Certificate in Medical Journalism

Understanding Medical Studies and Statistics

Absolute Risk – The probability that an individual will develop a specific health outcome over a defined period. Related terms: relative risk, incidence, prevalence. For example, if 5 out of 1,000 smokers develop lung cancer in ten years, the absolute risk is 0.5 %. Medical journalists use absolute risk to convey the real-world impact of exposures, avoiding exaggeration that can arise from relative measures. Practical application includes reporting the risk reduction from a new drug: “The medication lowers absolute risk of heart attack from 2 % to 1 %.” Challenges involve communicating small percentages clearly, especially when audiences misinterpret them as negligible, and ensuring that denominators are transparent.

Bias – Systematic error that skews study results away from the truth. Related terms: selection bias, information bias, confounding. Bias can arise at any stage, from participant recruitment to data analysis. For instance, a survey that only includes internet users may over-represent younger, tech-savvy individuals, creating selection bias. Journalists must identify potential bias when interpreting findings, noting statements such as “the authors acknowledge possible recall bias.” Practical steps include comparing study methods with established standards and seeking expert commentary. The main challenge is that bias is often subtle and may be hidden in methodological details that require careful reading.

Confidence Interval – A range of values, derived from sample data, that likely contains the true population parameter with a specified probability (commonly 95 %). Related terms: point estimate, margin of error, statistical significance. If a trial reports a mean reduction in blood pressure of 5 mmHg with a 95 % confidence interval of 2 to 8 mmHg, readers can be 95 % confident the true effect lies within that span. Journalists should report both the point estimate and its interval to convey precision. Practical use includes highlighting when intervals cross a null value (e.g., an odds ratio confidence interval that includes 1), signaling non-significant results. Challenges involve explaining the concept without resorting to technical jargon and avoiding misinterpretation that a confidence interval reflects the probability of the observed result.

Cohort Study – An observational design that follows a group of individuals sharing a characteristic (exposure) over time to assess incidence of outcomes. Related terms: prospective study, longitudinal study, incidence rate. A classic example is the Framingham Heart Study, which tracked participants for decades to identify cardiovascular risk factors. Medical journalists rely on cohort data to illustrate causal pathways, such as “long-term exposure to air pollution increased the risk of asthma by 30 %.” Practical application includes describing follow-up duration, attrition rates, and how exposure was measured. Challenges arise from loss to follow-up, which can introduce bias, and from the need to explain that causality is inferred, not proven, unlike randomized trials.

Cross-Sectional Study – A snapshot design that assesses exposure and outcome simultaneously in a defined population. Related terms: prevalence study, survey, point prevalence. For example, a national health survey may report that 12 % of adults are obese at the time of data collection. Journalists use cross-sectional

findings to highlight current health burdens, but must caution that temporal relationships cannot be established. Practical steps include noting the sampling method and response rate. Challenges include the inability to infer causation and the risk of prevalence-incidence bias, where chronic cases dominate the sample.

Data Censoring – The incomplete observation of a subject's event time, either because the event has not occurred by study end (right censoring) or because the exact timing is unknown (interval censoring). Related terms: survival analysis, Kaplan-Meier estimator, hazard. In oncology trials, patients still alive at the last follow-up are right-censored; their survival time contributes to the analysis up to that point. Journalists should explain that censored data are accounted for in statistical methods, preventing underestimation of survival. Practical application involves reporting median survival with appropriate confidence intervals. Challenges include conveying the concept to a lay audience and ensuring that the censoring mechanism does not differ between treatment groups, which could bias results.

Diagnostic Test – A procedure or measurement used to determine the presence or absence of disease. Related terms: sensitivity, specificity, predictive values. Common examples include PCR for viral detection or mammography for breast cancer screening. Reporting on diagnostic accuracy requires presenting both sensitivity (true-positive rate) and specificity (true-negative rate). For instance, a new COVID-19 antigen test may have 85 % sensitivity and 98 % specificity. Practical use includes discussing how test performance influences public health decisions, such as screening policies. Challenges involve explaining trade-offs: a highly sensitive test reduces missed cases but may increase false positives, affecting patient anxiety and healthcare costs.

