The Experimental Design Domain Explainer

Short readings that supply every idea needed for each lesson decision.

REQUIRED	OPTIONAL
Read the three short paragraphs and use the labeled case evidence.	**Open the publication links only if you want to go deeper.**

How every reading is structured

- Why it matters
- The question
- Prior takeaway
- Analogy and rules
- Biomedical example
- Case evidence
- Decision
- Take-home
- Glossary
- Optional sources

READING 1 OF 20 | DEVELOPMENT CAN GO WRONG

From an Observation to a Researchable Question

Out of everything we could ask about Mateo, which questions can science actually answer, and how do we sharpen one of them?

START WITH

A complete anatomical story can define structures, functions, repairs, and uncertainties while leaving treatment comparisons open to experiment.

REMEMBER

A researchable question names a population, a factor or intervention, a comparison, and a measurable outcome.

Why this matters

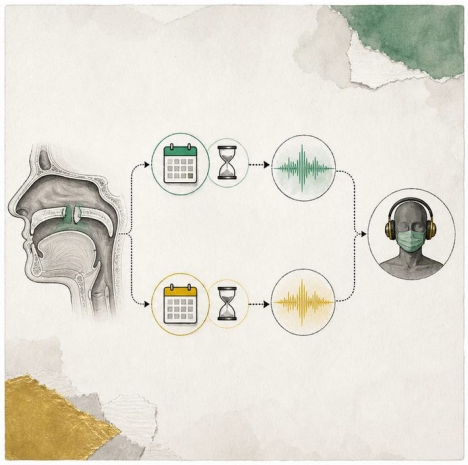
A focused question keeps a study from collecting many facts without answering a decision.

The biology in plain English

A clinical observation is a starting point, not yet a study question. Researchers must decide exactly who, what choice, and what outcome they will study.

PICO stands for population, intervention, comparison, and outcome. A time point is often added so the outcome has a clear deadline.

A useful question does not promise an answer. It tells the team which evidence could support or challenge each option.



Biomedical teaching illustration: Frame the palate-timing question with PICO

Check your understanding

- What does the C in PICO name?
- Why does the age 5 time point matter?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP01-E1	The open anatomical question is whether repair timing changes later speech function.	The question concerns a treatment choice, not the cause of Mateo's cleft.
EXP01-E2	A supplied candidate population is medically fit infants with isolated cleft palate.	The population is specific enough to set eligibility rules.
EXP01-E3	VPI can be scored at age 5 by trained assessors.	The outcome and time point can be defined before the study begins.

Concrete decision

You are the junior investigator writing the first protocol line. The team must choose one question before enrolling any participants.

A. Among eligible infants, does repair at 6 rather than 12 months reduce VPI at age 5?

B. What is the best cleft surgery?

C. Did Mateo's family cause his cleft?

Decision prompt: Choose the researchable question and label its population, intervention, comparison, outcome, and time.

Claim ceiling: You may define the study question. You may not predict which timing is better before data are collected.

Glossary

Term	Plain-English definition
researchable question	A question that is specific and measurable enough to answer with data through a feasible study.
PICO	A framework for building a clear research question by naming the Patient, Intervention, Comparison, and Outcome.
population	A defined group of individuals being studied or counted, such as all newborns in a region during one year.
comparison	The other group or condition you measure your treatment against, so you can judge whether the treatment made a real difference.
outcome	The measured result a study tracks to judge whether a treatment or condition makes a difference.
empirical	Based on direct observation, measurement, or experiment rather than on theory or opinion alone.

READING 2 OF 20 | DEVELOPMENT SCULPTS

From a Question to a Testable Hypothesis

How do we turn our PICO question into a hypothesis that an experiment could actually disprove, and what parts of the study must we control?

START WITH

A researchable question names a population, a factor or intervention, a comparison, and a measurable outcome.

REMEMBER

A falsifiable hypothesis links a defined change to a defined outcome and states what would count against it.

Why this matters

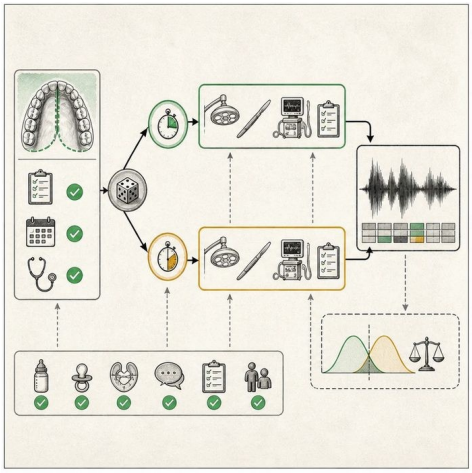
A hypothesis earns scientific value by taking the risk of being wrong.

The biology in plain English

A hypothesis predicts a relationship between variables. The independent variable is the factor assigned or compared. The dependent variable is the outcome measured.

Researchers also define constants and eligibility rules so the groups receive the same follow-up and outcome measurement.

The null hypothesis usually states that the groups do not differ. A good prediction says which findings would support it and which would count against it.



Biomedical teaching illustration: Convert the PICO question into a testable timing hypothesis

Check your understanding

- Which variable is repair timing?
- What finding would count against the earlier-is-better prediction?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP02-E1	The proposed study compares repair at 6 months with repair at 12 months.	Repair timing is the factor that differs between groups.
EXP02-E2	The primary outcome is VPI status at age 5.	The outcome must be measured the same way in both groups.
EXP02-E3	A valid hypothesis can be challenged by equal or higher VPI in the 6-month group.	The prediction is falsifiable rather than protected from failure.

Concrete decision

You are the methods lead completing the hypothesis card. A draft says, 'Earlier surgery will help somehow,' and the team must decide whether it can guide a study.

A. Rewrite it with timing, VPI at age 5, and a result that would oppose the prediction.

B. Keep it because every positive change can count as success.

C. Remove the comparison group so no result can disagree.

Decision prompt: Choose the revision and identify the independent variable, dependent variable, and falsifying result.

Claim ceiling: You may state a testable prediction. You may not claim the hypothesis is true before the comparison.

Glossary

Term	Plain-English definition
hypothesis	A testable, falsifiable prediction about how variables relate, written so an experiment can support it or prove it wrong.
null hypothesis	The default assumption that there is no real effect or difference, which a study tries to disprove with its data.
independent variable	The one factor a researcher deliberately changes in an experiment to test its effect on the measured outcome.
dependent variable	The outcome a researcher measures in an experiment, which may change in response to the variable being tested.
controlled variable	A factor kept the same across every group in an experiment so it cannot secretly affect the results.
control group	A comparison group set up to be like the treated group except for the one change being tested, so you can tell if that change did anything.
falsifiable	Able to be proven wrong by a real measurement, a key feature of a good scientific question or hypothesis.

READING 3 OF 20 | CELLS DIFFERENTIATE

Why We Trust Some Studies More Than Others

When two studies disagree, what about a study's design tells us which result to trust more?

START WITH

A falsifiable hypothesis links a defined change to a defined outcome and states what would count against it.

REMEMBER

The strongest design is the one that fits the question and controls its main threats.

Why this matters

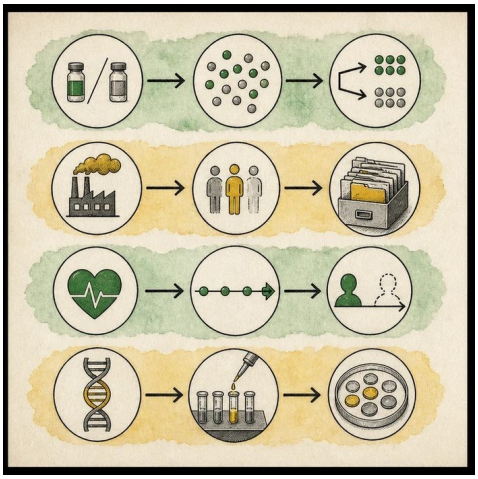
Students need a design map, not a memorized pyramid that ignores the question being asked.

