The Laryngoscope
Lippincott Williams & Wilkins, Inc., Philadelphia

Tutorials in Clinical Research: Part IV: Recognizing and Controlling Bias

James M. Hartman, MD; James W. Forsen, Jr., MD; Mark S. Wallace, MD; J. Gail Neely, MD, FACS

Objective: **This is the fourth of a series of Tutorials in Clinical Research.[1–3] The objectives of this article are to heighten reader awareness of biases and of methods to reduce their impact and to provide an easy reference document for the reader during future journal reading.** ***Study Design:*** **Tutorial.** ***Methods:*** **The authors met weekly for 4 months discussing clinical research articles and biases for which they might be at risk. Liberal use of reference texts and specific articles on bias were reviewed. Like the example by Sackett, biases were catalogued to create an easily understood reference. Articles were chosen to demonstrate how understanding bias might facilitate assessment of the validity of medical publications.** ***Results:*** **The article is organized into three main sections. The first section introduces specific biases. Two tables serve as rapid reference tools. The second section describes the most common biases linked to specific research approaches and reviews techniques to minimize them. The last section demonstrates the application of the information to an article in a manner that can be applied to any article.** ***Conclusions:*** **Assessing the validity of a medical publication requires an awareness of bias for which the research is inherently at risk. A review of the publication to determine what steps the authors did or did not undertake to minimize the impact of biases on their results and conclusions helps establish the validity. This article should be of assistance in this critical review task.** ***Key Words:*** **Bias, confounding factors, research/methods, research design, clinical protocols.**

Laryngoscope, **112:23–31, 2002**

INTRODUCTION

Bias means oblique or slanted.[4] Bias refers to the unintentionally random or systematic, or worse, the willful, distortion of truth. In comparative research, each component of the research model (baseline, maneuver, and outcome) must be similar or similarly measured, except for the variable being compared. If the components are not similar, the comparison is slanted, or biased.[5]

Additionally, biases sum to compound the distortion of truth as the research moves along the pathway from the baseline state to ultimate publication and its reading.[6]

Practicing medicine leads to more questions than one can answer; but the important questions must be pursued with efficiency. Understanding what types of studies are required to answer specific questions and how to search for key papers are great places to start. However, to avoid erroneous conclusions, the papers found must be evaluated for validity.

A rapid means of assessing validity begins by anticipating specific biases that might influence the results and by determining if the authors acknowledged and attempted to control these biases in the design, conduct, and analysis of their work. The significance of bias lies in its erosion of validity of data and conclusions; in turn this promotes falsehoods as reality, which if applied to patients, may inadvertently endanger them.

Each of the major parts of an investigation are at risk of bias, including selection of subjects, performance of the maneuver, measurement of the outcome, data collection, data analysis, data interpretation, and even reporting the findings.

The purpose of this article is to assist the busy practitioner in identifying common types of biases and the methods used to minimize them.

BIASES COMMON IN CLINICAL RESEARCH

It is worth emphasizing that the control of bias is a lifelong, progressive endeavor and that the literature describing methods to minimize biases abounds.[5] However, a brief awareness of specific biases and counter measures to control for them is helpful.

Many types of biases have been described and several names may label the same type of bias among different authors. To facilitate understanding of the different types of biases, it is helpful to progress through the research process (literature review, baseline state, maneuver, out-

From the Clinical Research Working Group, Department of Otolaryngology–Head and Neck Surgery, Washington University School of Medicine, St. Louis, Missouri, U.S.A.

Editor's Note: This Manuscript was accepted for publication August 15, 2001.

Send Correspondence to J. Gail Neely, MD, FACS, Professor and Director, Otology, Neurotology, Base of Skull Surgery, Director of Research, Department of Otolaryngology–Head and Neck Surgery, Washington University School of Medicine, 660 South Euclid Avenue, Box 8115, St. Louis, MO 63110, U.S.A. E-mail: jgneely@aol.com

come, analysis, publication); each component is susceptible to the inadvertent influence of biases. The classic paper by Sackett was used as a model for the following section.[7]

Publication, Literature Review, or Researching the Topic Biases

Reading the literature about a topic may promote an opinion that has arisen under the influence of several possible biases. With a *bias of rhetoric*, the conclusions may be based on opinion rather than evidence. A *one-sided reference bias* may occur when an author restricts his references to papers that support his opinions, thus the reader ends up with a skewed understanding of the topic. A *positive results bias* slants opinions because authors and editors are less likely to present or publish negative results (Table IA).

These are just some of the ways readers may obtain a slanted opinion about a specific diagnostic procedure, treatment, or other component in medical practice.

Selection or Susceptibility Biases

Selection biases are also called susceptibility biases. These occur when the groups to be compared are differentially susceptible to the outcome of interest, even before the experimental maneuver is performed. The resultant dissimilarities in the baseline attributes may influence the outcome independent of the experimental maneuver. For example, *popularity bias, centripetal bias*, and *referral bias* may skew the sample population in certain practices or centers so that these subjects might be different from the general population seen in other offices.

