



American Academy of Ophthalmology
P.O. Box 7424
San Francisco, CA 94120-7424
415.561.8500

Copyright © 2013 American Academy of Ophthalmology®. All Rights Reserved.

Portions of this guide have been adapted, with permission, from Alfred Sommer's book, *Epidemiology and Statistics for the Ophthalmologist* (New York: Oxford University Press, 1980).

Contents

- About This Guide 5
- 1. EVIDENCE-BASED MEDICINE 7
 - Defining Evidence-Based Medicine 7
 - Putting EBM into Practice 7
 - Grading Scales of Evidence 7
 - USPSTF Ophthalmology-Specific Recommendations 9
- 2. STATISTICAL CONCEPTS 11
 - Prevalence 11
 - Incidence 11
 - Prevalence vs Incidence 11
 - Cataract Surgery Rate (CSR) 11
 - Relative Risk 12
 - Absolute Risk 12
 - Absolute Risk Reduction (ARR) 12
 - Number Needed to Treat (NNT) 12
 - Odds Ratio 13
 - Sensitivity/Specificity 13
 - Positive and Negative Predictive Value (PPV/NPV) 14
 - Misclassification 14
 - P-values 15
 - Standard Deviation 15
 - Standard Error 16
- 3. EPIDEMIOLOGIC CONCEPTS 17
 - Bias 17
 - Direction of Bias 17
 - Types of Bias 17
 - Confounders 20
 - Validity 21
 - Internal Validity 21
 - External Validity 21
 - Clinical Significance 22
 - Efficacy vs Effectiveness 22
- 4. STUDY DESIGNS 23
 - Case Series 23
 - Cross-Sectional Study (Prevalence Survey) 23
 - Case-Control Study (Retrospective Study) 24
 - Cohort (Longitudinal Study) 24
 - Randomized Controlled Trial (RCT) or Randomized Clinical Trial 26
 - Systemic Review 27
- 5. ADVANCED STATISTICS 28
 - Types of Error 28
 - Confidence Intervals (CI): 28
 - Measurement 29
 - Statistical Tests 30
 - Survival Curves (eg, Kaplan-Meier) 31
 - Statistical Genetics 32
- 6. ADVANCED TOPICS IN EPIDEMIOLOGY 34
 - Ecologic Fallacy 34
 - Atomistic Fallacy 34
 - Missing Data 34
 - Propensity Scores 35
 - Multilevel Modeling 35
- 7. AN INTRODUCTION TO CONDUCTING CLINICAL RESEARCH 37
 - The Research Question 37
 - Study Design 37
 - Study Options 37
 - Nesting 38

Sample Size	38
Statistical Software	39
Data Management	39
Ethics and Institutional Review Broad Approval.....	39
Guiding Principles.....	39
Consent Forms	40
Questionnaires	40
Principles in Analyzing Data	40
Writing and Publishing.....	41
APPENDICES	42
APPENDIX I: OPHTHALMIC SURVEY METHODOLOGY	42
Goal.....	42
Important Terms	42
Interpretation.....	42
Design Issues.....	42
Methodology	43
Examples, Well-Designed Survey Methods	44
Andhra Pradesh Eye Disease Study	44
Baltimore Eye Survey	44
Barbados Eye Study	44
Beaver Dam Eye Study	45
Blue Mountains Eye Study	45
Melbourne Visual Impairment Study	45
Salisbury Eye Evaluation Project	45
APPENDIX II. CASE SERIES, CATEGORIES OF VISION LOSS	46
Case Series.....	46
Examples of Well-Designed Case Series.....	46
Aspects and Ranges of Vision, Vision Loss, Visual Acuity, and Reading Ability.....	47
Ranges of Vision From Normal (<20/25, 1.0) to Total Blindness (NLP, 0.0)	47
Visual Acuity Measurement	48
LogMAR (ETDRS-Type) Chart.....	48
Optotype Charts	48
Selected Articles, Vision Measurement.....	49
APPENDIX III. EPIDEMIOLOGY AND BIostatISTICS RESOURCES.....	50

About This Guide

Keeping up with ophthalmic research is a daunting task. Understanding the conclusion and implications, let alone the research methods and increasingly complex statistics, can be intimidating. Fortunately, however, there are many resources that can help one assimilate and synthesize research into clinical practice. Regardless of your background in epidemiology and statistics, the sections in this guide can help you make use of these resources and understand the science behind them.

Target Audience

- American and international ophthalmologists in practice or training who desire a fundamental understanding of epidemiology, statistics, and interpretation of ophthalmic research literature.

Objectives

- Review key concepts in epidemiology along with their applications to international ophthalmology.
- Describe common types of statistics found in ophthalmic literature.
- Review examples of outstanding clinical research.

Sections in This Guide

1. Evidence-Based Medicine

- Highlights the basics of evidence-based medicine and what it means to practice medicine in this manner.
- Includes a grading scale to highlight different types of evidence and how they should be implemented in practice.

2. Statistical Concepts

- Touches upon basic statistical concepts and definitions that must be mastered before one can appreciate different types of statistical analysis.
- Includes a standard definition for each concept and an ophthalmology related example from peer-reviewed literature.

3. Epidemiologic Concepts

- Highlights how to analyze study results in the context of different study designs.
- Considers limitations in study design, study conduction, and statistical analysis.

4. Study Designs

- Provides information on the different options one has in designing a study.
- Provides the definition, purpose, and potential limitations of the study design for each study type. Includes examples of each study design from peer-reviewed literature focusing on ophthalmology.

5. Advanced Statistics

- Focuses on advanced concepts in statistical analysis that address different types of statistical error and statistical analysis for specific scenarios.
- Explains which type of statistical test one should use given the type of data available.

6. Advanced Topics in Epidemiology

- Covers specific issues that may arise in an epidemiological study.
- Provides insight into dealing with potential fallacies and creating various epidemiological models.
- Addresses the issue of missing data and how to handle it during statistical analysis.

7. An Introduction to Conducting Clinical Research

- Provides a basic outline of what to expect when conducting clinical research.
- Addresses the research question, study design, statistical software, data management, ethics and the Institutional Review Board, data analysis, and data presentation.

Appendices

- Appendix I: Ophthalmic Survey Methodology
- Appendix II: Case Series, Categories of Vision Loss
- Appendix III: Epidemiology and Biostatistics Resources

1. EVIDENCE-BASED MEDICINE

This section highlights the basics of evidence-based medicine and what it means to practice medicine in this manner. A grading scale is included to describe types of evidence and how they should be implemented in practice.

Defining Evidence-Based Medicine

The [Cochrane Collaboration](#), an independent, international organization dedicated to providing accessible and up to date research evidence for clinical decisions, describes evidence-based medicine as follows:

“Evidence-based medicine (EBM) is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence-based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research.”

Putting EBM into Practice

Ophthalmologists have a responsibility to provide the safest and most effective patient care possible. Since the field of ophthalmology is often unfamiliar to patients and other care providers, the ophthalmologist must be prepared to understand the current standard of care as well as explain the EBM behind it. For example, some surgical procedures are offered because they are the only procedure available (ie, extracapsular extraction of a nuclear cataract in a small clinic without a phaco machine and no training in small-incision surgery). However, in another clinic with training in all 3 procedures, the ophthalmologist should know and be prepared to compare and contrast the risks and benefits of each procedure ([Gogate, et al](#)).

The American Academy of Ophthalmology produces an up-to-date set of [Preferred Practice Pattern guidelines](#). Similarly, using EBM the Academy has identified [5 common tests and treatments that ophthalmologists and patients should discuss](#).

Grading Scales of Evidence

Several groups have formalized the process of evaluating clinical recommendations. One is the United States Preventive Services Task Force (USPSTF). In addition to providing recommendation statements, the USPSTF shares the evidence behind each recommendation, thus allowing the clinician to understand and appropriately implement each recommendation. Several of their ophthalmology recommendations follow.

The current system grades the strength of evidence as follows:

- "A" (strongly recommends)
- "B" (recommends)
- "C" (no recommendation for or against)
- "D" (recommends against)
- "I" (insufficient evidence to recommend for or against) (Table 1)

Next, the certainty or confidence of each recommendation is graded separately as low, moderate, or high. See Table 1 and Table 2.

Table 1. What the Grades Mean and Suggestions for Practice		
Grade	Definition	Suggestions for Practice
A	The USPSTF recommends the service. There is high certainty that the net benefit is substantial.	Offer or provide this service.
B	The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial.	Offer or provide this service.
C	<i>Note: The following statement is undergoing revision.</i> Clinicians may provide this service to selected patients depending on individual circumstances. However, for most individuals without signs or symptoms there is likely to be only a small benefit from this service.	Offer or provide this service only if other considerations support the offering or providing the service in an individual patient.
D	The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits.	Discourage the use of this service.
I Statement	The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.	Read the clinical considerations section of USPSTF Recommendation Statement. If the service is offered, patients should understand the uncertainty about the balance of benefits and harms.

Source: [U.S. Preventive Services Task Force Grade Definitions](#). May 2008. Accessed October 10, 2013.

Table 2. Levels of Certainty Regarding Net Benefit	
Level of Certainty*	Description
High	The available evidence usually includes consistent results from well-designed, well-conducted studies in representative primary care populations. These studies assess the effects of the preventive service on health outcomes. This conclusion is therefore unlikely to be strongly affected by the results of future studies.
Moderate	The available evidence is sufficient to determine the effects of the preventive service on health outcomes, but confidence in the estimate is constrained by such factors as: <ul style="list-style-type: none"> • The number, size, or quality of individual studies. • Inconsistency of findings across individual studies. • Limited generalizability of findings to routine primary care practice. • Lack of coherence in the chain of evidence. As more information becomes available, the magnitude or direction of the observed effect could change, and this change may be large enough to alter the conclusion.
Low	The available evidence is insufficient to assess effects on health outcomes. Evidence is insufficient because of: <ul style="list-style-type: none"> • The limited number or size of studies. • Important flaws in study design or methods. • Inconsistency of findings across individual studies. • Gaps in the chain of evidence. • Findings not generalizable to routine primary care practice. • Lack of information on important health outcomes. More information may allow estimation of effects on health outcomes.
<p>Source: U.S. Preventive Services Task Force Grade Definitions. May 2008. Accessed October 10, 2013.</p> <p>* The USPSTF defines certainty as "likelihood that the USPSTF assessment of the net benefit of a preventive service is correct." The net benefit is defined as benefit minus harm of the preventive service as implemented in a general, primary care population. The USPSTF assigns a certainty level based on the nature of the overall evidence available to assess the net benefit of a preventive service.</p>	

Click [here](#) for more articles from the Grade Working Group regarding using the GRADE framework.

