*
- ❑ **4. Obtain resources and approvals.**
- ❑ *5. Gather your data.*
- ❑ *6. Analyze data.*
- ❑ *7. Interpret and disseminate your findings.*



AAMC

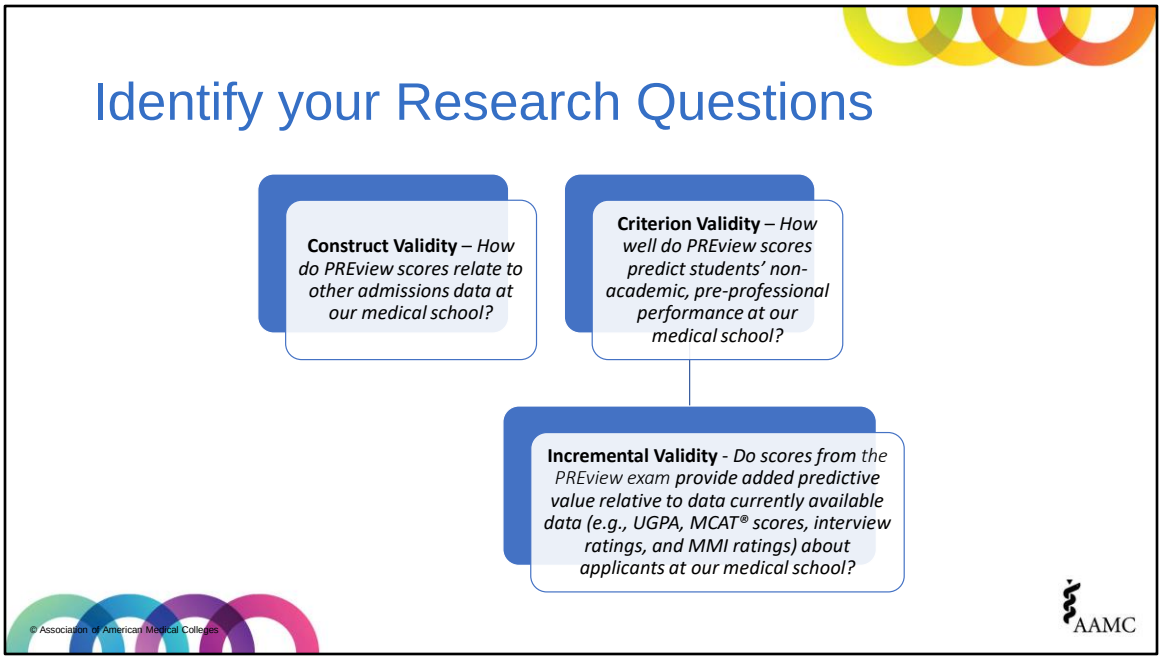
You're hear to learn how to set up a study to evaluate PREview scores in your admissions process…Here is a check-list of the things you'll need to gather and consider.
There are several foundational requirements for a successful validity study

- The first thing you'll need to do is identify your research questions, which we'll dive into on the next slide.
- If you determine you have research questions to answer, next you'll need to secure approval/leadership buy-in to conduct the research.
- After gaining approval, you'll next want to consider the availability, accessibility, conceptual relevance, and methodological appropriateness of data to answer your research questions; because your data needs will be driven by your research question, we will touch on this here, but cover it in more detail in the other webinars.
- Next, gathering the right human capital resources and IRB approvals will be important for the success of this research:

After gaining resources and approvals, you'll gather your data, and then analyze,

interpret, and disseminate it.

In the next few slides, we will provide more details on the first four steps in the checklist. We will touch on the last 3 concepts briefly here, and in more depth in the other courses, because these steps will be driven by your research questions.



Two critical questions when conducting research are "What do we want to know?" and "Why do we want to know it?" Your research questions represent the "What do we want to know?" part, that is, what questions are you attempting to answer by conducting this research.