Selection biases are perhaps one of the most important groups of biases. Multiple potential biases may occur at this point. See Table IB for the list of these biases capable of influencing subject selection (Table IB).

Exposure or Performance Biases

Exposure biases are also called performance biases. These biases refer to those that relate to exposures that are suspect of causing disease or interventions that might have an effect on disease. For example, in *proficiency* bias, the proficiency of the surgeons could bias the results. If treatment A was done by a group of skilled, experienced surgeons and treatment B was performed by junior residents, a bias in the comparison of treatments might be expected. Poor patient compliance, *compliance bias*, can yield results that are not the result of the intervention; however, subjects randomized to this treatment arm must be counted as having the treatment. Great care must be taken in the design of the project to determine compliance and to insure compliance is good. *Withdrawal bias* is a serious problem. If subjects withdraw or are lost from the study, there is no way to know what would have happened to them; it cannot be assumed that they will respond like those who stayed in the study. A well-designed and executed study will expend great effort to select reliable subjects and keep track of all subjects. Table IC lists additional exposure/performance biases (Table IC).

Detection or Measurement Biases

Detection biases are also labeled measurement biases. Bias in measurement of outcomes is a common and serious problem. Like selection of subjects, this group of biases is important.

Because of strong propensity of *expectation bias*, if the treating or intervention physician is also the one measuring the outcome, bias is almost assured. It is crucial that the outcome assessor is unaware of the intervention; it is also helpful that they be blinded to the hypothesis under study and concentrate only on measuring the outcome parameter(s) of interest. Some of these biases are reduced by the technique of double-blinding, meaning that both the subject and the treating physician are not aware of the intervention; in any case, it is crucial that the outcome assessor is blinded. The list in Table ID itemizes a number of ways the outcome measurement team or person can be biased (Table ID).

Analysis or Transfer Biases

Analysis biases have been called transfer biases.[5] After the implementation and completion of the investigation, the results must be compared and analyzed for significance. Transferring the data into an organized structure for that analysis creates opportunities for biases. *Data dredging bias*, searching through the data looking for anything that might account for differences or that might correlate with something, and *tidying-up bias*, excluding data points because they do not seem right, are all too common. An *a priori* design that includes definitions, hypotheses, and analyses to be done is crucial for a valid study (Table IE).

Interpretation Biases

Following data analysis, an interpretation of the results is typically offered. This represents yet another step in which biases can skew results away from the truth. For example, *cognitive dissonance bias* is all too common in the literature. In this bias, the investigator is convinced that a treatment, pathophysiology, diagnostic procedure, or other area of interest is better or worse, true, or not true, and despite findings to the contrary, will steadfastly contend that his or her point is still valid. This is often seen when an inadequate study concludes with admonitions for treatment or diagnostic maneuvers that are wholly unsupported by the results. All of the biases listed in Table IF are easily seen in the monthly literature offerings (Table IF).

Publication Biases

The process is complete, except for presenting and publishing the data. This too has inherent biases because authors are less likely to submit negative results and editors are also unlikely to publish them. These biases, seen in Table IA, apply to the process of presenting and publishing the work and in reading the literature after publication, as mentioned above (Table IA).

Confounding

In the prospective studies, such as clinical trials (randomized or nonrandomized) and prospective cohort stud-

