Clinical Chemistry 56:3
357–360 (2010)

The Title Says It All

Thomas M. Annesley

Crafting a good title is like bending a straight line into a seamless connected circle. Let me explain. Titles are the first thing readers see, yet they are often the final part of writing a manuscript, and just as often given the least attention. Guides on writing scientific papers recommend starting with the Methods section, followed by Results, Discussion, Introduction, Abstract, and Title. There is a certain degree of logic to this linear order, since authors already have detailed notes and documentation on the methods used, the results obtained, and the interpretation of the results when they begin to write a paper. The title, however, is the component that closes the circle. The title draws from the other sections of the paper and becomes the face of the paper—the descriptor, the advertisement, the pitch. Like a billboard, it is your 10-second opportunity to connect with the passerby (the reader). So it is important to make the title count. Here I provide some tips and examples to help you reach this goal.

Be Concise

When asked "How long should a title be?" my response is that the length should be just right. This is not to avoid the question, but to tell the author that a title should balance the number of words needed to describe the content of the article against losing or confusing the reader with too many words. How a title looks to the eye can mean as much as what it says. In most cases, titles end up being better if words are removed rather than added. Avoid using wasted words such as "a study of," "investigation of," "development of," or "observations on." Readers understand that you would not be writing the paper unless you had studied, investigated, developed, or observed something. Similarly, avoid including adjectives such as "new," "improved," "novel," "validated," and "sensitive." Why would a journal want to consider a study that was not new, not validated, or not sensitive? Many journals have strict limits on the number of words or characters that can be included in the title, so it always helps to look for words that can be removed from a title without affecting the clarity or message.

Be Clear

Consider the following titles:

- H1N1 Virus Testing on Mice Using Polymerase Chain Reaction
- Blood from Bone Marrow Donors Stored on Ice Yields Higher HLA-Match Percentages
- Treatment of Pediatric Melanoma Patients with Lasers
- Value of Amniotic Fluid Sphingomyelin Quantification in Fetuses with G1-α Gene Mutations of Unclear Significance

Those must be extraordinary mice in the first paper if they are capable of performing polymerase chain reaction analyses. I have donated blood in the past for a bone marrow drive, but I am thinking of dropping out in the future if the agencies are going to store me on ice just so they can get a better result. If I were a pediatrician, I would demand that someone disarm the lasers from those children before I would enter the treatment room. Finally, is the value of amniotic fluid sphingomyelin of unclear significance, or are the gene mutations of unclear significance? It is likely that readers will figure out the true meaning of an unclear title, but their subsequent mindset when reading of the remainder of the article (if they do not quit here) may already be negatively affected.

Now consider alternatives to the above titles:

- Polymerase Chain Reaction Testing of Mice for the H1N1 Virus
- Increased Bone Marrow Donor HLA-Match Percentages for Blood Stored on Ice
- Laser Treatment for Pediatric Melanoma
- Amniotic Fluid Sphingomyelin Quantification Identifies G1-α Gene Mutations of Unclear Significance

These hypothetical examples of odd titles and their modifications illustrate that the syntax (word order) in a title deserves more attention than it often receives. There should only be one meaning to your title. A good practice is to show the title to colleagues who are not coauthors and ask them to tell you what message they take away from your words.

University of Michigan Health Center, Ann Arbor, MI.
Address correspondence to the author at: University Hospital, Room 2G332, 1500 East Medical Center Drive, Ann Arbor, MI 48109-5054. Fax 734-763-4095; e-mail annesley@umich.edu.
Received December 1, 2009; accepted December 8, 2009.
Previously published online at DOI: 10.1373/clinchem.2009.141523

Table 1. Varying effectiveness of titles in providing information.

A New Method for Sirolimus
Analysis of Sirolimus by High-Performance Liquid Chromatography–Mass Spectrometry
Rapid Analysis of Whole-Blood Sirolimus by High-Performance Liquid Chromatography–Mass Spectrometry
Statins and Cholesterol
Effect of Statins on Serum Cholesterol
Reduction of Serum Cholesterol with Statin Therapy
Statin Therapy Reduces Serum Cholesterol in Patients with Cardiovascular Disease
Animal Testing for Flu Viruses
Animal Testing for the H1N1 Virus
Testing of Dogs for the H1N1 Virus
Polymerase Chain Reaction Testing of Dogs for the H1N1 Virus

Table 2. Change in emphasis through rearrangement of titles.

Amniotic-Fluid Sphingomyelin Quantification Can Identify G1-α Gene Mutations of Unclear Significance
Emphasis: *Amniotic fluid, sphingomyelin*
G1-α Gene Mutations of Unclear Significance Can Be Identified by Amniotic-Fluid Sphingomyelin Quantification
Emphasis: *G1-α gene mutations*
Blood Stored on Ice Yields Increased HLA-Match Percentages for Bone Marrow Donors
Emphasis: *Blood processing and storage*
Increased Bone Marrow Donor HLA-Match Percentages for Blood Stored on Ice
Emphasis: *Bone marrow HLA matching*
First-Trimester Maternal Pregnancy-Associated Plasma Protein A Is Influenced by Smoking
Emphasis: *First trimester maternal screening*
Smoking Influences First-Trimester Maternal Pregnancy-Associated Plasma Protein A
Emphasis: *Effect of smoking*
Proficiency Testing of Dried Blood Spot Trypsin for Newborn Screening for Cystic Fibrosis
Emphasis: *Proficiency testing, dried blood*
Newborn Screening for Cystic Fibrosis: Proficiency Testing of Dried Blood Spot Trypsin
Emphasis: *Newborn screening, cystic fibrosis*

Be Informative

Although it is considered a virtue for titles to be concise, sometimes titles are so short and incomplete that they tell the reader very little about the topic of the article. Table 1 illustrates variations in 3 titles, ranging from least to most informative. In the first example, the most informative title includes the specific advantage of the method (rapid), the specimen used (whole blood), the compound analyzed (sirolimus), and the method used (high-performance liquid chromatography–mass spectrometry). This complete description of the topic of the paper used 13 words (99 characters). The second example starts with a general statement that the paper is about statins and cholesterol, giving the reader little information. There is more information when the reader is told that statins affect cholesterol, but even more information when readers learn that statins reduce cholesterol. The most informative title tells readers the independent variable (statin therapy), the dependent variable (cholesterol), the observed effect (reduction), and the patient population studied (cardiovascular disease) using just 10 words (80 characters). In biomedical journals, the species studied is typically assumed to be humans unless otherwise specified. For all other studies (e.g., animals, bacteria, cell cultures), the title should clearly state the specific organism or biological system studied (e.g., dogs, *E. coli*, HeLa cells), as shown in the third example in Table 1.

Use Key Words and Terms Wisely

Key words and terms play an important role and should be chosen carefully. Since many readers first see your paper by scanning through a table of contents, as an author you want to capture the reader's attention with words and terms that highlight the content (or message) of the paper.

The key words or terms used in a title should also be the same ones used throughout the article. Using the titles in Tables 1 and 2 as examples, if you use *human leukocyte antigen (HLA) match* in the title, do not substitute other terms in the abstract or main text such as *histocompatibility match* or *major histocompatibility complex (MHC) match*. If you use *high-performance liquid chromatography* in the title, do not switch to *chromatographic assay* at various points in the article.

Key words and terms in a title are also important because they are the same terms that indexing services (e.g., PubMed) and search engines (e.g., Google) key on. Again, avoid using generic terms such as *animal*, *bacteria*, or *antibiotic* as your key words, each of which can carry several meanings. A potential reader seeking information on *Salmonella* poisoning will likely not enter *bacteria* as a search term in PubMed, since more than 1 million articles will show up. Your paper will be lost as a proverbial raindrop in the ocean of papers.

Search engines such as Google typically show only the first 6–7 words of a title, so there is a benefit to

placing the terms you most want associated with your paper early in the title (Table 2). This is also another reason to remove unnecessary phrases such as "development and validation of a sensitive" which already uses up the first 6 words.

Know the Journal and the Target Audience

A common error of authors is a failure to read the explicit instructions that outline formatting, style, word limits, etc., required by a given journal. In spite of access to this information, authors frequently submit papers with titles that do not meet the requirements of the journal. Some journals have a limit on the number of words or characters in a title, or may request that titles be written as phrases, not sentences (e.g., "Reduction in Cholesterol with Statin Treatment" vs "Cholesterol Concentrations Are Lowered with Statin Treatment"). Other journals prohibit subtitles or the designation of articles as part 1 or part 2. A good tip is to read past issues of the journal of interest to see titles that have been published.

The order in which words and terms are used in a title can also influence readers' interest in your paper. Therefore it is important to decide what you want to emphasize as the primary subject matter. Table 2 contains 4 pairs of titles in which the first half of the title in each pair emphasizes a different subject. The hypothesis of the study and the results are the same, but each of the 4 papers in the table will be seen in a different light and appeal to different audiences depending on how the words in the title are arranged.

Avoid Abbreviations

Abbreviations can cause several problems and should be avoided in a title. First, abbreviations confuse and lose readers if they are not experts in the subject matter of your paper. Second, unless an abbreviation in the title is an accepted standard abbreviation used by indexing services, your article may not get indexed properly and be missed by potential readers. Third, even if it is a standard abbreviation, indexing services generally provide readers with access to titles and abstracts only, not the main text of the paper, where abbreviations are defined.

In rare situations, the abbreviations for some nouns have become more widely used than the actual spelled-out names used in common speech. For example, the average person does not say that their family has longevity because of their good deoxyribonucleic acid—they say "DNA" instead. Similar common terms include RNA, AIDS, CDC, and FDA. In these circumstances, using the abbreviation instead of the full name may be clearer to the reader. When in question, however, check with the specific journal or seek out reference materials or style guides.

Learning Exercise

Knowing about their importance, you should be able to write clear, succinct, informative titles. The first 3 titles listed below can all be improved. Try to rewrite them using the tips and caveats discussed in this article. For the fourth title, try creating a running title. Some possible new titles are provided in a box after the list of selected additional reading materials.

- Development and Evaluation of a New ELISA for the Improved Detection of Lupus-Specific Antinuclear Antibodies
- A Validated Method for the Sensitive Quantification of Sirolimus in Whole Blood by the Use of Online Extraction Connected to Liquid Chromatography–Mass Spectrometry
- Evaluation of siRNA Molecules Reveals Them to Be Sensitive and Specific Biomarkers of Sepsis
- Reduction of Viral Load in Blood after Albinovir Treatment of HIV-Infected Patients

Final Thoughts

There is an old saying, "You don't get a second chance to make a first impression." The title of an article has the power to influence the first impression of your work by a reader, reviewer, or editor. The words selected for inclusion in a title describing the content of your paper must be clear, concise, informative, and relevant to the target audience. Scan the table of contents for several prominent journals and decide on the wording and style of titles that captured your attention. You will soon gain an appreciation for the importance of the title when you write your next paper for publication.

Additional Reading

Day RA, Gastel B. How to write and publish a scientific paper. Westport, CT: Greenwood Press, 2006.
Zeiger M. Essentials of writing biomedical research papers. New York: McGraw Hill, 2000.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures of Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:*

Employment or Leadership: T.M. Annesley, *Clinical Chemistry*, AACC.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Answers to Learning Exercise

Development and Evaluation of a New ELISA for the Improved Detection of Lupus-Specific Antinuclear Antibodies

Revised title:

ELISA with Improved Detection of Lupus-Specific Antinuclear Antibodies

A Validated Method for the Sensitive Quantification of Sirolimus in Whole Blood by the Use of Online Extraction Connected to Liquid Chromatography–Mass Spectrometry

Revised title:

Quantification of Whole-Blood Sirolimus by On-line Extraction Liquid Chromatography–Mass Spectrometry

Evaluation of siRNA Molecules Reveals Them to Be Sensitive and Specific Biomarkers of Sepsis

Revised title:

Plasma siRNA Are Biomarkers of Sepsis

Reduction of Viral Load in Blood after Albinovir Treatment of HIV-Infected Patients

Running title:

HIV Viral Load Reduction with Albinovir Treatment (49 characters)

Albinovir Treatment and HIV Viral Load (38 characters)

HIV Treatment with Albinovir (28 characters)

Clinical Chemistry 56:4
521–524 (2010)

The Abstract and the Elevator Talk: A Tale of Two Summaries

Thomas M. Annesley

What is an elevator talk, and what does it have to do with writing a paper? A lot. Imagine you are the president of the nonprofit Light Is The Solution Foundation. The board of directors is meeting at the New York Hilton, and you are waiting to ride the elevator from the 31st floor to the lobby. The doors open, and you suddenly find yourself standing with Bill Gates, whose philanthropic Gates' Foundation is meeting at the same hotel. Gates notices the logo on your shirt of a small child reading a book by the light of a lantern and asks you, "What is that? What do you do?"

Indeed, what do you do now? You have 30 floors, or about 1 minute, to get your message across. So you explain that normal living activities cease in many countries in the world after the sun goes down. Children have no light to read textbooks, mothers no light to cook, fathers no light to earn income. With this background, you then explain that the Light Is The Solution Foundation has addressed this problem by developing rechargeable lanterns that are low-cost, have a battery life of 30 hours, and put out light equivalent to three 60-W bulbs. You have given away 4500 lanterns in one country and have results showing that more children now share books, study together, and graduate at a higher rate. In fact, average incomes have risen by 20% for families who have received a lantern. You have concluded that this unique program could be expanded to any country that has even the crudest electrical grid or generators for recharging the lanterns.

This is the elevator talk. Your 1-minute opportunity to summarize what you do, how you do it, the results you produce, and the impact you make. A well-developed elevator talk entices the listener to want to learn more. In many professions, entire careers are made and lost as a result of elevator talks.

The elevator talk and the abstract of a scientific paper have a lot in common. Although written rather than spoken, the abstract also provides a summary of the important information an author wants to convey to the reader, with the goal of enticing the reader to want to learn more. Instead of a limitation on the amount of time, the author has a limitation on the number of words. The challenge is to make the most effective use of these words. Here I provide you with some basic information about the abstract and highlight the characteristics of a well-written abstract (Table 1).

Tell a Story by Answering Questions

An abstract is a summary or, more precisely, a condensed version of your paper. Its purpose is to tell the reader not just the basic information or data contained in the paper but also why the paper was written and what value it adds. For example, imagine you were writing a review article on pharmacogenomics. You would not immediately jump into descriptions of the current literature without first telling the reader why pharmacogenomics may be important to them. You want to provide the reader with some brief background of why the field exists to begin with. Nor would you simply stop the review without telling the reader where you believe this field of medicine is headed and what may lie ahead.

For papers of original research, the IMRAD format (Introduction, Methods, Results, and Discussion) is

Table 1. Characteristics of a well-written abstract.

Stands on its own without need to read the paper
States the hypothesis, question, or objective of the study
Completes the story by answering the hypothesis, question, or objective
Contains the same key words and terms as the title and the introduction
Follows the correct style and format
Follows the order of the main text (e.g., IMRAD)
Stays within the allowed word count
Does not contain information absent in the paper
Does not make conclusions unsupported by the data
Limits the use of abbreviations
Does not include references
Does not cite tables or figures

University of Michigan Health Center, Ann Arbor, MI.
Address correspondence to the author at: University Hospital, Room 2G332, 1500 East Medical Center Drive, Ann Arbor, MI 48109-5054. Fax 734-763-4095; e-mail annesley@umich.edu.
Received December 13, 2009; accepted December 22, 2009.
Previously published online at DOI: 10.1373/clinchem.2009.142026

commonly followed. Although different from the format of a review, each section contributes to the overall story by answering one or more questions:

Introduction—What problem, question, or hypothesis is being studied? Why would it be of interest to the reader?

Methods—How did you perform the study, test the hypothesis, or answer the question?

Results—What did you find? Did you solve the problem, prove the hypothesis, or answer the question?

Discussion—What do your results mean? What value do they add to the scientific literature?

A well-written paper tells a story, even if in scientific language, by answering important questions. A well-written abstract should tell the same story by answering the same questions. It should start with 1 to 2 sentences introducing the topic and the goal of the study and end with 1 to 2 sentences telling the reader what your results mean. In between are the most important answers for the reader, which are how you performed the study and what you found. Thus, the majority of the abstract should cover the methods used and the results obtained.

Use the Correct Style and Format

Abstracts can be written in 2 formats (simple/conventional and structured) and 2 styles (descriptive and informative). A simple abstract consists of a single narrative paragraph that may follow the IMRAD format without specifically associating the text or information with any of the IMRAD headers listed above; they can be descriptive or informative. Examples of journals that use the simple format are *Analytical Chemistry* and *The American Journal of Pathology*.

In a structured abstract, the text is divided under individual headings, almost like a miniversion of the paper. The author provides specific information under each heading. *Clinical Chemistry* requires structured abstracts with 4 headings: Background, Methods, Results, and Conclusions. *JAMA* uses a different set of headings: Context, Objective, Design, Setting, Patients/Participants, Main Outcome Measures, Results, and Conclusions.

Descriptive abstracts, as the name implies, describe the content of the paper, much like a summary paragraph found at the end of the paper. A descriptive abstract answers the questions discussed earlier but does so in general terms. It does not provide details about the design of the experiments or the resulting data and is often shorter in length (75–150 words). The descriptive abstract is appropriate for articles such as reviews that do not contain original research data; however, some high-visibility journals such as *Nature* and *Science* have lower word limits and publish descriptive abstracts for research papers as well.

Informative abstracts include actual details of the research study, such as study design, methods used, important results, and conclusions. An informative abstract includes enough material to serve as a surrogate for the full paper. Because most research studies follow detailed protocols, use sophisticated methodologies, and generate a substantial amount of data, the informative abstract is common in scientific publications.

Nearly every journal's information/instructions for authors state what type of abstract is required, so ignorance is no excuse for preparing an abstract improperly.