USPSTF Ophthalmology-Specific Recommendations

- **Glaucoma:**The USPSTF found insufficient evidence to recommend for or against screening adults for glaucoma in the primary care setting. Grade: I (source: [USPSTF](#)).

- **Vision screening for children and adults:**
 - The USPSTF recommends vision screening for all children at least once between the ages of 3 and 5 years, to detect the presence of amblyopia or its risk factors. Grade: B.
 - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of vision screening for children <3 years of age. Grade: I.
 - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of screening for visual acuity for the improvement of outcomes in older adults. Grade: I.
 - Source: [USPSTF](#)

2. STATISTICAL CONCEPTS

This section introduces basic statistical concepts and definitions that must be mastered before one can appreciate different types of statistical analysis. Each concept includes a standard definition and an ophthalmology-related example from peer-reviewed literature.

Prevalence

- **Definition:** The frequency with which a disease or trait is found in the population under study at one particular point in time.
- **Example:** According to the [World Health Organization](#) there are currently 39 million people blind worldwide.
- **Example:** The prevalence of neovascular AMD and/or geographic atrophy in the US population 40 years and older is 1.75 million or 1.47% (95% confidence interval, 1.38%–.55%) ([JAMA, 2004](#)).

Incidence

- **Definition:** The rate with which new cases of a disease or trait arise over a defined period of time.
- **Example:** A large, multicenter study in the United States found the incidence of retinopathy of prematurity in preterm infants (birth weights <1251g) was 68% ([Pediatrics, 2005](#)).
 - **Note:** Incidence and prevalence can be expressed in absolute numbers, such as 5 new cases per year or a total prevalence of 10 cases, but generally it is preferred that both values be written as rates to ease comparison with the denominator the population under study.
 - **Example:** Two cases of endophthalmitis may be a lot or a little, depending upon whether one talks about a single surgeon or an entire hospital, or over a single day or an entire year.
 - **Example:** The prevalence of blindness in Western Australia was 3384 (95% CI 2947 to 3983) or 0.15% of the population of 2.25 million ([BJO, 2012](#)).

Prevalence vs Incidence

- Incidence directly affects the prevalence; however, the relationship is not always as clear as one would imagine.
- **Example:** To decrease the prevalence of blindness secondary to cataracts in Sub-Saharan Africa, the cataract surgery rate (CSR) must be greater than the annual incidence of cataracts ([Informa Healthcare, 2013](#)).

Cataract Surgery Rate (CSR)

- **Definition:** the number of cataract extractions performed per million people per year in a given location.

- Achieving a CSR equal or greater than the incidence rate of cataracts is necessary to prevent the burden of untreated eye disease from increasing. As the world population ages and a greater percent of the population is over 60, the CSR will have always need to increase.
- References
 - [Authors not listed.] [Vision 2020: the cataract challenge](#). *Community Eye Health*. 2000;13:17–19.
 - West S. [Epidemiology of cataract: accomplishments over 25 years and future directions](#). *Ophthalmic Epidemiol*. 2007;14:173–178. Review.

Relative Risk

- **Definition:** Relative risk = incidence in group 1 /incidence in group 2.
- **Example:** The rate of postoperative endophthalmitis was 3% in patients who didn't receive antibiotics and 0.4% in patients who received gentamycin for a relative risk of $3/0.4 = 8$. Another way of saying this is that gentamycin reduces the risk by 86.5% ($1 - 1/8$) [BJO, 1977](#).

Absolute Risk

- **Definition:** The incidence or the risk of developing a condition when receiving a certain treatment.
- **Importance:** Often in studies only the relative risk is reported, with no clear, absolute risk. So although the study may be statistically significant, it is not practically or clinically significant. For example, a vitamin decreases the risk of getting X ocular disorder by 20%, a seemingly significant amount. But, if the incidence of the disease is only 1 in million per year than the risk is minimal regardless whether the drop is used.

Absolute Risk Reduction (ARR)

- **Definition:** The absolute risk of Y disease/outcome with X treatment minus without X treatment.

Number Needed to Treat (NNT)

- **Definition:** The number of patients who have to be treated with X treatment to prevent one case of Y disease/outcome.
- **Derivation:** $NNT = 1 / ARR$.
- **Example:** In the [Ocular Hypertension Study](#), patients with elevated IOP treated with topical ocular hypotensive drugs had a 4% risk of developing glaucoma while 9% of controls did. This 5% difference (absolute risk reduction) equates to a number needed to treat of 20 ($1/0.05$). Twenty individuals with ocular hypertension need to be treated to prevent one case of glaucoma.

Odds Ratio

- **Definition:** A ratio of the odds Y disease/outcome will occur with X treatment over the odds without X treatment. Odds are a difficult ratio to describe but easy to understand with illustrations.
 - For example, the odds that a single throw of a die will produce a 6 are 1 to 5, or 1/5. The odds is the ratio of the probability that the event of interest occurs to the probability that it does not. This is often estimated by the ratio of the number of times that the event of interest occurs to the number of times that it does not.
 - In a [cross-sectional study of the prevalence of hay fever and eczema in 11-year-old children](#), the probability that a child with eczema will also have hay fever is estimated by the proportion 141/561 (25.1%). The odds are estimated by 141/420. The denominator is the total minus the numerator.
 - Similarly, for children without eczema the probability of having hay fever is estimated by 928/14,453 (6.4%) and the odds is 928/13,525. We can compare the groups in several ways: by the difference between the proportions, $1410/561 - 928/14,453 = 0.187$ (or 18.7 percentage points); the ratio of the proportions, $(141/561)/(928/14,453) = 3.91$ (also called the *relative risk*); or the odds ratio, $(141/420)/(928/13,525) = 4.89$.
- **Example:** Diabetes was not associated with primary open-angle glaucoma (age-race-adjusted odds ratio, 1.03; 95% confidence interval, 0.85, 1.25). This was true for both types of diabetes. Persons whose primary open angle glaucoma had been diagnosed before the examination showed a positive association with diabetes (odds ratio, 1.7, 95% confidence interval, 1.03, 2.86), indicating that selection bias could explain the positive results of previous clinic-based investigations ([Ophthalmology, 2005](#)).

Sensitivity/Specificity

- **Sensitivity Definition:** the proportion of true positives a test is able to detect: $\# \text{ of cases detected} / \# \text{ of true cases}$
- **Specificity Definition:** the proportion of true negatives a test is able to detect: $\# \text{ of controls detected} / \# \text{ of true controls}$
- **Example:** Hollows and Graham reported that in individuals more than 40 years of age with IOP ≥ 21 mmHg on a single reading, only 0.4% had glaucomatous field loss. Second, of the 20 people with field loss, only 13 had an elevated pressure. Prevalence of IOP ≥ 21 mmHG: $397/4231 = 9.4\%$. Prevalence of glaucoma: $20/4231 = 0.4\%$.
 - Sensitivity of tonometric screening: $13/20$ (65%) ([BJO, 1966](#)). One-third of individuals with established field loss screened negative.
 - Specificity of tonometric screening: $(4231 - 397)/(4231 - 20) = 91\%$. 9% of non-glaucomatous eyes would screen positive.
 - Altering the cutoff: sensitivity and specificity can be increased and decreased by moving the cutoff value. For example, if an IOP of 20

mmHg was used instead of 21 mmHg, the sensitivity would increase slightly as a few more individuals with low pressure glaucoma would be detected, but the specificity would decrease as many true negatives (nonglaucomatous eyes) would be reassigned to the "positive group." Conversely, increasing the cutoff would decrease the sensitivity and increase the specificity.

Positive and Negative Predictive Value (PPV/NPV)

- **PPV Definition:** Proportion of positive test results that are **true positives**: = # of true cases detected / # of cases detected
- **NPV Definition:** Proportion of negative tests results that are true negatives: = # of true controls detected / # of controls detected
- **Note:** PPV and NPV are highly dependent upon the prevalence of cases, unlike sensitivity and specificity, which are independent of prevalence.
- **Example:** Using the previous example from sensitivity/specificity, the PPV for detecting a true case of glaucoma among those with a IOP ≥ 21 mmHg is 3.3% (PPV = 13/397). The NPV is 99.8% (NPV = 3834/3841). If the prevalence of glaucoma was greater in general, we would expect the PPV to increase and NPV decrease. Regardless, it is apparent that using a cutoff IOP of ≥ 21 mmHg in this population would result in many false positives and many patients unnecessarily examined for glaucoma.
- **Example:** Digital retinal photographs can be used to diagnose retinopathy of prematurity (ROP). One study evaluated this against indirect ophthalmoscopy in a longitudinal cohort study and found the PPV to 92% and the NPV to be 100%. The sensitivity was 100% and specificity 96% ([Ophthalmology, 2003](#)).

Misclassification

- **Definition:** The inappropriate assignment of cases and controls, or other experimental grouping.
- **Caution:** Too often one mistakes a test with good sensitivity and specificity as properly assigning true cases and controls. However, even "gold standard" measurements can be inaccurate, and a good statistician will adjust for this.
- **Example:** In a large, prospective cohort study on CMV retinitis, the determination of CMV retinitis was made a trained ophthalmologist with remarkably high sensitivity (0.96) and specificity (0.91). However, if the error in misclassification is adjusted for, the rate ratio of CMV retinitis and morality increased 29% from 2.4 to 3.1 ([Ophthalmology, 2005](#)).
- Click [here](#) for a diagram showing the relationship of positive predictive value, negative predictive value, sensitivity, and specificity.