Effect Size – A quantitative measure of the magnitude of a phenomenon, independent of sample size. Related terms: Cohen's d, odds ratio, risk difference. Effect size helps readers understand clinical relevance beyond statistical significance. For example, a drug that lowers systolic blood pressure by 2 mmHg may be statistically significant in a large trial, yet the effect size indicates limited clinical benefit. Journalists should contextualize effect size with guidelines or thresholds (e.g., a Cohen's d of 0.2 is considered small). Practical application includes comparing effect sizes across interventions to prioritize public health actions. Challenges include selecting the appropriate metric for the outcome type and avoiding misinterpretation that "large" effect sizes always imply better interventions.

Epidemiology – The study of distribution and determinants of health-related states in populations, and the application of this knowledge to control health problems. Related terms: incidence, prevalence, risk factor. Epidemiological research underpins public health policy, informing decisions such as vaccination campaigns. Journalists often translate epidemiological data into stories about disease trends, using terms like "incidence rates rose by 15 % over the past year." Practical steps include verifying the source, understanding the population studied, and noting any limitations. Challenges include the complexity of measures (age-standardized rates, attributable risk) and the need to avoid overstating causality from observational data.

Hazard Ratio – A measure used in survival analysis to compare the hazard (instantaneous risk) of an event occurring in one group versus another over time. Related terms: Cox proportional hazards model, relative risk, survival curve. A hazard ratio of 0.75 indicates a 25 % reduction in hazard for the treatment group.

Journalists should report the accompanying confidence interval and clarify that the ratio reflects risk over the study period, not a fixed time point. Practical application includes interpreting oncology trial results where median survival may not differ but hazard ratios do. Challenges involve explaining that hazard ratios assume proportional hazards, a condition that may not hold, and that they differ from risk ratios.

Incidence – The number of new cases of a disease occurring in a defined population during a specific time interval. Related terms: incidence rate, cumulative incidence, person-years. Incidence provides insight into disease risk, unlike prevalence, which includes existing cases. For example, an incidence of 10 per 100,000 person-years for Lyme disease indicates the disease's emergence in a region. Journalists use incidence to highlight emerging health threats. Practical steps include stating the denominator and time frame clearly. Challenges arise when incidence is low, leading to wide confidence intervals, and when comparing across populations with differing age structures, requiring standardization.

Kaplan-Meier Curve – A graphical method that estimates the probability of survival over time, accounting for censored observations. Related terms: survival function, log-rank test, median survival. The stepwise curve drops at each event (e.g., death) and remains flat during censored periods. Journalists can illustrate treatment benefits by showing two curves side by side, noting where one remains higher. Practical application includes reporting median survival times with confidence intervals. Challenges include explaining why curves may converge later, the impact of censoring, and the need for statistical tests (log-rank) to assess differences formally.

Likelihood Ratio – The ratio of the probability of a given test result in patients with the disease to the probability in patients without the disease. Related terms: diagnostic odds ratio, post-test probability, Bayes theorem. A positive likelihood ratio (LR+) greater than 10 greatly increases the odds of disease, while a negative likelihood ratio (LR-) less than 0.1 substantially decreases it. Journalists can translate LR values into changes in disease probability for readers, aiding decision-making. Practical use includes discussing how a new screening test improves diagnostic confidence. Challenges involve the need for pre-test probability estimates and the difficulty of conveying ratio concepts without mathematical jargon.

Meta-analysis – A statistical technique that combines results from multiple independent studies to produce a pooled estimate of effect. Related terms: systematic review, heterogeneity, forest plot. By aggregating data, meta-analysis increases power and can resolve conflicting findings. For instance, a meta-analysis of ten randomized trials may show a pooled relative risk of 0.85 for a cardiovascular drug. Journalists should report the overall effect size, confidence interval, and heterogeneity statistic (I^2). Practical application includes informing clinical guidelines. Challenges include dealing with publication bias, varying study quality, and heterogeneity that may limit the validity of pooled estimates.