Create an Abstract That Stands on Its Own

After the title, the abstract is the second most commonly read part of a paper. Like the title, the abstract must stand on its own. In the age of electronic publishing, fewer papers are being accessed by flipping through the pages of a printed copy of a journal. Internet sites such as PubMed display just the title and abstract for a published paper. Similarly, many journal Web sites allow nonsubscribers to see only the abstract for a published paper, with a pay-per-view option for access to the entire paper. Thus, the abstract must stand on its own. If readers are not impressed by the information in the abstract or perceive the study to be weak because the abstract is weak, they will simply move on to another paper. Similarly, editors and peer reviewers also form an initial opinion about a submitted paper from what they see in an abstract.

Regardless of the style or format, any abstract fails to achieve its goal if it lacks sufficient useful information. Authors sometimes assume that because no abstract can describe all of the information in a paper, individuals will take the time to read the full paper to find material missing from the abstract. Readers want as much detail as you can provide, given the word limit set by the journal. They also want to understand the rationale behind the study and what conclusions can be drawn from the results.

Write the Abstract after Completing the Main Text

Some authors find that drafting the abstract early in the process helps encapsulate the major points being considered for the paper. Since the abstract is a condensed version of the full paper, however, a logical time to write it is after the rest of the paper is complete. An abstract written too early in the process may end up containing information, data, or statements not found in the main text (or the reverse). I recall double-checking results in a table, finding that they needed

modification, and almost forgetting to make the same modification in the abstract.

Reviewers and editors often request the addition or deletion of text, more information about experiments, reanalysis of data, reinterpretation of results, modification of conclusions, and so forth. Thus, it is important to reevaluate the content of the abstract when you revise a manuscript to ensure that the abstract coincides with the revised text.

Another reason for writing the abstract after completing the main text is that it allows you to link the abstract to the title and the introduction, which is a subtle yet important aspect of a good paper. The main message about the study conveyed through the title should be conveyed again to the reader in the abstract. Similarly, the background information contained in the abstract should parallel the background information in the introduction. You can borrow sentences from the introduction and include them in the abstract. Use of the same nouns, verbs, or adjectives in the title, abstract, and introduction is not only perfectly acceptable but also potentially beneficial, because it allows you to use the key terms for inclusion in indexing services (e.g., PubMed) and search engines (e.g., Google) multiple times.

Avoid Abbreviations

One fairly common error authors make is to use abbreviations in the abstract and assume that the reader will refer to a separate list of abbreviations or the main text. Remember that journal Web sites, as well as PubMed, typically provide access to the abstract and not to the main paper. A few abbreviations are more widely used than the spelled-out names (e.g., DNA, RNA, AIDS) and can be used without confusing readers. If a disease with a long name, such as amyotrophic lateral sclerosis (ALS), is the focus of the paper and is mentioned multiple times in the abstract, the abbreviated form may be used after being defined the first time it is used in the abstract. Otherwise, try to avoid using abbreviations.

Learning Exercise

Below I have drafted an abstract for a hypothetical study. With the information presented in this guide, try to identify any weaknesses in the abstract and what suggestions you have for improvements. Then see if the revised version provided in the box after the list of selected additional reading materials corrects any of the problems you identified. The abstracts contain 205 and 203 words respectively.

BACKGROUND: Atherosclerotic disease is a major cause of death in the United States. We investigated which analyte, IL-6 or β-selectin, would be a better prognostic marker for atherosclerotic disease.

METHODS: We divided patients into 4 groups. Specimens from each patient were tested for interleukin-6 and β-selectin and matched against the patient's disease group. During the study period, these analytes were measured again to determine whether concentrations changed with disease severity. Mortality was also monitored for each group to investigate any relationship between IL-6 or β-selectin and the risk of death.

RESULTS: The IL-6 concentrations were different between groups, with the IL-6 concentrations significantly different between groups 1 and 3, and 1 and 4. Although IL-6 and β-selectin concentrations both changed, β-selectin changed by only 10% to 30%. Changes in disease severity were reflected in changes in IL-6. IL-6 values were the same for men and women and did not show any relationship with patient age. Intraindividual variation for IL-6 was much lower than that for β-selectin.

CONCLUSIONS: IL-6 and β-selectin concentrations change with a change in heart disease severity. Intraindividual variation of IL-6 was also much lower than β-selectin, further validating the use of IL-6 over β-selectin. Further work is needed to confirm this observation.

Final Thoughts

Former US President Woodrow Wilson once said, "If I am to speak ten minutes, I need a week for preparation; if fifteen minutes, three days; if half an hour, two days; if an hour, I am ready now." If he were writing a scientific paper, I predict he would have had the same thoughts about developing an abstract. Entire pages of information can be written in a short time; the process of condensing that information with well-chosen words takes much more time. But a good abstract is worth it.

Additional Reading

Katz MJ. From research to manuscript. New York: Springer; 2009.
Matthews JR, Matthews RW. Successful scientific writing. New York: Cambridge University Press; 2008.
Zeiger M. Essentials of writing biomedical research papers. New York: McGraw Hill; 2000.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 re-*

quirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors' Disclosures of Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:*

Employment or Leadership: T.M. Annesley, *Clinical Chemistry*, AACC.

Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Answer to Learning Exercise

BACKGROUND: Serum concentrations of the vascular inflammation marker β-selectin correlate with atherosclerotic disease severity, but β-selectin has a large intraindividual variation. We investigated whether interleukin-6 (IL-6), another marker of vascular inflammation, could predict disease severity and mortality risk.

METHODS: Consecutive outpatients undergoing evaluation for peripheral vascular disease (PVD) were divided into categories ranging from no functional impairment (group 1) to severe functional impairment (group 4). Blood was collected at baseline and quarterly over 3 years. Serum IL-6 and β-selectin were quantified to calculate intraindividual variation and to assess the relationships of these markers to disease severity and mortality.

RESULTS: Baseline median IL-6 concentrations were 12, 26, 96, and 144 μg/L for categories 1 to 4, respectively ($P < 0.001$ for categories 3 and 4 vs 1) and were not found related to age or sex. Median β-selectin concentrations increased 30% across the 4 categories. Increased disease severity and mortality were associated with higher IL-6 concentrations ($P < 0.01$ for both), but not β-selectin. Intraindividual variation for group 1 was 14% for IL-6 and 36% for β-selectin.

CONCLUSIONS: IL-6 appears to be a better marker of disease severity and mortality than β-selectin in patients with PVD, with lower intraindividual variation and significant concentration changes with increasing disease severity.

Clinical Chemistry 56:5
708–713 (2010)

"It was a cold and rainy night": Set the Scene with a Good Introduction

Thomas M. Annesley

In theatrical productions, there is a process called *setting the scene*, which is the act of describing a situation so that the audience understands what is happening. Setting the scene lays the groundwork for what to expect during the remaining acts in the production. Similarly, a well-written introduction in a scientific paper sets the scene for the reader. It starts by telling the reader what is happening or has happened (the context), and ends by giving the reader a glimpse of what follows in the remainder of the article (the plot).

Introductions seemingly should be easy to write, since they do not require details about methods and results or a discussion of the results. Besides, the introduction is usually found right after the abstract, where you already summarized the content for the reader. In actuality, however, writing a good introduction requires considerable time and thought. Here I provide information about the structure of a good introduction and how to avoid common problems that editors see with submitted manuscripts.

The Conical Introduction

Introductions have shapes. Some individuals see them as funnels, others as cones or inverted pyramids. Whatever image you choose should go from large to small, broad to narrow. This is how the information in the introduction should flow as well (Fig. 1). Begin by providing the reader with background information on the topic of the paper. Describe what is known about a disease, technique, or compound and why it is an important topic. Do not be concerned if this takes several sentences. If there is a certain amount of background information that you believe the reader must have to follow the remainder of the article, include it. But make sure that the background information directly relates to your specific study. For example, if you are reporting on a new marker for pancreatic cancer, do not devote unnecessary text to the epidemiology, therapy, life expectancy, medical costs, etc., of cancer in general. Get to the known information about pancreatic cancer as soon as possible.

Having presented relevant background information, the next step is to narrow the introduction and focus the reader's attention on the importance of continued research on particular aspects. Tell the reader about needed but unknown information, an unsolved problem, a knowledge gap, or limitations of prior studies. There may be a lack of a good analytical technique or the availability of a new animal model. Perhaps no one recognized the problem before now or tied the literature together to identify a possible solution. The important goal here is to demonstrate to the reader that there are important missing pieces of the puzzle that need to be filled in. Using the analogy of a theatrical production, you should set the scene by putting the necessary background information into the proper context.

Now the introduction can be narrowed again by focusing on the goal of your study (the plot). From this point on, the text must provide a clear rationale for why you initiated the study. The reasons for doing research are limited. You test a hypothesis, answer a question, solve a problem, or fulfill a purpose. The text should include something like the following:

- *We hypothesized that . . .*
- *We tested the hypothesis that . . .*
- *We asked whether . . .*
- *To answer this question, . . .*
- *This prompted us to investigate whether . . .*
- *To resolve this apparent difference . . .*
- *We solved this problem by . . .*
- *The purpose of our study was . . .*

Importantly, this type of presentation tells the reader to expect a clear answer by the end of the article regarding the study goals or hypothesis—i.e., true/false, yes/no, works/doesn't work.

Optionally, some writers choose to add a short concluding sentence or two telling the reader something about the approach taken, the plan of attack, or the proposed solution in the paper and its importance. If included, however, my recommendation is that you not provide method details, results, or conclusions. The reader should already have had a brief exposure to these items through the abstract.

Department of Pathology, University of Michigan Health System, Ann Arbor, MI. Address correspondence to the author at: Department of Pathology, University of Michigan Health System, Room 2G332, 1500 E. Medical Center Dr., Ann Arbor, MI 48109-5054. E-mail annesley@umich.edu.
Received January 13, 2010; accepted January 19, 2010.
Previously published online at DOI: 10.1373/clinchem.2010.143628

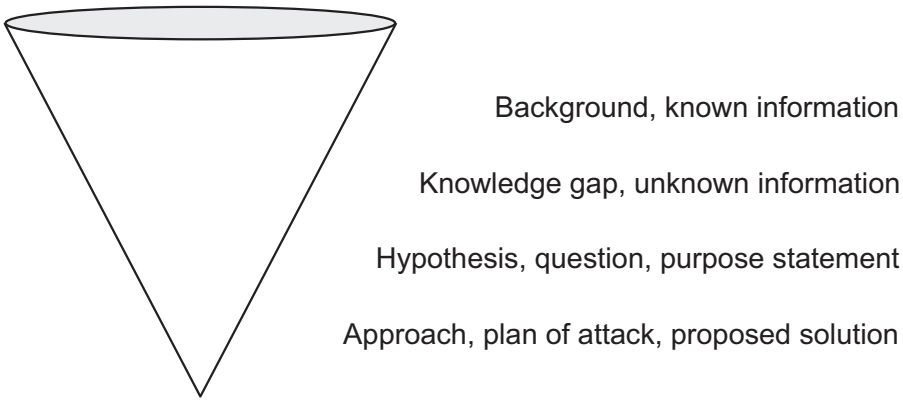


Fig. 1. The introduction: a cone or funnel.

Example 1 provides the introduction for a hypothetical study of a biomarker in vascular inflammation. Compare the format of this introduction to the cone concept in Fig. 1. The first sentence (top of the cone) tells the reader that the study relates to the broader topic of cardiovascular disease, which is an important health problem. The next sentence narrows the topic to chronic inflammation, which is linked to cardiovascular disease, followed by a sentence that focuses the topic further to β-selectin, a marker of inflammation that is increased in the serum of patients with peripheral vascular disease. Besides referenced association studies,

Example 1

Cardiovascular disease is a public health problem worldwide. Chronic inflammation has been linked to cardiovascular disease and sudden cardiac death *(1–3)*. Recent studies have demonstrated that a strong association exists between β-selectin, a recognized systemic marker of inflammation, and cardiovascular disease *(4–6)*, and that patients with peripheral vascular disease have increased serum concentrations of β-selectin that correlate with the degree of functional impairment *(7,8)*. Additionally, prospective studies have presented data regarding the prognostic value of β-selectin in predicting the severity of underlying cardiovascular disease and risk of mortality. The Vascular Inflammation Prediction (VIP) Study found a positive correlation between concentrations of β-selectin and the risk of developing cardiovascular disease *(9)*. The Canadian All-Cause Mortality Study revealed that individuals with serum β-selectin concentrations >90 μg/L are 4.5 times as likely to die within 5 years as those with concentrations ≤90 μg/L *(10)*.

Whereas these association and prospective studies indicate that β-selectin is a predictor of cardiovascular disease and risk of mortality, they provide little information about the underlying pathophysiology of vascular inflammation and the contributory role, if any, of β-selectin.

We therefore investigated in an animal model whether β-selectin is a cause or just a marker of vascular inflammation associated with cardiovascular disease.

Using a herpes simplex virus type 2 infection protocol to stimulate continuous β-selectin production in mice, we investigated the effects of β-selectin production on the development of atherosclerotic lesions, life span, and potential mechanisms of β-selectin–induced inflammation.

Example 2[1]

Because of the speed and selectivity that it affords compared to other techniques, liquid chromatography (LC) coupled with electrospray ionization–tandem mass spectrometry (ESI-MS/MS) is being increasingly used in clinical laboratories to quantify steroids *(1)*, therapeutic drugs *(2–4)*, vitamins *(5)*, biogenic amines *(6)*, and metabolic intermediates *(7–9)*. One use of LC-ESI-MS/MS in our laboratory is for quantifying immunosuppressants in whole blood and serum. We use methanol for sample preparation and chromatography because it is readily available and less costly than acetonitrile. Methanol is a mobile-phase component in multiple published methods for immunosuppressants *(2, 10–14)*.

When quantifying immunosuppressants by this approach, we encountered the problem of a slow loss of 32-desmethoxyrapamycin, the internal standard for sirolimus, if the methanolic working solution was stored at ambient temperature. We presumed that this loss resulted from degradation of 32-desmethoxyrapamycin in the methanol being used, an effect similar to that reported for the internal standard ascomycin in some brands or grades of acetonitrile *(15)*. In the course of investigating whether alternate commercial sources and grades of methanol would correct the loss of 32-desmethoxyrapamycin, and also be suitable for use in the mobile phase, we noted large differences in the ionization of not just 32-desmethoxyrapamycin, but also other immunosuppressants and their internal standards when different sources and grades of methanol were evaluated.

Coeluting components originating from biological matrices have previously been shown to negatively affect (ion suppression) or positively affect (ion enhancement) the analyte signal in ESI-MS analyses. This report describes the phenomenon on ionization changes related to the organic solvent used in the LC-ESI-MS/MS analysis.

[1]Modified from Clin Chem 2007;53:1827–34.

the introduction next emphasizes that 2 major prospective studies have found a positive correlation between β-selectin and actual cardiovascular risk. The first paragraph has provided background and known information from citable work, all the way demonstrating to the reader the importance of β-selectin as a subject of research. The second paragraph (narrower section of the cone) presents the unknown information (knowledge gap) that previous work has failed to address. Even without the question being explicitly stated, the reader can begin to deduce what the study question will be. The third paragraph (even closer to the tip of the cone) narrows the focus to the question itself and the purpose of the study. Is β-selectin a contributing factor or just a marker of cardiovascular disease? At the minimum, there will be a yes/no answer. The last paragraph of the introduction gives the reader some clue as to how the study was performed, i.e., using a viral infection model in mice. No method details are provided, no results provided, no conclusions stated. Overall, this introduction follows the model in Fig. 1.

Transition Phrases

In the introduction, the story becomes clearer if transition phrases are used. Transition phrases allow the author to emphasize important points, and also help the reader differentiate the known, the unknown, the question, and the experimental approach. I previously listed some examples of ways to lead into the question or hypothesis. Examples of transition phrases that can be used to highlight the known, or link the known to the unknown, are shown below:

- *These prior studies show that . . .*
- *Supporting the theory that . . .*
- *These studies are important because . . .*
- *Interestingly, . . .*
- *More importantly, . . .*
- *Using this information, . . .*
- *Yet, . . .*
- *Unlike . . .*
- *Whereas it has been shown that . . .*
- *On the other hand, . . .*
- *It is unclear . . .*
- *The question remains, however, . . .*
- *Although prior studies demonstrated . . .*

Different Study Types, Same Model

Many published articles describe a new method or a secondary finding that sheds new light on a topic.

These types of studies were not initiated to directly answer a question or test a hypothesis, yet they had a purpose that should be described in the introduction. Regardless of the type of study, the same process of honing down to the problem to be addressed, as illustrated in Fig. 1, can be followed when writing the introduction. Example 2 and the learning exercise at the end of the article illustrate ways that this can be done. Example 2 is a modified introduction from an article describing the finding that, in addition to biological matrices, solvents can impact the performance of mass spectrometric assays. The introduction starts with the broad topic of electrospray ionization mass spectrometry, describing what advantages it has and how clinical laboratories are successfully using this technique, subsequently narrowing the subject to a specific assay that served as the origin for the study. This is the known information. The second paragraph, while not directly describing a knowledge gap or problems with previous studies, brings a previously unknown problem to the attention of the reader. It also indirectly introduces the question: Does the quality of solvents have any significant impact on ionization efficiency in electrospray mass spectrometry? The last 2 sentences (paragraph 3) close the introduction by stating the purpose of the paper and what new information is going to be provided. This introduction, although different in style, narrows from known information to a previously unknown problem to the specific purpose of the paper.

Length, Detail, and Overlap

Introductions tend to be too long rather than too short. A long introduction reminds me of the courtroom scene in a television show, where an attorney keeps feeding statements to a witness until the frustrated judge asks, "Counselor, is there a question in there somewhere?" Similarly, in sifting through what was intended to be an impressive overview of the topic and the issues, the reader may fail to appreciate the question when it finally arrives.

There are several ways to avoid giving too much information. One is to characterize the audience of the selected journal. Ask yourself, "If I were the reader, how much information would I really need to understand the study question and why it matters?" Another way to avoid excessive length is to go back in time only as far as needed to bring the reader up to speed. Unless being cited as influential work in the field, is mention of older work or an older reference necessary? A third way is to set a target word limit before linking the known information to the unknown information, and then similarly the unknown information to the study question. The fourth way is to consider whether some of the information or associated references might fit better into the discussion section, where you are interpreting your results and their relevance.