TABLE I.
List of Biases and Definitions.*

Named Biases	Definition
A Publication Bias (researching the topic)	
1. Bias of rhetoric[7]	Fallacious arguments presented in a compelling manner but without appealing to reason; the conclusions are often based on opinion rather than evidence
2. All's well literature bias[7]	Bias against controversy; publications may exclude or play down reports which are contentious
3. Reference bias[14]	Authors may cite more references that support their position than contradict it; this serves to skew the understanding of the topic
4. Positive results bias[15]	Investigators and publishers are more likely to submit and accept manuscripts with positive results
5. Hot stuff bias[7]	Topics receiving unusual levels of attention may tempt authors or editors to rush additional results to publication regardless of the quality of work
6. Pre-publication bias[9, 16]	Errant research previously published may be used to support a particular investigation
7. Post-publication bias[9, 16]	Published results are extrapolated to a different population
8. Sponsorship bias[17]	Funding sources may influence the probability specific results get published
9. Meta-analysis bias[17]	Meta-analyses based only on published manuscripts miss unpublished possibly relevant studies which may mislead the conclusions
B Selection Bias (susceptibility bias)	
1. Popularity bias[7]	Heightened interest in a condition may attract patients to a specific center when they wouldn't otherwise present for inclusion
2. Centripetal bias[7]	The reputation of a clinician or center may draw individuals with specific disorders to them skewing their population
3. Referral filter bias[7]	Referral of difficult cases to tertiary centers increases the concentration of rare disorders in the center's patient mix
4. Diagnostic access bias[7]	People vary in their ability to procure diagnostic testing which establishes the presence of a given disease
5. Diagnostic suspicion bias[7]	Knowledge of a pre-existing exposure may influence the intensity with which the diagnostic process is applied
6. Unmasking bias[7]	Signs or symptoms sometimes associated with a certain disease may be caused by an innocent exposure, which results in an unusual search for the disease
7. Mimicry bias[7]	An innocent exposure is errantly considered causative when it produces a similar but unrelated condition, which is mistaken for the disease of interest
8. Previous opinion bias[7]	Knowledge of pre-existing evaluation results might alter the subsequent diagnostic process on the same patient or his relatives
9. Wrong sample size bias[7]	Too small a sample might miss significant differences while too large a sample might establish significant differences that are not clinically relevant
10. Admission rate bias (Berkson)[18]	The relative frequency of a disease in hospitalized patients may not reflect the frequency of the population as a whole and thus may bias the study
11. Prevalence-incidence bias (Neyman)[7]	The late look at the prevalence of a disease in a population may underestimate the prevalence because of fatal, mild, or resolved cases
12. Diagnostic vogue bias[7]	The same condition may be known by multiple monikers which if unknown could impact the acquisition of subjects
13. Diagnostic purity bias[7]	When groups become very specifically defined they may not be representative of the broader population
14. Procedure selection bias[7]	Patients may or may not be offered a treatment because of coexisting morbidities or poor prognosis
15. Missing clinical data bias[7]	Clinical data might be missing yet considered normal when it is negative or not ascertained
16. Non-contemporaneous control bias[7]	Over time, changes in definitions (including staging), exposures, diagnostic abilities, and treatments render historical controls noncomparable
17. Starting time bias[7]	Failing to identify the start of exposure or disease may lead to errant stratification
18. Unacceptable disease bias[7]	Socially unacceptable conditions tend to be underreported
19. Migrator bias[7]	Migrants may differ from the population as a whole with respect to their susceptibility to disease, procurement of healthcare, and availability for follow-up
20. Membership bias[7]	Participation in a group may imply a level of health that is different from the general population, i.e., construction workers or runners
21. Nonrespondent bias[7]	The group who fails to respond to a survey may not reflect the entire population surveyed
22. Volunteer bias[8]	Volunteer subjects that differ from the broader target population with respect to exposures, disease status, comorbidities, etc
23. Allocation bias[5]	Nonrandom assignment to a specific study group based on prognostic variables can create incomparable groups
24. Vulnerability bias[5]	Some subjects may be particularly resistant to or overly susceptible to an experimental maneuver
25. Authorization bias[19]	Inability to obtain authorization for release of medical records from some subjects may eliminate them from the study and thus bias the results
C Exposure Bias (performance bias)	
1. Contamination bias[7]	The control group subjects may mistakenly receive the maneuver of interest or be affected by an extrinsic maneuver, which diminishes the differences in outcomes of the experimental and control groups
2. Withdrawal bias[7]	Subjects that drop out may differ significantly from those that remain
3. Compliance bias[7]	The measured efficacy of a treatment can be confounded by non-adherence to the regimen

(Continues)

TABLE I.
(Continued.*)