The biology in plain English

Randomized trials are strong for fair comparisons of assignable treatments. They are not the right design for diagnosis, prognosis, rare harms, or many mechanism questions.

Observational studies are needed when researchers cannot or should not assign an exposure. Laboratory studies can test how a gene or cell process works.

Every design has threats. Good planning names the largest threat and adds a safeguard that fits it.



Biomedical teaching illustration: Match four cleft questions to four study designs

Check your understanding

- Why can a suspected harmful exposure not be randomized?
- Which design family can test a gene mechanism directly?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP03-E1	Repair timing is an assignable clinical choice when both options are acceptable.	A randomized trial can compare the treatment options under equipoise.
EXP03-E2	A suspected harmful pregnancy exposure cannot be assigned.	An observational design is required for that cause question.
EXP03-E3	A gene mechanism requires perturbation and molecular or tissue outcomes.	A bench experiment answers a different question than a treatment trial.

Concrete decision

You are chairing the design-selection meeting. A teammate declares that an RCT is always the best design for every cleft question.

A. Match each design to the treatment, harm, prognosis, or mechanism question.

B. Use an RCT even when it assigns suspected harm.

C. Use a case report for every question because it is fastest.

Decision prompt: Choose the design rule and match the three supplied questions to appropriate designs.

Claim ceiling: You may select a design family. You may not call any design free of bias or automatically decisive.

Glossary

Term	Plain-English definition
evidence hierarchy	A ranking of research designs by how trustworthy their evidence is, placing systematic reviews and trials above observational studies and expert opinion.
levels of evidence	A ranking of how trustworthy a study's findings are, with reviews of many strong trials at the top and single expert opinions near the bottom.
randomized controlled trial	A study that randomly assigns participants to a treatment or a control group, the strongest design for showing a treatment truly works.
observational study	A study that watches and records what happens to groups without assigning any treatment, so it can show links but not prove cause.
case report	A detailed written account of one patient's diagnosis, treatment, and outcome, useful for sharing rare or unusual findings.
bias	A systematic error in how data is collected or interpreted that tilts results in one direction, making conclusions less accurate or unfair.
confounding	When a hidden third factor influences both the suspected cause and the outcome, making a link look real when it may be misleading.

READING 4 OF 20 | DEVELOPMENT CAN GO WRONG

Studying a Cause You Cannot Assign

If we cannot assign a suspected harmful exposure, how can we still gather evidence about whether it raises the chance of a cleft?

START WITH

The strongest design is the one that fits the question and controls its main threats.

REMEMBER

Case-control studies efficiently study rare outcomes, but their odds ratios remain vulnerable to selection, recall, and confounding.

Why this matters

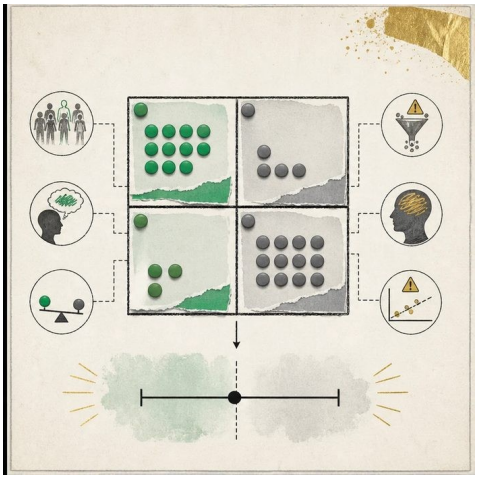
Rare outcomes often require looking backward, but the shortcut comes with clear limits.

The biology in plain English

A case-control study begins with people who have an outcome and comparable people who do not. Researchers then compare earlier exposures.

The odds ratio compares exposure odds in the two groups. A confidence interval shows the range of values reasonably compatible with the data and model.

Selection, recall, and confounding can change an odds ratio. An association is evidence to investigate, not proof of cause.



Biomedical teaching illustration: Test a suspected exposure with a supplied 2 by 2 table

Check your understanding

- Why is this design efficient for a rare outcome?
- What does an interval that includes 1.0 mean here?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP04-E1	The table has 30 exposed and 70 unexposed cases, plus 15 exposed and 85 unexposed controls.	The odds ratio is about 2.4.
EXP04-E2	The reported 95 percent confidence interval is 0.9 to 3.4.	The interval includes 1.0, so the data remain compatible with no odds difference.
EXP04-E3	Parents of affected infants may recall an exposure differently while searching for an explanation.	Recall bias could inflate or reduce the observed association.

Concrete decision

You are the epidemiologist reviewing a suspected prenatal exposure. The team must decide whether the supplied result proves that the exposure caused clefts.

A. Report an uncertain association and plan better exposure measurement plus confounder control.

B. Declare the exposure proven causal because the odds ratio is above 1.

C. Declare no possible effect because the interval crosses 1.

Decision prompt: Choose the interpretation and cite the odds ratio, interval, and one bias risk.

Claim ceiling: You may report the observed association and uncertainty. You may not claim individual blame or causation.

Glossary

Term	Plain-English definition
case-control study	A study that starts with people who have a condition (cases) and people who do not (controls), then looks back to compare past exposures.
exposure	Contact with a substance, agent, or condition that could affect health, such as a chemical, pathogen, or environmental factor.
odds ratio	A number comparing the odds of having a factor in affected versus unaffected people; above 1 suggests the factor is associated.
95 percent confidence interval	A range of values, built from sample data, that would capture the true value about 95 times out of 100 if the study were repeated.
recall bias	A study error where people with a condition remember past exposures differently than people without it, distorting the results.
matching	Pairing study participants with similar traits like age or sex across groups so those traits cannot bias the comparison.
teratogen	An agent such as a drug, chemical, or infection that can disrupt development and cause birth differences if exposure happens during pregnancy.

READING 5 OF 20 | DEVELOPMENT SCULPTS

Measuring Inheritance Over Time and in Twins

How can a study measure risk forward in time, and how can comparing twins separate inherited risk from environmental risk?

START WITH

Case-control studies efficiently study rare outcomes, but their odds ratios remain vulnerable to selection, recall, and confounding.

REMEMBER

Cohorts establish time order, while twin comparisons estimate population patterns under assumptions rather than one person's cause.

Why this matters

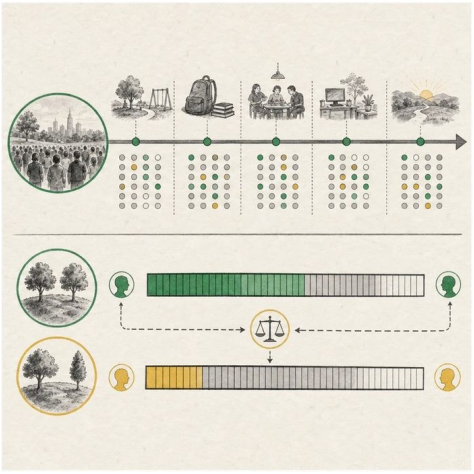
These designs answer different parts of the risk question and expose different sources of uncertainty.

The biology in plain English

A cohort starts before the outcome and follows people over time. It can establish that an exposure came first, but confounding can still remain.

Twin studies compare how often both twins share a trait. Greater similarity in identical twins can support genetic contribution, but genes and environments are not perfectly separated.

Heritability describes variation in a particular population under particular conditions. It does not explain a fixed percentage of one person's condition.



Biomedical teaching illustration: Compare a prospective cohort with a twin concordance study

Check your understanding

- What time-order strength does a cohort add?
- Why is heritability not a personal cause score?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP05-E1	A cohort records an exposure before birth outcomes are known.	This supports time order and reduces some recall problems.
EXP05-E2	Monozygotic twins share more DNA than dizygotic twins on average.	Higher concordance can support a genetic contribution at the population level.
EXP05-E3	Twin estimates rely on assumptions, and heritability changes across populations and settings.	The estimate does not assign a percentage cause to Mateo.