Odds Ratio – The ratio of the odds of an event occurring in an exposed group to the odds in an unexposed group. Related terms: logistic regression, relative risk, case-control study. An odds ratio of 2 suggests the exposure doubles the odds of the outcome. Odds ratios are common in case-control studies where incidence cannot be directly measured. Journalists must clarify that odds differ from probabilities, especially when the outcome is common, where odds ratios can overstate risk. Practical use includes reporting associations between lifestyle factors and disease. Challenges involve translating odds ratios into intuitive terms for lay readers and noting that they approximate relative risk only when outcomes are rare.

P-value – The probability of observing data as extreme as, or more extreme than, those observed, assuming the null hypothesis is true. Related terms: statistical significance, alpha level, Type I error. A p-value below 0.05 is traditionally deemed "statistically significant," indicating that the observed effect is unlikely due to chance alone. Journalists often highlight p-values, but must also discuss effect size and confidence intervals to avoid overemphasis on binary significance. Practical application includes stating, "The difference was statistically significant ($p = 0.03$)." Challenges include explaining that a p-value does not measure clinical importance, nor does it provide the probability that the hypothesis is true, and that misuse can lead to "p-hacking" concerns.

Peer Review – The process by which experts evaluate a manuscript's methodology, data interpretation, and relevance before publication. Related terms: editorial board, blind review, impact factor. Peer review aims to enhance scientific rigor and credibility. Journalists rely on peer-reviewed studies as trustworthy sources, but must recognize that review is not infallible and that retractions still occur. Practical steps include checking the journal's reputation and noting whether the article underwent peer review. Challenges include the time lag between study completion and publication, potential reviewer bias, and the rise of predatory journals that claim "peer review" without rigorous standards.

Placebo – An inert substance or intervention used as a control to assess the true effect of an active treatment. Related terms: double-blind, randomization, placebo effect. In a randomized controlled trial, participants may receive either the experimental drug or a placebo pill that looks identical. The placebo effect can produce measurable improvements due to expectations alone. Journalists should explain that a drug's benefit is determined by comparing outcomes between the active and placebo groups. Practical application includes reporting the magnitude of improvement beyond placebo. Challenges involve ethical considerations when withholding proven therapy, and explaining to audiences why a "sugar pill" can still yield real-world benefits.

Randomized Controlled Trial (RCT) – An experimental study design where participants are randomly assigned to an intervention or control group, minimizing selection bias. Related terms: allocation concealment, intention-to-treat, blinding. RCTs are the gold standard for evaluating therapeutic efficacy. For example, a double-blind RCT may compare a new antihypertensive drug to a placebo over six months. Journalists should highlight the randomization process, blinding status, and primary outcomes. Practical use includes informing readers about the level of evidence supporting a new treatment. Challenges include reporting complex designs (e.g., crossover or adaptive trials) in an accessible way and recognizing that even well-conducted RCTs may have limited generalizability due to strict inclusion criteria.

Sensitivity – The proportion of true positives correctly identified by a diagnostic test. Related terms: false-negative rate, clinical sensitivity, screening. A test with 95 % sensitivity will detect 95 % of individuals who truly have the disease, missing 5 % (false negatives). Journalists often pair sensitivity with specificity to convey overall test performance. Practical application includes discussing why a highly sensitive test is preferred for ruling out disease ("SnNout"). Challenges arise when sensitivity varies across disease stages, and when audiences misinterpret high sensitivity as meaning the test is always accurate, ignoring false-positive considerations.