Named Biases	Definition
4. Therapeutic personality bias[7]	Unblinded treatment allows the practitioner to influence the patient's perception of benefit
5. Bogus control bias[7]	When subjects allocated to the experimental maneuver group expire or withdraw before the maneuver is administered and are reallocated to the control group or are omitted, the experimental and control groups are no longer matched and the differences between may be biased toward the experimental group
6. Misclassification bias[8]	Subjects may be mislabeled and subsequently be allocated to the wrong group for proper comparison
7. Proficiency bias[5]	Experimental maneuvers or treatments may not be equally administered to subjects
DDetection Bias (measurement bias)	
1. Insensitive measure bias[7]	The outcome measure may not be capable of detecting clinically significant differences
2. Underlying cause bias (rumination bias)[7]	Case subjects may more intensely review prior events or exposures searching for a cause, thus different levels of recall of exposures may occur compared with control subjects
3. End-digit preference bias[7]	The natural tendency to convert analogue data to digital data using a preferred ending digit; for example, an even digit rather than an odd digit
4. Apprehension bias[7]	Certain measures can be artificially affected by interaction with investigators, e.g., blood pressure in epistaxis patients
5. Unacceptability bias[7]	Measurements themselves can be painful, intrusive, or embarrassing resulting in avoidance
6. Obsequiousness bias[7]	Subjects may alter their response toward the direction they perceive desired by the investigator
7. Expectation bias[7]	Observers may err in measuring data toward the expected outcome, e.g., normal temperatures in postoperative patients with fever
8. Substitution game bias[7]	The substitution of a risk factor that has not been established as causative may broaden the study population inappropriately
9. Family information bias[7]	Family communication regarding exposures or illness is stimulated by a case subject in its midst, thus more intense remembrance may occur compared with control subjects
10. Exposure suspicion bias[7]	Knowledge of a subject's disease status may affect the intensity of a search for the cause
11. Recall bias[7]	Cases may be questioned about exposures more vigorously than control subjects intensifying their recollections
12. Attention bias[7]	Subjects may alter their behavior when they know they are under scrutiny thereby altering outcomes
13. Instrument bias[7]	Calibration errors may lead to inaccurate measurements being recorded, e.g., an unbalanced weight scale
14. Surveillance bias[5]	When a new treatment is under study, the subjects may be monitored with greater scrutiny than control subjects to guard against adverse reactions
15. Comorbidity bias[5]	Diagnostic procedures investigating a sign or symptom may discover an unrelated previously undetected disease which is then falsely associated with the original sign or symptom, e.g., CT or MRI imaging for evaluation of headache discovers an asymptomatic maxillary sinus mucus retention cyst
16. Nonspecification bias[5]	The maneuver of interest is not well-delineated so unwitting, inadvertent exposures that effect the maneuver might occur, e.g., measuring bleeding times without identifying for the subjects products that commonly contain aspirin
17. Verification bias[20] (work-up bias)	If testing the efficacy of a diagnostic process is restricted only to patients proven to have the disease, the sensitivity of the test can be overestimated
EAnalysis Bias (Transfer Bias)	
1. Post-hoc significance bias[7]	Levels of significance determined after data collection may lead to errant conclusions
2. Data dredging bias[7]	Data reviewed for any and all possible associations without a predetermined hypothesis may yield faulty conclusions
3. Scale degradation bias[7]	Collapsing of measurement scales may obscure differences; for example, collapsing dimensional data, such as specific age, into ranges, such as 0–40, 41–80, etc, may miss important relationships
4. Tidying-up bias (deliberate elimination bias)[7]	Excluding outlying data that seems implausible cannot be statistically justified and contaminates the analysis
5. Repeated peeks bias[7]	Repetitive analysis of the accumulating data may lead to premature conclusions and inappropriate termination of the trial
F Interpretation Bias	
1. Mistaken identity bias[7]	Efforts to improve patient compliance in a trial may instead cause the investigator to manage more aggressively, on determining a treatment outcome it may be errantly attributed to the original prescribed treatment
2. Cognitive dissonance bias[7]	Belief in a hypothetical mechanism may increase, rather than diminish, in response to contradictory evidence; for example, a subject or physician may believe more strongly than ever in antibiotics for a viral sore throat, despite strong proof to the contrary
3. Magnitude bias[7]	Interpretation of data may be affected by the way the data is presented, using a measurement scale that creates a preferred illusion, biases the interpretation; for example, an increase of 10,000 may actually represent only an increase of 0.001%
4. Significance bias[7]	Confusing statistical significance with clinical significance may result in pointless conclusions
5. Correlation bias[7]	Equating a correlation with causation risks making fallacious conclusions; for example, histologic membranous ectasia (hydrops) seems to be correlated with symptoms of vertigo and hearing loss, but is not synonymous with, nor necessarily causative
6. Under-exhaustion bias[7]	Failure to reject the null hypothesis because of a limited investigation may lead to authoritarian speculation rather than truthful conclusions

* The general organization and many of the named biases were taken from the classic article by Sackett.[7]

ies, innate distortions, i.e., biases, are generally located in and effect specific sites along the research pathway. As mentioned above, these biases are broadly named for the site in which they occur; for example, inappropriate admission bias, susceptibility bias, performance bias, detection bias, and transfer bias. However, in retrospective studies, such as retrospective cohort studies and case-control studies, some biases may be harder to name and identify. Some of these biases are external to the pathway and may affect both the maneuver and the outcome. These biases are called confounding variables, or confounders. Because these are vague in definition and location, the investigator must perform "a review of systems" on the research pathway for the specific project to search for any additional variables that might distort the true relationship between the exposure (maneuver) and the outcome.[5]

These confounding variables are extraneous to the research design, but may interfere with the association between the exposure and the outcome. A confounding variable: 1) is associated with the exposure; 2) although independent of the exposure, it is a risk factor for the disease, sometimes in an occult or previously unrecognized way; and 3) is not a direct link between the exposure and the disease. For example, if exposure to A is being investigated to determine if it causes disease B, however older people are much less exposed to A but are much more prone to get disease B, age may be considered a confounder. If the study does not control for age, the confounder, exposure to A may be falsely assumed to be the cause of disease B.[8]

BIASES LINKED TO SPECIFIC RESEARCH TECHNIQUES

The different types of research approaches can be arranged in a hierarchical order, which is based on the ease in which biases are minimized by the structural approach per se. In this arrangement, the randomized, controlled trial is considered the least subject to bias. It is followed by the cohort study and finally the case-control study. All, however, are subject to serious biases.