"Why do we want to know it?" pertains to the value of having the answer to a research question, that is, what you plan do with that information. There can be many research questions you could ask, and multiple ways to cut the data. Prioritizing research questions based on their value can help you and your data analysis team avoid getting bogged down running scores of low value analyses.

Something that is nice to know but that nobody will act on should have a lower priority compared to something that will have a big impact and on which leadership is committed to act. If answering a research question will require extra time and effort to obtain the needed data, this consideration can help you determine if it is worth it.

Above are what we think is likely to be the most common research questions:

Construct Validity – *How do PREview scores relate to other admissions data at our*

medical school?

Criterion Validity – *How well do PREview scores predict students' non-academic, pre-professional performance at our medical school?*

And finally, another criterion-related validity question centers around:
Incremental Validity - *Do scores from the PREview exam provide added predictive value relative to data currently available data (e.g., UGPA, MCAT scores, interview ratings, and MMI ratings) about applicants at our medical school?*

Gain Institutional Buy-In and Support

- Articulate the benefits as well as the resource costs.
- Gain faculty participation buy-in and support.
- Be transparent with students.

AAMC

Because this research will require resources (for example, your and other staff time to collect and analyze data) leadership support and buy-in will be critical. It will likely be helpful to have one or more influential individuals at your medical school champion the value of this research and then communicate its importance to any individuals you may need to leverage to complete the research. For example, if you decide to actively collect performance ratings from faculty, faculty may be reluctant to take on the additional burden of rating students. Having an internal champion can help combat this potential roadblock

Gain leadership buy-in and support by articulating the benefits of the research for your school, as well as being honest about the resource costs. Success likely will depend on your ability to get cooperation and access to resources – information and otherwise - from other people. This is easier when organizational leaders are informed and onboard.

Gain faculty participation buy-in and support. Lack of faculty buy-in or support could be a major barrier – especially if you need to collect information such as research-only outcome ratings from them. Poor rates of responding can kill an otherwise well-designed study. Leadership support and effective communication are key. So is

making sure what you ask of others is reasonable. Consider using a brief standardized measure (e.g., the AAMC Research-only Performance Rating Tool) or some modified version of a standardized measure that does not take much effort or time to complete, especially if faculty members will be asked to rate multiple students.

Transparency with students will be important, particularly if performance ratings of students will be collected for research purposes only. It is possible, even likely, that students will catch wind of this research. As a result, the purpose of the study should be shared, and students should be informed that their grades and academic status will be unaffected.

Identify Relevant Data

Research Question	Validity Type	Data Needed	Data Pools	Will Help Us Understand
How do PREview scores relate to other admissions data at our medical school?	Construct	• Demographics • PREview scores • Other admissions data	• Applicant pool	What the PREview exam measures and how it overlaps with or is distinct from other measures
How well do PREview scores predict students' non-academic, pre-professional performance at our medical school?	Criterion	• Demographics • PREview scores • Performance Outcomes	• Matriculant pool	How to incorporate PREview scores into your school's holistic review process
Do scores from the PREview exam provide added predictive value relative to currently available data?	Incremental	• Demographics • PREview scores • Other admissions data • Performance outcomes	• Applicant and matriculant pool	The value of adding PREview scores to your admissions process

AAMC

Relevant data will be dependent on your research question, so we will dive deeper into this in the other courses.

Here, we touch on the data needs at a high-level based on the research question.

Note that conceptual relevance (e.g., of performance outcomes to PREview scores) and methodological appropriateness (e.g., data reliability, data variability) will be very important factors to consider at this stage.

As you identify different possible data sources, they will be less useful to you if they don't meet the conceptual and methodological appropriateness factors discussed on the next few slides.

Consider Conceptual Relevance of Data

If conducting a criterion and/or incremental validity study, it will be critical to the success of the research to pick the right outcomes to avoid drawing erroneous conclusions.