Concrete decision

You are selecting a design for a grant proposal on cleft risk. The proposal claims a twin study can cleanly separate genes from environment for one infant.

A. Use cohort and twin evidence with explicit assumptions and population-level wording.

B. Treat twin concordance as a personal genetic diagnosis.

C. Ignore shared environment because identical twins share DNA.

Decision prompt: Choose the defensible claim and name one strength plus one assumption for each design.

Claim ceiling: You may discuss time order and population variance. You may not assign Mateo's cleft a heritability percentage or single cause.

Glossary

Term	Plain-English definition
cohort study	A study that follows a group of people over time, comparing those with and without an exposure to see who develops an outcome.
prospective	Describing a study that follows participants forward in time from the start, recording events as they happen rather than looking back.
relative risk	A number comparing how likely an outcome is in an exposed group versus an unexposed group, where one means no difference.
twin study	Research that compares identical and fraternal twins to estimate how much of a trait is due to genes versus environment.
concordance	How often both members of a pair, such as twins, share the same trait or condition.
zygosity	Whether twins came from one fertilized egg that split (identical) or from two separate fertilized eggs (fraternal).
heritability	A measure of how much of the variation in a trait across a population is due to genetic differences rather than environment.

READING 6 OF 20 | CELLS DIFFERENTIATE

Finding a Risk Gene Among Millions of Bases

How do scientists find a single risk gene hidden among millions of DNA bases?

START WITH

Cohorts establish time order, while twin comparisons estimate population patterns under assumptions rather than one person's cause.

REMEMBER

Gene discovery combines broad scans, family structure, quality control, and independent replication.

Why this matters

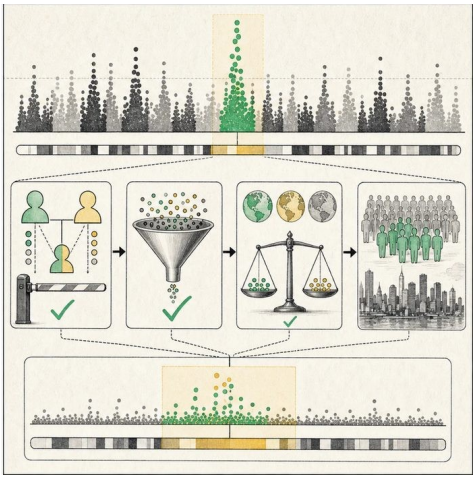
Gene discovery becomes trustworthy through converging checks, not one dramatic peak.

The biology in plain English

A genome-wide association study scans many common variants for frequency differences between groups. A peak points to a region linked with risk.

Family trios test transmission from parents to an affected child. Other family designs can also identify new or rare variants.

Quality control removes poor data and addresses ancestry structure. Replication asks whether the signal appears in an independent group.



Biomedical teaching illustration: Move from a GWAS signal to a replicated cleft-risk locus

Check your understanding

- What does a GWAS peak identify first?
- Why is replication independent?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP06-E1	A GWAS compares millions of common variants across many people.	It can reveal risk-associated regions without selecting one candidate in advance.
EXP06-E2	A case-parent trio can test whether one allele is transmitted to affected children more often than expected.	Family structure can reduce some population-stratification problems.
EXP06-E3	Genotyping quality, ancestry structure, sample size, and independent replication affect credibility.	A single peak is not yet a proven causal gene.

Concrete decision

You are the statistical geneticist reviewing a new genome-wide peak. One sample shows a strong signal near a developmental gene, but no second cohort has tested it.

A. Call it a candidate locus and require quality checks plus independent replication.

B. Name the nearest gene as Mateo's proven cause.

C. Ignore the signal because GWAS can never find useful loci.

Decision prompt: Choose the next step and distinguish a locus, an associated variant, and a causal mechanism.

Claim ceiling: You may nominate a replicated risk locus. You may not name a causal gene or patient diagnosis from proximity alone.

Glossary

Term	Plain-English definition
SNP	A single-nucleotide polymorphism, a one-letter difference in DNA at a specific spot that varies between people.
GWAS	A genome-wide association study, which scans DNA from many people to find genetic variants linked to a trait or disease.
transmission disequilibrium test	A family-based test that checks whether a gene variant is passed from heterozygous parents to affected children more often than chance.
case-parent trio	A study design that compares the DNA of an affected child with that of both parents to find gene variants passed down more often than expected.
over-transmission	When a particular gene version is passed from parents to affected children more often than the fifty-fifty rate expected by chance.
population stratification	A bias in genetic studies that appears when groups differ in both ancestry and trait rates, falsely linking variants to the trait.

READING 7 OF 20 | DEVELOPMENT SCULPTS

Is the Association Real, or Just Chance?

How do we know an association is real and not just chance?

START WITH

Gene discovery combines broad scans, family structure, quality control, and independent replication.

REMEMBER

Effect size and confidence interval show magnitude and precision; a p-value measures model compatibility, not truth.

Why this matters

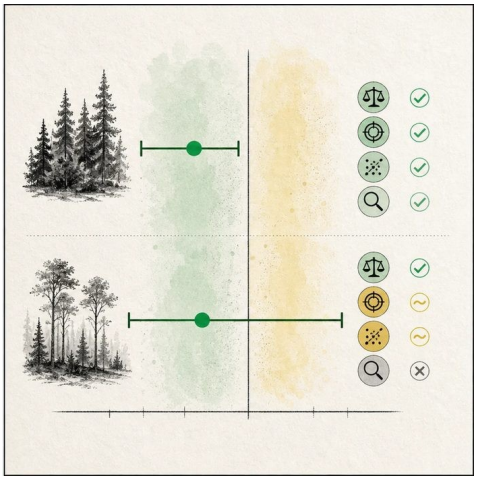
Students can make better decisions when statistics stay connected to size, uncertainty, and study quality.

The biology in plain English

An effect size tells how large a difference or association is. A confidence interval shows a range of values compatible with the data and model.

A wide interval signals low precision. An interval that crosses the null value includes both no difference and effects in more than one direction.

A p-value asks how unusual the data would be under a stated null model. It does not measure importance, bias, or the chance that a claim is true.



Biomedical teaching illustration: Interpret two treatment estimates without a significance shortcut

Check your understanding

- Which result is more precise, and why?
- What does a p-value not tell you?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP07-E1	Study A reports a risk ratio of 0.60 with a 95 percent interval from 0.37 to 0.98.	The estimate suggests lower risk, with values near no difference still close to the interval edge.
EXP07-E2	Study B reports a risk ratio of 0.55 with an interval from 0.18 to 1.67.	The result is much less precise and includes possible benefit, no difference, or harm.
EXP07-E3	A p-value is calculated under a statistical model and null hypothesis.	It is not the probability that the scientific claim is true or false.

Concrete decision

You are the data reviewer preparing a one-slide result summary. A teammate wants to label Study A true and Study B false based only on whether p is below 0.05.

A. Compare effect sizes, intervals, designs, and bias before stating the claim.

B. Use p below 0.05 as proof and ignore magnitude.

C. Use the larger-looking effect as proof even when its interval is wide.

Decision prompt: Choose the interpretation and explain why the two intervals support different precision claims.

Claim ceiling: You may describe magnitude, precision, and statistical compatibility. You may not convert a threshold into scientific truth.

Glossary

Term	Plain-English definition
association	A statistical link where two things tend to occur together, which suggests a relationship but does not by itself prove one causes the other.
odds ratio	A number comparing the odds of having a factor in affected versus unaffected people; above 1 suggests the factor is associated.
p-value	A number showing how likely a result could happen just by chance; a smaller p-value makes luck a less believable explanation.
confidence interval	A range of values that likely contains the true result, showing how precise an estimate is; a narrow range means more certainty.
statistical significance	A result unlikely to be due to chance, often shown by a p-value below a set threshold such as 0.05.
null result	A finding that shows no meaningful effect or difference, which is still valuable evidence and worth reporting.