The first step to minimize bias is to have a clear idea of the question, and what approach is required and feasible to achieve the answer. The second step is to prospectively design the study in detail before, or *a priori*, the investigation is undertaken.

As has been mentioned in the preceding Tutorials, even searching the literature is best done with a clear question in mind, a prospectively designed approach for the search, and a clear idea of the specific research approach required to answer the question. This prospective approach to literature searching immediately helps minimize succumbing to the publication, literature review, researching the topic biases.[9]

Randomized Clinical Trials

A randomized clinical trial (RCT) is usually the best approach to determine a new treatment efficacy and safety. It prospectively compares effects of a new interventional maneuver on the experimental group with a placebo or standard treatment on the control group. By using strategies like randomization, blinding, and a control group, these studies guard against biases threatening a study's validity.[8] Yet despite these techniques, some types of bias may still impact RCTs. Risks even exist after the experiment is complete during data analysis and interpretation. Bias is more likely to occur at this juncture if the hypothesis has not been generated *a priori* or if the conclusions extend beyond the question addressed in the hypothesis.

An example demonstrating the strengths of the RCT is the article "Bells Palsy Treatment with Acyclovir and Prednisone Compared With Prednisone Alone: A Double-Blind, Randomized Controlled Trial."[10] After reading the article, it is obvious that selection bias, execution of the treatment, and measuring its outcome have been managed to a degree by using a control group, randomization, and double blinding. Blinding examiners minimized bias in detection. Finally, the authors' conclusion cites the findings of the study. In summary, the authors applied a regimen to minimize bias and made evidence-based conclusions, successfully minimizing some of the obvious biases that might have seriously threatened the validity of their study. This is not to infer that this is a perfect example, but it is a reasonable one for this survey.

Cohort Studies

A cohort study, by definition always prospective unless otherwise stated, follows one or more groups forward over a period of time. The purpose is often to identify exposures occurring in subjects that might result in a specific outcome of interest. By noting a temporal sequence between exposure and outcome, inferences about causation can be made.

A cohort study reduces bias relative to assessing the potential cause(s) of a disease because the outcome (disease status) is unknown to the examiner at the time of exposure documentation. However, during attempts to look retrospectively at the same question of cause by looking at medical records in a retrospective cohort, this protection may be lost unless the exposure examiner is blinded to the outcome.

Significant biases may effect cohort studies, including sampling, measurement, data analysis, and interpretation biases. These may be minimized by selecting a sample of subjects that is representative of the larger population at risk of the disease and by selecting a control group, which will not receive the exposure in question but is similar in all other factors. Unbiased and appropriate selection of these comparison groups is difficult. Additionally, limiting attrition of subjects is essential, as is the *a priori* hypothesis regarding the relationship between exposure and outcome.

An example of a cohort study with some good bias control and bias errors can be found in "Development of Tympanosclerosis in Children With Otitis Media With Effusion and Ventilation Tubes."[11]

The cohort reflects the at-risk target population: young children with otitis media with effusion. The investigators established an effectively matched control group of ears by only treating one ear; all the subjects had bilateral effusions. Through careful selection, bias at this step was avoided.

Thereafter, however, the study was subject to a number of biases in detection, data analysis, interpretation, and publication. The observers were not blinded to at least the hypothesis and the measurement of degree of tympanosclerosis was subjective. Data analysis became difficult to follow after the first set of tubes. By the end of the 5-year study, 37.8% of subjects had withdrawn or were lost to follow-up; no statements accounting for them were offered and their missing evaluations were not considered in the data analysis. The interpretation of data suggested several biases in the Discussion section. For example, "it seems that tube insertion on only one occasion can induce changes which are as severe as those caused by insertion of tubes on several occasions." No such comparison was made in the paper, so no such inference can be made. Another example of biased interpretation suggested cognitive dissonance bias. "Although from this study, tube reinsertion may not significantly affect the development of tympanosclerosis, the insertion may increase the other damage to the tympanic membrane." The authors reinforced their position even in the face of contradictory or missing evidence.

In deference to the authors, this is a difficult study to do in an unbiased manner; however, with a tight design, such a study could be valid.

The purpose of this example is to illustrate some of the subtle but important biases that can occur during the conduct of a study and during the reading of the literature to answer questions.

Case-Control Studies

A case-control study identifies two groups of subjects who are similar but differ with respect to the presence or absence of a particular disorder. Cases have the disorder and controls do not. Then a look at characteristics or exposures is undertaken to see if the two groups differ by some putative cause of the disorder. This allows inferences about a possible relationship between exposure and outcome.