Specificity – The proportion of true negatives correctly identified by a diagnostic test. Related terms:

false-positive rate, clinical specificity, confirmatory test. A specificity of 90 % means that 90 % of disease-free individuals are correctly identified, while 10 % are incorrectly labeled as positive. High specificity is valuable for confirming disease (“SpPin”). Journalists must explain that a test with high specificity reduces unnecessary follow-up procedures. Practical use includes reporting the trade-off between sensitivity and specificity in screening programs. Challenges include conveying that no test is perfect, and that specificity can be influenced by disease prevalence and test thresholds.

Statistical Power – The probability that a study will detect a true effect of a specified size, usually set at 80 % or 90 % during sample-size calculation. Related terms: Type II error, beta, sample size. Low power increases the risk of false-negative findings, potentially overlooking beneficial interventions. Journalists should note when a study is “under-powered” and interpret non-significant results cautiously. Practical application includes discussing whether a trial’s sample size was adequate to answer its primary question. Challenges involve explaining abstract concepts like “beta error” and ensuring that readers understand that a non-significant result does not prove absence of effect.

Survival Analysis – A set of statistical methods for analyzing time-to-event data, accounting for censored observations. Related terms: hazard ratio, Kaplan-Meier estimator, Cox model. Survival analysis is essential in oncology, cardiology, and epidemiology where outcomes such as death or disease recurrence occur over varying periods. Journalists may report median survival times or hazard ratios to convey treatment benefits. Practical steps include describing the follow-up duration and the proportion censored. Challenges include the need to explain why traditional averages are inappropriate for time-to-event data and to avoid misinterpretation of survival curves as simple percentages.

Systematic Review – A rigorous, methodical synthesis of all relevant studies on a specific research question, following predefined criteria. Related terms: meta-analysis, PRISMA, evidence hierarchy. Systematic reviews aim to minimize bias by searching multiple databases, assessing study quality, and summarizing findings transparently. Journalists rely on systematic reviews to present balanced overviews, such as “A systematic review of 25 trials found no clear benefit of supplement X.” Practical application includes highlighting the number of studies, total participants, and overall conclusions. Challenges involve explaining heterogeneity among studies, the possibility of publication bias, and the distinction between a systematic review with and without meta-analysis.

Validity – The extent to which a measurement or study accurately reflects the concept it intends to assess. Related terms: internal validity, external validity, construct validity. Internal validity concerns the credibility of causal inferences within the study, while external validity addresses generalizability to broader populations. For example, a tightly controlled RCT may have high internal validity but limited external validity if participants are not representative. Journalists should comment on both aspects when evaluating evidence. Practical steps include noting inclusion criteria, setting, and whether results were replicated. Challenges include balancing technical explanations with readability and recognizing that high validity does not guarantee practical significance.

Variable – Any characteristic, attribute, or measurement that can take on different values among study participants. Related terms: independent variable, dependent variable, covariate. Variables can be categorical (e.g., gender) or continuous (e.g., blood pressure). In statistical models, independent variables

are predictors, while dependent variables are outcomes. Journalists must identify key variables that drive study conclusions, such as “smoking status” as an exposure variable. Practical application includes describing how variables were measured (self-report vs. laboratory). Challenges involve clarifying complex variable coding (e.g., ordinal scales) and ensuring that readers understand the difference between correlation and causation.

Z-score – A standardized value representing the number of standard deviations a data point is from the mean of its distribution. Related terms: standardization, normal distribution, percentile. A z-score of +2 indicates a value two standard deviations above the mean, corresponding roughly to the 97.5th percentile. Medical journalists may use z-scores to contextualize lab results, such as “Your cholesterol is 1.5 z-scores above the age-adjusted average.” Practical use includes comparing results across different tests or populations. Challenges include explaining the concept without heavy statistical language and noting that z-scores assume an underlying normal distribution, which may not hold for all clinical data.