- **Challenges**
  - Few well-established measures of non-academic competencies.
  - Possibly limited opportunities to observe in year 1 or 2.
- **Conceptual Relevance**
  - Outcomes data (for addressing research questions 2 and 3) must be conceptually related to the non-academic, pre-professional competencies assessed via PREview scores.

AAMC

Let's start with conceptual relevance of data:

This may be hardest part of the research to nail down because there is a shortage of well-established medical school performance outcomes conceptually related to professionalism. Furthermore, these data may not be readily available on all students, may be available in a form that needs to be transformed to numeric score or may require active data collection from others, for example, a survey to capture performance feedback from faculty or peers.

That said, it will be critical to the success of the research to pick the right outcomes to avoid drawing erroneous conclusions about the utility of PREview scores in your admissions process.

Consider Conceptual Relevance of Data

If conducting a criterion study – the two most likely outcomes schools might include are:

Outcome Category	Archival vs. Active Data Collection	Description	Pros (+)	Cons (-)
Course Outcomes/ Grades (e.g., in Patient-centered Medicine course)	Archival	Final grade in first- or second-year course in which students learn clinical and patient interaction skills	• Represents performance in initial year(s) of med school. • Easy to access and available at most medical schools.	• Risk of only partial content relevance (grade often comprised of unrelated assignments/exams).
Evaluation Using the AAMC-provided Research Only Performance Tool	Requires Collection	Faculty (or clinical preceptors) provide an evaluation of student performance	• Standardized direct observation of student performance. • Focus on relevant criteria (e.g., professionalism).	• Active data collection requires more resources. • Faculty participation required.

AAMC

Most outcomes will have pros and cons that should be carefully weighed. The two most common types of outcomes we envision schools leveraging for a PREview exam criterion study are existing course-based outcomes, like grades, and the AAMC **developed Research Only Performance Tool.**

In the first row, we have listed a course grade in a patient-centered medicine course. This outcome would typically include a final class grade in the first or second year. Though this outcome is typically easy to access, the final course grade likely reflects more than professionalism competencies (in other words, reflect some competencies not conceptually related to PREview scores). If you are interested in using existing data to studying PREview criterion validity, check out Course 103.

Conversely, in the second row, we have listed AAMC's performance rating tool that could be used to collect medical student performance outcomes for research purposes only. It would require asking faculty to provide an evaluation of student performance. This outcome benefits from being highly conceptually related to the competencies assessed on the PREview exam, but requires a greater level of effort to collect, including requiring faculty participation. If you are interested in using this approach to studying PREview criterion validity, please check out the separate course

on the Research Only Performance Tool.

Consider Methodological Issues that can Affect Data Analysis & Interpretation

Research studies with small samples or samples that lack variance are less likely to find differences or degrees of relationships that exist.

Sample Size (n):
The number of people who participate in your study.

Lack of Variance:

Scores in a distribution that are either very similar or the same (e.g., everyone passing a pass/fail course).

Example: Range Restriction - Limiting range of scores in a sample by removing people on a decision rule (e.g., having data on only those who scored above a 5 on **the** PREview **exam).**

AAMC

Sample Size and Lack of Variance are important issues that can affect your ability to find a correlation (relationship between two variables, like PREview scores and student performance) that exists that you should bear in mind at the design phase of your research.

Something to consider early on when **determining whether a local study is right for you is sample size**; that is, how many people need to participate in the study, and whether this number is possible to attain at your school. Small samples represent a risk to validity research because, with them, you are less likely to find statistically significant results unless the true underlying relationship between your two variables of interest is very strong. This has to do with a consideration called *statistical power*.