This method of study is prone to many sources of bias and is less able to defend against them, yet offers a practical means of answering many clinical questions. The potential biases include selection bias, observation bias, and biases from analyzing and interpreting the data. The selection of the controls is particularly subject to subtle, but important biases. Bias can be minimized by carefully matching controls to cases, applying the appropriate observational technique identically to both groups, and gathering information from both groups in the same fashion. A careful evaluation for potential confounding variables is important.

An example of a case-control study is demonstrated in "Pharyngeal pH Monitoring in Patients With Posterior Laryngitis."[12] The cases consisted of consecutive patients diagnosed with posterior laryngitis after suggestive symptoms were noted and confirmatory physical findings were seen on videostroboscopy. Healthy age-matched volunteers were recruited by advertisement and were free of any reflux symptoms or signs.

Both groups were examined by pharyngoesophageal pH monitoring. By matching controls with cases, and by eliminating from controls anyone with reflux symptoms on a questionnaire, the authors minimize selection biases from this study, if the issue was to differentiate between asymptomatic normal subjects from the cases and if the questionnaire was validated for that purpose. On the other hand, if the issue were to differentiate people with symptoms, but without objective evidence of laryngitis, this strategy would not work.

Data recorded in both groups included pH exposure in the pharynx, proximal and distal esophagus. Analysis of the data was performed only on pH measures to determine statistical significance. Conclusions were then limited to original objectives for which evidence had been collected, thereby minimizing interpretation bias.

In summary, the investigators established an *a priori* question, collected information in identical fashion from cases and controls, and drew specific conclusions based on their data. This is not a perfect example; however, it does serve to illustrate some of the methods to reduce biases.

METHODS TO MINIMIZE THE IMPACT OF BIASES ON RESEARCH

A list of some of the specific methods to avoid or minimize bias is displayed in Table II.

Searching for both positive and negative published studies and trial registries for unpublished investigations can help reduce *publication biases*.

Using strict eligibility, inclusion, and exclusion criteria and randomization for the allocation of maneuvers can minimize *selection biases*.

Exposure biases are diminished by prospectively establishing criteria for performing the experimental maneuver and blinding the investigator and subjects when possible. Rigorous maintenance of contact with subjects helps to avoid noncompliance, withdrawal, and loss to follow-up from the study.

Measurement biases can be reduced by prospectively establishing detailed methods for data collection, by blinding interviewers to subjects' diagnoses, or by soliciting a history of exposure before a diagnosis is determined, and by applying detection methods equally to both groups. Additionally, it is helpful to seek exposure information from independent sources.[7] Requiring response rates of over 80% can minimize *nonrespondent bias*. Matching as many confounding variables in cases and controls in case-control studies helps minimize biases affecting the interpretation of putative causal exposures.

Analysis biases are decreased by prospectively choosing statistical methods best suited to evaluate the data and analyzing the association between confounding variables and the results.

Interpretation biases can be avoided by using one or more control groups and by basing conclusions on the hypothesis-driven data collection.

Example Article for Bias Assessment

Common examples include the latest news cited by the press, the monthly journals that arrive on our desks, or the articles handed out by pharmaceutical representatives supporting their products.

TABLE II.
Categories of Bias and Techniques for Control.

Category of Bias	Control Techniques
A Publication Bias	
	Establish similar numbers of positive and negative published studies[17]
	Investigators should present negative studies for publication and editors should publish negative studies[17]
	Evaluate validity of studies based on their methods not their conclusions; then weight their conclusions accordingly[17]
	Search trial registries for unpublished data[17]
	Identify funding sources and possible conflicts of interest
B Selection Bias	
	Select the most rigorous study design feasible to address the hypothesis[5, 8]
	A priori create explicit criteria defining methods to be used throughout the study[5, 8]
	Randomization removes human judgment from allocation to the groups and should be used whenever possible[5, 8]
	Create specific criteria for admission to evaluate the maneuver[5]
	Ascertain subjects' baseline state and classify them based on that state[5]
	Control groups should be similar to the experimental group with respect to variables that are predictors of disease incidence[21]
	In case-control studies both groups should only be comprised of subjects who have undergone identical diagnostic testing and there should be no difference in how exposure or disease information is gathered[8, 7, 20]
	Create a list of all possible means of exposure to the maneuver of interest prior to the performance of the maneuver, and then screen subjects for these external exposures[5]
	Match or adjust for confounding variables in the two groups[20, 7]
	More than one control group can be established[21]
C Exposure Bias	
	Select the most rigorous study design possible to address the question[5]
	A priori establish specific criteria for performing the maneuver[5, 7]
	Blinding used in experimental trials prevents knowledge of the maneuvers from influencing the outcome measurement[5, 8]
	Hide the hypothesis from the interviewer and subjects[5]
	Consider creating decoy hypotheses to disguise the question of interest[5]
	Divide the labor, i.e., someone other than the person to ascertain the outcomes must performance of the maneuver[5, 8]
	Maintain aggressive contact with subjects to maintain compliance with the protocol and to minimize attrition from the study[8, 7]
D Detection Bias	
	A priori establish explicit criteria for collecting data on exposures and outcomes[7]
	Limit any differences between groups in how information is obtained about exposure, disease, or outcome[8]
	Double blinding subjects and investigators when possible prevents knowledge of exposures from influencing the detection of outcome events[5, 7]
	Blind the interviewer to the hypothesis[5]
	Establish rigorous, rigid, format for data acquisition, i.e., phrasing of questions, methods for recording answers, etc[5]
	Detection procedures should be applied equally to both groups[5]
	Hide the identity of the subjects from the data collector when possible[5]
	Use archived data[5]
	Acquire data about exposure from independent sources[7]
	Create a division of labor by having a different person record data than performs the maneuver[5, 8]
	Examine for potential confounding variables affecting the outcome[8]
	Maintain aggressive contact to avoid attrition from the study[8]
	For questionnaires, obtain response rates of $\geq$ 80%[7]
E Analysis Bias	
	Establish *a priori* the statistical methods best suited to evaluate the data[5, 8]
	Do not use unknown data, but do report how it was managed[5]
	Determine the significance of the association between confounding variables and exposures and outcomes[8]
	Do not exclude outlying difficult to explain data[7]
F Interpretation Bias	
	Use control groups for comparison[8]
	Base conclusions on the data and limit them to the hypothesis[5]