The last option above calls attention to a couple of problems that editors encounter in submitted papers: (*a*) unnecessary overlap of the introduction and discussion sections and (*b*) inconsistencies between these 2 sections. A few brief sentences at the beginning of the discussion help reorient the reader to the purpose of your study and your findings, but you should try to keep background or reference material in one section or the other, not both. Repetition between the sections not only wastes words, but also can create the impression that you had little to discuss in the paper and thus reused background information to fill space. As you try to interpret your study results and put them in context with other studies, you may find that some background or reference material fits better in the discussion than in the introduction. This gives you the opportunity to link specific results or points of discussion to others' work, citing their work where it makes most sense.

Consistency with Other Sections

Although you want to minimize repetition, it is important that the text be consistent among all of the sections of the paper. Background information, knowledge gaps, purpose statements, and proposed solutions in the abstract should be consistent with the introduction. The methods used must reflect any mentioned in the introduction. The results must relate to the study question, hypothesis, or problem first presented in the introduction. The discussion, or summary if separately written, must answer the question posed in the introduction. Sometimes reviewers and editors request changes to the text, restatement of the question or problem, reinterpretation of results, or modification of the conclusions. Thus, it is a good idea to take a fresh look at the introduction after the final draft is written or after any revision to be certain that it is still accurate and consistent with the rest of the article.

Learning Exercise

Below I have provided 10 sentences that together make up an introduction for a paper describing a new method. Using the concept for writing an introduction shown in Fig. 1, rearrange the sentences to create an introduction. Compare your final product with the one provided in the box after the list of selected additional reading materials.

- Iohexol is not bound to serum proteins and is filtered through the glomerulus, with no identifiable reabsorption or tubular secretion, making it an ideal marker for estimating GFR.

- Ultraperformance liquid chromatography (UPLC), a recently introduced modification of LC, allows rapid chromatography owing to faster gradient curves, as well as the potential to use smaller particles and higher flow rates.
- Protocols have been developed that involve a single intravenous injection of iohexol followed by timed blood collections.
- Iohexol is an iodinated contrast dye that has been shown to be useful in clearance studies for the determination of GFR.
- We have combined these 2 techniques to develop a UPLC-MS/MS assay for iohexol in human serum that uses a simple sample preparation, a structural analog internal standard with the same retention time, and a ballistic gradient for rapid chromatographic analysis.
- Both of these techniques require lengthy run times to separate iohexol from endogenous interfering compounds and the internal standards.
- In the subset of patients with suspected renal insufficiency for whom it is important to have an accurate assessment of glomerular filtration rate (GFR), clearance measurements provide the best information.
- No urine collection or quantification in urine is necessary, an advantage over iothalamate, the other agent used for GFR studies.
- By comparison, the high selectivity of tandem mass spectrometry (MS/MS) as a detector generally allows simpler specimen cleanup and shorter chromatographic times compared with UV detection.
- The majority of published methods for quantifying iohexol have used capillary electrophoresis or gradient elution liquid chromatography (LC) coupled with ultraviolet (UV) detection.

Final Thoughts

When you introduce an important speaker, you want to give a "proper introduction." This usually entails telling the audience about the speaker's background, area of research, and the topic to be presented. If you go on and on about the speaker's background and accomplishments, or spend too much time talking about how you came to know this individual, or forget to reinforce the topic of the lecture, by the time the speaker says a word the audience may have trouble recalling why they were there in the first place. We have all suffered through such introductions. Your own work must be important to you; otherwise you would not want others to read it. So give it a proper introduction as well, using the tips and ideas that have been presented.

Suggested Additional Reading

Friedman GD. Please read the following paper and write this way! Am J Epidemiol 2005;161:405.

Katz MJ. From research to manuscript. New York: Springer, 2009.

Zeiger M. Essentials of writing biomedical research papers. New York: McGraw Hill, 2000.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures of Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:*

Employment or Leadership: T.M. Annesley, *Clinical Chemistry*, AACC.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Answer to Learning Exercise[1]

In the subset of patients with suspected renal insufficiency for whom it is important to have an accurate assessment of glomerular filtration rate (GFR), clearance measurements provide the best information. Iohexol is an iodinated contrast dye that has been shown to be useful in clearance studies for the determination of GFR *(1)*. Iohexol is not bound to serum proteins and is filtered through the glomerulus, with no identifiable reabsorption or tubular secretion, making it an ideal marker for estimating GFR. Protocols have been developed that involve a single intravenous injection of iohexol followed by timed blood collections *(2–4)*. No urine collection or quantification in urine is necessary, an advantage over iothalamate, the other agent used for GFR studies.

The majority of published methods for quantifying iohexol have used capillary electrophoresis or gradient elution liquid chromatography (LC) coupled with ultraviolet (UV) detection *(5–9)*. Both of these techniques require lengthy run times to separate iohexol from endogenous interfering compounds and the internal standards. By comparison, the high selectivity of tandem mass spectrometry (MS/MS) as a detector generally allows simpler specimen cleanup and shorter chromatographic times compared with UV detection. Ultraperformance liquid chromatography (UPLC), a recently introduced modification of LC, allows rapid chromatography owing to faster gradient curves, as well as the potential to use smaller particles and higher flow rates. We have combined these 2 techniques to develop a UPLC-MS/MS assay for iohexol in human serum that uses a simple sample preparation, a structural analog internal standard with the same retention time, and a ballistic gradient for rapid chromatographic analysis.

Comment: Following the cone shape model, this introduction narrows from known information to a problem to a solution to the problem. The first paragraph provides a general overview of iohexol and why it has advantages in GFR evaluations. The second paragraph narrows the focus to published methods for quantifying serum iohexol and their drawbacks. Although unknown information or an unsolved problem is not directly stated, the need for an improved assay is implied. The introduction closes with a proposed solution to the problem. Is there a need to propose a hypothesis or ask the question of whether mass spectrometry could be used to quantify iohexol? One could do so, but the answer must be yes, otherwise there would be no need to report the new assay.

[1] Modified from Clin Chem 2009;55:1196–202.

Clinical Chemistry 56:6
897–901 (2010)

Who, What, When, Where, How, and Why: The Ingredients in the Recipe for a Successful Methods Section

Thomas M. Annesley*

In a prior article on abstracts I discussed the need, when writing a paper, to tell the story by answering questions. If I extended that concept by asking what question(s) the Methods section of a paper answers, the first response that comes to my mind is "How did I perform the study?" Yet "how" is just one of the main ingredients in the recipe for a successful Methods section. An informative Methods section also starts with 1 part *what*, 1 part *when*, 1 part *where*, 1 part *who*, and 1 part *why*. As with any recipe, the proportions of each can be modified to taste, depending on the study type and journal format, but each must be added lest someone notice that something seems to be missing from the final product.

The Methods section is also called the Materials and Methods, Patients and Methods, Study Design, or Experimental section. The goals of this section are to allow readers to (*a*) understand how and why the experiments were performed, (*b*) better understand the remainder of the paper and how the results and conclusions derived from the experiments, (*c*) be able to reproduce the study with an expectation of success, and (*d*) acknowledge that the results and conclusions are valid based on the strength of the methods and study design. Making sure to include the important details about *who*, *what*, *when*, *where*, *how*, and *why* in the study can help achieve these goals. I have listed in Table 1 some example questions that, depending on the study, might be important to answer for the reader. In the remainder of this article I discuss other important ingredients that can help you develop a winning recipe for your Methods sections.

Length and Detail

Although the Methods section should not read like a procedure manual or cookbook, it is the one part of a research paper for which length (word count) is a secondary consideration after clarity and adequate detail. As long as you help readers reach the goals listed above, your Methods section should be as long as necessary to describe the important experiments in your study.

The importance of each question in Table 1 and the amount of detail required can vary depending on the type of study and the target audience. For example, if you are comparing 2 analytical methods for quantifying serum human chorionic gonadotropin and need to identify specimens from healthy individuals, pregnant women, or patients with renal failure or cancer, you might rely on the existing medical record and the prior judgment of several clinicians. Detailing where the diagnosis was made (e.g., clinic vs hospital) or who made the diagnosis (e.g., attending physician vs resident) becomes less important to the reader than detailing what protocols were followed to compare the assays, where the analyses were performed, and what instruments were used.

By comparison, if you are doing a clinical study for which a diagnosis, histopathology interpretation, or response to treatment is the major outcome, then who made the diagnosis or what diagnostic criteria were used become key details compared with who in the central laboratory tested the patient's blood or what was the analytical principle behind the commercial method used. In the second example you still want to state that the testing was performed in the central laboratory on a specified analyzer, but no additional details are needed.

Errors of omission (insufficient detail) are common in Methods sections. Experimental conditions and details sometimes become self-obvious to authors and may be unintentionally left out. One way to avoid leaving out important detail is to treat the first draft as if it were a standard operating procedure used for training individuals about the analyses, diagnostic criteria, drug preparation, or even surgery used in the research study. If you consider what details would cause the experiment to fail if left out, you may decide that some details (e.g., wearing latex gloves, the brand of pipette, where reagents are stored in the laboratory, type of sutures) are relevant only to your facility and do not need to be included in the final paper. But you may discover that you forgot to include something as simple, yet crit-

University of Michigan Health System, Ann Arbor, MI.
* Address correspondence to the author at: University of Michigan Health System, Room UH2G332, 1500 East Medical Center Drive, Ann Arbor, MI 48109-5054. E-mail annesley@umich.edu.
Previously published online at DOI: 10.1373/clinchem.2010.146589

Table 1. Who, what, when, where, how, and why questions to consider when writing the Methods section.

Who
Who maintained the records? Who reviewed the data? Who collected the specimens? Who enrolled the study participants? Who supplied the reagents? Who made the primary diagnosis? Who did the statistical analyses? Who reviewed the protocol for ethics approval? Who provided the funding?
What
What reagents, methods, and instruments were used? What type of study was it? What were the inclusion and exclusion criteria for enrolling study participants? What protocol was followed? What treatments were given? What endpoints were measured? What data transformation was performed? What statistical software package was used? What was the cutoff for statistical significance? What control studies were performed? What validation experiments were performed?
When
When were specimens collected? When were the analyses performed? When was the study initiated? When was the study terminated? When were the diagnoses made?
Where
Where were the records kept? Where were the specimens analyzed? Where were the study participants enrolled? Where was the study performed?
How
How were samples collected, processed, and stored? How many replicates were performed? How was the data reported? How were the study participants selected? How were patients recruited? How was the sample size determined? How were study participants assigned to groups? How was response measured? How were endpoints measured? How were control and disease groups defined?
Why
Why was a species chosen (mice vs rats)? Why was a selected analytical method chosen? Why was a selected experiment performed? Why were experiments done in a certain order?

ical, as the pH of a buffer, the need to perform sample preparation in glass vs plastic, or an antibiotic drip during the surgery.

Of course, errors of commission (irrelevant detail) can detract from the Methods section as well, and you must make sure to avoid adding information that can be cited and found elsewhere. For example, if you used a previously published method without modification, it is sufficient to reference the method and its principle (e.g., "We used the LC-MS/MS method of Anderson for quantifying testosterone."). However, if you modified the published method in any way, then it is imperative that you include details of any modifications that were made (and why). Or if in a clinical study you compared continuous secondary endpoints with use of the Wilcoxon rank-sum test, there is no need to provide a detailed description that this test is a nonparametric alternative to the 2-sample *t*-test for assessing whether 2 independent samples of observations come from the same distribution. Readers can access information about this statistical test elsewhere.

Style and Format

The Methods section should be divided into subsections with associated subheadings. The use of subheadings helps organize the material in the reader's mind. When the materials used in the study are described, 3 formats are appropriate—as a listing under the subheading of reagents and supplies, as part of the description of an individual experiment, or both. Generic reagents, such as solvents, chemicals, and buffers, that are used throughout the study or in multiple places in the study protocol can be listed in a subheading labeled as Materials. However, if a reagent is specific to an individual experiment or method, such as a PCR experiment, then the reagents, enzymes, etc. used solely for the PCR should be listed in the paragraph detailing the PCR experiment, thereby helping the reader associate the importance of specific reagents with specific experiments. Be sure to include the source or vendor for all chemicals, reagents, animals, and instruments used in the study. Some journals also request that the location of the vendor be included the first time that the material is mentioned.

The Methods section should be written in the past tense because you are describing experiments and protocols that you did in the past:

- *The experiment was performed at room temperature.*
- *We quantified the drug by immunoassay.*
- *Assay correlation was determined by Spearman rank correlation.*
- *We performed a 2-way ANOVA.*
- *Study participants were recruited from the blood donor service.*
- *Human embryonic kidney cells were cultured in Dulbecco's modified Eagle's medium.*
- *Tissue release of C-reactive protein was monitored after blood flow restriction.*

The 2 exceptions to the use of the past tense are the presentation of summarized data and the introduction of figures and tables within the Methods section:

- *Our protocol is summarized in Figure 1.*
- *Figure 1 illustrates the steps in the procedure.*
- *The data are summarized as median and interquartile range.*
- *The results from these analyses are presented as relative risk ratios with 95% CIs.*

Table 2. All passive voice, all active voice, and combined use of both.

All passive voice:
Study participants were recruited from the blood donor clinic. Each participant was asked to fill out a questionnaire, which was then used to classify each individual. Heart rate, blood pressure, and temperature were evaluated by a nurse for each individual, at which time a 10-mL blood sample was obtained from each participant by a phlebotomist for routine chemistry testing. This information was compiled and used to create a database of patient characteristics vs laboratory results.
All active voice:
We recruited study participants from the blood donor clinic. We asked each participant to fill out a questionnaire, which we then used to classify each individual. A nurse evaluated each individual for heart rate, blood pressure, and temperature, at which time a phlebotomist obtained a 10-mL blood sample from each participant for routine chemistry testing. We compiled this information and created a database of patient characteristics vs laboratory results.
Combined passive and active voice, irrelevant information removed:
We recruited study participants from the blood donor clinic. Each participant was asked to fill out a questionnaire, which we then used to classify each individual. Heart rate, blood pressure, and temperature were evaluated, and a 10-mL blood sample obtained for routine chemistry testing. We compiled this information and created a database of patient characteristics vs laboratory results.

Resources on writing differ in their preference for either the passive or the active voice for the Methods section. Either style is acceptable if used properly, but a combination of both provides the reader with some variation in presentation of the experiments. Whatever form you choose, be sure to avoid monotony in the presentation, as shown in Table 2. The first 2 examples in Table 2 also include potentially irrelevant information such as who measured heart rate and blood pressure, and who obtained the blood samples. Another way to avoid monotony when using predominantly the passive or active voice is to use transition phrases:

- *After mixing for 1 minute, we added 7 mL methylene chloride . . .*
- *To evaluate ion suppression, extracts were prepared . . .*
- *Because the cells did not adhere to polypropylene, fibroblasts were grown . . .*
- *Because of its signal-enhancing properties, cinnamic acid was added . . .*
- *Based on previous reports of calcitrol-induced inhibition, we added calcitrol . . .*

In addition to avoiding monotony and improving the flow of the text, transition phrases can be used to start a new paragraph, introduce a new experiment, or describe for the reader why the experiment was performed.

The Methods section should present the experimental procedures in chronological order. An exception can be made to using this format if all of the experiments were performed independently of one another, with no clear ordering of their performance. In this situation the experiments can be ordered from most to least important, helping to ensure that the most important experiment is brought to the attention of, and retained by, the reader.

It is important that the details for specific experiments be presented chronologically as well. I remember one set of instructions for an instrument repair that stated something like, "Turn the luer lock counterclockwise to unlock and remove the valve. But first, bleed the gas to remove the backpressure so the valve does not fly outward and potentially cause injury." No kidding! Similarly, chronology in sentence structure serves both the author and reader well. Instead of writing "the supernatant was transferred to another tube after centrifugation at 8800*g* for 10 minutes," the actual sequence of events should be written as "after centrifugation at 8800*g* for 10 minutes, the supernatant was transferred to another tube." Similarly, "we obtained a 3-mm punch biopsy sample after the patient gave informed consent" is better stated as "after the patient gave informed consent, we obtained a 3-mm punch biopsy sample."

Tables and figures should be included in the Methods only if they will save a large amount of text, and be of clear benefit in helping the reader understand the experiment being described. For example, in a methods paper you may have a large number of assay parameters that must be summarized (e.g., gradient conditions, mass transitions, voltage settings, detector settings, and programmed instrument changes). Describing these one after another in the text may result in a cumbersome paragraph of numbers and terms that would be more easily understood if summarized in a table. In another situation you may be describing a complex workflow protocol that is better understood as a schematic diagram. Nevertheless, the circumstances justifying a table or figure in the Methods section are few in number.

The decision of whether to put information in the Methods section vs the Results section can be confusing. The general rule is that anything known or planned at the beginning of the study goes in the Methods section, and anything that was not known or planned goes in the Results section. In some types of studies, however, the initial experiments described in the Methods section may yield data that lead to a change in subsequent experiments or to additional experiments. Because these later experiments are driven by data ob-

tained during the course of the study, the description of these experiments may make more sense if included in the Results section along with the corresponding results:

When we evaluated the data, we noted an apparent bimodal distribution related to sex. Because the original patient data set included 13 women and 47 men, we increased the number of samples obtained from women to 45 to confirm whether there was indeed a sex difference. Statistical analysis of the expanded data set (45 women and 47 men) confirmed a bimodal distribution [median (interquartile range) of 36 (14) mg/L for women and 61 (23) mg/L for men].

Lastly, make sure that the Methods section is consistent with all of the other sections in the final version of your paper. Is there an important method or experiment that is missing in the Abstract? Is there a method or experiment listed in the Abstract that is missing in the Methods section? Are there corresponding results in the Results section to match each method or experiment included in the Methods section? Is there an explanation, either in the Methods section or the Discussion, as to why each experiment was performed? As stated at the beginning of this article, you don't want a missing ingredient or the wrong ingredient to affect your final product.