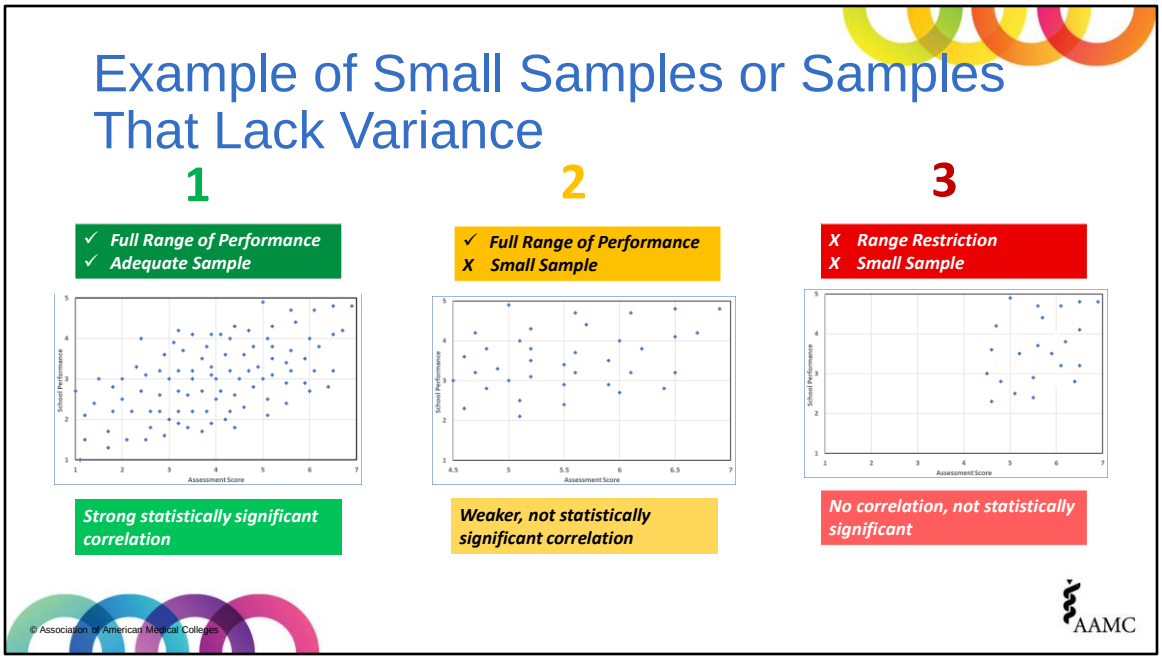
Stated simply, power is the likelihood of correctly finding a relationship that exists. Power is affected by the strength of the relationship and the sample size in your study. The flip side of power is called "Type II error." That is, erroneously concluding there is no relationship when in fact a relationship exists. So, with small sample sizes,

you will have less power and are more likely to conclude that no relationship exists between PREview scores and other variables, even when it does.

Lack of Variance/Range Restriction is a pervasive threat to validity researchers and can result in reduced correlations when there is a restriction on the possible range of PREview scores, on the range of outcome data, or both. For example, a school only accepts applicants who have PREview scores of 8 or 9 will result in range restricted correlations.

Similarly, if faculty or peers provide little variance in their ratings, for example, all 5-ratings, then any correlations will be restricted.
There are corrections for attenuation; however, these corrections can be sophisticated and deserve their own section.

Let's look at an example.



In each of these scatterplots, we're examining the same two variables and pulling samples from the same population.