Recently a pharmaceutical representative delivered the article "Onset of Action and Efficacy of Terfenadine, Astemizole, Cetirizine, and Loratadine for the Relief of Symptoms of Allergic Rhinitis."[13] This article was an attractive example to use in this manuscript because it was unsolicited, came from a journal outside of otolaryngology, and yet was germane to most otolaryngologists. The format used to discuss this example article was to emphasize the positive aspects of the study design and study conduct, and to illustrate the measures to control biases. It was not the intention to formally critique the work.

A systematic approach to assess the authors' attention to and control of biases begins with the stated research question. They proposed to compare onset and efficacy of the medications. A quick look at the menu of research approaches to best answer clinical questions and some reflection on the feasibility of studying this question suggests that a randomized, double-blind, placebo-controlled design would be best.[3] The authors did use this design.

The randomized design specifically attempted to make the groups comparable by hoping to randomly distribute all of the baseline variables equally between groups, thus helping control for *selection bias*. Sometimes even randomization fails to make the groups exactly alike. Analyzing the baseline variables between groups to determine if significant differences were found could check this; Table I in an article is often devoted to this exercise. In this article, Table III assessed some of these variables and found no significant differences.

Double blinding, meaning the subjects and the investigator were unaware or "blinded" to the interventional maneuver, helps control *detection bias*. This minimizes the natural tendency of the subjects and of the investigator to skew results toward some preconceived "better medicine." Because neither knows, it gives each medication and the placebo an equal chance to relieve symptoms, if indeed they do. They did this.

The placebo control was an important way of determining if any of the medications were more effective than no medication. This helped avoid *therapeutic personality bias, expectation bias,* and *measurement biases*.

Another appropriate early question is did the funding source influence the results? Were there conflicts of interest? This could not be clearly determined.

Having ascertained these design issues quickly from the Materials and Methods section of the article, greater scrutiny revealed additional efforts to minimize biases.

The selection process is a high-risk endeavor therefore attention was focused there. Did the authors *a priori* define the selection method? Was that selection method uniformly applied? In the example, subjects were solicited from allergic patients and by advertisement. This potentially risked a *volunteer bias*, in which the study sample might differ from non-volunteers or the population at large. This was dealt with by confirming the subjects were representative of the larger target population. In this case, eligible subjects were required to have a documented clinical history of seasonal allergic rhinitis for the previous 2 years, positive skin tests to ragweed antigen, and response to pre-study priming exposure to ragweed pollen in a ventilation-controlled room. Exclusion criteria and the security of the exposure room reduced *contamination bias*.

Executing the experimental maneuvers was the next step in the research method and, of course, the next point in which biases may have had impact. Critical information centers on whether the different maneuvers were applied in identical fashion to each group under investigation. In the example study, each group was administered the test drug after a 1-hour induction of symptoms by ragweed pollen in the experiment room. In each group, the maneuvers were administered in identical fashion. This helped control *exposure bias*.

Of 111 primary subjects, 19 did not complete the study. All were accounted for in the Results section where it was noted that 14 left for nonmedical reasons and 5 quit secondary to symptoms that were too intolerable to continue. However, the impact of this loss on each of the final groups was not well delineated, increasing the risk that *withdrawal bias* may have some opportunity to skew the results.

Measuring outcomes is also at risk for biases. In the example, the subjects had to rate their symptoms and response to treatment. This could introduce *attention bias (Hawthorne effect)* and *confounding bias* from intersubject variability in interpreting the severity of symptoms and the degree of response to treatment. The authors did attempt to minimize the impact by "educating subjects as to standardized methods of scoring severity levels." These severity levels were determined *a priori*, seemed to be sufficiently sensitive, and seemed appropriate to the study question; however, no mention was made of previous validation of the scoring model.