Learning Exercise

Answer the following questions about the Methods section:

1. What are the questions that are answered in a Methods section?
2. Should the Methods section be written in the past, present, or future tense?
3. How does sentence structure differ between the passive and active voice?
4. In what ways are transition phrases helpful?
5. In what order are subsections organized in the text?
6. Are figures and tables allowed in the Methods section?

Final Thoughts

In Act II of William Shakespeare's *Hamlet*, Polonius states, "Though this be madness, yet there is method in it." This statement has evolved into the modern phrase, "method to one's madness," meaning a rational plan that is hidden by a mysterious action, or a strange plan that manages to yield results. This strategy may have worked for Polonius, but will not work in a scientific paper. Poorly described experiments will trump the credibility of your results. If readers cannot understand how and why the experiments were performed, they will be hesitant to acknowledge the results and conclusions as valid. So make your Methods section work for you, not against you.

Resources and Additional Reading

Foote MA. Materials and methods: a recipe for success. Chest 2008;133:291–3.
Iles RL. Guidebook to better medical writing. Olatha, KS: Iles Publications; 2003.
Katz MJ. From research to manuscript. New York: Springer; 2009.
Lang TA. How to write, publish, and present in the health sciences. Philadelphia: ACP Press; 2010.
Van Damme H, Michel L, Ceelen W, Malaise J. Twelve steps to writing an effective materials and methods section. Acat Chir Belg 2007;107:102.
Zeiger M. Essentials of writing biomedical research papers. New York: McGraw Hill; 2000.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures of Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:*

Employment or Leadership: T.M. Annesley, AACC.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Answers to Learning Exercise

1. Who, what, when, where, how, and why.
2. Past tense, except for presenting summarized data and the introduction of figures and tables.
3. Active voice—the subject of the sentence performs the action (acts upon something). Example: Harold delivered the flowers.

   Passive voice—the subject of the sentence receives the action (is acted upon). Example: The flowers were delivered by Harold.
4. Help the flow of the text, introduce a new experiment, or describe why an experiment was performed.
5. Chronological order or order of importance.
6. Yes, if they save a large amount of text and help the reader understand the experiment being described.

Clinical Chemistry 56:7
1066–1070 (2010)

Show Your Cards: The Results Section and the Poker Game

Thomas M. Annesley*

In 5-Card Draw, one of the most popular versions of poker, you start with a specific question: "Can I win with the cards I have decided to play?" The final answer is yes or no. After looking at your initial cards (initial findings) you can be satisfied with what you have (preliminary data) or seek some new cards (new experiments). But in the end you must openly "show your cards" (results). Your cards give you the answer. You cannot hide a card, nor can you add an undealt card to make your hand look better. Playing poker and writing the Results section of a scientific paper have similarities, as I will point out in this article.

Presenting Your Results

In poker, how you present your cards affects how your competitors grasp the importance of the cards. One winning set of cards in poker is the *straight*, defined as 5 consecutively sequenced cards (e.g., 6, 7, 8, 9, 10). You may have this group of cards, but if you present them as 6, 10, 8, 7, 9, your straight is not immediately evident. The worth of cards when presented in a logical manner is clearer and easier to grasp. The same holds true for your Results section. Your important results may be better understood if presented in a certain order.

There are several options for the presentation order of results (Table 1); one may work better than another for the type of study being reported. The most straightforward approach is to use a chronological order with subheadings that parallel the methods and their sequence presented earlier in the paper. This order allows readers to more easily go back and refer to the methods associated with a given result.

A second approach is to group results by topic/study group or experiment/measured parameter. An example of this format is a comparison of the diagnostic and analytical performance of 3 assays for serum prostate-specific antigen. If grouped by assay as the *topic*, the results for diagnostic accuracy, analytical performance, interference testing, and cost analysis for assay 1 would be presented first, followed by a separate presentation of the same results for assay 2 and then assay 3. This order allows the reader to see the results for each assay as a packet of information, which is a logical way to remember information. By comparison, if the results are grouped by *measured parameter*, important similarities or differences in assay performance may be clearer and can be emphasized as important findings.

Grouped by topic:
Assay 1: diagnostic accuracy, performance, interferences, cost.
Assay 2: diagnostic accuracy, performance, interferences, cost.
Assay 3: diagnostic accuracy, performance, interferences, cost.

Grouped by measured parameter:
Diagnostic accuracy: assay 1, assay 2, assay 3.
Performance: assay 1, assay 2, assay 3.
Interferences: assay 1, assay 2, assay 3.
Cost: assay 1, assay 2, assay 3.

In clinical studies that involve multiple groups of individuals or patients receiving different treatments, it is common to order the results from general to specific. The characteristics of the overall study population, such as sex and age distribution, initial and final numbers in each group, and dropouts are first presented. This information is followed by the data and results for each specific group, i.e., starting with the control group or the group receiving the standard treatment, followed by the results for the disease group or the group receiving the experimental treatment. Lastly, if you undertook a study for which the order in which the results are presented is not critical to their being understood, presenting the results from most to least important immediately highlights the results you want to emphasize.

Results should be presented in the past tense. The Results section usually ends up heavier on the passive voice, but some conscious use of the active voice can help the flow and readability of the text (e.g., "we observed that the 2 groups" versus "it was observed that the 2 groups").

University of Michigan Health System, Ann Arbor, MI.
* Address correspondence to the author at: University of Michigan Health System, Rm. UH2G332, 1500 East Medical Center Drive, Ann Arbor, MI 48109-5054. Fax 734-763-4095; e-mail annesley@umich.edu.
Received April 7, 2010; accepted April 16, 2010.
Previously published online at DOI: 10.1373/clinchem.2010.148148

Table 1. Options for presentation order of results.

1. Chronological order
2. Grouping by topic or experiment
3. General to specific
4. Most to least important

Data and Results Are Not the Same

One valuable lesson I learned about writing a well-crafted Results section came from Zeiger's book, *Essentials of Writing Biomedical Research Papers*. The same concept—namely, that data and results are not the same—was discussed more recently in an article by Foote in the journal *Chest* (see Resources and Additional Reading). Authors can err by offering the reader results but no data, or data but no results. Data are facts and numbers. Data are usually presented in tables and figures as raw data (individual data points) or summarized data (mean, percent, median and range). Results are statements in the main text that summarize or explain what the data show. As an example, let's use a hypothetical study comparing the effectiveness of radiation treatment, chemotherapy with an existing drug (Blasteride), and a new monoclonal antibody-based therapy (Neuroxomab) for the treatment of neuroblastoma. One of the endpoints in the study is survival rate after diagnosis and initiation of treatment (Figure 1). Four ways to present the information in Figure 1 for the reader might be as follows:

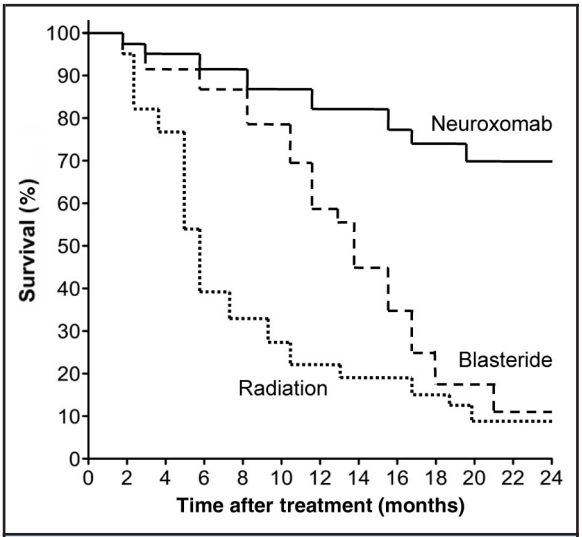


Fig. 1. Two-year survival rates for patients with neuroblastoma treated with Neuroxomab, Blasteride, and radiation.

Figure 1 shows the survival rates following diagnosis and initiation of treatment in the 3 treatment groups. At 6 months the survival rates were 95% for the Neuroxomab group, 91% for the Blasteride group, and 39% for the radiation-treated group. At 12 months the rates were 83%, 69%, and 23%;, at 18 months 74%, 17%, and 15%; and at 24 months were 70%, 11%, and 9%.

Figure 1 shows the survival rates following diagnosis and initiation of treatment in the 3 treatment groups. At 6 months the survival rates were significantly higher in the Neuroxomab and Blasteride treatment groups compared with the radiation-treatment group. At 12, 18, and 24 months the survival rates in the Neuroxomab group exceeded those of both the Blasteride and radiation-treatment groups.

Six months after diagnosis and initiation of treatment, the survival rates for the Neuroxomab and Blasteride groups were 2.4 and 2.3 times higher, respectively, than the radiation treatment group (both $P < 0.001$), but survival rates were not found to differ between the Neuroxomab and Blasteride groups ($P = 0.56$) (Figure 1). By 12 months, however, patient survival in the Neuroxomab group was 1.2 times higher than in the Blasteride group ($P = 0.031$), and 4.3 and 6.4 times higher at 18 and 24 months (both $P < 0.001$).

Six months after diagnosis and initiation of treatment, survival rates in the Neuroxomab and Blasteride groups (95% and 91%, respectively) were significantly higher than in the radiation treatment group (39%, $P < 0.001$ for both), but survival rates were not found to differ between the Neuroxomab and Blasteride groups ($P = 0.56$) (Figure 1). By 12 months, however, the patient survival rate in the Neuroxomab group was significantly higher than in the Blasteride group (83% vs 69%, $P = 0.031$), a difference that became even greater at 18 and 24 months (74% vs 17% and 70% vs 11%; both $P < 0.001$).

The first paragraph above provides data but no results. What do the data show? What is the point? Are the treatment groups statistically different at 6 months? The second paragraph contains results but no data. Is it clear from the figure how much higher the survival rates for patients in the Neuroxomab and Blasteride groups were compared with patients in the radiation group and with each other? What is the level of significance of any differences?

Paragraphs 3 and 4 above contain both data and results. They describe the important treatment differences and report when the differences occurred and whether they were statistically significant. Paragraph 3 states the magnitude (e.g., 2.4 times higher) of the most important differences between the treatments, and whether the differences were statistically significant. The reader must look at the figure to see the percent survival data, but this is perfectly fine as long as the

Table 2. Neuroblastoma survival rates over time for Neuroxomab, Blasteride, and radiation-therapy patient groups.

Time, months	Survival, %		
	Neuroxomab	Blasteride	Radiation
6	95[a,b]	91[a]	39
12	83[a,c]	69[a]	23
18	74[a,d]	17[e]	15
24	70[a,d]	11[e]	9

[a] $P < 0.001$ vs radiation group.
[b] $P = 0.56$ vs Blasteride.
[c] $P = 0.031$ vs Blasteride.
[d] $P < 0.001$ vs Blasteride.
[e] Not significant vs radiation group.

Table 3. Reporting guidelines for various types of studies.

Consolidated Standards of Reporting Trials (CONSORT: www.consort-statement.org/)
Enhancing the Quality and Transparency of Health Research (EQUATOR; www.equator-network.org/home/)
Metaanalyses of Observational Studies in Epidemiology (MOOSE; JAMA 2000;283:2008–12)
Minimum Information about a Microarray Experiment (MIAME; www.mged.org/Workgroups/MIAME/miame_2.0.html)
Minimum Information for Biological and Biomedical Investigations (MIBBI; mibbi.org/index.php/Main_Page)
Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE; Clin Chem 2009;55:611–22)
Preferred Reporting Items for Systematic Reviews and Metaanalyses (PRISMA; www.prisma-statement.org/)
Standards for the Reporting of Diagnostic Accuracy (STARD; www.stard-statement.org/)
Strengthening the Reporting of Observational Studies in Epidemiology (STROBE; www.strobe-statement.org/)

reader can fairly easily estimate the percentages at 6, 12, 18, and 24 months.

Paragraph 4 includes the actual survival rates (e.g., 95%, 91%, and 39% at 6 months) rather than the relative magnitudes of any differences. The inclusion of these survival-rate data in this paragraph is acceptable because the figure contains a lot of information and you are highlighting selected important differences. However, let's now say that the survival data and *P*-values had been provided in a table (Table 2). Because Table 2 contains the same information included in paragraph 4, you need not repeat this information in both places:

Six months after diagnosis and initiation of treatment, the Neuroxomab and Blasteride groups showed significantly higher survival rates compared with the radiation-treatment group (Table 2), but survival rates in the Neuroxomab and Blasteride groups were not found to differ. By 12 months, however, patient survival in the Neuroxomab group was significantly higher than in the Blasteride group, a difference that became even greater at 18 and 24 months.

This rule about nonrepetition of data is not absolute, but is a rule that should be broken only in rare circumstances. If a table or figure supplies a large amount of data, it is acceptable to restate a key piece of data in the text, such as the 2 groups in the table with statistically significant differences, if this helps the reader zero-in on an important result without having to plow through a long list of data.

State the Result, the Whole Result, and Nothing but the Result

In the American judicial system witnesses are sworn in by asking if they will tell the truth (the facts), the whole truth (tell everything), and nothing but the truth (no lies, conjecture, or interpretation). A complete Results section in a scientific paper also satisfies these requirements. Telling the facts is the easy part, because this is the goal of this section: to tell the reader what you found during the study. Requirements 2 and 3 above are areas in which authors can run into problems.

Satisfying the second requirement involves an intentional effort to include all data. There are well-crafted guidelines and checklists available that can help you meet the minimal standards for reporting data and results for many types of studies (Table 3). As an author you should use the checklists and flow diagrams in these guidelines when appropriate for your study. Doing so not only helps make the strengths, weaknesses, and sources of bias clear to the reader, but also helps you remember to include key data that otherwise inadvertently might have been omitted. For example, how many patients were excluded from the study? How many were lost to follow-up? How many dropouts were there? How many patients finished the study? How many individuals had an inconclusive result or diagnosis? These are all data and results and belong in the Results section.

Including all results also means not leaving out a negative result (hiding a card) or a result relevant to the report because it serves some other purpose for you as the author. Anyone who chooses to repeat your work or use your methods will likely encounter the same type of negative results that you did, and the fact that these were not acknowledged in your paper will not serve

you well. Referring to "unpublished results" annoys most editors and peer reviewers unless you can present a good argument for not including them. Trying to stake a claim to a future study by presenting an attention-grabbing preliminary result, but then not showing any corresponding data, can make readers question your motive.

The Results section is just that: results. To satisfy the third requirement above, this section should contain nothing but the results. No methods, no discussion. There is a temptation to remind the reader about the details of the experiment performed or the method used to generate the results, especially if it has been several pages since the Methods section ended. Method, study, and experimental details should not be restated in the Results section. Of course, you can refer to a specific experiment or method when describing the corresponding results; just do not repeat experimental details already described in the Methods section, as exemplified below. Although well intended as a link between a method and a result, the first 2 sentences of the next paragraph are unnecessary:

We compared the death rates for the 262 healthy controls with those of the 203 congestive heart failure patients over a 2-year period. Survival curves were generated with the Masterson mortality index formula. The congestive heart failure group was found to have a significantly higher short-term mortality rate.

However, this example is a good opportunity to illustrate how a transition phrase can serve as a link between a previously described experiment and a result without repeating what was in the Methods section:

When the 2-year survival curves for healthy controls and congestive heart failure patients were compared, the congestive heart failure group was found to have a significantly higher short-term mortality rate.

The only time that experimental details are appropriate for the Results section is when the initial experiments (rightly described in the Methods section) yield data that lead to additional experiments, not part of the original protocol, but which became necessary later on. The description of these experiments may make more sense if included in the Results section with the corresponding results.

When reporting results, authors feel an urge to comment on the results, e.g., how the results compared with prior work, were consistent with what was predicted in another paper, or explained the reason that a marker is increased in a disease. The interpretation or analysis of the results, however, belongs in the Discussion section. In the Results section you can describe what the data show, in the Discussion section you describe what the data mean.

"Significance" Is Misused a Significant Amount of the Time

The purposely incorrect heading here is meant to emphasize the fact that the terms *significant*, *significance*, and *significantly* are used erroneously in many submitted papers. In biomedical publications these terms are intended to identify relationships that have been statistically tested and determined unlikely to have occurred by chance. These terms should also be followed by a mathematical value or limit (e.g., $P = 0.067$ or $P < 0.001$). Unless you have such proof of statistical significance, you should use other terms such as *substantial*, *considerable*, or *noteworthy*. Similarly, authors like to draw unwarranted attention to nonsignificant findings by stating that the data "trended toward" or "tended to show." If the findings are not clear, don't try to imply something about them that cannot be supported.

Consistency of Results with Other Sections

Lastly, make sure that the Results section is consistent with all of the other sections in the final version of your paper. Is there a result that does not have a corresponding method or experiment in the Methods section? Conversely, is there a method or experiment for which you have reported no results? Is there a result not covered in the Discussion section, or discussion of a result not contained in the Results section? Are the most important results the same as those highlighted in the Abstract? Do the results relate to the study question, hypothesis, or problem first presented in the Introduction?

Learning Exercise

1. Point out which information is data and which is a result in the following paragraph:

Baseline median IL-6 concentrations were 12, 26, 96, and 144 μg/L for categories 1 to 4, respectively, and were not found related to age or sex. Median β-selectin concentrations increased 30% across the 4 categories. Increased disease severity and mortality were associated with higher IL-6 concentrations, but not β-selectin. Intraindividual variation for group 1 was 14% for IL-6 and 36% for β-selectin.

2. Choose whether the presentation of results in the following sentence is chronological, grouped by topic/study group, grouped by experiment/measured parameter, general to specific, or most to least important:

The mean (SD) admission interleukin concentrations were 13.6 (1.4) μg/L, 10.3 (1.1) μg/L, and 3.6 (05) μg/L in the coronary bypass graft, percutaneous inter-

Table 4. Serum antiproxin concentrations in patients with congestive heart failure.