P-values

- **Definition:** P-value is the possibility of obtaining a test statistic value equal or more extreme than the value reported assuming the null hypothesis is true. The null hypothesis is commonly that there is no difference between comparison groups. It is also referred to as the *alpha error* and typically and arbitrarily set at .05 or .01.
- **Example:** $P = .05$ means the likelihood is less than 5 out of 100 that the value observed (such as comparing 2 group means) would be equal or larger than what was observed if there was truly no difference.
- **Example:** To test whether trachoma on a population level could be controlled with mass azithromycin treatment, a community in Tanzania received a dose en masse and tested for recurrence at multiple time points over 2 years. The study found that the overall prevalence of active trachoma was significantly lower at each follow-up point than it had been at baseline ($P < 0.001$ for each comparison). In other words, the chance that they would statistically, by accident, get such an extreme finding that there was no difference in trachoma prevalence between the baseline and at follow-up time points is less than 1 in 1000 ([NEJM, 2004](#)).

Standard Deviation

- **Definition:** standard deviation (SD) (represented by the symbol sigma, σ) is a measure of the variation or "dispersion" from the mean.
- **Interpretation:** The greater the SD, the more data points tend to spread away from the mean. A "normal" bell-shaped distribution as reflected below contains 68% of values within 1 SD, 95% within 2 SD, and 99% within 3 SD.
- Click here for

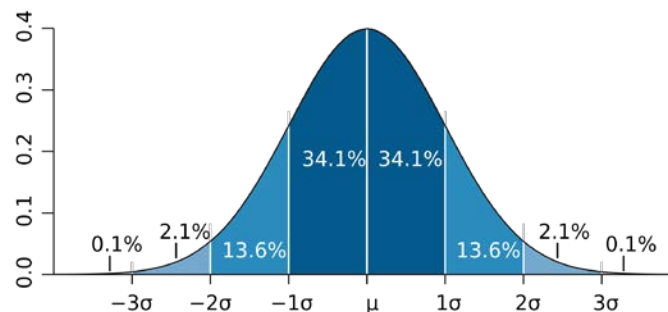


Figure1. Standard deviation. (Courtesy Mwtoews [[CC-BY-2.5](#)], via Wikimedia Commons.)

- **Example:** A possibly standard deviation of IOP among normal eyes is 2. Assuming a mean population pressure of 12, we would expect that 68% of the population to fall within 10–14 mmHg (assuming a normal distribution as shown in the graph above). Within 2 standard deviations (8–16 mmHg) one would expect to find 95% of the population.

Standard Error

- **Definition:** The standard error (SE) is the standard deviation of the sampling distribution of a statistic:

$$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

Where:

S = SD

n = number of observations in the sample

- **Interpretation:** The standard error is commonly confused with standard deviation, which is understandable because they are directly related. However, while the standard deviation describes the variation in values in a sample, the standard error describes the variation in the population mean depending upon the sample of the population chosen.
- **Example:** The mean IOP for the first examination of 100 eyes might be 16.5. The second series of measurements on the same eyes might have a mean of 17.4, and the third 15.9. The standard error of the mean (SE) would describe the variation of these individual study means (16.5, 17.4, and 15.9) about the mean for all 3 studies together (16.6). It should be intuitively obvious that the variation of sample or study means about the "true" mean (ie, the mean of all the samples or studies together) will be smaller than the variation of individual measurements about a single study's mean. Since each study's mean has already averaged out the extreme highs and lows of the individual measurements within that study, the study means are less likely to vary by a large amount from one another, or the overall mean. As with the SD, the overall sample mean ± 1 SE will include two-thirds of all individual sample means, and the overall sample mean ± 2 SE will include 95%.
- Confidence intervals (see "Confidence Intervals" in Section 5, "Advanced Statistics") are determined by the mean (\bar{x}) and the standard error (SE):

$$\text{Upper/lower 95\% Limit} = \bar{x} \pm (SE * 1.96)$$

For more information, see "Confidence Intervals" in Section 5, "Advanced Statistics."

3. EPIDEMIOLOGIC CONCEPTS

This section highlights how to analyze study results in the context of different study designs. Limitations in study design, study conduction, and statistical analysis are also considered.

In an ideal world, study results would always reflect the true nature of things and be easily applicable to clinical medicine. Unfortunately, there are many reasons why one must be cautious in interpretation of study results. Considerations include limitations in study design, conduction, and analysis both for the sample of people the study was conducted in, the greater population it reflects, and the potential extrapolation to different populations. The accuracy of a study depends on the degree of error that enters into the study.

Bias

- **Definition:** Any systemic error in the design, conduct or analysis of a study that influences the association of an exposure on the outcome.
- **Importance:** Selection bias can produce an artificial association when none actually exists or hide one that truly does
- **Note on designing studies:** Minimizing bias is a crucial step in the design of a study and is often more difficult than a researcher appreciates. Consultation of an epidemiologist is often necessary and always recommended. For an excellent analysis of the effects of bias and how to measure the impact on study findings, see [Intl J Epi, 1996](#).

Direction of Bias

- **Positive bias:** Overestimate true value of association between exposure and outcome.
- **Negative bias:** Underestimate true value of association between exposure and outcome.

Types of Bias

Selection bias

- **Definition:** The selection of subjects between groups or overall differs (systematically) from an ideal sample of the source population (not an accurate sample)
- **Note:** Selection bias can be minimized with random sampling and with larger same size. It may also affect only one group within a study, but still influence the overall conclusion.
- **Note:** Selection bias affects internal validity of a study, the ability to draw true inferences about the sample selected. Generalizability of the study conclusions to the larger population, or external validity is a separate concept.

- **Example:** Selection of only men into a study would limit the ability of a study to be generalized to a population of men and women, but it would not influence the internal validity about whether a drug improves AMD outcomes in a sample of men. On the other hand, the assignment of more men than women to the experimental arm in a co-ed study would introduce potential bias into the study because gender may influence outcome independently, and thus internal validity of the study would be affected. Hence, when selecting and assigning subjects it is important to keep in mind the ultimate goal of the research.
- Types of selection bias:
 - Nonresponse bias
 - a. Demographic, socioeconomic, cultural, vocational, and health can all influence response or participation in studies.
 - b. Control or adjust for this by comparing information that is known about nonresponders to responders.
 - Loss to follow-up (dropouts) bias
 - a. When dropouts are not random and sample at end of study is characteristically different from sample that started
 - b. **Red flags:** High dropouts; no mention in analysis of how dropouts were treated statistically. (See “Missing Data” in Section 6 for appropriate methods of handling this common but complicated issue.)
 - Berkson’s bias
 - a. Case-control study in a hospital setting. Patients with 2 medical disorders are more likely to be hospitalized than those with only one. Therefore, an association of the 2 diseases can be found when one doesn’t really exist in other words study exposure and outcome increase probability of selection (being in hospital).
 - Neyman bias (prevalence-incidence bias)
 - a. When one wants to study incident cases but ends up studying prevalent cases because incident cases of short duration (eg, asymptomatic, clinical resolution, fatality) are excluded.
 - Membership bias (health worker effect)
 - a. Study participants are healthier than the general population on average because they are healthy enough to participate (eg complete paperwork, attend clinic visits), while the very sick are unable to participate. Similarly, occupation can affect participation.
 - Immortal time bias
 - a. **Definition:** A study design problem that overestimates exposure time by counting time when the outcome of interest cannot have occurred (or the participant would not be included in the study).

Information bias

- Misclassification
 - Differential
 - a. **Definition:** Probability of misclassification is different between study groups; this can easily happen if we pay more attention to the exposed than the controls.
 - b. **Direction of bias:** Positive, negative, or inverse.
 - c. **Example:** Ocular hypertensive patients are more likely to be incorrectly diagnosed with glaucoma than non-hypertensive patients because they possibly receive more frequent eye care.
 - Nondifferential
 - a. **Definition:** Probability of misclassification is equivalent between groups. It's similar to random error (precision) because it occurs when one is not perfect at classifying by group or outcome.
 - b. **Direction of bias:** Generally negative (weakens association between exposure and outcome).
 - c. **Example:** The tonometer occasionally returns an incorrect value.
 - Specific types of misclassification bias (can be both differential and nondifferential):
 - a. **Recall bias (ruminant bias):** Cases are more likely to remember an exposure or falsely draw an association between an exposure and an outcome (disease).
 - b. **Reporting bias:** Social undesirability (eg, illicit drug use) or desirability (eg, medical drug compliance)
 - Preventing information bias:
 - a. Mask subjects and researchers, standardize data collection with non-subjective measurements, take multiple measurements, and verify subjective responses in subset of sample by objective measurements (eg, a diabetes study asking about diet and blood sugar control could measure A1c in a subset of participants to verify whether their subjective responses correlated with a clinical value).

Analytical bias

- A third type of bias ("fishing") that can enter in the analytical stage if the analyst is aware of a desired study outcome and multiple tests are performed looking for a particular statistical relationship without adjusting for the additional tests run.
 - **Example:** Bonferroni adjustment for multiple tests

Publication bias

- A fourth type of bias happens at the publication phase. Often studies of positive results are preferentially published or accepted to higher impact journals than studies with negative results. However, both

could be of the same research strength and hence we will find in the literature a tendency for stronger associations than what probably exist in the real world. For instance, if 100 studies examining drug X for treatment of age-related macular generation (AMD) were performed in the past decade, those that show a more exciting effect of the drug on treating AMD are more likely to be published than those with equivocal results regardless of study design or sample size.

Red flags: bias in scientific papers

- **Selection bias:** No mention of inclusion criteria and how participants were selected, especially if procedure different between study groups.
- **Information bias:** No mention of how data was collected; were researchers masked to study groups during data collection? Was there a standardized collection procedure? Was it the same between groups?
- **Analytical bias:** Were the primary hypothesis and statistical methods declared before analysis was begun?
- **Publication bias:** Was the study funded or performed by a party of interest? Is it possible similar studies with negative results have not been published?