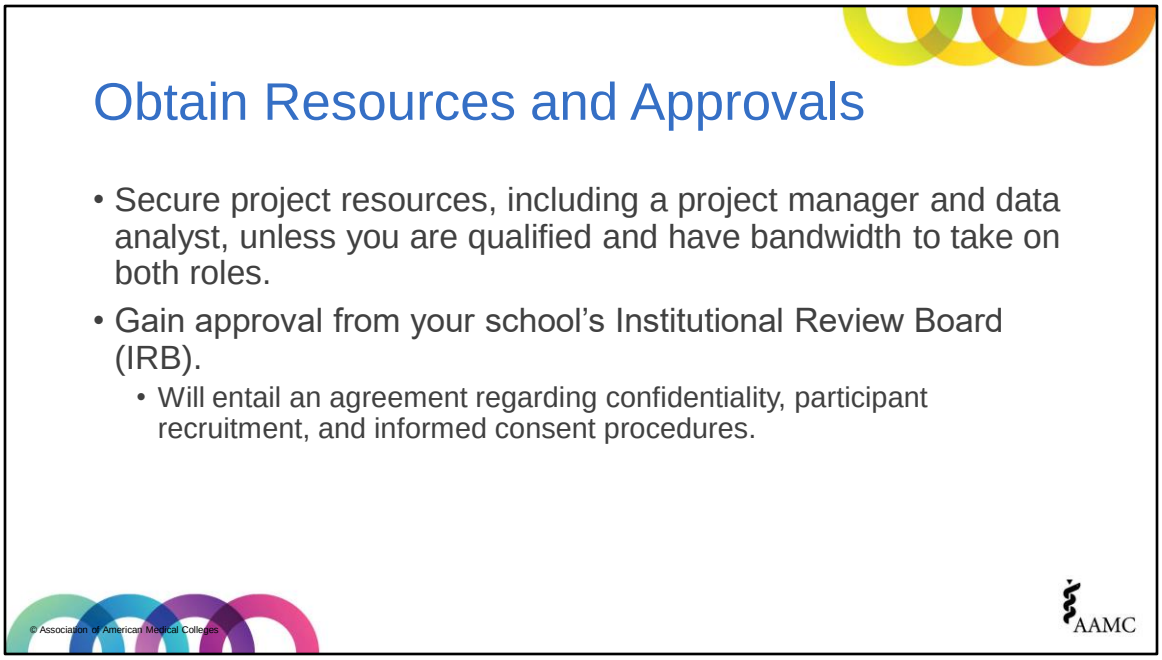
In the first scenario, we have data from the full range of performance on both variables, and an adequate sample size. Here, we'd conclude that we have a strong, statistically significant positive correlation.

In the second scenario, we still have the full range of performance, but observations from fewer people. Because we have fewer people, we observe a smaller correlation, and we are less confident that it is meaningful, or, in other words, statistically significant.

In the third scenario, we still have the smaller sample size, and now we have a restricted range on the assessment: only people with scores above 4.5 are in the sample. With the range restriction, we lose information about people who scored poorly on the test. Between the range restriction and smaller sample, we conclude that there is no correlation between the two variables (though one exists).

Sample size and data that inherently lacks variance can't be solve for and thus, should

be considered at the study design step.



Since we've covered data issues at a high level, now we're going to cover the resources and approvals you'll need to secure:

First, you'll likely need a project team. Your project team should likely include a data analyst with knowledge/experience in data collection, statistical analyses, and interpretation. This might include, for example, medical school staff in your evaluation or educational exam group, or medical researchers.

- Your data analyst will need access to data analysis software (e.g., Excel, SPSS, SAS, R).
- You also may need support from your school's IT department to identify data sources and pull the data.

Next, you will very likely need to get your research approved through your school's Institutional Review Board or IRB.

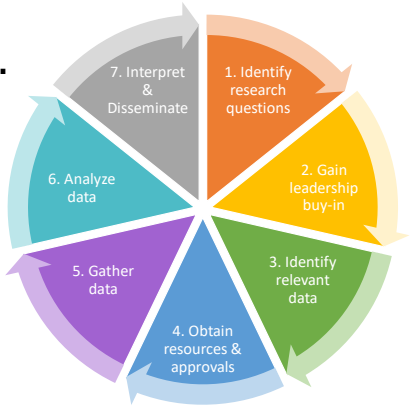
- The IRB is formally designated to approve (or reject), monitor, and review biomedical and behavioral research involving humans.
- The IRB will review the methods proposed for research to ensure that

they are ethical.