Stage/classification	Antiproxin, ng/L, median (interquartile range)
I/healthy	99 (36–144)
II/asymptomatic heart failure	216 (147–296)[a]
III/symptomatic heart failure	556 (328–791)[b,c]

[a] $P = 0.019$ vs healthy patients.
[b] $P < 0.001$ vs healthy patients.
[c] $P = 0.017$ vs asymptomatic heart failure.

vention, and congestive heart failure patient groups, respectively.

3. Pretend that a journal editor has decided that you must remove Table 4 from your paper and place the information it contains into the main text. How might you write a paragraph that presents the data and results in this table?

Final Thoughts

A Results section that clearly presents your results, makes effective use of both data and results, includes all the important results, and does not wander off into discussion of the results, will result in a better paper and a greater chance of its acceptance for publication. In the end, isn't that the result you are looking for?

Resources and Additional Reading

Foote M. The proof of the pudding: how to report results and write a good discussion. Chest 2009; 135:866–8.
Huth EJ. Writing and publishing in medicine. Baltimore: Williams and Wilkins; 1999.
Katz MJ. From research to manuscript. New York: Springer; 2009.
Lang TA. How to write, publish, and present in the health sciences. Philadelphia: ACP Press; 2010.
Zeiger M. Essentials of writing biomedical research papers. New York: McGraw Hill; 2000.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures of Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:*

Employment or Leadership: T.M. Annesley, AACC.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Answers to Learning Exercise

1. Baseline median IL-6 concentrations were 12, 26, 96, and 144 μg/L for categories 1 to 4, respectively [DATA], and were not found related to age or sex [RESULT]. Median β-selectin concentrations increased 30% across the 4 categories [RESULT]. Increased disease severity and mortality were associated with higher IL-6 concentrations, but not β-selectin [RESULT]. Intraindividual variation for group 1 was 14% for IL-6 and 36% for β-selectin [DATA].

2. The presentation is grouped by experiment/measured parameter, which is the mean admission interleukin concentration. Even though the data are presented from the highest (13.6 μg/L) to the lowest (3.6 μg/L) value, the higher value is not necessarily the most important finding.

3. Median (interquartile range) serum antiproxin concentrations were 99 (36–144), 216 (147–296), and 556 (328–791) ng/L in healthy individuals, asymptomatic heart failure patients, and symptomatic heart failure patients, respectively. The median concentrations in asymptomatic and symptomatic heart failure patients were 2.2-fold higher ($P = 0.019$) and 5.6-fold higher ($P < 0.001$), respectively, than in healthy individuals, and symptomatic patients had significantly higher serum antiproxin concentrations compared with asymptomatic patients ($P = 0.017$).

Clinical Chemistry 56:8
1226–1228 (2010)

If an IRDAM Journal Is What You Choose, Then Sequential Results Are What You Use[1]

Pamela A. Derish[2] and Thomas M. Annesley[3*]

Sequential results? IRDAM journals? Both of these entities really exist, and both are relevant to how results are reported for some types of studies. How and when they apply become clearer if we start with a refresher on the organization of the scientific paper. An earlier article in this series *(1)* introduced the IMRAD[4] format (Introduction, Methods, Results, and Discussion), a standard set in 1972 by the American National Standards Institute and the most commonly used format today. This format works well for studies in which the experiments are planned in advance or performed in a predefined order. It therefore includes a study design subsection in the Methods section, usually at the beginning. Examples of these types of studies include method development and validation, randomized controlled trials, studies of diagnostic test performance, intervention trials, and observational studies.

In contrast, basic research studies often begin with a hypothesis to be tested, but beyond the initial experiment or starting point, the experiments performed throughout the study are not necessarily planned in advance. In fact, the results of one experiment typically set the direction for subsequent experiments. Because the results, not a preset series of experiments or methods, drive the study, articles written for many basic research journals tend to emphasize the Results section and subordinate the Methods section to the Results. The format used by many high-impact basic research journals, such as *Nature*, *Proceedings of the National Academy of Sciences*, *Journal of Clinical Investigation*, and *Journal of Cell Biology*, is arranged so that the Results section immediately follows the Introduction. The Methods section is placed at the end, or it may even be published as a supplemental data file. This is the IRDAM format (Introduction, Results, Discussion, and Methods) *(2)*.

The IRDAM format requires a substantial change in how the Results section is organized. Because the methods are listed at the end of the paper, or online, the reader is not exposed to details of the experimental protocols and methods before the results are presented. Therefore, the rationale for why these experiments were performed, how they were performed, and how the data were analyzed has to be presented in the Results section to pull the story back together for the reader. This goal is accomplished through the presentation of *sequential results*. In addition, the last paragraph of the Introduction section may be expanded so that instead of ending with the study question or hypothesis, it goes on to state the experimental approach and the answer to the study question or take-home message of the results.

Sequential Results Format

A sequential Results section generally consists of a series of subsections of 1 to 2 paragraphs. Each of these subsections contains its own subheading and pertains to a separate experiment. The basic format, repeated in each subsection, includes 4 elements: the question, the overview of the experiments, the results, and the answer to the question *(3)* (Table 1).

The question is usually stated as the purpose or goal of that particular part of the study. Sometimes it is helpful to add some background information that puts the question into the context of the overall study. Given that describing how the results were obtained is essential for the reader to assess their validity, the next element in the sequential results format is an overview of how the experiment was designed, what methods were used and perhaps why they were used, any modifications that were made, and what control experiments were included. Remember that this element consists of an overview of the experiment only; specific details about the experiment belong in another part of the paper, usually the Methods section and figure legends, but sometimes in an online data file. The results or findings of the experiment come next, usually through a descriptive presentation that includes summarized data and reference to tables and figures. The answer is provided in 1 to 2 sentences. Where appropriate, a con-

[1] Scientific research papers should avoid the use of nonstandard abbreviations in titles (exceptions being well-known ones or long expanded forms); however, educational or special feature articles may take liberties to attract the reader's attention.

[2] Department of Surgery, University of California, San Francisco, San Francisco, CA; [3] Department of Pathology, University of Michigan Health System, Ann Arbor, MI.
* Address correspondence to this author at: Department of Pathology, University of Michigan Health System, Rm. 2G332, 1500 East Medical Center Dr., Ann Arbor, MI 48109-5054. Fax 734-763-4095; e-mail annesley@umich.edu.

Received May 20, 2010; accepted May 26, 2010.
Previously published online at DOI: 10.1373/clinchem.2010.150961

[4] Nonstandard abbreviations: IMRAD (format), Introduction, Methods, Results, and Discussion; IRDAM (format), Introduction, Results, Discussion and Methods.

Table 1. Elements of a sequential results format.

Question:
The hypothesis, purpose, or goal
May include background information
Experiment:
Overview of design, methods, controls
Full details provided elsewhere in the paper
Results:
The findings from the experiment
Can include summarized data
Reference to figure or table
Answer:
What the results show, prove, confirm
Can add conclusion or comment
Other interpretation saved for Discussion

clusion or a comment about the results can be added as well.

Each of these sequential Results subsections becomes a kind of miniversion of a paper, in which some background is provided about why the experiment was performed or what question was being addressed, how it was tested, the results obtained, and the answer to the question. This format also parallels the format used in the Abstract at the beginning of the paper, but in this case summarizing an individual experiment rather than the overall study.

We illustrate an expanded Introduction section and the first 2 subsections of a sequential Results section in Example 1. In the Introduction paragraph of the example, sentence ***A*** states the question of the study, which would be the usual way to end an Introduction in an IMRAD paper. Given that it is an IRDAM paper, the Introduction is expanded by having sentence ***B*** state the experimental approach used to address the question and having sentence ***C*** answer the question.

In the 2 subsections of sequential results, each of which consists of 1 paragraph, note that each subsection is preceded by a specific, rather than a generic, subheading that states the result. Each experiment gets its own new subheading and paragraph (or paragraphs, if required). If more experiments were done, the sequence would continue accordingly. In the first subsection, sentence ***A*** provides background; sentence ***B*** states the question or purpose of the experiment and describes the experiment; sentences ***C***, ***D***, and ***E*** state the results; and sentence ***E*** also gives the answer. In the second subsection, sentence ***F*** states the next question, ***G*** describes the experiment, ***H*** states the results, and the last sentence gives the answer.

Emphasizing Results in a Paragraph That Contains Other Information

We have just seen in the example that the results wind up buried in the middle of the paragraph. To emphasize results when so much else is happening in the paragraph, you can "signal" to the reader that the sentence contains results by beginning that sentence with the phrase "We found," "We observed," or "We detected" *(3)*.

In addition to signaling the results, you can also signal other elements of the sequential results format and thereby keep the story line clear. In the example, the question for each experiment is signaled slightly differently, by the use of "To investigate whether" in sentence ***B*** of subsection **1** and by "to determine the extent" in sentence ***F*** of subsection **2**. The answer is signaled clearly in subsection **2** by "Thus" at the beginning of sentence ***I***. Other options would be to start the answer sentence with "Therefore," "This result shows," or "These findings indicate." In the example, the answer to the question in sentence ***E*** illustrates another technique for making the answer clear. The sentence begins with a results statement noting that this result is especially important ("Of more interest") but ends by stating that the result is "consistent with IL-6γ contributing to vascular inflammation." The phrase "consistent with" links the result back to the question.

Final Thoughts

A well-known idiom asks whether the ends justify the means, but in scientific studies, reviewers and readers expect to know whether the means (the methods) justify the ends (the results). Because the Results section of an IRDAM scientific paper is bracketed by the Introduction and the Discussion, it has a different narrative structure than that of an IMRAD paper. By providing the results together with the rationale for the experiments, how they were performed, and how the data were analyzed, the sequential results format is key to revealing enough of the story for the reader to decide whether the means justify the ends in studies published in journals that follow the IRDAM format.

Resources and Additional Reading

Lang TA. How to write, publish, and present in the health sciences. Philadelphia: ACP Press; 2010.
Tacker M. Writing papers on "sequential" research: sources of guidance sought. Sci Ed 2004;27:208.
Zeiger M. Essentials of writing biomedical research papers. New York: McGraw-Hill; 2000.

Example 1. A Typical Last Paragraph of the Introduction Section and 2 Subsections of the Sequential Results Section for a Hypothetical Study

Introduction (last paragraph of the section)[1]

1 [A]We sought to answer the question of whether interleukin-6γ (IL-6γ) is a cause or just a marker of vascular inflammation associated with atherosclerotic disease. [B]Using a herpes simplex virus type 2 (HSV2) infection model to bring about continuous production of IL-6γ, we investigated the effects of IL-6γ production on the development of vascular inflammation and atherosclerotic lesions in mice. [C]Our results show that in contrast with continuous production of β-selectin or vascular lipoprotein-binding molecule (VLM), 2 other markers of vascular inflammation, IL-6γ appears to play a direct role in the formation of atherosclerotic lesions in mice.

Results (first 2 subsections, each consisting of 1 paragraph)

1 ***Increased plasma IL-6γ, β-selectin, and VLM after gene injection.*** [A]Viral vectors have been successfully used to generate the in vivo production of ferritin and transcobalamin *(10, 11)*. [B]To investigate whether protein markers of vascular inflammation and atherosclerosis could be generated via a virus-infection protocol, we prepared plasmid HSV2-IL-6γ, HSV2-β-selectin, and HSV2-VLM by cotransfection of the virus with cDNA encoding IL-6γ, β-selectin, and VLM, respectively *(12)*. [C]Four weeks after percutaneous injection with 1×10^7 infectious units, we detected no IL-6γ, β-selectin, or VLM in plasma from control mice injected with non-cotransfected HSV2. [D]Only IL-6γ was present in plasma from mice injected with HSV2-IL-6γ (Figure 1), only β-selectin was present in mice injected with HSV2-β-selectin (Figure 1B), and only VLM was present in mice injected with HSV2-VLM (Figure 1C). These results show that the HSV2 model could stimulate continuous production of these 3 proteins. [E]Of more interest was the fact that at 24 weeks, the mice injected with HSV2-IL-6γ demonstrated the vascular inflammation markers β-selectin and VLM in plasma as well, a finding consistent with IL-6γ contributing to vascular inflammation.

2 ***IL-6γ increases development of atherosclerotic lesions.*** [F]We next sought to determine the extent to which overexpression of IL-6γ, β-selectin, or VLM contributes to the formation of atherosclerotic lesions. [G]Twenty-four weeks after injection, the mice were killed, and the areas of the atherosclerotic lesions in the aortic roots were assessed by morphologic analysis and immunohistologic staining with eosin or Sudan IV. [H]We observed that the mean areas of lesions in mice injected with HSV2-IL-6γ were 2.4-fold and 2.2-fold larger than the lesions in mice injected with HSV2-β-selectin or HSV2-VLM, respectively (Figure 2), and 2.7-fold larger then the control mice. [I]Thus, IL-6γ appears to contribute to the formation of atherosclerotic lesions.

[1] The references and figures mentioned in this example do not correspond to any actual references or figures in the present article.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures of Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:*

Employment or Leadership: T.A. Annesley, AACC.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

References

1. Annesley TM. The abstract and the elevator talk: a tale of two summaries. Clin Chem 2010;56:521–4.
2. Day RA, Gastel B. How to write and publish a scientific paper. Westport (CT): Greenwood Press; 2006. p 21.
3. Zeiger M. Essentials of writing biomedical research papers. New York: McGraw-Hill; 2000. p 158.

Clinical Chemistry 56:8
1229–1233 (2010)

Put Your Best Figure Forward: Line Graphs and Scattergrams

Thomas M. Annesley*

There is an old saying that "a picture is worth a thousand words." In truth, only a well-prepared, self-explanatory picture is worth a thousand words. The same holds true for research studies, for which 1 of the main methods we use to communicate our message is in figures and graphs. Figures and graphs tell much of the story by giving readers a visual anchor to help them see, understand, and remember information. Think about a report that you recently read and found useful. You likely do not remember the text used to state the results, or even the actual numbers, but you can recall much about the trends, relationships, outcomes, categories, or general experimental parameters shown in a graph. Despite the fact that you no longer have much recollection of the text, you can draw a reasonable representation of a graph from the published report and tell what you remember from it.

In this educational article I discuss line graphs and scattergrams and use examples to illustrate how to put your best *figure* forward so readers will remember you and your message.

Basics of a Good Graph

The components of a graph include axes, labels, scales, an origin, tick or reference marks, symbols, and a legend. Beyond these basics, however, a *good* graph has several attributes:

1. It draws attention to the data and not the graph itself.
2. The data points (symbols) and connecting lines are easy to read and distinguish.
3. Both the numbers and labels for the axes are readable and their meaning is clear.
4. The lengths of the 2 axes are visually balanced (ratio of x axis to y axis = 1.0 to 1.3).
5. The scales used on each axis match the range of the data.
6. Tick marks are used appropriately.
7. The legend is clear and concise.
8. The reader can understand the message without referring back and forth to the main text.
9. The data deserve to be graphed.

Line graphs and scattergrams make use of a horizontal and a vertical axis, typically called the x and y axis, respectively, to illustrate the relationship between 2 or more variables. By convention, the variable plotted on the x axis is referred to as the *independent variable*. The independent variable is the variable that is manipulated or changed by the investigator. The variable plotted on the y axis is the *dependent variable*. This variable is called the dependent variable because its value responds to (depends on) the value of the independent variable. It changes when the independent variable changes.

For example, one may study serum phenytoin concentration versus prescribed dose. The dose is the independent variable and the resulting serum concentration is the dependent variable because it depends on (or is caused by) a change in the independent variable. Think of it as asking a question: Does changing the dose (cause) result in a change in the circulating phenytoin concentration (effect)? This way of identifying a cause and effect relationship may often help you to determine whether the study involves independent and dependent variables and how you should design a figure to show the experimental results.

Another example is a study of serum prostate-specific antigen (PSA) as a noninvasive predictor of tumor staging. In this case the known (independent) variable is the tumor stage (a predefined variable or reference point), and the unknown (dependent) variable is the concentration of PSA in a patient's serum. It is possible to plot more than 1 dependent variable in a graph (e.g., total and free PSA), but there should be only 1 independent variable, in this example the tumor stage, plotted in a graph.

Although many studies have an independent variable, its presence is not a strict requirement. In some cases the study involves looking at an association of 2 variables, without any underlying proof of causation. A comparison of 2 analytical methods for quantifying troponin is a good example. In this case neither method has an effect on the other, and therefore no independent variable exists. The data for either method could

University of Michigan Health System, Ann Arbor, MI.
* Address correspondence to the author at: University of Michigan Health System, Room UH2G332, 1500 East Medical Center Drive, Ann Arbor, MI 48109-5054. E-mail annesley@umich.edu.
Received April 30, 2010, accepted May 6, 2010.
Previously published online at DOI: 10.1373/clinchem.2010.150060

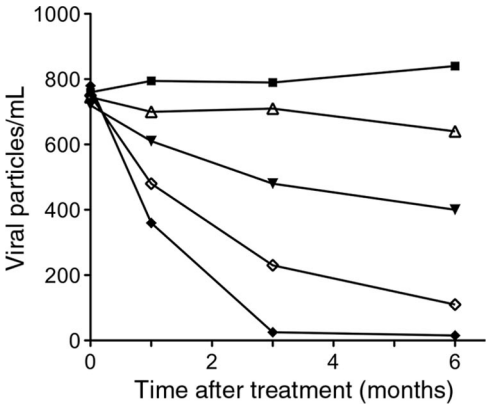


Fig. 1. Change in blood viral load during daily oral treatment with albenovir; ■, 0 mg/kg; △, 2 mg/kg; ▼, 5 mg/kg; ◇, 10 mg/kg; ◆, 20 mg/kg.

be plotted on the *y* axis. This having been said, however, it is still important in any study to determine if the data you are analyzing and graphing has an independent variable.