Attributable Risk – The difference in incidence of disease between an exposed group and an unexposed group, reflecting the portion of risk that can be linked to the exposure. Related terms: risk difference, population attributable fraction, exposure. If smokers have an incidence of lung cancer of 30 per 1,000 person-years and non-smokers have 5 per 1,000, the attributable risk is 25 per 1,000. Journalists can use this metric to illustrate public health impact, stating that “smoking accounts for 25 additional cases per 1,000 individuals.” Practical application includes informing policy decisions such as taxation or smoking bans. Challenges include communicating absolute numbers versus percentages and acknowledging that attributable risk does not imply inevitability.

Case-Control Study – An observational design that compares individuals with a disease (cases) to those without (controls) to assess prior exposure differences. Related terms: odds ratio, retrospective study, matching. Cases and controls are often matched on age, sex, or other variables to reduce confounding. For example, a case-control study may find that 40 % of lung cancer cases were exposed to asbestos compared with 10 % of controls, yielding an odds ratio of 5. Journalists should highlight the retrospective nature and potential recall bias. Practical steps include describing how controls were selected and whether exposure assessment was blinded. Challenges involve explaining that causality cannot be firmly established and that selection of appropriate controls is critical.

Confidence Level – The proportion of times that a confidence interval would contain the true parameter if the same study were repeated many times. Related terms: confidence interval, alpha level, coverage probability. A 95 % confidence level means that, in the long run, 95 % of such intervals would capture the true value. Journalists often conflate confidence level with probability, but it reflects the reliability of the interval-building method, not the probability that the specific interval is correct. Practical application includes reporting both the interval and its confidence level. Challenges include clarifying this subtle distinction without confusing readers.

Diagnostic Odds Ratio (DOR) – A single indicator of test performance that combines sensitivity and specificity, calculated as (sensitivity × specificity) / [(1 – sensitivity) × (1 – specificity)]. Related terms: likelihood ratio, ROC curve, test accuracy. A DOR greater than 1 indicates that the test discriminates between disease and non-disease; higher values reflect better performance. Journalists may use DOR to compare competing

diagnostic tools when space is limited. Practical steps include providing the DOR value with its confidence interval. Challenges involve explaining that DOR does not distinguish between false-positive and false-negative errors, and that clinicians may prefer likelihood ratios for decision-making.

Effect Modification – The alteration of the effect of a primary exposure on an outcome by a third variable, indicating that the relationship differs across subgroups. Related terms: interaction, stratified analysis, moderator. For instance, a medication may reduce heart attack risk more in women than in men, suggesting gender as an effect modifier. Journalists should note when studies report subgroup analyses and whether interaction terms were statistically significant. Practical application includes tailoring health messages to specific populations. Challenges include avoiding over-interpretation of subgroup findings, which can be prone to false-positive results if not pre-specified.

Forest Plot – A graphical display used in meta-analysis to show individual study effect estimates and their confidence intervals, alongside the pooled estimate. Related terms: meta-analysis, heterogeneity, weight. Each study appears as a horizontal line with a square representing its weight; the pooled result is depicted as a diamond. Journalists can use forest plots to visually convey consistency across studies. Practical steps include describing the overall effect size, the width of the diamond (confidence), and any notable outliers. Challenges involve explaining technical symbols to non-specialist readers and ensuring that visual complexity does not obscure key messages.

Heterogeneity – The degree of variation in study outcomes beyond what would be expected by chance alone. Related terms: I^2 statistic, random-effects model, subgroup analysis. High heterogeneity (e.g., $I^2 > 75\,\%$) suggests that studies differ in participants, interventions, or methodology. Journalists should report heterogeneity metrics when presenting meta-analysis results, noting whether a random-effects model was used. Practical application includes exploring reasons for variability, such as dosage differences. Challenges include interpreting heterogeneity without oversimplifying, and conveying that substantial heterogeneity may limit the applicability of pooled findings.

Intention-to-Treat (ITT) Analysis – A principle whereby all participants are analyzed in the groups to which they were originally assigned, regardless of adherence or protocol deviations. Related terms: per-protocol analysis, crossover, compliance. ITT preserves the benefits of randomization and provides a conservative estimate of treatment effect. Journalists should highlight when an RCT reports ITT results, as they reflect real-world effectiveness. Practical steps include noting any deviations from ITT and any supplementary per-protocol analyses. Challenges involve explaining why participants who stopped treatment are still counted, and how this approach can dilute observed benefits.