Confounders

- **Definition:** A characteristic that meets the following criteria:
 - Associated with exposure
 - Cause (or surrogate) of disease
 - Not affected by exposure
- Evaluating confounders
 - Limiting the impact of confounders is incredibly important in study design and analysis and the reader should be aware of the techniques used and evaluate whether the researchers truly controlled for confounding.
 - Similar to bias, a confounder can have a positive influence (stronger association than what is actually present in sample population) or negative influence (weakens association).
- Handling confounders
 - Study design techniques: randomization, restriction, matching
 - a. **Randomization:** Ideal; has ability to eliminate confounding both known and unknown variables. However, randomization is not perfect, especially in smaller studies.
 - b. **Restriction:** Where study populations are limited. If gender will influence outcomes in a study, was it restricted to women?
 - c. **Matching:** If including both sexes where group assignments paired by sex?
 - Matching on multiple characteristics can limit study group sizes and analytical power. One way of handling this is by

- matching subjects with different characteristics but similar risk profiles by a propensity score. (See “Propensity Scores” in Section 6, “Advanced Topics in Epidemiology.”)
- Frequency matching: Guarantees similar frequency of matched variables in study and comparison groups.
 - Overmatching or improper matching: Can induce confounding where it didn’t exist.
- Study analysis techniques: stratification, multivariable analysis
 - **Stratification:** Results are calculated and displayed by strata (eg by gender, if one assumes gender may influence the relationship between the exposure and the outcome).
 - **Multivariate analysis:** A statistical technique that can control for the influence of multiple variables on each other. (See “Statistical Tests” in Section 5, “Advanced Statistics.”)
 - **Residual confounding:** Unknown confounders or mistreatment of known confounders that leaves misclassified or opens backdoor pathways.
 - **Misclassification:** IOP categories should be continuous or proxy variables; sunlight exposure: time outside isn’t the same, doesn’t account for intensity of sun, time of day, or use of protective clothing.
 - **Note:** Mediators are distinct from confounders because they are affected by exposure and on the causal pathway between exposure and disease.
 - a. In other words, they are intermediate variables, translating at least part of the effect of exposure on disease. While it is important to control for confounders, controlling for mediators is only necessary if one is interested in the direct effect of an exposure on disease outside of any indirect effect through a mediator.

Validity

Internal Validity

- **Definition:** The extent to which study results truly reflect the relationship of the exposure and the outcome in the source population.

External Validity

- **Definition:** External validity, often called *the generalizability of study*, is the extent to which association of exposure and outcome in the study population can be extrapolated to other populations different in demographic makeup, health, socioeconomic status, historical time, or other characteristic. The external validity is affected both by the precision of measurement and accuracy in relation to the true value.
- **Example:** IOP management in adults with elevated IOP decreases risk of developing glaucoma. Different population could be children, or something less obvious such as patients without elevated IOP.

Clinical Significance

- **Definition:** The extent to which study findings impact clinical practice. Often even well-designed studies, with minimal bias and confounding, and good internal and external validity, may have results that are statistically significant, but not clinically significant. For instance, a large study that finds drug X decreases IOP 1mmHg more than drug Y may have a low p-value, but has little clinical impact because drug decisions are based on a myriad of factors including cost and side effects.
- **Example:** The Age-Related Eye Disease Study was an 11-center, double-masked random clinical trial that enrolled 3640 participants to investigate age-related macular degeneration (AMD).
 - Researchers found that patients at high risk for developing advanced AMD (Categories 3 and 4) reduced their risk of developing advanced stages of AMD by 25% when treated with the combination of antioxidants and zinc.
 - However, there was no statistically significant benefit in many patients with early AMD (Category 2) ([Arch Ophthalmol, 2001](#)).
 - Clinicians thus had to decide how to integrate these findings into their practice.
 - If supplements were given to all Americans with intermediate and advanced (8 million), 300,000 (95% confidence interval, 158,000-487,000) of them would avoid advanced AMD and any associated vision loss during the next 5 years ([Arch Ophthalmol, 2003](#)). But, at the same time, they may be increasing their risk of cancer and other medical conditions by taking vitamins.
- Association vs causation
 - Determining the true cause of an event is one of the most difficult aspects of epidemiology. For instance, establishing conclusively that cigarette smoke causes lung cancer took over 40 years since the original case-control study was published. Differentiating exposures that are associated with the outcome versus causing the outcome can be difficult.
 - Epidemiologists are particularly hesitant to make early conclusions as they too often see researchers drawing inferences that are not supported by their data and the media generalizing results beyond the scope of the study. One set of criteria commonly used to judge causation is Hill's criteria for causation: temporal relationship, strength of association, dose-response relationship, consistency, plausibility, consideration of alternative explanations, reversibility, specificity, and coherence.

Efficacy vs Effectiveness

- Efficacy is the effect of a drug in the ideal study setting, while effectiveness is the effect in the real world and is evaluated in a pragmatic trial.

4. STUDY DESIGNS

This section provides information on the different options one has in designing a study. Each study type reviewed includes the definition, purpose, and potential limitations of the study design. Also included are examples of each study design from peer-reviewed literature focusing on ophthalmology.

Case Series

- **Definition:** A publication of several interesting, similar clinical cases that serve to educate, but not provide statistical testing.
 - Single case reports or case-series can act as an important early step in the process of generating hypotheses in clinical research.
 - Uncontrolled case series are one of the most common types of studies reported in ophthalmic literature.
- Problems associated with uncontrolled case series:
 - No estimate of effectiveness (in terms of effect size) can be determined from a case series, and no explicit mechanisms for controlling bias exist.
 - Uncontrolled case series are useful only when an effect is dramatic.
- For more information regarding the appropriate use of uncontrolled case series in ophthalmology, see Kempen JH. Appropriate use and reporting of uncontrolled case series in the medical literature ([Am J Ophthalmol, 2001](#)).

Cross-Sectional Study (Prevalence Survey)

- **Definition:** The recording of observations on a population, or a subset sample, at one point in time. By definition they are descriptive studies that are not longitudinal and do not directly compare 2 treatments.
- **Purpose:** Often performed as initial study to determine burden of disease (prevalence) and associated factors with outcome for more thorough analysis by a case-control, cohort, or randomized control study.
- **Example:** The relationship between central corneal thickness (CCT) measured by ultrasonic pachymetry and intraocular pressure was evaluated by a single ophthalmologist ([Ophthalmology, 1999](#)). Of patients attending a general ophthalmic clinic, mean central corneal thickness (CCT) was 553.9 μm (95% confidence intervals 549.0–558.8 μm) in the clinically normal eyes, 550.1 μm (95% CI, 546.6–553.7 μm) in the eyes with primary open-angle glaucoma, 514.0 μm (95% CI, 504.8–523.3 μm) in the eyes with normal-tension glaucoma, eyes, 530.7 μm (95% CI, 511.2–550.1 μm) in the eyes with pseudoexfoliative glaucoma, 559.9 μm (95% CI, 546.8–573.0 μm) in the eyes with chronic angle-closure glaucoma, and 579.5 μm (95% CI, 574.8–584.1 μm) in the glaucoma-suspect eyes. These differences of mean CCT are statistically not equivalent ($P < 0.001$ analysis of variance).
- Cluster sampling

- **Definition:** A method of dividing a population into "natural" groups to improve the facility and ease of surveying. These natural groups (eg, towns, schools, classrooms) are then randomly selected and all or a random sample of the individuals in the cluster surveyed.
- **Example:** Refractive error study in children: results from Shunyi District, China ([AJO, 1999](#)).
- For analysis of clustered data, see "Multilevel Modeling" in Section 6, "Advanced Topics in Epidemiology."

Case-Control Study (Retrospective Study)

- **Definition:** An observational study comparing 2 already existing groups (2 different outcomes) to identify associations that could have contributed to the outcome of interest in either a risk or protective manner. These studies often require minimal resources and are easier to conduct than cohort or randomized controlled trials.
 - **Cases:** Individuals with the outcome of interest (eg cataract).
 - **Controls:** Individuals matched on criteria to closely resemble the cases. Common variables matched on: age, sex, race
- **Example:** The Lens Opacities Case-Control Study evaluated risk factors for age-related cataracts among 1380 ophthalmology outpatients, by comparison of the following groups: posterior subcapsular only, 72 patients; nuclear only, 137 patients; cortical only, 290 patients; mixed cataract, 446 patients; and controls, 435 patients. Factors associated with an increased risk of cataracts included low education (odds ratio [OR]= 1.46) and diabetes (OR = 1.56), while the regular use of multivitamin supplements was associated with a decreased risk (OR =0.63) of cataract ([Arch Ophthalmol, 1991](#)).

Cohort (Longitudinal Study)

- **Definition:** A longitudinal, observational study that follows a group of individuals (cohort) measuring the occurrence of various exposures and specified outcome(s). Cohort studies can be both prospective and retrospective. Unlike a case-control study where individuals are matched, a cohort is a usually random selection of people from a specific group.
- **Example:** An extensively studied cohort is the Nurses' Health Study, which began in 1976 and has followed more than 100,000 female nurses for the occurrence of cancer and cardiovascular disease. Other famous examples include the Framingham Heart Study, a group of roughly 5,000 individuals in Framingham, Massachusetts, United States, about whom more than 1,000 papers have been published, and the EPIC cancer and nutrition study that spanned 10 European countries and 500,000 participants. Cohort studies of this size often have multiple substudies of case-cohort and case-control nested inside of them.
- Prospective cohort study

- **Other terms:** Concurrent cohort study or concurrent prospective study.
- **Definition:** A cohort study that begins in present time and follows individuals into the future
- **Classic prospective cohort studies:** Framingham Heart Study, Nurses Study, EPIC
- **Example:** A prospective, cohort study was designed to determine the 5-year outcomes of patients with cytomegalovirus (CMV) retinitis and AIDS in the era of highly active antiretroviral therapy (HAART); 503 patients were followed every 3 months for 5 years. Overall mortality was 9.8 deaths/100 person-years (PY). The rate of retinitis progression was 7.0/100 PY and the rate of retinal detachment was 2.3/100 eye-years (EY) ([Ophthalmology, 2010](#)).
- **Example:** Predictors of Long-term Progression in the Early Manifest Glaucoma Trial followed more than 250 individuals every 3 months for up to 11 years ([Ophthalmology, 2007](#)).
- **Example:** Fifteen-Year Cumulative Incidence of Age-Related Macular Degeneration: [The Beaver Dam Eye Study](#)
- **Example:** Risk Factors for Incident Open-angle Glaucoma: [The Barbados Eye Studies](#)
- Retrospective cohort study
 - Other terms: Nonconcurrent prospective study, historical cohort study.
 - **Definition:** A cohort study using already collected data and selecting a group to follow longitudinally through in the past-time. Just like a prospective study, both multiple exposure and outcomes can be measured.
 - **Drawbacks:** A temporal relationship can be tougher to establish between exposure and outcome, because the two may not be measured regularly or appropriately. Other drawbacks include selection bias, because the available cohort may not be the ideal study population, and measurement error (such as misclassification error) if the classification protocol changes during the study period. Environmental and temporal differences between historic populations can limit application to present day populations.
 - **Example:** A study on pediatric glaucoma outcomes was designed as a retrospective medical record review of 20 years of pediatric glaucoma patients at a single hospital in Quebec. A total of 163 patients (254 eyes) were found. Average follow-up was roughly 8 years. During this time, 113 (69.3%) of children had at least 1 surgical procedure ([Can J Ophthalmol, 2009](#)).
- Retrospective review of prospective cohort
 - Temporal division is often blurred, in which case it is best to think that both retrospective or prospective studies are longitudinal studies. What is important is whether the cohort was established to evaluate the reported exposure-outcome association. Prospective cohorts with one question can sometimes be examined retrospectively with a different question.