Fig. 1 shows an example of a line graph with the desired attributes mentioned earlier. This graph represents data from a hypothetical study of the efficacy of a new antiviral drug, albenovir. In this study albenovir was given orally daily to randomized groups of patients at 5 doses (0, 2, 5, 10, and 20 mg/kg). Blood samples were collected from patients at selected time points the beginning of treatment, and the samples were analyzed for circulating viral particles. The change in viral load vs time is plotted in the graph. In this figure the symbols representing the different doses are large and easily differentiated from one another, which allows them to be easily understood. The connecting lines are also clear and wide enough to draw attention to the data. A general rule is that the symbols and any lines or curves inside of the 2 axes are the most prominent features, the wording in the axes labels somewhat less prominent, and the axes and tick marks the least prominent. In this graph the 2 axis lines are proportional in length and narrow enough that they do not draw attention away from the data. The font size for the wording of the axis labels, which again highlights more important information, is larger than that of the numbers and tick marks on the axes. The tick marks are on the outside of the axes because they are associated with the numbers on the axes and not the plotted data points inside the axes. The scales are also proportional to the range of values and there is minimal wasted space throughout the graph.

The legend for this figure is concise, and the message can be understood even without having access to the main text. The graph in Fig. 1 does not include a title (often included in PowerPoint slides) because there is a legend that conveys the important information.

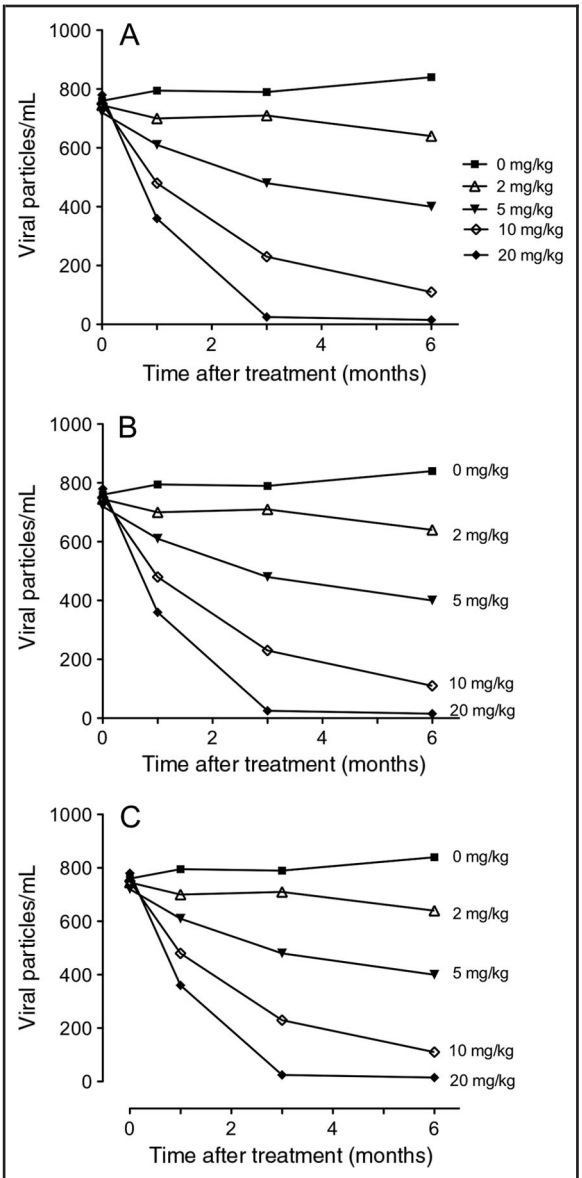


Fig. 2. Change in blood viral load during daily oral treatment with albenovir.
(A–C), different representations of the same data.

Fig. 2 shows several style options for improving the layout of the data plotted in Fig. 1. Although Fig. 1 is simple and clean, the reader must refer back and forth to the legend to associate the symbols and lines with the different treatment protocols. If there is extra space within the graph, or to the right, then one can consider

adding a key to the symbols, as illustrated in Fig. 2A. If a key is added, however, it is important that the order of the symbols (top to bottom or left to right) in the key be the same as the order in which the symbols and lines are plotted in the actual graph, as was done here. A benefit of this approach is that it can simplify the message in the legend. If space allows, it is possible to consider an even a more effective design and place individual labels next to each line or set of data (Fig. 2B).

Sometimes data points have numerical values that fall directly on (or very close to) the *x* or *y* axis, as they do for these albenovir data. When this is the case, data points may be visually distorted or obscured by the axis line, especially when a graph is reduced to print size. In this situation (and only in this situation), one or both axes can be offset to allow a clearer visualization of the data (Fig. 2C). As an exercise, compare the data presentation and legends for Fig. 1 and 2 and evaluate how each influences what you see and read.

Common Mistakes

The next 3 examples highlight common mistakes authors make when preparing graphs. Fig. 3 shows the relationship between plasma and serum sodium for paired specimens from 150 patients. Sodium concentrations, even in critically ill patients, fall within a fairly narrow range from 125 to 165 mmol/L. Because many computer programs automatically default to *x*- and *y*-axis intercepts of 0, the graph may look like the one shown in Fig. 3A. There are 3 problems with this type of data presentation. First, the data points are compressed close together, making it difficult to see any scatter, or abnormal high and low values. Second, even with a correlation line, it can be difficult to see whether 1 or 2 outliers may have exerted undue influence on the overall correlation data. Third, because such a plot fails to properly convey the information in the data, it wastes space, which editors dislike for both economic and aesthetic reasons. These same data can be presented more clearly by narrowing the ranges of the axis scales to fit the true range of the data (Fig. 3B). An even better representation of the data can be obtained by creating a Bland–Altman plot (Fig. 3C), in which differences outside of the 95% limits of agreement are easily seen.

Similarly, editors often see results graphed as shown in Fig. 4. In this hypothetical example, an investigator developed a new HPLC assay for plasma alanine to support a collaborative study of rats undergoing high-stress experiments. To validate the stability of alanine in blood during courier transport across the university, specimens were collected into 4 different anticoagulant-containing tubes and stored at room temperature for selected time periods before centrifugation and freezing of the plasma. An obvious time-dependent loss of alanine in heparin specimens can be seen in Fig. 4A, even though the figure wastes a lot of usable space and compresses the data. Expanding the *y*-axis scale, as in Fig. 4B, not only makes much better use of space, it also shows a time-dependent increase in plasma alanine of almost 10% in oxalate tubes over an 18-hour period.

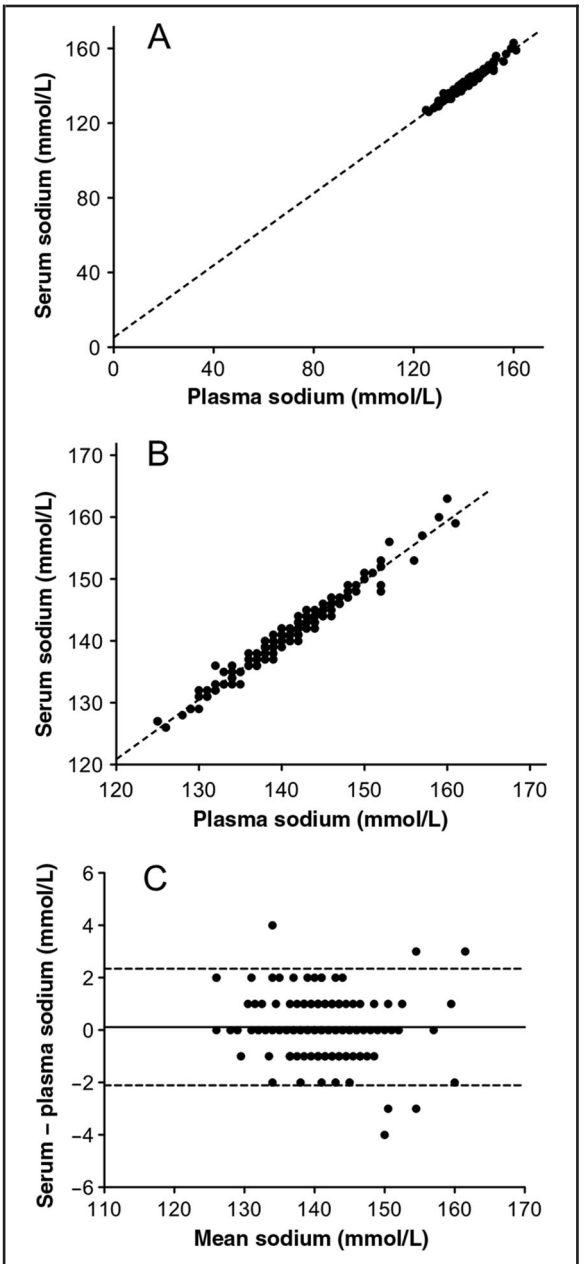


Fig. 3. Plasma vs serum sodium for paired specimens from 150 patients.

(A), *x*- and *y*-axis scales of 0–165 mmol/L; (B), *x*- and *y*-axis scales of 120–170 mmol/L; (C), Bland–Altman plot.

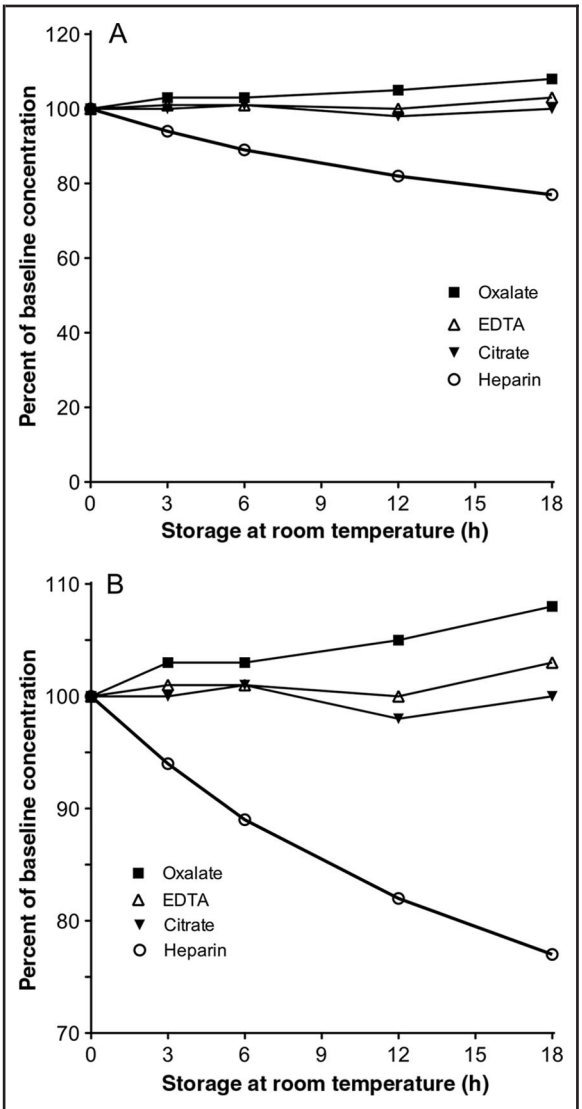


Fig. 4. Percent change in plasma alanine concentration after storage of whole blood at room temperature.

(A), *y*-axis scale of 0–110 μg/L; (B), *y*-axis scale of 70–110 μg/L.

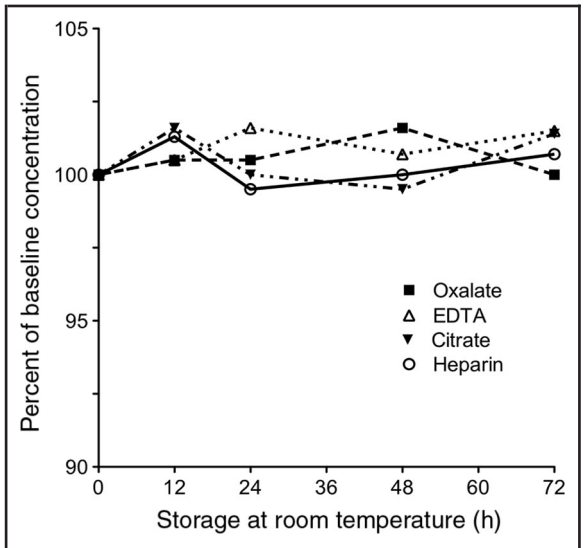


Fig. 5. Percent change in whole-blood tacrolimus concentration after storage at room temperature.

The technique shown in Fig 4B is sometimes referred to as a *suppressed zero*. Although not necessarily bad if it increases the clarity of data presentation, one can see how a suppressed-zero scale can be misleading if used to exaggerate what would otherwise be small differences. The news media are often criticized for making graphs of economic data that have a restricted *y*-axis range, thus artificially magnifying the significance of any changes. So if you decide to use an expanded scale, bring this to the attention of the reader by stating directly in the figure legend that the scale has been expanded or does not start at 0.

Fig. 5 shows a graph that meets all of the criteria for a good graph except one. Can you guess what it is? It is a good example of results that do not need to be presented as a graph. They are useful results and should be reported, but the message can be conveyed just as easily in the main text: "When whole blood specimens were collected into oxalate, EDTA, citrate, or heparin-containing tubes, and stored at room temperature for up to 72 hours, no statistically significant change in the tacrolimus concentration was observed for any of the tube types."

Learning Exercise

Using the information presented here about the characteristics of good and bad graphs, you should be able to identify features that add to or detract from the visual impact of a graph. The example shown in Fig. 6 has at least 12 problems. Can you identify these? Answers are provided in a box after the list of selected additional reading materials.

Final Thoughts

People are visual and expressive by nature, and authors (including this one) want to *show* what they have done. A picture can be worth a thousand words, but a few well-chosen words also can replace a picture. The key is to know when to use one or the other to most effectively state your message.

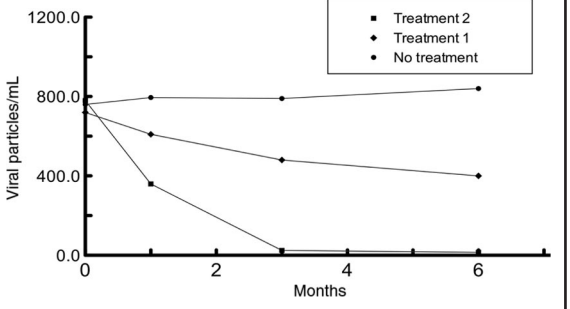


Fig. 6. Change in viral load during treatment; ●, 0 mg/kg; ◆, 5 mg/kg; ■, 20 mg/kg.

Resources and Additional Reading

Day RA, Gastel B. How to write and publish a scientific paper. Westport (CT): Greenwood Press; 2006.
Freeman JV, Walters SJ, Campbell MJ. How to display data. Malden (MA): Blackwell Publishing; 2008.
Gustavii B. How to write and illustrate a scientific paper. New York: Cambridge University Press; 2008.
Lang TA. How to write, publish, and present in the health sciences. Philadelphia (PA): ACP Press; 2010.
Zeiger M. Essentials of writing biomedical research papers. New York: McGraw Hill; 2000.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures of Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:*

Employment or Leadership: T.M. Annesley, AACC.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Answer to Learning Exercise (Problems with Fig. 6)

The symbols are too small.

The symbols are too similar (solid box, solid circle, solid diamond) and are difficult to distinguish.

The data-connecting lines are narrow and do not draw attention to the data.

The text in the labels is small.

The *x* and *y* axes are too wide and draw the focus away from the data.

The numbers on the axes are proportionately too large.

The numbers on the axes are 2 different font sizes.

The *y*-axis numbers have an unnecessary decimal point.

The scale for the *y* axis is too large and creates wasted space.

The *x* axis says "months" and a fuller description may alleviate the need for the reader to refer to the main text.

The tick marks are on the inside of the axes and hide the symbols.

The ratio of the *x* axis to the *y* axis is too large (ideally 1.0 to 1.3)

The symbol legend within the graph identifies different treatments, whereas the figure legend identifies milligram per kilogram doses.

The symbol order (top to bottom) in the legend within the graph is different from the order (top to bottom) of the actual symbols in the figure.

Clinical Chemistry 56:9
1394–1400 (2010)

Bars and Pies Make Better Desserts than Figures

Thomas M. Annesley*

In a previous article on figures *(1)* I discussed line graphs and scattergrams, 2 of the most widely used approaches for presenting data and results in scientific papers. In line graphs and scattergrams, each axis is of a continuous variable. For example, the *x* axis may show a continuous range of phenytoin doses, and the *y* axis may show the corresponding range of the resulting serum phenytoin concentrations. Or, the *x* axis may be a range of months after chemotherapy, and the *y* axis may be the percentage of surviving patients. Assay comparisons, chromatograms, ROC curves, and PCR amplification curves are all examples of line graphs or scattergrams.

There are, however, situations in which the variables are discontinuous (also called "discrete" or "nominal" variables), meaning that they are categorically different (e.g., eye color); combined within an interval (e.g., ages 21–30 years, 31–40 years, 41–50 years); or numerically scaled (ordinal variables) (e.g., tumor stages 1, 2, 3, and 4). When a visual representation of the results for discontinuous variables is desired, 2 commonly used approaches are bar graphs and pie charts. In this article, I discuss the pros and cons of these 2 types of figures.

Bar Graphs

A bar graph, also known as a column graph when the bars are plotted vertically *(2)*, is a 2-dimensional figure in which a set of discontinuous independent variables is plotted versus a continuous dependent variable. Fig. 1 is a plot of the dollars spent per capita on healthcare, which is a continuous variable, for 5 industrialized countries with medium-sized populations, each of which is a discontinuous variable. The length of the rectangular bar (or column) represents the dollars spent by each country.

If you are considering a bar graph, the use of certain style elements can help you create an effective figure. First, the space (gap) between the bars should be narrower than the width of the bars so that the gaps do not dominate the figure and pull the focus away from the bars. A good starting point is a gap that is 50% of the width of the bars. Second, the shading or pattern within the area of the bars should be pleasing to the eye and easy to distinguish from other bars when multiple sets of data are plotted in the same figure.