Why Gene Studies Use Such Tiny P-Values

Why do gene studies demand a far tinier p-value than 0.05?

START WITH

Effect size and confidence interval show magnitude and precision; a p-value measures model compatibility, not truth.

REMEMBER

Many tests inflate false positives, so thresholds and independent replication must be planned before results are seen.

Why this matters

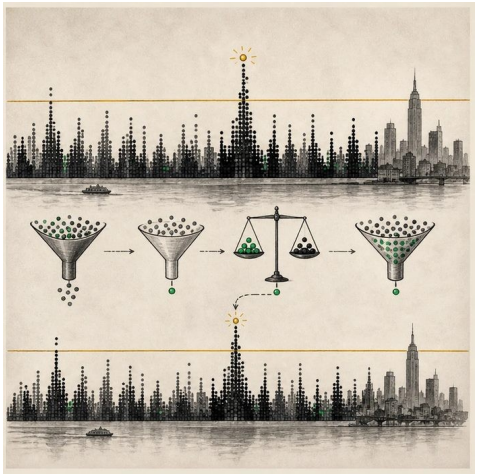
Large searches are powerful only when their many chances to be wrong are counted.

The biology in plain English

When a study runs many tests, some small p-values can appear by chance. Multiple-testing methods reduce that false-positive risk.

A threshold near 5 x 10^-8 is widely used for many common-variant GWAS analyses. Different designs may need different corrections.

Replication tests the signal in new participants. Researchers also check data quality, ancestry structure, effect direction, and biological follow-up.



Biomedical teaching illustration: Control false positives in a common-variant GWAS

Check your understanding

- Why is p below 0.05 too loose for one million tests?
- What does replication add beyond a smaller p-value?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP08-E1	Testing one million variants at p below 0.05 would produce many chance hits under the null.	A usual single-test threshold is not adequate for a genome-wide scan.
EXP08-E2	For many common-variant GWAS analyses, p below 5 x 10^-8 is a conventional threshold.	The convention is not universal for every genetic analysis.
EXP08-E3	The candidate peak repeats in an independent cohort with the same direction of effect.	Replication lowers concern that the first signal was sample-specific chance or error.

Concrete decision

You are the GWAS quality-control lead. A peak has p equal to 2 x 10^-6 in the discovery sample and no replication result.

A. Label it suggestive and require the planned threshold, quality checks, and replication.

B. Declare it significant because it is below 0.05.

C. Use 5 x 10^-8 as a magic rule for every genetic test.

Decision prompt: Choose the report language and explain how multiplicity and replication change the claim.

Claim ceiling: You may apply the supplied common-variant GWAS convention. You may not treat that number as universal or proof of mechanism.

Glossary

Term	Plain-English definition
multiple testing	Running many statistical tests at once, which raises the chance of a false positive and requires a stricter cutoff to stay reliable.
false positive	A test result that signals a condition is present when it actually is not, a kind of error that can lead to needless worry or treatment.
Bonferroni correction	A math adjustment that makes the cutoff for significance stricter when you run many tests, so chance results are not mistaken for real ones.
genome-wide significance	A very strict p-value cutoff (about 5 in 100 million) used in genome-wide studies so that testing millions of spots does not produce false hits.
false discovery rate	The expected share of results called significant that are actually false alarms, a key check when many tests are run at once.
replication	Repeating a study to see if the result holds; a key test of whether a finding is real.

READING 9 OF 20 | CELLS DIFFERENTIATE

Taking a Gene to the Bench

How do we test a gene's actual job at the bench?

START WITH

Many tests inflate false positives, so thresholds and independent replication must be planned before results are seen.

REMEMBER

Orthogonal assays answer different mechanism questions: where RNA or protein is and what changes after perturbation.

Why this matters

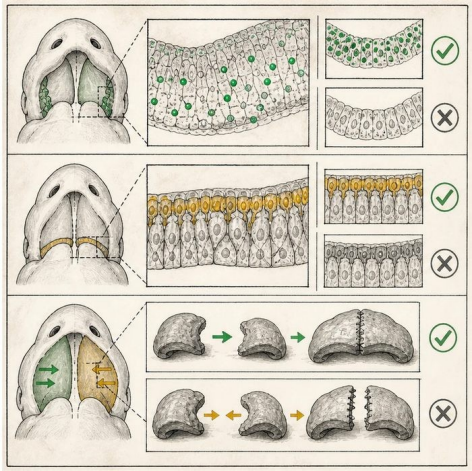
Mechanism becomes clearer when different methods agree while answering different questions.

The biology in plain English

In situ hybridization can show where RNA is present. Immunohistochemistry uses antibodies to show where a protein is located.

A knockout or knockdown asks what changes when gene function is reduced. Each method needs controls for background signal and handling.

Expression shows opportunity for a gene to act. Perturbation adds functional evidence, but alternative effects and model limits still matter.



Biomedical teaching illustration: Test a candidate developmental gene with three assays

Check your understanding

- What does in situ hybridization locate?
- Why is expression alone not proof of function?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP09-E1	RNA signal appears in developing palatal shelves at the relevant stage.	The gene is transcribed in a useful place and time.
EXP09-E2	Protein staining appears in palatal epithelium and is absent in the no-primary-antibody control.	The protein location is supported by a specificity control.
EXP09-E3	Perturbed embryos show altered shelf elevation more often than matched controls.	Changing the gene is associated with a phenotype, but other effects still need checks.

Concrete decision

You are planning the bench follow-up to a replicated locus. The team has funding for three assays and must decide whether one expression image alone proves function.

A. Combine RNA location, protein location, perturbation, and controls.

B. Use one bright stain as proof the gene causes cleft palate.

C. Skip controls because the candidate came from GWAS.

Decision prompt: Choose the assay set and state the specific question each assay answers.

Claim ceiling: You may connect expression and perturbation evidence. You may not treat location alone as causal proof.

Glossary

Term	Plain-English definition
knockout	An organism or cell engineered to have a specific gene switched off, used to reveal what that gene normally does.
conditional knockout	A genetic tool that switches a gene off only in certain cells or at a chosen time, so researchers can study its role without affecting the whole animal.
in situ hybridization	A lab method that uses a labeled probe to show exactly where a specific gene's RNA or DNA sits inside a tissue or cell.
immunohistochemistry	A lab technique that uses antibodies tagged with stain to show exactly where a specific protein sits within a tissue slice.
penetrance	The fraction of people carrying a disease-linked genotype who actually show the trait.
control	The comparison condition that isolates the effect of the variable being tested by keeping everything else the same.

READING 10 OF 20 | DEVELOPMENT CAN GO WRONG

CRISPR as an Experimental Tool

How do we edit a specific gene to test exactly what it does?

START WITH

Orthogonal assays answer different mechanism questions: where RNA or protein is and what changes after perturbation.

REMEMBER

CRISPR is a controlled perturbation only when the edit, mosaicism, phenotype, and off-target controls are verified.

Why this matters

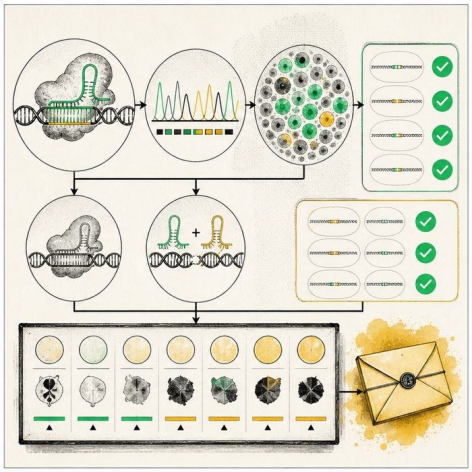
A programmable tool is only as informative as the verification wrapped around it.

The biology in plain English

CRISPR uses a guide RNA to bring a cutting enzyme near a chosen DNA sequence. Repair can create a small change or install a planned edit.

Researchers sequence the target to confirm what changed. In early embryos, mosaicism means different cells may carry different edits.