- **Example:** To estimate the rate of visual field progression in individuals with open-angle glaucoma (OAG) researchers retrospectively reviewed the data from 9 prospective cohort studies. According to the study, “The rate of progression was the mean of all subjects' damage in the worse eye divided by an average time since onset. The mean duration of disease was lowest among Chinese persons at 10.5 years (95% CI: 8.8-12.6) and was highest in African-derived subjects at 15.4 years (95% CI: 14.6-15.9). By combining disease duration and progression rate, the model predicted that 15% or fewer of the worse eyes would reach the end of the field damage scale in the patient's lifetime” ([IOVS, 2008](#)).

Randomized Controlled Trial (RCT) or Randomized Clinical Trial

- **Definition:** An experimental (not observational) prospective study testing a hypothesis by randomly assigning participants to different study arms.
- A double-masked (both participants and researcher) RCT is the gold-standard for clinical research.
- Ophthalmologists prefer the term *masked* over *blinded*.
- **Example:** Ranibizumab versus verteporfin for neovascular age-related macular degeneration. In this 2-year, multicenter, double-masked study, patients were randomly assigned in a 1:1:1 ratio to receive either monthly intravitreal injections of ranibizumab (0.3 mg or 0.5 mg) plus sham verteporfin therapy or monthly sham injections plus active verteporfin therapy. The primary endpoint was the proportion of patients losing fewer than 15 letters from baseline visual acuity at 12 months. Of the 423 patients enrolled, 94.3% of those given 0.3 mg of ranibizumab and 96.4% of those given 0.5 mg lost fewer than 15 letters, as compared with 64.3% of those in the verteporfin group ($P < 0.001$ for each comparison). Thus, the authors concluded ranibizumab was superior to verteporfin as intravitreal treatment of predominantly classic neovascular AMD ([NEJM, 2006](#)).
- **Red flag:** Studies that do not include intention to treat (ITT) analysis. Many study participants may be randomized to one study drug or arm, but end up in another for a variety of reasons. There are several ways of handling this change in classification. The most conservative way is ITT analysis that treats each person in analysis as if the person had stayed in the group to which he or she was originally assigned.

Systemic Review

- **Definition:** A methodological review and summary of relevant clinical research. While a single study may through random chance come to a false conclusion (eg, Type 1 or Type II error), by combining multiple studies and effectively increasing sample size, researchers statistically decrease such a chance and come closer to the truth. In the analysis, one assumes that the individual studies are samples from a single, larger population. Hence, it is important that in the selection of studies, one ensures that the study “populations” resemble each other. Well-defined selection criteria are critical.
- **Red flag:** If systematic review or meta-analysis study selection methods section is not clear enough to allow the reader to personally repeat.
 - A systematic review that incorporates advanced statistical technique for combining study results is called a *meta-analysis*.
 - **Example:** ROP - high or low oxygen saturation and severe retinopathy of prematurity: a meta-analysis. This was a systematic review and meta-analysis to find the association between severe ROP incidence of premature infants with high or low target oxygen saturation measured by pulse oximetry. Ten publications were included, using a random-effects model of analysis. Low oxygen saturation (70%–96%) in the first several postnatal weeks was associated with a reduced risk of severe ROP (0.48, 95% CI:0.31–0.75). High oxygen saturation (94%–99%) at ≥32 weeks' postmenstrual age was associated with a decreased risk for progression to severe ROP (0.54, 95% CI: 0.35–0.82) ([Pediatrics, 2010](#)).
 - **Example:** Efficacy and cost of ambulatory cataract surgery: a systemic review. This review found 10 observational studies and 5 randomized clinical trials that met selection criteria. It compared inpatient to ambulatory cataract surgery and found that over the 15 studies, there was no difference in vision, but an increased odds of elevated IOP (OR = 2.3, 95% CI: 1.3–3.9) ([Med Clin \(Barc\), 2000](#)).

5. ADVANCED STATISTICS

This section focuses on advanced concepts in statistical analysis that address different types of statistical error and statistical analysis for specific scenarios. This section also explains which type of statistical test to use given the type of data available.

Types of Error

- **Type I:** The possibility of incorrectly concluding that there is a significant difference between 2 groups when one really doesn't exist (in statistical jargon, it is the probability of erroneously rejecting the null hypothesis).
 - The alpha value is the probability of a Type I error. Hence, when we decide that the cutoff point or alpha level (α) for a statistical test is 0.05, we understand there is a 5% chance of Type I error.
 - The p-value of a statistical test is then compared against the alpha value to determine statistical significance.
- **Type II:** The risk of not reporting a difference when one truly exists, a false negative, (in statistical jargon, failing to reject a truly false null hypothesis).
 - The amount of risk is the beta level (β).
 - The power of a study is the chance of detecting a difference if one truly exists and equal $1-\beta$.
 - Similar to the common cutoff of a p-value of .05, studies are typically designed to have at least 80% power or a beta level of 0.2.

Confidence Intervals (CI):

- **Definition:** An estimated range that contains with a defined probability the true value (in statistical jargon, "We're 95% confident that the interval covers the truth.")
 - 95% is generally used because it is equivalent of ± 2 standard deviations of the mean.
 - a. A 99% confidence interval is ± 3 standard deviations of the mean.
 - b. The alpha-level and CI are related in that: $1 - \alpha = CI$
 - If the confidence intervals of 2 outcomes do not overlap, they are considered statistically different.
 - a. CI significance and p-value significance generally go together.
 - **Example:** The mean serum vitamin A level of 268 randomly sampled children was 20 ug/ml. The S.E. of this mean was 0.5. The 95% confidence limits are $20.0 \pm 2(.5)$, or 19.0 and 21.0. The interval between 19.0 and 21.0 has a 95% chance of including the true mean serum vitamin A level of 6000 children. A 99% confidence interval would be 18.5 to 21.5 ug/ml. A 99% CI in comparison to a 95% CI increases the confidence interval contains the true value, but lowers the precision of the estimate.

Measurement

- Measurement error
 - **Definition:** the incorrect determination or classification of a value
 - Why worry about measurement error?
 - a. Inaccurate values, small or large, and whether due to problems in accuracy or precision, on a global level or within a particular group, can greatly influence statistical tests and statistical significance.
 - b. **Example:** Is CMV retinitis associated with mortality?
 - The diagnosis of CMV retinitis is clinically made with high sensitivity and specificity:
 - Sensitivity (Se) = 0.96
 - Specificity (Sp) = 0.91
 - Nonetheless, there is still some misclassification error as 4% of positive cases are missed and 9% of negative cases are not rejected. This small difference can have large effect on statistical tests.
 - In the study, the observed rate ratio association between CMV retinitis and mortality was 2.40. If, however, the error was adjusted for using Greenland's formula, the association would be 3.10, an increase of 29%.
 - References:
 - a. Jabs DA, et al. [Risk factors for mortality in patients with AIDS](#). *Ophthalmology*. 2005;112:771–779.
 - b. Greenland S. [Basic methods for sensitivity analysis of biases](#). *Int J Epidemiol*. 1996;25:1107–1116.
 - c. Sechrest L. [Validity of measures is no simple matter](#). *Health Serv Res*. 2005;40:1584–1604.
- Measurement agreement
 - Whether it be multiple measurements with the same device, different devices, or multiple measurers, the agreement in these measurements is often important to measure. Several terms are of importance:
 - a. Receiver operating characteristic (ROC) curve: a graphical representation of discrimination - on the Y-axis sensitivity and on the X-axis 1-specificity
 - b. Intra-rater reliability
 - **Definition:** The precision of repeated measurements made by a single observer.
 - c. Inter-rater reliability
 - **Definition:** The precision of single or multiple measurements made by multiple observers.
 - d. Kappa statistic
 - **Definition:** The agreement between 2 observers beyond what would be expected from chance.
 - **Equation:** $K = (\text{Observed \%} - \text{Expected \%}) / (100\% - \text{Expected Agreement \%})$.

- e. Coefficient of variation
 - Definition: A measurement of dispersion relative to the mean.
 - **Equation:** $CV = \text{Std Dev} / \text{Mean}$
 - **Example:** Macular thickness measurements in healthy eyes using 6 different optical coherence tomography instruments ([IOVS, 2009](#)).
- f. Reliability
 - **Definition:** The consistency of a measure.
 - **Equation:**
subject variation / total variation = (1 – measurement error) / total variation
 - Strength: $\leq .8$ = weak; .81-.9 = moderate; and $> .9$ = strong.

Statistical Tests

- A well-written introduction to choosing a statistical test with a decision-tree flowchart and glossary of terms can be found [here](#).
- To choose the appropriate statistical test, you first have to know the type of data you are working with:
 - **Nominal:** Categories that have no quantitative relationship to each other (eg, brown, black, blue)
 - a. **Dichotomous:** is a subtype (eg yes, no)
 - **Ordinal:** Categories that can be ranked (eg, category of vision impairment)
 - **Continuous:** Numerical (eg, intraocular pressure)
- Second, you must determine whether data points are paired or independent. For example, the same eye measured before and after surgery would be paired or IOP readings in both eyes from the same person would be paired.
 - **Example:** paired measurements with continuous data requires a Paired-Sample t-test and paired measurements with categorical data requires a McNemar test
 - **Red flag:** Clinical trials involving paired-eyes should always address the related-nature of the data, however a 2012 study (Karakosta et al) found only 26% of studies involving 2 eyes mentioned the possible correlation and only 7% included the appropriate statistical tests ([Am J Ophthalmol, 2012](#)).
 - **Advanced:** Analysis of multiple paired data points over time is commonly performed with a generalized estimating equation (GEE) or mixed-effects regression modeling techniques.
 - a. **Example:** Longitudinal analysis of IOP in both eyes.
- Knowing the type of data and whether it is paired are first steps in choosing the appropriate statistical test, summarized in the following table.