Third, it is important, as with other types of figures, to avoid the use of a suppressed-zero scale (i.e., a scale that does not include 0), because this practice can exaggerate differences among groups *(1)*. Although all of the bar graphs in Fig. 1 present the same data, 3 have presentation styles that detract from the graph. In Fig. 1A, the width of the gap is the same as the bars. The color within each bar is white. This combination produces a figure that not only is initially difficult to grasp but also looks more like jail bars than data bars.

Fig. 1B illustrates the opposite extreme. The gap here is only 15% of the width of the bars, so that when the ratio of the horizontal and vertical axes is set at 1, the bars become so wide that they more closely resemble a histogram than a bar graph. The use of lines or crosshatches within the bars to distinguish the data for the 5 countries makes the figure look busy. Now compare Fig. 1, A and B, with Fig. 1C, which has the 50% gap width mentioned above and a gray-scale shading that works well and is easy to distinguish from the white background. This bar graph has maximal clarity and would make a good published figure or slide in a presentation. The same data look much different (Fig. 1D) when a suppressed-zero scale is used. Note how much larger the differences in healthcare expenditures appear when the scale does not include 0. Look at newspaper or television reports, and you will find examples of the use of a suppressed-zero bar graph to magnify changes that might not otherwise appear so large.

Current graphics software packages allow scientists to plot the same data in many different formats, some quite fancy. Three-dimensional bar graphs (Fig. 2, A and B), can look impressive, but they rarely add value and can actually be less clear to the reader. Note how the mortality differences in age groups 21–30 years and 31–40 years are more difficult to assess in Fig. 2A than in the other 2 panels of this figure. You will find, however, that presenting the same results as a 2-dimensional graph (Fig. 2C) will nearly always be easier to read. Note, too, that all 3 panels present the 2

University of Michigan Health System, Ann Arbor, MI.
* Address correspondence to the author at: University of Michigan Health System, Rm. UH2G332, 1500 East Medical Center Dr., Ann Arbor, MI 48109-5054. Fax 734-763-4095; e-mail annesley@umich.edu.
Received June 21, 2010; accepted June 24, 2010.
Previously published online at DOI: 10.1373/clinchem.2010.152298

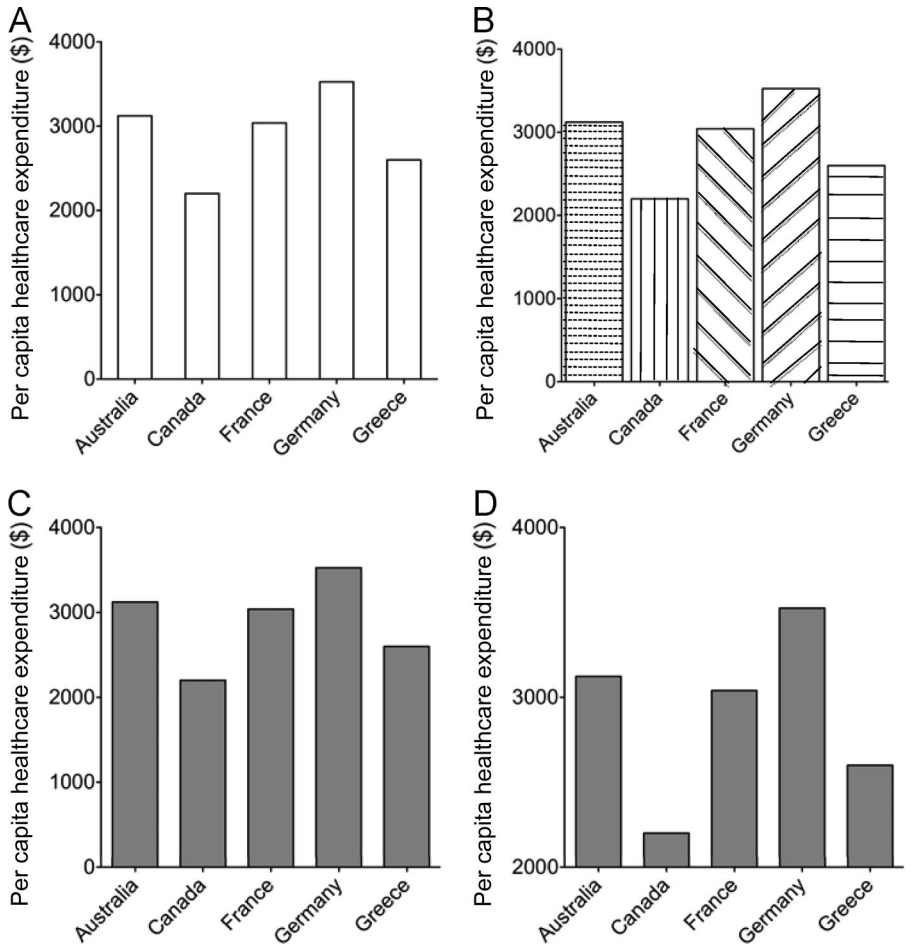


Fig. 1. Per capita healthcare expenditures for 5 industrialized countries with medium-sized populations.
All expenditure amounts have been converted to US dollars. Panels (A), (B), and (C) are different representations of the same data. (D), Same data plotted on an expanded scale.

sets of results (women vs men) in complementary gray tones that stand out from the background. When multiple sets of data are plotted on the same graph, a gradation of shades from white to gray to black will be easier to read than patterns.

Bar graphs are useful for visual comparisons of data (Fig. 1) or for showing trends in the data (Fig. 2) and are most informative when you are more interested in the actual value of a variable than its CI *(3)*. This feature is why bar graphs are popular in slides for presentations. They focus the audience's attention on a single data value. In scientific publications, however, the distribution of the data also is critical for interpreting the data and results. The display of information regarding the data distribution is an area in which bar graphs have potential limitations. One can create a summary-data chart by adding the SD (Fig. 3, left), the 95% CI, or the interquartile range, but the bar usually remains the major visual element and therefore can mask the distribution of the data. A better alternative is either a line graph (mean and SD, or median and 95% CI), in which a symbol represents the mean (Fig. 3, middle), or an individual-value graph that shows the mean value as a horizontal line, as well as all of the data points and their spread (Fig. 3, right).

Fig. 4 illustrates another important point to remember: Bars do not have to be plotted vertically. This figure shows the per capita healthcare expenditures for 19 countries that participated in an economic survey. The format that many individuals would choose (Fig. 4A) is the placement of the country names horizontally and the bars vertically. If the number of variables is small or the category name (the country here) is short in length (e.g., Fig. 1), then the reader may be able to

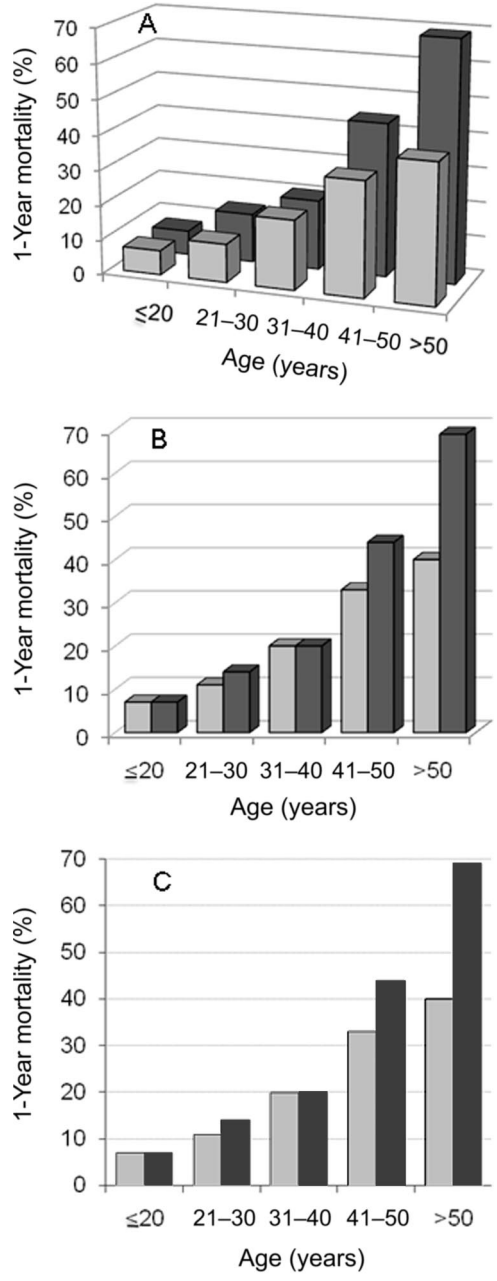


Fig. 2. One-year mortality vs age for patients diagnosed with Preuter disease.

Indicated are data for men (□) and women (■). Panels (A), (B), and (C) are different representations of the same data.

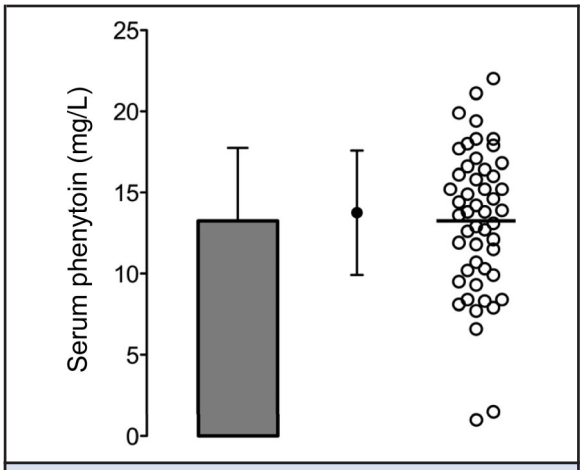


Fig. 3. Trough serum phenytoin concentrations for 36 pediatric patients receiving a dosage of 6 $mg \cdot kg^{-1} \cdot day^{-1}$.

(A), Bar graph showing the mean and SD. (B), Line graph showing the mean and SD. (C), Individual-value graph with the mean value represented as a horizontal line.

read and understand the graph readily. If the category names are long and many, however, they can become difficult to read unless the reader rotates the page. The same data become easier to read if the format is reversed and the per capita expenditures are shown on the horizontal axis (Fig. 4B). This version is acceptable if you are interested more in an individual expenditure for a single country than a comparison with other countries. If the reader is going to see only a series of apparently unrelated bars, however, with each bar representing a single data point, you should consider whether the data might be more informative in a table, where the actual numerical information for each country could be listed. This argument reiterates the earlier point that bar graphs are effective for presentations but are not always so for scientific papers. If you are going to use a bar graph, the best representation of the data is that of Fig. 4C. The relationships among the countries are much more apparent, and the graph shows the differences and trends from the highest to the lowest values.

Pie Charts

A pie chart is a circular drawing that is divided into segments, with each segment representing a data category or group. The size of each segment reflects its percentage or proportion of the total area of the pie. Pie charts are popular but are not useful in most scientific papers. They are used more frequently in magazines and newspapers to illustrate a specific difference between selected groups, often to draw attention to large differences or to express a viewpoint. Pie charts are most understandable if the number of categories is limited to 6 or fewer. Presenting more than 6 categories

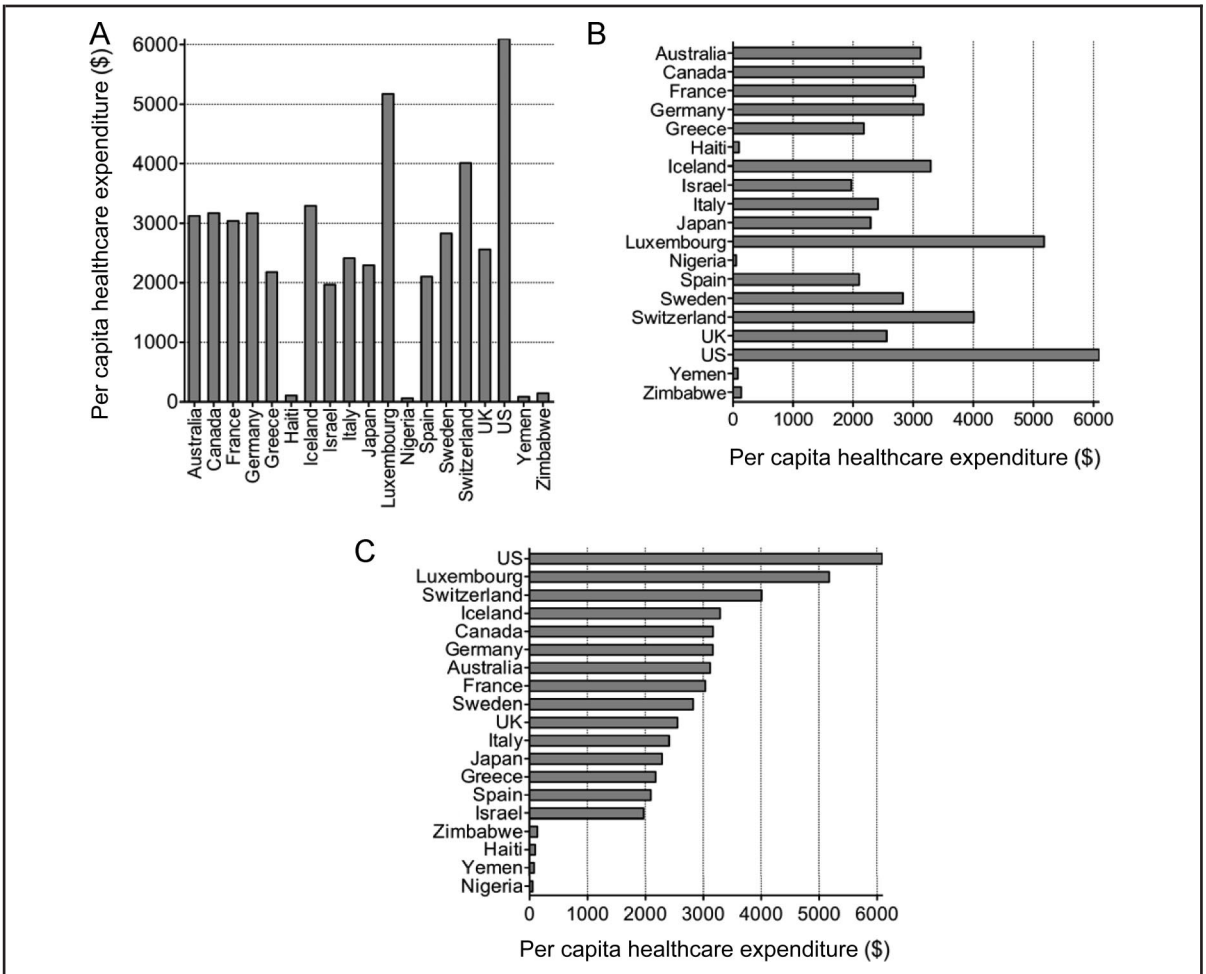


Fig. 4. Per capita healthcare expenditures for 19 countries or unions participating in an economic survey.
All expenditure amounts have been converted to US dollars. Panels (A), (B), and (C) are different representations of the same data.

not only makes the pie chart appear busy and confusing but also makes finding usable, nonclashing color hues, shading, or background patterns more difficult. Although a pie chart is good for displaying the relative size or percentage contribution of each included piece of data, a potential problem with a pie chart is that readers may infer that the circle represents 100% of all possible data or all possible outcomes, which may not be the case.

For example, one could take the per capita healthcare expenditure data of Fig. 4, select 6 countries from the bar graph, and plot them as a pie chart (Fig. 5). Because the pie chart in Fig. 5A contains a selected subset of the data, it shows that the per capita healthcare expenditure in the US is roughly 45 times that of Zimbabwe. Is the intended message that the US spends too much? That Zimbabwe spends too little? One might even conclude that the US spends 75% of all per capita healthcare dollars, which is not true. Readers might be drawn to a different conclusion if 6 other countries, all with higher per capita expenditures, were compared in a pie chart (Fig. 5B). In this case, US expenditures do not look so out of proportion.

A pie chart is most accurate when all available data or possible outcomes are included. For example, in a hypothetical clinical study of the efficacy of a new chemotherapeutic regimen, the 4 possible outcomes or end points could be (*a*) complete remission, (*b*) partial remission, (*c*) no improvement, or (*d*) death due to treatment complications. All patients in the study should fit into one of these categories. A pie chart representation of the results might look similar to Fig. 6. All of the available data

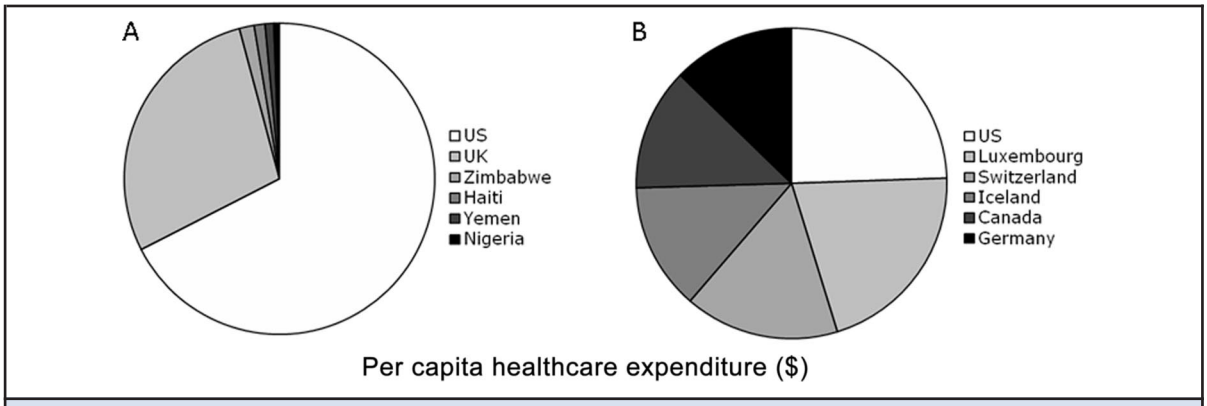


Fig. 5. Proportional per capita healthcare expenditures for selected countries or unions participating in an economic survey.
All expenditure amounts have been converted to US dollars.

are included, and the total number in each group is provided.