Non-targeting controls, a second guide, off-target checks, blinded scoring, and rescue can help separate a gene effect from artifacts.



Biomedical teaching illustration: Verify a CRISPR perturbation before reading the palate phenotype

Check your understanding

- What does mosaicism mean?
- Why use a second guide?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP10-E1	Sequencing confirms a frameshift in 78 percent of reads from the target tissue.	The edit occurred, but the sample is mosaic.
EXP10-E2	A second guide produces a similar phenotype while a non-targeting guide does not.	Independent perturbation and a negative control strengthen specificity.
EXP10-E3	The top predicted off-target sites show no detectable edits in the supplied assay.	The tested concern is reduced, not eliminated everywhere in the genome.

Concrete decision

You are the genome-editing quality reviewer. One embryo has a palate defect after injection, but the target was not sequenced and no control guide was used.

A. Withhold the gene-function claim and require edit plus specificity checks.

B. Call the gene causal because CRISPR was injected.

C. Ignore mosaicism because some cells were edited.

Decision prompt: Choose the claim status and cite the minimum verification steps needed next.

Claim ceiling: You may interpret the supplied edit and control checks. You may not claim all off-target effects are absent or a human treatment is ready.

Glossary

Term	Plain-English definition
CRISPR-Cas9	A gene-editing tool that uses a guide RNA to steer the Cas9 protein to a chosen DNA site and cut it, so a sequence can be changed.
guide RNA	A lesson term defined by the surrounding model and case evidence.
off-target editing	An unwanted DNA change made by a gene-editing tool at the wrong spot in the genome, away from its intended target.
mosaicism	When a person's body contains two or more genetically different sets of cells that arose from a single fertilized egg.
editing efficiency	The percentage of treated cells in which a gene-editing tool such as CRISPR actually made the intended change to the DNA.
sequence verification	Reading a stretch of DNA again to confirm that a detected variant is real and not an error from the test.

READING 11 OF 20 | DEVELOPMENT SCULPTS

When Is a Mouse a Good Stand-In for Mateo?

When does studying a mouse actually teach us something true about Mateo, and when does it fool us?

START WITH

CRISPR is a controlled perturbation only when the edit, mosaicism, phenotype, and off-target controls are verified.

REMEMBER

A model earns relevance by matching the mechanism and outcome needed for the question while following Replace, Reduce, and Refine.

Why this matters

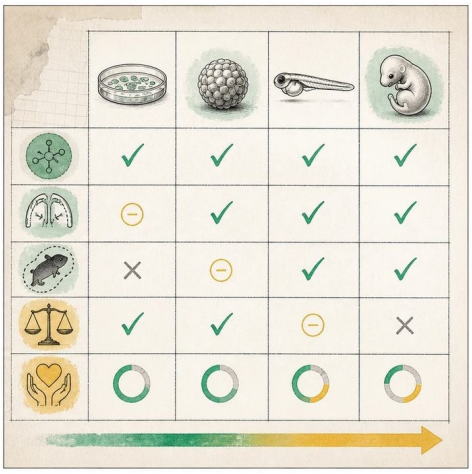
Model validity and research ethics belong in the same decision, not in separate checklists.

The biology in plain English

A useful model matches the biological process and outcome needed for a question. No model copies every part of a human condition.

Replace means use a non-animal method when it can answer the question. Reduce means use no more animals than needed. Refine means reduce pain and improve care.

Researchers should plan sample size, controls, humane endpoints, randomization, and blinded scoring before a study begins.



Biomedical teaching illustration: Select a model for palatal shelf elevation

Check your understanding

- What should determine model choice?
- What do the three Rs mean?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP11-E1	The question asks whether a pathway changes mammalian palatal shelf elevation and fusion.	A model must represent the relevant tissue movement and timing.
EXP11-E2	A cell culture can test signaling but cannot reproduce whole-palate elevation.	It can replace animals for some early questions, not the full outcome.
EXP11-E3	A mouse study uses prior data for sample size, shared controls, analgesia, humane endpoints, and blinded scoring.	The design applies Reduce and Refine while protecting validity.

Concrete decision

You are the model-selection and animal-welfare reviewer. The team wants to use the largest animal available because it seems more human-like.

A. Choose the least burdensome model that can measure the required mechanism and outcome.

B. Choose by body size alone.

C. Reject all nonhuman and cell models as useless.

Decision prompt: Choose the model sequence and justify it with mechanism fit plus the 3Rs.

Claim ceiling: You may judge fitness for the stated question. You may not call any model a complete stand-in for Mateo.

Glossary

Term	Plain-English definition
animal model	A lab animal, such as a mouse, used to study a human disease because its biology is similar enough to test causes and treatments safely.
face validity	Whether a test or measure looks, on the surface, like it really assesses what it claims to assess.
construct validity	How well a test or measurement actually captures the abstract idea it claims to measure, rather than something else.
the 3Rs	An ethics framework for animal research, replace, reduce, and refine, that minimizes animal use and suffering while keeping good science.
ARRIVE guidelines	A standard checklist for reporting animal research clearly and completely, so other scientists can judge the study and repeat it.

Knock It Out, Then Put It Back: Proving a Gene Causes a Defect

What extra experiment turns 'knock out the gene and a cleft appears' into proof the gene itself caused the cleft?

START WITH

A model earns relevance by matching the mechanism and outcome needed for the question while following Replace, Reduce, and Refine.

REMEMBER

Loss plus targeted rescue strengthens causation when the rescue is specific, verified, and compared with controls.

Why this matters

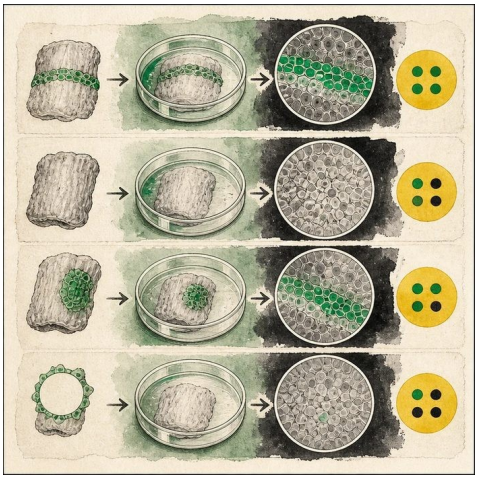
Rescue asks whether restoring the missing function recovers the outcome.

The biology in plain English

A knockout tests what happens when gene function is lost. A rescue restores the gene or pathway in a planned tissue and time.

If the phenotype improves after targeted rescue but not after an empty control, the causal explanation becomes stronger.

Rescue may be partial, and it does not prove that the gene works alone. Researchers still check dose, timing, background, and delivery.



Biomedical teaching illustration: Test knockout and epithelial rescue of a palate phenotype

Check your understanding

- What does the empty-vector group control?
- Why is rescue not automatic proof of strict sufficiency?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP12-E1	Wild-type cultures fuse in 18 of 20 samples, while knockout cultures fuse in 4 of 20.	Loss of function is associated with a large defect in this model.
EXP12-E2	Targeted epithelial rescue restores fusion in 15 of 20 knockout samples.	Restoring function in the relevant tissue recovers much of the outcome.
EXP12-E3	Empty-vector rescue produces fusion in 5 of 20 samples.	The delivery procedure alone does not explain the targeted rescue.

Concrete decision

You are the mechanism-review panelist. The team must decide whether knockout alone is enough for its strongest causal claim.

A. Use knockout plus targeted rescue and controls to strengthen the model-specific causal claim.

B. Claim the gene is sufficient for normal human palate development from knockout alone.

C. Ignore the empty-vector group because rescue improved.

Decision prompt: Choose the claim and compare all four supplied groups.

Claim ceiling: You may claim that targeted restoration strengthens causation in this model. You may not claim strict sufficiency or a patient-specific cause.