- It will likely entail an agreement regarding what will and won't be done with the data, including confidentiality as well as participant recruitment and informed consent procedures

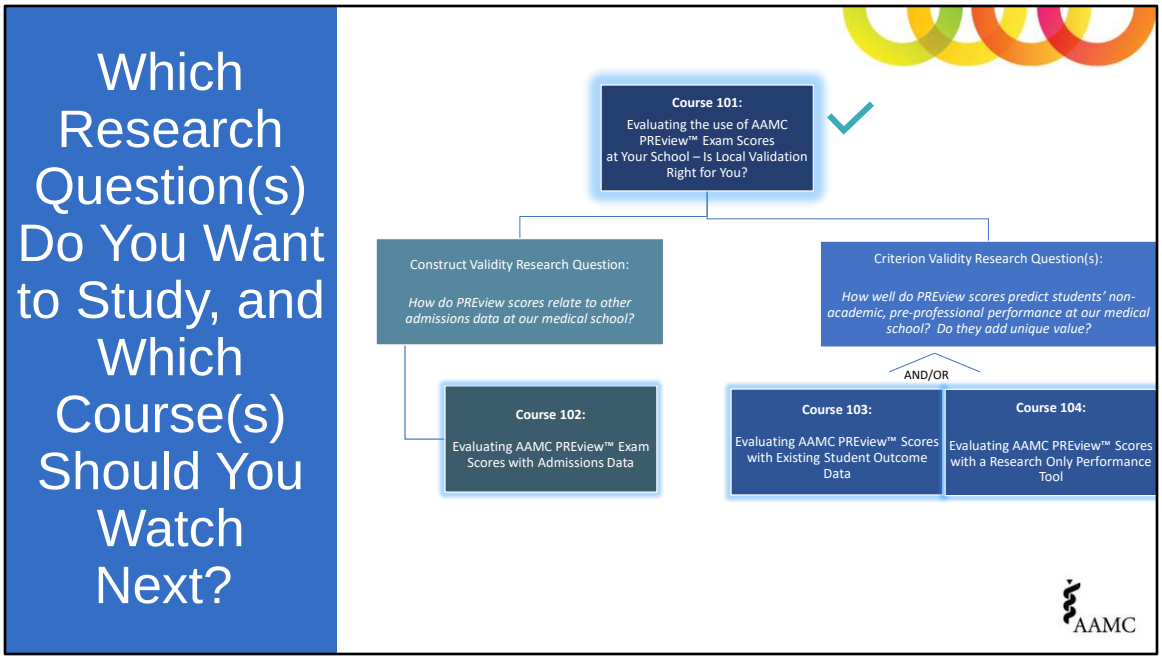
Summary – A Review of the Study Checklist

- ❑ **1. Identify your research questions.**
- ❑ **2. Gain leadership buy-in.**
- ❑ **3. Identify relevant data.**
  - ❑ *Conceptual considerations*
  - ❑ *Methodological considerations*
- ❑ **4. Obtain resources and approvals.**
- ❑ *5. Gather your data.*
- ❑ *6. Analyze data.*
- ❑ *7. Interpret and disseminate your findings.*



AAMC

Today we covered Topics 1 – 4. In future courses, we will dive into more detail on all topics highlighted in bold. Other courses are organized around possible research questions, and thus these will be discussed in more detail at the beginning of the next step of courses. Then, identifying, gathering, analyzing and interpreting the data based on your research question will be discussed.



Given all that we've discussed today, to understand the value of PREview scores at your school, most likely one of these local validation approaches will make sense for your school. In the next three courses, we offer a detailed look at how each one of these studies might be implemented, cautions and considerations when designing and conducting this study at your school, and guidelines for analyses and interpretation of results.

Course 102 advises you on conducting research to understand how the PREview exam relates to other admissions data. Even if this is not a primary concern to you, this course could be helpful if you are interested in exploring the incremental validity of the PREview test over other admissions data.

Course 103 addresses how to conduct research to understand how the PREview exam relates to and predicts important student outcomes. It includes how to conduct a criterion validity study using existing student performance outcomes at your school, whereas Course 104 describes how and why you might want to investigate outcomes using a research only performance tool.