Logistic Regression – A statistical model used to predict a binary outcome (e.g., disease present/absent) from one or more predictor variables. Related terms: odds ratio, multivariate analysis, model fit. Logistic regression yields adjusted odds ratios that account for confounding factors. For example, a study may report that after adjusting for age and BMI, exposure to pollutant X has an odds ratio of 1.8 for asthma. Journalists can convey that the association remains after controlling for other variables. Practical application includes summarizing adjusted results in lay language. Challenges include describing the concept of adjustment without technical jargon and ensuring that model assumptions (e.g., linearity of log odds) are not violated.

Median Survival – The time point at which 50 % of the study cohort has experienced the event of interest (e.g., death). Related terms: Kaplan-Meier, survival curve, censoring. Median survival is often reported when the full distribution cannot be captured due to limited follow-up. For instance, a cancer trial may state a median overall survival of 24 months for the experimental arm versus 18 months for control. Journalists should pair median survival with confidence intervals and note any censoring that may affect interpretation. Practical use includes comparing treatment benefits in a tangible timeframe. Challenges arise when median survival is not reached, requiring alternative reporting methods such as 2-year survival rates.

Meta-regression – An extension of meta-analysis that explores the relationship between study-level characteristics (e.g., dose, year of publication) and effect size. Related terms: moderator analysis, heterogeneity, covariate. Meta-regression can identify sources of variability across studies, such as higher efficacy in younger participants. Journalists may cite meta-regression findings to explain why results differ, noting that these analyses are observational and may be prone to ecological fallacy. Practical steps include describing the predictor variable and its impact on the pooled estimate. Challenges involve the limited statistical power when few studies are available and the potential for over-interpretation of spurious associations.

Negative Predictive Value (NPV) – The proportion of individuals with a negative test result who truly do not have the disease. Related terms: sensitivity, prevalence, false-negative rate. NPV depends heavily on disease prevalence; in low-prevalence settings, even tests with modest sensitivity can achieve high NPV. Journalists can use NPV to reassure readers about the reliability of a negative screening result. Practical application includes explaining that a negative COVID-19 antigen test in a low-risk population has a high NPV, reducing the likelihood of infection. Challenges include conveying that NPV is not a fixed property of the test and that changing prevalence alters its value.

Positive Predictive Value (PPV) – The proportion of individuals with a positive test result who truly have the disease. Related terms: specificity, prevalence, false-positive rate. PPV increases with higher disease prevalence. For a rare disease, even a highly specific test may yield many false positives, lowering PPV. Journalists should contextualize PPV when discussing screening programs, especially in asymptomatic populations. Practical steps include providing both PPV and NPV to give a balanced picture. Challenges involve explaining that PPV is not intrinsic to the test and that it varies across settings, which can be confusing for audiences accustomed to fixed performance metrics.

Publication Bias – The tendency for studies with positive or significant results to be published more frequently than those with null or negative findings. Related terms: funnel plot, selective reporting, gray literature. Publication bias can distort the evidence base, inflating perceived treatment effects. Journalists must be aware of this bias when summarizing research, especially meta-analyses that may not include unpublished data. Practical application includes mentioning whether a review searched trial registries or gray literature. Challenges include detecting bias without access to unpublished studies and communicating its impact without undermining confidence in the entire scientific enterprise.

Random-Effects Model – A statistical approach in meta-analysis that assumes true effect sizes vary between studies, incorporating both within-study and between-study variance. Related terms: heterogeneity, fixed-effects model, DerSimonian-Laird. When heterogeneity is substantial, a random-effects model yields

wider confidence intervals, reflecting uncertainty. Journalists should note when such a model is used, as it influences the pooled estimate's interpretation. Practical steps include reporting the model choice and its rationale. Challenges involve explaining why a more conservative model may produce different conclusions than a fixed-effects approach, and ensuring readers understand that variability among studies is not necessarily a flaw.