First Variable	Second Variable	Example	Appropriate Test
Continuous	Continuous	IOP and age	Pearson correlation coefficient (r)
	Dichotomous	IOP and gender	Student's t-test
	Dichotomous paired	IOP and eye (left vs right of same patient)	Paired t-test
	Nominal	IOP and eye color	AVOVA
Dichotomous	Dichotomous	Cataract and gender	Chi-square test (Fischer exact test if sample sizes are small)
	Dichotomous paired	Cataract and (left vs right of same patient)	McNemar Chi-square test
Ordinal	Dichotomous	Category of vision impairment and cataract	Mann-Whitney U Test
	Dichotomous paired	Category of vision impairment and (left vs right of same patient)	Wilcoxon matched-pairs rank test
Nominal	Nominal	Eye color and race	Chi-square test

Survival Curves (eg, Kaplan-Meier)

- **Definition:** a graphical representation of the percent survival (survival being the absence of the outcome of interest) over time.
- Survival curves are a function of the survival rate at each time interval and the number of observations at each time. For example, if in between time points 3 study participants drop out of the study, their absence will be reflected at the following time point by a decreased denominator in the equation ($\#$ surviving / $\#$ at time point).
- Types:
 - Grouped – uses time bins
 - a. log-linear model (Poisson regression)
 - Ungrouped: Uses exact time.
 - a. Kaplan-Meier estimates
 - b. Cox regression model
 - **Example:** [Clinical Evaluation and Risk Factors of Time to Failure of Ahmed Glaucoma Valve Implant in Pediatric Survival Analysis](#)

Statistical Genetics

- **Importance:** The study of genetics has become increasingly complex with advancements in genotyping and new understandings of the regulation of gene expression. Ophthalmology is at the front of this field. While a review of ophthalmic genetics is beyond the scope of this section, presented here a few of the more common genetic analysis and statistics.
- Concepts:
 - Mendelian inheritance
 - a. **Law 1:** Every individual possess 2 alleles of every gene, one of which is randomly passed on to offspring.
 - b. **Law 2:** Separate genes are passed down independently of each other (Mendelian trait).
 - This law can generally be considered true for genes located on separate chromosomes or opposite ends of the same chromosome. However, for genes located nearer each other, the principle of genetic linkage holds that the closer genes are located to each other on a chromosome the more likely that they will be inherited together. Thus, genes adjacent to each other are almost always passed on together and are considered “genetically linked.”
 - **Mendelian trait:** A trait that is controlled by a single gene and shows a simple Mendelian inheritance pattern. Examples include Marfan syndrome (autosomal dominant), sickle-cell anemia (autosomal recessive), and xeroderma pigmentosa (autosomal recessive).
 - a. **Multifactorial and polygenic disorders:** These depend on the expression of multiple genes and environmental influences. These are the most common and most difficult to study, like glaucoma and age-related macular degeneration.
- Basic study designs
 - **Familial aggregation:** Identifies families with a high incidence of the disorder of interest and tracks inheritance patterns; best for Mendelian traits.
 - Twin concordance: Compares allelic expression in identical or fraternal twins and is helpful in distinguishing the influence of genetics over environment
- Advanced study designs
 - **Genetic linkage studies:** Map disease susceptibility loci to intervals that are several megabases long (a range including hundreds of genes).
 - a. **LOD score (logarithm (base 10)):** Compares the likelihood of the test result if 2 loci are indeed linked, to the likelihood of observing the same result by chance. A LOD score $> +3$ is generally considered significant.
 - **Genome-wide association study (GWAS):** Survey common genetic variants across the genome; this extensive of a study has

only recently become feasible due to advances in genotyping technology.

- a. **Single nucleotide polymorphisms (SNPs):** Minor differences in the nucleotide sequence of a gene that are often conserved through generations and populations and can be mapped as markers of disease.
 - **Example:** Age-related macular degeneration. Genetic variants near CFH and ARMS2/HTRA1 are the strongest genetic contributors to AMD susceptibility ([Proc Natl Acad Sci U S A, 2005](#)).
 - **Example:** Myopia. Genetic linkage study of high-grade myopia in a Hutterite population from South Dakota ([Molecular Vision, 2007](#)).
 - **Example:** Myopia. Genome-wide association analysis identified a novel susceptible locus for pathological myopia at 11q24.1. GWAS-based case/control association analysis using 411,777 markers with 830 Japanese patients and 1,911 Japanese controls (297 cases and 934 controls in the first stage, and 533 cases and 977 controls in the second stage). ([PLOS Genetics, 2009](#)).

For more information on advanced statistics, see the external page, ["Application of Advanced Statistics in Ophthalmology."](#)

6. ADVANCED TOPICS IN EPIDEMIOLOGY

This section covers specific issues that may arise in an epidemiological study. Insights are offered on dealing with potential fallacies and creating various epidemiological models. This section also addresses the issue of missing data and its treatment during statistical analysis.

Ecologic Fallacy

- **Definition:** Drawing inferences about individual-level associations based on group-level data.
- **Example:** Populations with more exposure to the sun are at a greater risk of developing cataract; hence, if individuals limit exposure to the sun, they will lessen their risk of developing cataracts.

Atomistic Fallacy

- **Definition:** Drawing inferences about group-level association from individual-level data.
- **Example:** Individuals with narrow angles are at an increased risk of glaucoma; hence, populations with narrow angles have a higher prevalence of glaucoma.

Missing Data

- **Question:** What to do when the data for a participant is incomplete or completely missing? How does a statistical program handle missing data?
 - **Examples of missing data:** Nonparticipation, loss to follow-up, error in data entry, participant refusal to answer question.
- **Significance:** How missing data is handled can change statistical significance and the strength of association between an exposure and an outcome.
- Statistical methods of dealing with missing data
 - List-wise deletion
 - a. **Definition:** Removal from analysis of any participant with any missing data. This is the most extreme form of excluding missing data.
 - Model-wise deletion
 - a. **Definition:** Removal from analysis of any participant with missing data points in only the variables in the analysis of interest
 - b. **Pros:** Simple; technique commonly applied by default in statistical programs.
 - c. **Cons:** Reduced sample size (reduced power and precision; introduces selection bias as the missing data may not be at

- random; eg, participants who drop out may be characteristically different from those who remain in the study).
- Multiple imputation (MI)
 - a. **Definition:** A statistical method that fills in missing data in a dataset by prediction using existing data in the dataset. By imputing (predicting) multiple times, a mean and confidence interval can be reported.
 - b. MI is gaining popularity in large data analysis.
 - c. For more information: [Review: A gentle introduction to imputation of missing values.](#)
 - Inverse probability weighting (IPW)
 - a. **Definition:** a statistical method of reweighting variables of a study population to balance losses of missing data due to selection bias, confounding, and explicit missing data.
 - **Red flag:** A study that does not describe how missing data is treated.

Propensity Scores

- **Definition:** A statistical method of pooling patients with different exposures (risks for an outcome of interest) into groups of similar overall risk profiles to allow for a more robust analysis. The score is the probability that a participant would have a certain exposure given the value of measured confounders.
- **Use:** Alternative to regression and stratification for handling of selection bias, particularly useful when more is known about the exposures than the outcome.
- Examples:
 - [Use of propensity scores in ophthalmology.](#)
 - [Cataract - AREDS Report 21: A Propensity Score Approach.](#)
 - [Use in analyzing large databases.](#)

Multilevel Modeling

- **Definition:** An epidemiological concept and category of statistical analysis when data are layered and there is non-independence of observations.
- **Synonyms:** Multilevel model, mixed model, random effects model.
- **Concept:** Multiple measurements in the same person over time would be related and more dependent upon each other than a single measurement in multiple people over time. Similarly, members of the same household or community or treated by the same doctor may be more alike than a random sampling of individuals.
 - In ophthalmology, not only are longitudinal measurements common, but there are often 2 measurements (2 eyes) per participant. These measurements are separate, but related. Similarly, people in a group, such as children in a classroom, school, or community, are related. When one adds in 1 eye per person with multiple measurements over time and includes other

groupings of data (like geographical area, hospital floor, or family unit) in analysis there are multiple levels of data to adjust for. If such observations are not treated as partially related, but instead completely at random from a population, then the statistical precision (confidence intervals) would be overestimated.

- Example: [Eye drops](#).

7. AN INTRODUCTION TO CONDUCTING CLINICAL RESEARCH

This section gives a basic outline of what to expect when conducting clinical research. It also addresses the research question, study design, statistical software, data management, ethics and the Internal Review Board, data analysis, and data presentation.

The Research Question

- **Question:** What are you trying to prove?
- **Principles of a good research question:** Is it feasible, interesting, novel, ethical, and relevant?
- **Specificity:** Is the question specific? An achievable narrow question is much more rewarding than an unachievable broad question.

Study Design

- **Considerations:** Evaluate not only the purpose of the study, but also the resources available, access to study populations, duration of the study, and how relevance of the study question may change over time or in different populations

Study Options

- Observational
 - Single time point: Cross-sectional (single and multi-site).
 - Longitudinal:
 - a. Case-control: Better than cohort study when want to study multiple exposures, or the outcome of interest is rare (eg, endophthalmitis), generally requires fewer resources and is quicker to perform than a cohort study or RCT.
 - b. Cohort:
 - Better than case-control when want to study multiple outcomes; can perform prospectively or retrospectively and can nest multiple studies inside if cohort is large.
 - Fixed vs dynamic study population: Do subjects enter the study at any time or do they all start at a single time point?
 - c. Case-control vs cohort?
 - Rare outcome? Case-control.
 - Limited resources? Case-control or retrospective cohort.
 - Direct estimate of population prevalence? Cohort.
 - Multiple outcomes? Cohort.
 - Multiple exposures? Case-control.
- Experimental
 - RCT
 - Nonrandomized clinical trial: Generally considered inferior to RCTs and are used out of ease or allowance of participant to choose treatment.