Glossary

Term	Plain-English definition
knockout	An organism or cell engineered to have a specific gene switched off, used to reveal what that gene normally does.
rescue experiment	Adding a working gene back into a mutant to see if it restores the normal trait, which proves that gene was responsible.
causation	When one factor actually brings about another, shown by controlled evidence rather than by the two simply appearing together.
redundancy control	Repeating measurements or samples in an experiment so a single fluke does not drive the conclusion.
necessity and sufficiency	Two tests of a cause: necessary means the result fails without it, and sufficient means the factor alone can produce the result.

READING 13 OF 20 | CELLS MOVE

The Fairest Test: How to Compare Two Treatments in Real Children

When you cannot control a child's biology, what makes a comparison of two treatments actually fair?

START WITH

Loss plus targeted rescue strengthens causation when the rescue is specific, verified, and compared with controls.

REMEMBER

Fair treatment trials require equipoise, concealed randomization, unbiased outcome assessment, and analysis by assigned group.

Why this matters

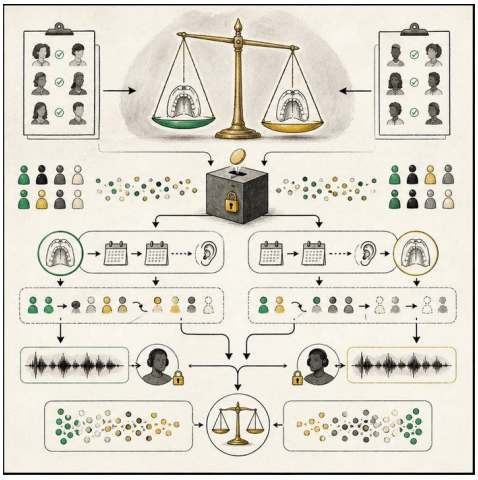
Randomization is an ethical and statistical tool, not a badge that automatically makes a trial fair.

The biology in plain English

Equipoise means the expert community has genuine uncertainty between acceptable options. Without it, random assignment can be unethical.

Allocation concealment prevents recruiters from predicting the next group. Blinded outcome assessment reduces expectation effects.

Intention-to-treat analyzes participants in their assigned groups. It protects the original comparison even when schedules change.



Biomedical teaching illustration: Design a fair comparison of two accepted palate-repair timings

Check your understanding

- What must exist before randomization is ethical?
- Why keep crossovers in their assigned group?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP13-E1	Both timings are accepted options and experts remain uncertain about their balance of outcomes.	Clinical equipoise can support randomization.
EXP13-E2	A central system reveals assignment only after an eligible infant is enrolled.	Allocation concealment reduces selection manipulation.
EXP13-E3	Speech assessors receive coded recordings and analysis keeps participants in assigned groups.	Blinding and intention-to-treat protect the comparison.

Concrete decision

You are the clinical-trial methods chair. A surgeon wants to choose timing after seeing the next assignment and remove crossovers from analysis.

A. Use concealed randomization, coded assessment, and intention-to-treat.

B. Let the recruiter predict and change assignments.

C. Analyze only participants who perfectly followed the schedule.

Decision prompt: Choose the fair-trial plan and explain the threat controlled by each safeguard.

Claim ceiling: You may design the comparison when both options are acceptable. You may not randomize repair versus knowingly harmful no repair.

Glossary

Term	Plain-English definition
randomized controlled trial	A study that randomly assigns participants to a treatment or a control group, the strongest design for showing a treatment truly works.
randomization	Assigning participants to study groups purely by chance, so the groups start out similar and the comparison stays fair.
blinding	Keeping participants, researchers, or both unaware of who got which treatment so expectations cannot bias the results.
intention-to-treat	A way to analyze a trial that keeps every patient in their first-assigned group, even if they switched or quit, to give an honest result.
equipoise	Genuine uncertainty in the expert community about which treatment is better, which is what makes randomly assigning patients in a trial ethical.

READING 14 OF 20 | DEVELOPMENT SCULPTS

The TOPS Trial: How Surgeons Tested the Best Time to Repair a Palate

How did one real randomized trial test the best timing of cleft palate surgery, and what exactly did its main result prove?

START WITH

Fair treatment trials require equipoise, concealed randomization, unbiased outcome assessment, and analysis by assigned group.

REMEMBER

TOPS supports a narrow claim: earlier repair lowered VPI in its eligible population, with uncertainty and limits.

Why this matters

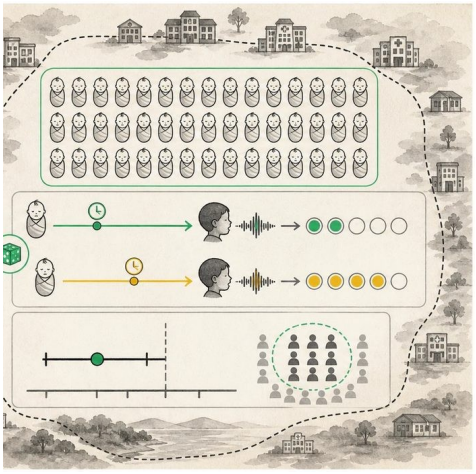
A real trial becomes useful only when students read both its finding and the boundary around it.

The biology in plain English

TOPS compared palate repair at 6 months with repair at 12 months in medically fit infants with isolated cleft palate. It used many centers and blinded speech assessment.

At age 5, VPI was recorded in 8.9 percent of the earlier group and 15.0 percent of the later group. The risk ratio was 0.59, with a 95 percent interval from 0.36 to 0.99.

The result supports earlier repair for the trial's eligible population and settings. It does not answer every cleft type, center, safety, or growth question.



Biomedical teaching illustration: Read the TOPS trial from eligibility to claim ceiling

Check your understanding

- Who was eligible for TOPS?
- Why can the result not be applied to every cleft case?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP14-E1	TOPS randomized 558 medically fit infants with nonsyndromic isolated cleft palate at 23 centers.	The study population did not include every cleft type or setting.
EXP14-E2	VPI at age 5 occurred in 8.9 percent of the 6-month group and 15.0 percent of the 12-month group among analyzable participants.	The earlier group had fewer primary-outcome events.
EXP14-E3	The risk ratio was 0.59 with a 95 percent interval from 0.36 to 0.99.	The estimate favors earlier repair, but other outcomes and uncertainty matter.

Concrete decision

You are the evidence lead briefing a cleft team. The team asks whether TOPS proves every infant with cleft lip and palate should have repair at 6 months in every center.

A. Apply the result to similar eligible infants while discussing uncertainty, setting, and other outcomes.

B. Generalize the result to every cleft type and center.

C. Ignore the trial because its interval approaches the null.

Decision prompt: Choose the briefing and cite population, event rates, risk ratio, and one limit.

Claim ceiling: You may state the primary TOPS finding for its eligible population. You may not prescribe timing for Mateo or generalize to all clefts and settings.

Glossary

Term	Plain-English definition
primary outcome	The single main result a study is designed to measure, chosen ahead of time to answer its central question.
risk ratio	A number comparing the chance of an outcome in an exposed group versus an unexposed group; above 1 means higher risk with exposure.
confidence interval	A range of values that likely contains the true result, showing how precise an estimate is; a narrow range means more certainty.
power calculation	A planning step that estimates how many subjects a study needs to reliably detect a real effect if one exists.
generalizability	How well a study's findings hold true for people or settings beyond the specific group that was studied.

READING 15 OF 20 | CELLS DIFFERENTIATE

How Do You Measure Something as Fuzzy as Speech?

How do you turn a fuzzy idea like 'good speech' into something you can measure fairly and identically for hundreds of different children?

START WITH

TOPS supports a narrow claim: earlier repair lowered VPI in its eligible population, with uncertainty and limits.

REMEMBER

An operational definition makes speech scoring repeatable, while patient-reported outcomes capture effects clinician scores can miss.

Why this matters

A study cannot compare treatment effects unless its outcome means the same thing across groups.

The biology in plain English

An operational definition states exactly how a concept will be measured. For speech, it names the task, recording method, scale, and scoring rule.