Risk Ratio (Relative Risk) – The ratio of the probability of an event occurring in the exposed group to that in the unexposed group. Related terms: incidence, odds ratio, absolute risk reduction. A risk ratio of 0.7 indicates a 30 % risk reduction for the exposed group. Relative risk is intuitive for prospective studies where incidence can be directly measured. Journalists can translate risk ratios into everyday language, such as "the vaccine cuts the chance of infection by nearly one-third." Practical application includes comparing interventions across studies. Challenges arise when the baseline risk is low, where relative reductions may appear large despite small absolute differences.

Standard Deviation (SD) – A measure of the spread of data points around the mean in a distribution. Related terms: variance, normal distribution, dispersion. In a normally distributed set, about 68 % of values fall within ±1 SD of the mean. Journalists often report SD alongside means to indicate variability, e.g., "average blood pressure was 120 mmHg ± 15 mmHg." Practical use includes comparing groups; overlapping SDs suggest similar distributions. Challenges include explaining that SD describes variability but not the range of extreme values, and that for skewed data, median and interquartile range may be more appropriate.

Statistical Significance – The determination that an observed effect is unlikely to have arisen by chance alone, based on a pre-specified alpha threshold (commonly 0.05). Related terms: p-value, confidence interval, Type I error. Statistical significance does not equate to clinical importance. Journalists must balance reporting significance with effect size and real-world relevance. Practical steps include stating both the p-value and the magnitude of the effect. Challenges include preventing readers from assuming that non-significant results mean "no effect," and avoiding over-reliance on the arbitrary p-value cutoff in interpreting findings.

Surrogate Endpoint – A biomarker or intermediate outcome used as a substitute for a clinically meaningful endpoint (e.g., mortality). Related terms: primary endpoint, clinical outcome, validation. Surrogate endpoints, such as cholesterol level for cardiovascular risk, can accelerate trials but may not predict actual health benefits. Journalists should note when studies rely on surrogates and discuss whether the surrogate has been validated. Practical application includes explaining that lowering blood pressure is a surrogate for reducing stroke risk. Challenges involve conveying uncertainty about surrogate-to-outcome translation and avoiding premature endorsement of interventions based solely on surrogate improvements.

Time-to-Event Data – Data that record the duration until an event of interest occurs, incorporating censored observations. Related terms: survival analysis, hazard function, Kaplan-Meier. Time-to-event analysis is essential for outcomes like time to relapse or time to death. Journalists can present results as median time to event or as hazard ratios. Practical steps include clarifying the start point (e.g., randomization) and the event definition. Challenges include describing censoring in an accessible way and ensuring that readers understand that not all participants experience the event within the study period.

Unblinded Study – A trial in which participants, investigators, or both are aware of the treatment assignments. Related terms: open-label, bias, placebo effect. Unblinded designs can introduce performance and detection bias, potentially inflating treatment effects. Journalists should highlight lack of blinding as a limitation, especially when outcomes are subjective (e.g., pain scores). Practical application includes discussing how blinding was or was not feasible (e.g., surgical vs. drug interventions). Challenges involve explaining why some studies cannot be blinded and how researchers attempt to mitigate bias through objective outcome measures.

Variance – The average squared deviation of each data point from the mean, reflecting data dispersion. Related terms: standard deviation, statistical power, heterogeneity. Variance is the square of the standard deviation and is used in many statistical formulas, such as ANOVA. Journalists rarely report variance directly, but understanding it helps interpret confidence intervals and sample-size calculations. Practical steps include noting that higher variance requires larger sample sizes to detect a given effect. Challenges include avoiding confusion between variance and standard deviation, and ensuring that readers recognize variance as a theoretical concept rather than a directly observable metric.