- Single- vs multisite
 - If multisite analysis, should include multilevel modeling. (See “Multilevel Modeling” in Section 6, “Advanced Topics in Epidemiology.”)
- Frequency of follow-up
 - How often are subjects measured for exposures or outcomes? Can exposure status change over time or is it fixed?

Nesting

- Embedding a case-control or cohort (case-cohort) study into an ongoing or historic cohort study can save time and money
 - Example of recommended study guidelines: [Glaucoma](#).

Sample Size

- Principles:
 - Larger the sample size, greater the study power.
 - Smaller the sample size, easier the study to conduct.
 - Plan ahead.
 - Dropouts happen.
- Estimating sample size
 - Depends on the study design and analytical method
 - Simple example for comparison of 2 groups with a binary outcome:

$$n = \frac{(z_{\alpha} + z_{\beta})^2(p_1q_1 + p_2q_2)}{(p_2 - p_1)^2}$$

Where:

n = sample size

z_{α} = 1.96 for an alpha error of .05 (two-tailed)

z_{β} = 0.84 for a beta error of 0.2 (one-tailed)

p_1 = estimate of the proportion of group 1 that will experience the outcome

p_2 = estimate of the proportion of group 2 that will experience the outcome

$q_1 = 1 - p_1$

$q_2 = 1 - p_2$

- Estimating “ p ” or the effect size can be done by consulting already published values from similar studies, performing a pilot study, or choosing a value that would be clinically meaningful.
- **Example:** If we are interested in whether panretinal photocoagulation can reduce the rate of neovascularization by 50% in patients with diabetes and we know that 10% of patients with diabetes go blind within a year, we can estimate the study sample size by: $n = (1.96 + 0.84)^2 \times (.10 \times .90 + .05 \times .95) / (.05 -$

$.10)^2 = 431$. Assuming a loss of follow-up of around 20%, the required sample size would be 500–550 per group (treated and untreated)

- Techniques for reducing sample size:
 - Continuous variables
 - More precise measurements
 - Paired measurements
 - Unequal group sizes
 - a. Multiple controls per case (if cases are difficult to come by)
- Use common outcomes.

Statistical Software

Multiple statistical packages exist, ranging from free to thousands of dollars a year. Some of the most popular include:

- Statistical Analysis System (SAS)
- Stata (StataCorp)
- IBM SPSS Statistics (IBM)
- R (free)
- OpenEpi (free)
- Epi Info (Centers for Disease Control; free)

Data Management

Various free and costly programs exist for open-access and secure data management across single or multiple users and computers. Examples include Microsoft Excel, Microsoft Access, and RedHat.

Ethics and Institutional Review Broad Approval

Guiding Principles

- Respect for persons
 - The right of autonomy and protection of individuals who cannot practice this right: Informed consent.
- Beneficence
 - Risk of research: Be acceptable in light of benefit.
- Justice
 - Risks of research: Do not be disproportionately spread, especially on vulnerable populations (children, seniors, prisoners, people with limited access to health care or impaired decision-making capacity)
 - Benefits of research: Do not be disproportionately spread so that the access to trial drugs or free treatment or the application of research findings may be only for certain populations or groups.
 - Children and ethnic minorities are underrepresented in clinical research.
 - The application of these principles in the form of study design, populations studied, consent forms, confidential data management,

and disclosure are the responsibility of the author and are reviewed by an Internal Review Board (IRB).

- For general research ethics training, see the [National Institutes of Health online training course](#).

Consent Forms

- Informed consent is fundamental research ethics principal; however, the form is often considered a formality with little thought put into the actual design besides meeting the requirements of the IRB.
- Writing and administering a consent form that is understandable to the participant is a moral responsibility of the researcher.
- Consulting a more extensive resource on this topic is recommended, especially in the international setting.

Questionnaires

- The importance of forethought and trial in creation of a questionnaire cannot be underestimated.
- Creating a well-designed questionnaire is beyond the scope of this section, but underlying principles include the following.
 - Construct simple, straightforward questions that avoid ambiguous language and have mutually exclusive answers.
 - Consult local community members to ensure questions and answer choices are understandable and relevant.
 - Realize that the formatting, question order, answer order, presentation, and administration of the survey will influence responses.

Principles in Analyzing Data

- Employ a priori hypothesis and analysis methods.
- Keep records of all data manipulations and store an unedited raw data file.
- Establish before looking at the data the analysis methods and statistics that will be used (don't "go fishing.").
- When multiple tests are performed, the chance of a type I error (concluding a significant result when one doesn't actually exist) increases. For example, with a p-value of 0.05, if we test whether ten different variables are related to the outcome there is a 40% that we will erroneously conclude that one or more are statistically significant when in reality there is no significance. One method of adjusting for this is the Bonferroni correction.

Writing and Publishing

- The art of this craft comes easier to some, but fortunately everyone improves with practice and experience.
- For detailed instructions and advice on writing and publishing specific to ophthalmology, refer to the following series of articles:
 - Ophthalmology and vision science research (J Cataract Refract Surg)
 - a. Part 1: [Understanding and using journal impact factors and citation indices](#)
 - b. Part 2: [How to commence research – Eureka or that’s a little unusual?](#)
 - c. Part 3: [Avoiding writer's block—understanding the ABCs of a good research paper](#);
 - d. Part 4: [Avoiding rejection—structuring a research paper from introduction to references](#)
 - e. Part 5: [Surfing or sieving: using literature databases wisely](#)
 - [Clinical Research: A Primer for Ophthalmologists](#), by Alfred Sommer.
- Finally, take heart in knowing that even monumental papers sometimes go decades without being noticed ([Sleeping Beauties in Ophthalmology](#)).

APPENDICES

APPENDIX I: OPHTHALMIC SURVEY METHODOLOGY

Goal

- To describe the various methods of collecting epidemiologic data on ophthalmic diseases using survey methods that will provide accurate estimates in order to:
 - Define the prevalence, severity, incidence, and progression of eye diseases
 - Investigate associated risk factors
 - Plan appropriate eye health care services
 - Develop prevention and rehabilitation programs
 - Project costs of these programs

Important Terms

- **Survey:** An observational epidemiologic study based on examination of all persons in a given population or a specifically defined subgroup of the population
- **Prevalence:** An estimate by a cross-sectional study that assess the presence of disease in a specific population at a point in time.
- **Incidence:** An estimate by follow-up of a population cohort to assess the probability of developing a disease over a period of time.

Interpretation

- Observational studies can provide unique information regarding prevalence, incidence, and the impact of eye disease in the general population.
- Controlled clinical trial interventions may have a different impact on the incidence or progression of a disease than they would in the general population of persons with this condition.
 - Clinical trials and observational incidence studies complement each other in providing information to assess the potential public health impact of intervention.

Design Issues

A critical step in conducting a survey is to obtain a representative sample of the population to be studied. Nearly complete participation of the cohort members is important to ensure the validity of the estimates obtained and governs the ability to generalize findings beyond the population study. The study design must attempt to minimize the influence of biases that may result from the methods used to measure endpoints and risk factors. Several potential biases of surveys include the following:

- **Nonresponsive bias:** If participants are not representative of the proposed study population due to selective nonparticipation.
 - This bias can be minimized if attempts are made to ensure the participation of all those eligible.
- **Recall bias:** When participants with a disease are questioned about historical information, as they may tend to remember and report such information more readily than persons without the condition.
- **Observation bias:** When the examiner is not masked as to a specific condition that may affect the measurement of an endpoint or a risk factor.
 - Observation bias can be minimized by evaluating intra- and inter-observer reproducibility and to conduct rigorous training, pretesting, and quality assurance of measurements throughout the study.
- **Sampling bias:** When observations and conclusions based on a sample of people are generalized to dissimilar groups.
 - Worldwide comparisons within and among surveys provide important information, but they depend on using standardized protocols that are comparable and observations with internal and external validity.
 - a. Validity refers to whether a clinical observation reflects the true state of the parameter being measured. For the observation to be valid, it must be neither biased nor incorrect due to chance.

Methodology

- Ethical Standards and Informed Consent
 - Informed consent must be voluntarily given and obtained prior to participating in a study, and participants must have the right to withdraw from the study at any time.
 - a. Guidelines on the general conduct of biomedical research are published in the [World Medical Association Declaration of Helsinki](#) and in *Ethics and Epidemiology: International Guidelines* ([website](#)).
- Sample selection
 - Observational population-based studies are “representative” in that persons are selected from a sampling frame according to specific procedures because they reside in the population and meet the criteria for the study.
 - a. The eligibility criteria for the study must be defined prior to recruitment and examination, and must be adhered to strictly.
 - b. A large sample size is necessary if the observations in a survey are to provide meaningful conclusions.
- Identification of the population sample
 - A *sampling frame* is required to identify eligible persons in the population.
 - a. Some studies have used a household census conducted by trained interviewers to identify eligible residents.

- b. Other studies have successfully used large population databases to identify the population.
- Participation
 - Participation rates of at least 80% should be achieved whenever possible, as high participation rates reduce selection bias.
 - Biases may also result if responders differ from non-responders in an observational study.
 - a. The initial visit for a household census provides the opportunity to encourage participation, either by scheduling appointments for further examination at a convenient site or, when appropriate, scheduling home examination visits.
- Data collection
 - The study should establish a detailed protocol for examination, measurements, disease gradings, and disease classification.
 - All protocols should be detailed in a manual of procedures, and examiners should be standardized in all measurements prior to the beginning of the study.

Examples, Well-Designed Survey Methods

Andhra Pradesh Eye Disease Study

Dandona I, Dandona R, Naduvilath TJ, et al. [Burden of moderate visual impairment in an urban population in southern India](#). *Ophthalmology*. 1999;106:497–504.

Dandona R, Dandona L, Naduvilath TJ, Nanda A, McCarty CA. [Design of a population-based study of visual impairment in India: The Andhra Pradesh Eye Disease Study](#). *Indian J Ophthalmol*. 1997;45: 251–257.

Baltimore Eye Survey

Katz J, Tielsch JM, Quigley HA, Javitt J, Witt K, Sommer A. [Automated suprathreshold screening for glaucoma: the Baltimore Eye Survey](#). *Invest Ophthalmol Vis Sci*. 1993;34:3271–3277.