Training, coded recordings, and agreement checks reduce rater drift. Language and dialect must be considered so normal differences are not mislabeled.

Patient-reported outcomes ask how speech affects communication or daily life. They add information that a clinician score may miss.

Biomedical teaching illustration: Operationalize VPI and speech participation

Check your understanding

- What belongs in an operational definition?
- Why add a patient-reported outcome?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP15-E1	The protocol defines VPI using specified ratings from standardized speech samples.	Every group is measured with the same rule.
EXP15-E2	Assessors are trained, blinded to timing, and agreement is checked on duplicate coded recordings.	The process tests whether scoring is reliable.
EXP15-E3	A validated child or caregiver measure records speech participation and burden.	The study includes an outcome that clinical resonance scores may miss.

Concrete decision

You are the speech-outcomes lead finalizing the protocol. The draft outcome is simply 'speech sounds better,' with no task, rule, or patient voice.

A. Define the task, rubric, training, blinding, reliability check, and reported outcome.

B. Keep the phrase because every listener knows better speech.

C. Use only one surgeon's impression.

Decision prompt: Choose the measurement plan and explain what each measure adds.

Claim ceiling: You may define reliable study outcomes. You may not claim one score represents every language, dialect, or lived experience.

Glossary

Term	Plain-English definition
outcome measure	The specific, defined result a study tracks to judge whether a treatment worked, such as speech quality or healing rate.
operational definition	A precise statement of how a concept will be measured or observed in a study, so anyone can apply it the same way.
standardization	Making methods, measurements, or definitions the same across a study or site, so results can be fairly compared.
patient-reported outcome	A result measured directly from what the patient says about their own health, symptoms, or quality of life.
validated instrument	A survey or test that has been formally checked to make sure it measures what it claims to, consistently and accurately.

What Could Fool Us Into the Wrong Conclusion?

What could fool us into thinking we found the right cause, when we did not?

START WITH

An operational definition makes speech scoring repeatable, while patient-reported outcomes capture effects clinician scores can miss.

REMEMBER

Bias is directional design or measurement error; confounding mixes effects; each threat needs a matched safeguard.

Why this matters

Precision cannot rescue a study that measures or compares the wrong things in a consistent direction.

The biology in plain English

Bias is systematic error from selection, measurement, missing data, or study conduct. A larger sample can make a biased estimate more precise without making it correct.

A confounder is linked to both the factor being studied and the outcome, but is not simply a step in the causal pathway.

Randomization, blinding, retention, matching, measurement, and adjusted analysis address different threats. No safeguard is perfect.

Biomedical teaching illustration: Red-team a speech study for selection, measurement, and confounding

Check your understanding

- Why can a large sample still be biased?
- How is a confounder different from measurement bias?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP16-E1	Earlier-repair participants are more likely to receive care at high-volume centers with strong speech services.	Center resources could confound an observational comparison.
EXP16-E2	Assessors know each child's group and expect earlier repair to work.	Expectation could create directional measurement bias.
EXP16-E3	Follow-up is 92 percent in one group and 71 percent in the other, with access-related reasons.	Differential attrition could change the comparison.

Concrete decision

You are the adversarial methods reviewer before data lock. The team says a sample of 5,000 makes bias and confounding disappear.

A. Match each supplied threat with a design, measurement, or analysis safeguard.

B. Accept the claim because large samples remove systematic error.

C. List every possible bias without changing the protocol.

Decision prompt: Choose the red-team response and propose one matched safeguard per evidence card.

Claim ceiling: You may reduce named threats. You may not claim adjustment removes all unmeasured bias or confounding.

Glossary

Term	Plain-English definition
bias	A systematic error in how data is collected or interpreted that tilts results in one direction, making conclusions less accurate or unfair.
confounding	When a hidden third factor influences both the suspected cause and the outcome, making a link look real when it may be misleading.
blinding	Keeping participants, researchers, or both unaware of who got which treatment so expectations cannot bias the results.
recall bias	A study error where people with a condition remember past exposures differently than people without it, distorting the results.
surgeon-as-confounder	A study bias where differences in patient outcomes come from which surgeon operated rather than from the treatment being tested.

READING 17 OF 20 | DEVELOPMENT SCULPTS

How Do We Combine Many Studies Into One Answer?

How do we combine many separate studies into one trustworthy answer, without cherry-picking?

START WITH

Bias is directional design or measurement error; confounding mixes effects; each threat needs a matched safeguard.

REMEMBER

Systematic reviews reduce cherry-picking, while meta-analysis is credible only when studies and pooling assumptions support it.

Why this matters

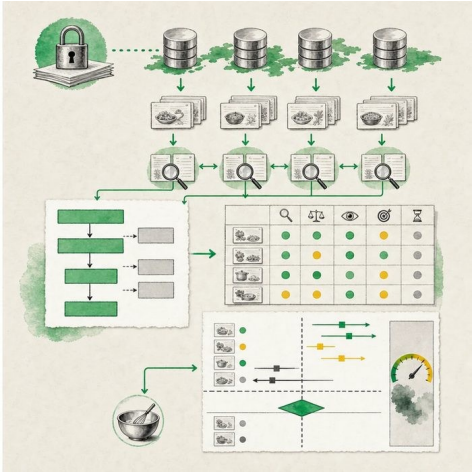
A systematic process protects the answer from selective attention, but synthesis still requires judgment.

The biology in plain English

A systematic review uses a planned question, broad search, clear eligibility rules, duplicate screening, and risk-of-bias assessment. PRISMA guides transparent reporting.

Meta-analysis combines numerical estimates when studies are similar enough for the average to mean something. It is not required for every review.

Heterogeneity means results or study features differ. Researchers examine populations, methods, outcomes, and settings before pooling.



Biomedical teaching illustration: Move from a PRISMA flow diagram to a defensible forest plot

Check your understanding

- What should be planned before screening?
- When might a review avoid one pooled estimate?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP17-E1	The protocol was registered before screening and lists databases, eligibility, outcomes, and analysis.	The process is harder to change after seeing results.
EXP17-E2	Two studies use blinded age-5 VPI ratings, while one uses an unvalidated parent question at age 2.	Outcome differences may limit a meaningful pooled estimate.
EXP17-E3	The forest plot shows effects in different directions and major clinical differences across centers.	Heterogeneity must be explained rather than hidden by one average.

Concrete decision

You are the systematic-review editor deciding whether to approve pooling. The authors want one dramatic number even though one study measures a different outcome and timing.

A. Present the review and pool only a justified compatible subset.

B. Combine every number because meta-analysis is always strongest.

C. Discard studies that disagree with the preferred answer.

Decision prompt: Choose the synthesis plan and cite protocol, outcome compatibility, and heterogeneity evidence.

Claim ceiling: You may synthesize and pool compatible studies when justified. You may not claim a pooled estimate repairs weak or incompatible evidence.

Glossary

Term	Plain-English definition
systematic review	A study that follows a planned, transparent method to gather, judge, and combine all relevant research on one question.
meta-analysis	A study that statistically combines the results of many separate studies to reach a more reliable overall conclusion.
heterogeneity	Wide variation within a group, such as many different genetic causes leading to the same outward condition.
PRISMA 2020	A widely used checklist and flow diagram that guides researchers to report systematic reviews clearly, completely, and transparently.
publication bias	The tendency for studies with positive or exciting results to get published while studies with no effect stay hidden, skewing the evidence.

READING 18 OF 20 | CELLS DIFFERENTIATE

What Makes Research on Children Ethical?

What has to be true before scientists are allowed to study a baby like Mateo?

START WITH

Systematic reviews reduce cherry-picking, while meta-analysis is credible only when studies and pooling assumptions support it.

REMEMBER

Child research requires favorable risk and benefit, IRB review, parental permission, and affirmative assent when the child is capable.

Why this matters

Children can contribute important evidence only when the study respects their welfare and developing autonomy.