Analysis of the data posed the next threat from biases. The analysis should be prospectively designed and hypothesis-driven to avoid *data dredging bias*. The authors defined in advance the level of statistical significance, the required power, and the statistical tests that were to be used in the analysis of the implicitly hypothesis-driven study.

Finally, interpreting the data also is susceptible to biases. In the Discussion section of the article, the efficacy of the antihistamines was ranked, based on statistically significant data. The article implied clinically significant differences, which was not fully defined by this study; this may be an example of *significance bias*.

The very act of publication itself has some inherent potential for biases, such as *positive results bias*. The authors did not state whether they would have published negative results alone. However, to their credit, they did emphasize both statistically significant and nonsignificant data.

CONCLUSION

Clinical experience is derived from both the practice of medicine and the assimilation of the literature. That literature must be assessed for validity before being assimilated; otherwise, it may promote mistruths as reality which, when applied to patients, may inadvertently endanger them.

Therefore, an awareness of biases and their causation should facilitate their recognition in an article. Uncontrolled biases may render a manuscript invalid.

The text and tables here provide a rapid reference tool for future application by the reader. The example section at the end of the paper illustrates the application of this tool for the validity assessment of an article.

The next tutorial will address outcomes research.

BIBLIOGRAPHY

1. Neely JG, Hartman JM, Wallace MS. Building the powerful 10-minute office visit. Part I: introduction to the new section. *Laryngoscope* 2000;110:1595–1601.
2. Neely JG, Hartman JM, Wallace MS. Tutorials in Clinical Research: building the powerful 10-minute office visit. Part II: beginning a critical literature review. *Laryngoscope* 2001;111:70–76.
3. Neely JG, Hartman JM, Wallace MS, Forsen JW Jr. Tutorials in Clinical Research: Part III: selecting a research approach to best answer a clinical question. *Laryngoscope* 2001;111:821–830.
4. *Webster's Third New International. Dictionary of the English Language Unabridged.* Chicago: Encyclopaedia Britanica, Inc, 1976.
5. Feinstein AR. *Clinical Epidemiology: The Architecture of Clinical Research*. Philadelphia: W.B. Saunders Co, 1985.
6. Maclure M, Schneeweiss S. Causation of bias: the episcope. *Epidemiology* 2001;12:114–122.
7. Sackett D. Bias in analytic research. *J Chron Dis* 1979;32: 51–63.
8. Portney LG, Watkins MP. *Foundations of Clinical Research. Applications to Practice*. Norwalk, CT: Appleton and Lange, 1993.
9. Neely J. Clinical epidemiology/outcomes research. Literature review articles as a research form. *Otolaryngol Head Neck Surg* 1993;108:743–748.
10. Adour K, Ruboyianes J, VonDoersten P, et al. Bell's palsy treatment with acyclovir and prednisone compared with prednisone along: a double-blind, randomized, controlled trial. *Ann Otol Rhinol Laryngol* 1996;105:371–378.
11. Maw A. Development of tympanosclerosis in children with otitis media with effusion and ventilation tubes. *J Laryngol Otol* 1991;105:614–617.
12. Ulualp S, Toohill R, Hoffman R, Shaker R. Pharyngeal pH monitoring in patients with posterior laryngitis. *Otolaryngol Head Neck Surg* 1999;120:672–677.
13. Day J, Briscoe M, Clark R, Ellis A, Gervais P. Onset of action and efficacy of terfenadine, astemizole, cetirizine, and loratadine for the relief of symptoms of allergic rhinitis. *Ann Allergy Asthma Immunol* 1997;79:163–172.
14. Gotzsche P. Statistics in medicine. Reference bias in reports of drug trials. *BMJ* 1987;295:654–656.
15. Dickersin K. The existence of publication bias and risk factors for its occurrence. *JAMA* 1990;263:1385–1389.
16. Chalmers TC, Frank CS, Reitman D. Minimizing the three stages of publication bias. *JAMA* 1990;263:1392–1395.
17. Thornton A. Publication bias in meta-analysis: its causes and consequences. *J Clin Epidemiol* 2000;53:207–216.
18. Roberts R, Spitzer W, Delmore T, Sackett D. An empirical demonstration of Berkson's bias. *J Chron Dis* 1978;31: 119–128.
19. Jacobsen S, Xia Z, Campion M, et al. Potential effect of authorization bias on medical record research. *Mayo Clin Proc* 1999;74:330–338.
20. Begg C. Biases in the assessment of diagnostic tests. *Stat Med* 1987;6:411–423.
21. Rothman K. *Modern Epidemiology*. Boston/Toronto: Little, Brown & Co, 1986.