Tielsch JM, Sommer A, Witt K, Katz J, Royall RM. [Blindness and visual impairment in an American urban population](#). The Baltimore Eye Survey. *Arch Ophthalmol*. 1990;108:286–290.

Barbados Eye Study

Leske MC, Connell AM, Schachat AP, Hyman L. [The Barbados Eye Study. Prevalence of open angle glaucoma](#). *Arch Ophthalmol*. 1994;112:821–829.

Leske MC, Connell AM, Wu SY, et al. [Incidence of open-angle glaucoma](#). *Arch Ophthalmol*. 2001;119:1310–1315.

Beaver Dam Eye Study

Klein R, Klein BE, Lee KE. [Changes in visual acuity in a population](#). The Beaver Dam Eye Study. *Ophthalmology*. 1996;103:1169–1178.

Klein R, Klein BE, Linton KL, De Mets DL. [The Beaver Dam Eye Study: visual acuity](#). *Ophthalmology*. 1991;98:1310–1315.

Blue Mountains Eye Study

Attebo K, Mitchell P, Smith W. [Visual acuity and the causes of visual loss in Australia. The Blue Mountains Eye Study](#). *Ophthalmology*. 1996;103:357–364.

Melbourne Visual Impairment Study

Livingston PM, Guest CS, Bateman A, Woodcock N, Taylor HR. [Cost-effectiveness of recruitment methods in a population-based epidemiological study: the Melbourne Visual Impairment Project](#). *Aust J Publ Health*. 1994;18: 314–318.

VanNewkirk MR, Weih L, McCarty CA, Taylor HR. [Cause-specific prevalence of bilateral visual impairment in Victoria, Australia: The Visual Impairment Project](#). *Ophthalmology*. 2001;108:960–967.

Weih LM, VanNewkirk MR, McCarty CA, Taylor HR. [Age-specific causes of bilateral visual impairment](#). *Arch Ophthalmol*. 2000;118:264–269.

Salisbury Eye Evaluation Project

Muñoz B, West S, Rubin GS, Schein OD, Fried LP, Bandeen-Roche K. [Who participates in population based studies of visual impairment? The Salisbury Eye Evaluation Project experience](#). *Ann Epidemiol*. 1999;9:53–59.

Review

Review of 30 years of international studies, found a paucity studies from Central Africa, Central and Eastern Europe, the Caribbean, and Latin America, and among children or regarding near vision and mild distance VI. Bourne R, Price H, Taylor H, et al. [New systematic review methodology for visual impairment and blindness for the 2010 Global Burden of Disease Study](#). *Ophthalmic Epidemiol*. 2013;20:33–39.

APPENDIX II. CASE SERIES, CATEGORIES OF VISION LOSS

Case Series

- **Definition:** Initial small-scale investigations of new treatments acting as a pilot study for a larger trial.
 - Hypotheses in clinical epidemiology derive from clinical observations, and single case or a series of cases can act as an important early stop in the process of investigating new treatments.
 - a. Kempen JH. [Appropriate use and reporting of uncontrolled case series in the medical literature](#). *Am J Ophthalmol*. 2011;151:7–10.
- Uncontrolled case series are one of the most common types of studies reported in ophthalmic literature.
 - Such series are commonly found throughout the surgically dominated specialties, and are normally very specific.
- Benefits of uncontrolled case series:
 - In the past, presenting a successful series of a new surgical technique was thought sufficient evidence of both its safety and effectiveness.
- Problems associated with uncontrolled case series:
 - No estimate of effectiveness (in terms of effect size) can be determined from a case series, and no explicit mechanisms for controlling bias exist.
 - Only when an effect is so dramatic and obvious can this sort of evidence be admissible for evidence-based practice.

Examples of Well-Designed Case Series

Bonini S, Bonini S, Lambiase A, et al. [Vernal keratoconjunctivitis revisited: a case series of 195 patients with long-term followup](#). *Ophthalmology*. 2000;107:1157–1163.

Chan WM, Lam DS, Lai TY, et al. [Photodynamic therapy with verteporfin for symptomatic polypoidal choroidal vasculopathy: one-year results of a prospective case series](#). *Ophthalmology*. 2004;111:1576–1184.

Ciulla TA, Starr MB, Masket S. [Bacterial endophthalmitis prophylaxis for cataract surgery: an evidence-based update](#). *Ophthalmology*. 2002;109:13–24.

Jaissle GB, Leitritz M, Gelisken F, et al. [One-year results after intravitreal bevacizumab therapy for macular edema secondary to branch retinal vein occlusion](#). *Graefes Arch Clin Exp Ophthalmol*. 2009;247:27–33.

Ryan SJ, Maumenee AE. [Birdshot retinochoroidopathy](#). *Am J Ophthalmol*. 1980;89:31–45.

Stahl A, Agostini H, Hansen LL, Feltgen N. [Bevacizumab in retinal vein occlusion-results of a prospective case series.](#) *Graefes Arch Clin Exp Ophthalmol.* 2007;245:1429–36.

Aspects and Ranges of Vision, Vision Loss, Visual Acuity, and Reading Ability

Ranges of Vision From Normal (<20/25, 1.0) to Total Blindness (NLP, 0.0)

Normal Vision

- The standard for visual acuity (20/20, 1.0) is defined as the ability to recognize a standard letter (1 M-unit) at a standard distance (1 meter).
- In normal reading, newsprint (about 1 M) is read at about 40 cm.
- Normal adult vision ranges from 20/12.5 (1.6) to 20/25 (0.8).

Minimal and Mild Vision Loss

- The minimal and mild vision loss range is a transitional range between normal vision and more pronounced vision loss, classified as “Low Vision”.
- Print may be held slightly closer, but general reading performance is still adequate.
 - Minimal Vision Loss
 - a. Minimal vision loss ranges from 20/32 (0.63) to 20/40 (0.5).
 - Mild Vision Loss
 - a. Many functional criteria fall within this transitional range (whether for a driver’s license or for cataract surgery).
 - Cataract surgeons may consider “less than 20/40” an indication for surgery.
 - Refractive surgeons may consider “20/40 or better” a satisfactory outcome.
 - b. Mild vision loss ranges from 20/50 (0.4) to 20/63 (0.32).

Moderate Vision Loss

- In the moderate vision loss range, support from large-print materials or moderate power magnifiers is required to enhance available vision.
- In the U.S., children in this range qualify for special-education assistance.
- Moderate vision loss ranges from 20/80 (0.25) to 20/160 (0.125).

Severe Vision Loss

- When visual acuity drops to 20/200 (0.1) or below, reading performance is compromised.
 - Reading distance for newsprint becomes 10 cm or less, which is possible with appropriate magnifiers.

- With severe vision loss, reading endurance is limited and reading speed is reduced because of the small field of strong magnifiers.
- In the U.S., individuals in this range are considered “legally blind” and qualify for a tax break and for disability benefits.
- Severe vision Loss ranges from 20/200 (0.1) to 20/400 (0.05).

Profound Vision Loss

- In the profound vision loss range (<20/400, <0.05) visual reading becomes marginal and recreational reading is extremely difficult.
 - In Europe, many benefits do not start until this level is reached.
 - The WHO includes this range in its “blindness” category.
- In rehabilitation, the emphasis shifts from vision enhancement aids to aids that substitute the use of other senses.
- Profound vision loss ranges from 20/500 (0.04) to 20/1000 (0.02).

Near-Blindness and Total Blindness

- Visual reading is no longer possible and talking books, Braille, or other non-visual sources must be relied upon.
 - Near-total vision loss
 - a. In the near-blindness range (<20/1000, CF 1 m), vision becomes unreliable and use of vision substitution skills becomes dominant.
 - Total blindness
 - a. Use of vision substitution skills is the only option.

Visual Acuity Measurement

LogMAR (ETDRS-Type) Chart

- Visual acuity should always be tested with the LogMAR (ETDRS) chart if it is available and the patient is literate.
- The LogMAR chart is the standard for visual acuity measurement, consisting of letters.
 - The LogMar chart is sometimes referred to as the Snellen Chart, which is actually an older and less accurate way to measure visual acuity.
 - Bailey and Lovie updated the Snellen chart by creating standard spacing, an equal number of letters on each line, and a geometric progression of letter sizes.

Optotype Charts

- Definition: The term *optotype* refers to symbols that are specifically designed for visual acuity measurement.
- Several types of Optotype Charts exist
 - Landolt Cs

- a. Landolt Cs (or Landolt broken rings) are the standard against which the recognizability of other optotypes is calibrated.
- b. The Landolt Broken rings chart consists of the letter "c" facing left, right, up, and down.
- Numbers
 - a. Used for illiterate adults that can understand numbers.
- Lea Symbols
 - a. Lea symbols consist of 4 symbols (house, square, apple, circle).
 - b. Lea symbols give the most reliable results for young children and in developmentally delayed individuals.
- HOTV test
 - a. The HOTV test also uses 4 symbols with the letters H, O, T, and V.
- Tumbling Es
 - a. The tumbling E chart consists of the letter "E" facing left, right, up, and down.
- Pictures
 - a. Pictures are hard to standardize and should only be used as a last resort if none of the other tests are available or workable.

The Foundation of the American Academy of Ophthalmology offers a ["home eye test"](#) that may help in discovering a vision problem that requires professional attention. It is not a substitute for a complete medical eye examination.

Selected Articles, Vision Measurement

Chaplin PK, Bradford GE. [A historical review of distance vision screening eye charts: what to toss, what to keep, and what to replace](#). *NASN Sch Nurse*. 2011;26(4): 221-228

Hered RW, Murphy S, Clancy M. [Comparison of the HOTV and Lea Symbols chart for preschool vision screening](#). *J Pediatr Ophthalmol Strabismus*. 1997;31(1):24-28.

APPENDIX III. EPIDEMIOLOGY AND BIOSTATISTICS RESOURCES

- Johns Hopkins Open Courseware:
 - [Epidemiology](#)
 - [Biostatistics](#)
- Open Courseware Consortium:
 - [Epidemiology](#) (multiple languages)
- University of Alabama:
 - [Biostatistics](#)
- University of Pittsburg:
 - [Epidemiology](#)