The biology in plain English

In the United States, federal Subpart D rules add protections for children in research. An IRB reviews risk, benefit, privacy, permission, and assent.

Parental permission is similar to informed consent from a parent or guardian. Information must be understandable, voluntary, and honest.

Assent is a child's affirmative agreement when capable. Researchers explain the study at an appropriate level and respect dissent within the approved plan.

Biomedical teaching illustration: Apply Subpart D protections to a pediatric speech study

Check your understanding

- What extra review protects children?
- Why is silence not assent?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP18-E1	The study involves children, coded recordings, and no change from accepted care.	The IRB must assess risk, privacy, and added protections.
EXP18-E2	The protocol gives caregivers plain-language information and time for questions.	Permission should be informed and voluntary.
EXP18-E3	Children who can understand receive an age-appropriate explanation and are asked for affirmative agreement.	Assent is more than silence or failure to object.

Concrete decision

You are the pediatric IRB reviewer. A protocol says parental signature alone is enough and children should not be told because that is faster.

A. Require IRB-approved protections, permission, and capability-based assent.

B. Use only the parent signature for every capable child.

C. Treat a child's silence as affirmative assent.

Decision prompt: Choose the ethics correction and explain the role of the IRB, permission, and assent.

Claim ceiling: You may apply the supplied educational ethics framework. You may not replace local IRB review or legal guidance.

Glossary

Term	Plain-English definition
informed consent	Agreeing to a treatment or study only after truly understanding its purpose, risks, benefits, and alternatives; the understanding, not the signature, is the ethics.
assent	A child or teen's own agreement to take part in a study, given alongside a parent's legal permission, after the study is explained at their level.
equipoise	Genuine uncertainty in the expert community about which treatment is better, which is what makes randomly assigning patients in a trial ethical.
Institutional Review Board (IRB)	A committee that reviews research involving people to make sure it is ethical and that participants are protected from harm.
incidental findings	Unexpected health results discovered by a test that was looking for something else, raising questions about whether to report them.

READING 19 OF 20 | CELLS MOVE

How Does the World Check That a Study Is Trustworthy?

Once a study is finished, how does the rest of the world check whether it is actually trustworthy?

START WITH

Child research requires favorable risk and benefit, IRB review, parental permission, and affirmative assent when the child is capable.

REMEMBER

Transparent reporting lets others appraise and reproduce methods, but it cannot repair a weak design.

Why this matters

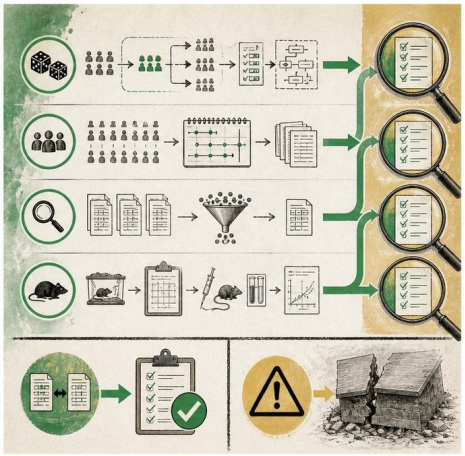
Science becomes checkable when the route from question to result is visible.

The biology in plain English

Reporting guidelines list information readers need. CONSORT supports randomized trials, STROBE observational studies, PRISMA systematic reviews, and ARRIVE animal research.

Transparent reports include participants, methods, outcomes, missing data, harms, analyses, and changes from the plan. Protocols and code can add detail.

Good reporting reveals what was done. It does not fix poor randomization, biased measurement, weak models, or selective data collection.



Biomedical teaching illustration: Match CONSORT, STROBE, PRISMA, and ARRIVE to four studies

Check your understanding

- Which guideline fits a systematic review?
- Why does reporting not repair weak design?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP19-E1	The trial report includes eligibility, allocation, participant flow, outcomes, harms, and analysis.	Readers can check what happened from enrollment to result.
EXP19-E2	The study type determines the matched guideline: CONSORT, STROBE, PRISMA, or ARRIVE.	One checklist does not fit every design.
EXP19-E3	A fully reported study still has unconcealed allocation and severe missing data.	Transparency exposes the weakness but does not remove it.

Concrete decision

You are the journal reproducibility editor. An author says completing a reporting checklist proves the study is trustworthy.

A. Use the matched guideline, then judge design, conduct, analysis, and replication separately.

B. Accept the checklist as proof of validity.

C. Reject reporting guidelines because they cannot fix a study.

Decision prompt: Choose the editorial decision and match each supplied study to its guideline.

Claim ceiling: You may judge reporting completeness and exposed strengths or weaknesses. You may not call a study reproducible or valid from a checklist alone.

Glossary

Term	Plain-English definition
reproducibility	The ability to get the same results when an experiment is repeated using the same methods, which builds trust in findings.
reporting standard	An agreed set of items a study must describe so readers can judge and reproduce the work, like the PRISMA checklist for reviews.
peer review	The process where independent experts evaluate a research study for quality and accuracy before it is published.
CONSORT	A standard checklist and flow diagram that researchers use to report every key detail of a randomized controlled trial clearly and honestly.
STROBE	A standard checklist that guides researchers on how to clearly and fully report observational studies in medicine.

READING 20 OF 20 | DEVELOPMENT SCULPTS

What New Question About Mateo Would You Investigate, and How?

What new question about Mateo's cleft is worth investigating, and how would you design a sound, ethical study to answer it?

START WITH

Transparent reporting lets others appraise and reproduce methods, but it cannot repair a weak design.

REMEMBER

A defensible study plan aligns question, design, variables, sampling, analysis, ethics, reporting, and a claim ceiling.

Why this matters

The final lesson asks students to build one coherent study, not recite disconnected method terms.

The biology in plain English

A protocol connects one focused question to a design that can answer it. Eligibility, variables, sample, outcomes, safeguards, and analysis are chosen before results.

This capstone uses a supplied coded dataset. Students do not recruit people, search for missing facts, or invent measurements. Ethics and privacy still shape the work.

The final claim must match the design. An observational cohort can estimate an adjusted association and uncertainty, but it cannot guarantee cause.

Biomedical teaching illustration: Defend one complete Mateo-related study protocol

Check your understanding

- Why can students finish this study inside the lesson package?
- What is the strongest supported claim?

Case evidence supplied for this lesson

ID	Finding supplied in the case	Why it matters
EXP20-E1	The question asks whether a structured speech follow-up program is associated with improved age-8 communication after palate repair.	The outcome and population are specific, but the program is not randomly assigned.
EXP20-E2	The supplied synthetic CSV contains 24 coded records with program group, center, baseline score, age-8 score, and missing-data flags.	Students can complete the design without outside recruitment or invented measurements.
EXP20-E3	The analysis compares adjusted estimates, reports intervals and missingness, protects coded data, and follows STROBE.	The strongest result is an adjusted association, not proof of cause.

Concrete decision

You are the principal investigator at the protocol defense. The board must select one complete plan from three proposals using only the supplied case file and dataset.

A. Approve the aligned coded-records cohort with its observational claim ceiling.

B. Approve a plan that recruits unnamed children and invents measurements.

C. Call the program causal because the adjusted p-value is small.

Decision prompt: Choose the defensible protocol and justify question, design, variables, sample, analysis, ethics, reporting, and claim ceiling.

Claim ceiling: You may defend an adjusted observational association using the supplied dataset. You may not claim causation or require outside data.

Glossary

Term	Plain-English definition
study protocol	The detailed written plan for a study that sets the question, methods, and steps before any data are collected.
primary outcome	The single main result a study is designed to measure, chosen ahead of time to answer its central question.
multifactorial	Caused by several genes acting together with environmental factors, not by a single gene alone.
nonsyndromic	Describing a condition, such as a cleft, that occurs on its own without a set of other linked birth differences.
convergent evidence	Support for a conclusion that comes from several independent methods or studies pointing to the same answer, which makes it more trustworthy.