The common convention in creating pie charts is to consider the circle as a clock face *(4)*, starting with filling the largest section (wedge) in at 12:00, plotting subsequent sections in clockwise manner, and ending with the smallest section approaching 12:00 (Fig. 5). In some cases, the categories have a natural order or association, as in Fig. 6, that is best understood if they are plotted in a specified order (e.g., the best outcome to the worst outcome).

Of course, a pie chart is not necessary in most scientific papers. The same data can be presented in a table or even in the text. For example, the data in Fig. 6 could be stated in the text as: "In the men's group, 21 patients achieved complete remission, 52 patients achieved partial remission, 40 patients experienced no change, and 6 patients died from complications believed to have been due to the chemotherapy. In the women's group, 44 patients achieved complete remission, 41 patients achieved partial remission, 22 patients experienced no change, and 4 patients died from complications believed to have been due to the chemotherapy." Authors must decide on the best use of page space and word count.

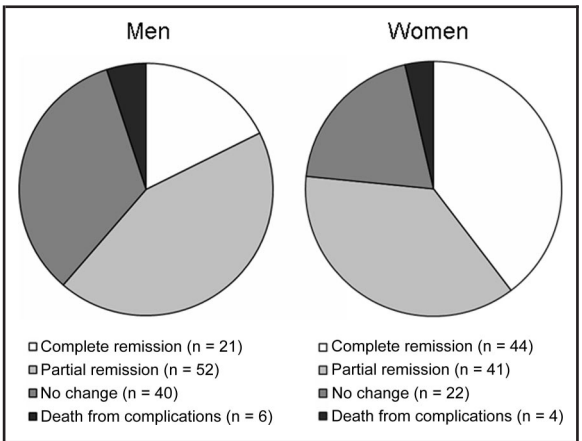


Fig. 6. One-year follow-up status of patients with acute lymphoblastic leukemia treated with a combination of vincristine and cytarabine.

Learning Exercise

With the data provided in Fig. 6, transform this figure from a pie chart to a bar/column graph. Be sure to add an appropriate legend to the new figure as well. After you have finished this exercise, compare your graph with the examples shown after the list of resources and additional reading materials.

Final Thoughts

Bar graphs and pie charts can be effective for summarizing data in a slide presentation or poster. They serve as a visual anchor for the audience while you explain the data and can highlight important differences or trends that might be missed if the data were presented only in text or a table. In scientific papers, however, a bar graph or pie chart must not only present the data but also be easily understood without having to refer repeatedly back to the main text. Authors can easily confuse readers with graphs that are unnecessarily complicated or that potentially misrepresent or underrepresent the data. In many cases, bars and pies make better desserts than figures.

Additional Reading

Day RA, Gastel B. How to write and publish a scientific paper. Westport, CT: Greenwood Press; 2006.
Lang TA. How to write, publish, and present in the health sciences. Philadelphia: ACP Press; 2010.
Zeiger M. Essentials of writing biomedical research papers. New York: McGraw-Hill; 2000.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures of Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:*

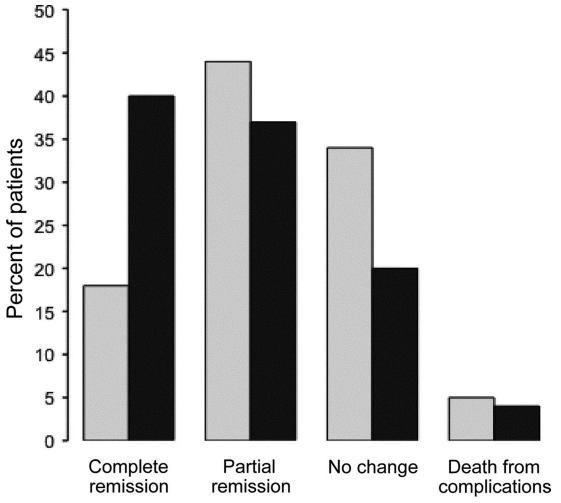
Employment or Leadership: T.M. Annesley, *Clinical Chemistry*, AACC.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

References

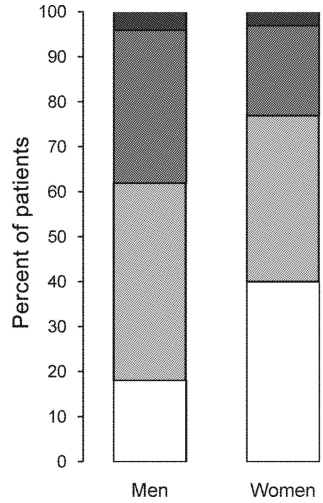
1. Annesley TM. Put your best figure forward: line graphs and scattergrams. Clin Chem 2010;56:1229–33.
2. Gustavii B. How to write and illustrate a scientific paper. New York: Cambridge University Press; 2008. p 24.
3. Browner WS. Publishing and presenting clinical research. Philadelphia: Lippincott Williams & Wilkins; 2008. p 101.
4. Freeman JV, Walters SJ, Campbell MJ. How to display data. Malden, MA: Blackwell Publishing; 2008. p 17.

Answer to Learning Exercise



Example 1. One-year status of patients with acute lymphoblastic leukemia treated with vincristine plus cytarabine.

Men (■), n = 119; women (■), n = 111.



Example 2. One-year status of patients with acute lymphoblastic leukemia treated with vincristine plus cytarabine.

□, complete remission; ■, partial remission; ■, no change; ■, death from complications. Men, n = 119; women, n = 111.

Commentary

Because the numbers of men (n = 119) and women (n = 111) differ, the best way to compare outcomes is to plot the percentages of men or women in each response category. It is also important to include the number of patients on the graph or in the legend. Example 1 is a clustered bar graph, in which the *categories* are plotted on the horizontal axis. The pattern of response rates is easy to see and compare for both sexes. When fewer than 3 groups are included, clustered bar graphs are better for showing trends and allow group comparisons. Example 2 is a stacked bar graph, in which the *groups* are plotted on the horizontal axis. Because stacked bar graphs must add up to 100%, they have the same characteristics as pie charts. When >3 groups are compared, a stacked bar graph may be easier to understand, especially if there is a natural order to the categories. This consideration is a good reason to plot your data several ways and then decide on the format that most clearly presents your message.

Clinical Chemistry 56:10
1528–1534 (2010)

Bring Your Best to the Table

Thomas M. Annesley*

Sometime in the past you were likely taught about the importance of "bringing your best to the table"—in other words contributing your best ideas and results. Indeed, it is an important key to success. But *what* ideas and results you bring to the table may not be enough. Sometimes it is *how* you present your ideas and results that distinguishes you and determines your success.

This consideration also holds true while writing a scientific paper, when you want to bring your best to another type of table—the table within your manuscript that shows your data. With this type of table, it is also important to understand that how you present your data can affect the success of your paper. In this article, I provide some basic tips on how to create a table that presents your data in a clear and visually appealing manner.

Tables vs Graphs

Data can be presented in either a graph or a table, so how do you choose which is best when writing a scientific paper? Let's compare the two. Graphs have an immediate visual impact and are good for showing trends or patterns or for highlighting differences between sets of data. Although the data in a graph are quantitative, a graph does not work well when the exactness or precision of the data is important.

Tables, on the other hand, are better when the individual or summarized values are more important than trends. Tables can be used for presenting both quantitative and qualitative data *(1)*. A table can contain words, symbols, numbers, or a combination of all three *(2)*. Tables allow side-by-side comparison of data *(3)*, such as the imprecision of 2 assays at several analyte concentrations. Although there is often a direct relationship between the variables listed in a table (e.g., death or myocardial infarction vs the tertile for a cystatin C concentration), the variables in a table do not need to have a direct relationship. Tables are also good for presenting large amounts of information that would be too cumbersome or confusing to place in the text (e.g., mass transitions for 20 different drugs being monitored in a toxicology screen).

A good table, although used for a purpose different from that of a good graph, has many of the same attributes *(4)*. A table should draw attention to the data and not to the table itself. In other words, the data within the table should be easy to distinguish and not be lost among poorly arranged words and numbers. Every table should have a clear and concise title. The table should stand alone without the need to refer repeatedly back to the main text. Lastly, the data must deserve to be in a table rather than in the main text.

Table Components

Scientific tables contain 5 major elements (Fig. 1): a title, column headings, stubs (row headings), data fields (the spaces in the columns that contain the data), and footnotes *(2)*. Sometimes a table may contain a spanner, a heading separated by a horizontal line that sits above secondary column headers. A spanner indicates that the column headings below it should be considered as part of a common group. Copy editors and printers will sometimes query an author about one of these table elements, so it is important to understand how each of these elements fits into the presentation of the data.

The title of a table should be sufficiently informative for the reader to understand the experimental data being presented without having to refer to the paper. Consider the following 3 hypothetical titles:

Example 1: *Statin therapy and cancer recurrence.*

Example 2: *Effect of daily oral primvastatin or dorvastatin on the 4-year odds ratio for the recurrence of prostate and breast cancer.*

Example 3: *The effect of daily oral primvastatin or dorvastatin on the 4-year odds ratio (OR) for the recurrence of prostate and breast cancer shows a 3-fold lower ($P = 0.002$) OR for the recurrence of breast cancer for patients receiving primvastatin ($OR = 2.3$) versus dorvastatin ($OR = 6.8$).*

Unless the headers or footnotes in the table include the names of the statin drugs tested, the type of cancer, the time period of the study, and the end point being measured, the title in example 1 may force the reader to look back at the experimental methods section of the paper to understand the context of the data being presented in the table. Example 2 is an informa-

University of Michigan Health System, Ann Arbor, MI.
* Address correspondence to the author at: University of Michigan Health System, Rm. 2G332, 1500 East Medical Center Dr., Ann Arbor, MI 48109-5054. Fax 734-763-4095; e-mail annesley@umich.edu.
Received July 14, 2010; accepted July 22, 2010.
Previously published online at DOI: 10.1373/clinchem.2010.153502

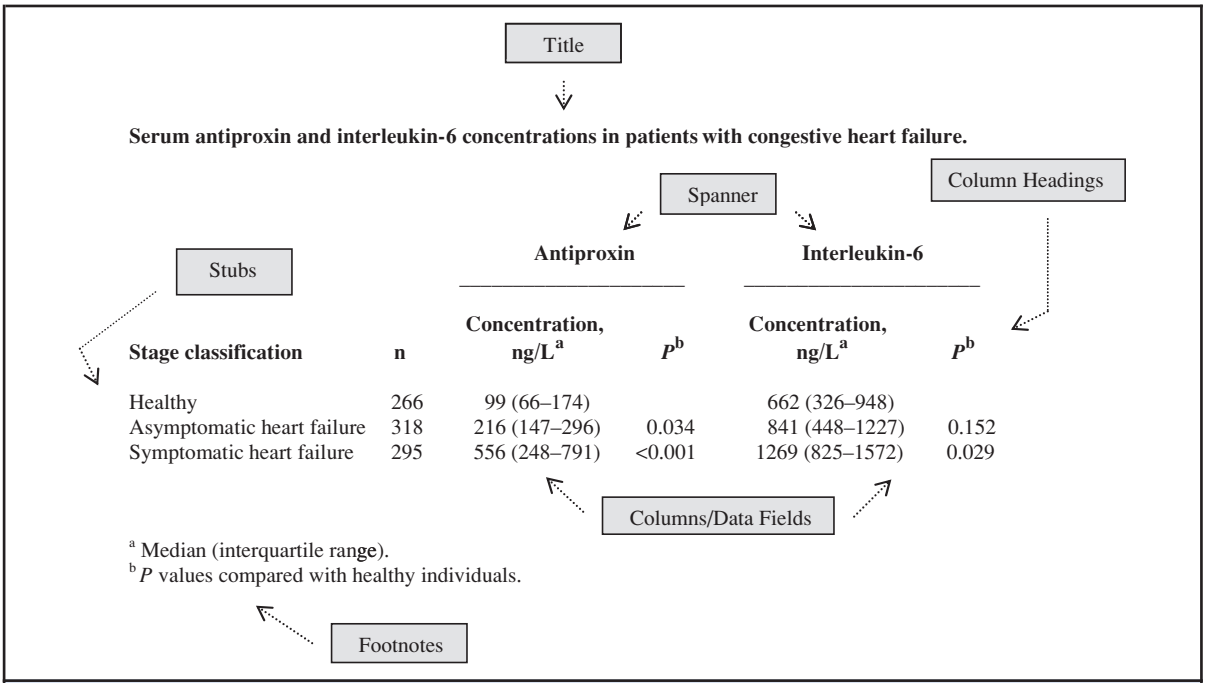


Fig. 1. Major components of a scientific table.

tive title that tells the reader much more about the study design (a 4-year follow-up study), the measured end point (odds ratio for recurrence), the specific diseases being studied (prostate and breast cancer), and the therapeutic agents being tested. The title in example 3 contains the same useful information as in example 2 but makes the mistake of providing the data and results as well. Titles should not include experimental details, data, or results. The data belong in the data fields of the table. Results and experimental details belong in the text.

The far left column containing the stubs lists one or more variables, usually the independent variable(s). All of the remaining columns to the right of the stubs must link back to the rows. This first column may not require a column header if the meaning of the stubs is evident (Table 1). If the variables have units (e.g., time, concentration, percent), these units should be added in parentheses or after a comma. If the variables all have the same units, the common unit can be designated after the main header above the stubs, if one is present.

When listing units of measure, be sure to use the journal's required format. Many journals, including *Clinical Chemistry*, request that SI units be used throughout a scientific paper. In the text, conventional units may be added in parentheses after SI units, but a different format should be used in a table. In a table, one should use the primary units in the body of the table and provide conversion factors to the secondary

Table 1. Options for displaying annual per capita healthcare expenditures.

A. Annual per capita healthcare expenditures.	
	Expenditure, $
Israel	1971
Madagascar	36
Sweden	2828
Yemen	82
Zimbabwe	149

B. Annual per capita healthcare expenditures.	
	Expenditure, $
Israel	1971
Madagascar	36
Sweden	2828
Yemen	82
Zimbabwe	149

C. Annual per capita healthcare expenditures.	
	Expenditure, $
Sweden	2828
Israel	1971
Zimbabwe	149
Yemen	82
Madagascar	36

Table 2. Long-term outcomes according to cystatin C tertile.[a,b]

	Cystatin C tertile			
	First tertile: <0.86 mg/L (n = 378)	Second tertile: 0.86–1.01 mg/L (n = 365)	Third tertile: >1.01 mg/L (n = 385)	Log-rank *P*
Death within 4 years	3.4% (12)	6.2% (21)	13.5% (48)	<0.001
Spontaneous MI within 3 years	5.5% (19)	7.5% (22)	9.8% (36)	0.03
Procedure-related MI within 3 years	8.0% (30)	11.8% (43)	7.9% (30)	0.10
MI (spontaneous or procedure-related) within 3 years	12.6% (46)	18.1% (61)	16.3% (61)	0.17

[a] Data are presented as percentages from Kaplan–Meier curves at long-term follow-up; the number of events is in parentheses.
[b] Table from Clin Chem 2009;55:1118–25. Used with permission.

(i.e., conventional) units in footnotes, which are discussed below.

The remaining columns, which usually represent the dependent variables, should each have a concise heading. Because the units in each data field of a column are typically the same, it is acceptable to state the units just once in the column header, either in parentheses or after a comma, rather than to repeat the units after each value. If you have missing data or no applicable entry for a data field, do not leave the space blank, because the reason for the missing entry may not be clear to the reader. If there are no data for that field, you can insert an em dash (—) or an ellipsis (. . .) to designate that no data are available. The abbreviation NA can also be used, but it must be defined in a footnote, because NA can be interpreted as "not applicable," "not available," or "not analyzed" *(2)*.

Each footnote should be placed on a separate line at the bottom of the table. Most journals recommend that superscripted letters rather than numbers be used to designate footnotes, because a superscripted number can be confused with an exponent. Some journals prefer the use of symbols to numbers for designating footnotes, so it is wise to read the selected journal's submission requirements. Letters (or numbers, or symbols) designating footnotes should be ordered alphabetically (or numerically), starting with the title of the table and then working left to right and down, just as in reading text. The symbol designating a footnote that applies to the entire table should be placed after the title. Symbols for footnotes that apply to an entire row or column should be placed after the row or column heading (Fig. 1), and those that apply to a single data field should be added after the entry for that field *(2)*.

Table Alignment

As an author, you can help your tables make a positive visual impression and make them easy to read if you follow some general rules for formatting and alignment:

- The stubs should all be left justified.
- In the columns/data fields, words should be left justified, and whole numbers should be right justified.
- Data fields containing decimal points, plus/minus symbols, slashes, hyphens, or parentheses should be aligned on these elements.
- When the text in a stub wraps to a second line, the corresponding data field should align with the top line of the stub.

Some examples show how these alignment rules help improve the clarity of a table. Table 1 shows the annual per capita healthcare expenditures for a set of countries. Note how the names of the countries in the stubs and the numbers in the column are easier to read when they are justified appropriately (Table 1A), compared with the centered spacing often seen in drafts of papers (Table 1B). Because the meaning of the stubs is evident, it is not necessary to add a heading for the stubs. Also note that the unit of measure ($) for the data is placed after the column heading because the same unit applies to all of the data fields.

Long-term outcomes according to cystatin C tertile are shown in Table 2. Because the meanings of the stubs are evident, no heading is necessary. Entries in the stub column are left justified. The numbers (percentages) in the data fields contain a decimal point; therefore, the numbers in each column, including the *P* values in the last column, are aligned with the decimal point. Note that to fit within the width of the stub column, the text in the fourth row has wrapped to a new line, but the corresponding data fields in this row have remained aligned with the top line of the stub.

Table 3 also illustrates many of the concepts described earlier for creating an effective table. The title is sufficiently informative that a reader can understand

Table 3. Phenytoin concentrations measured by immunoassay for matrices supplemented with 10 mg/L phenytoin.

	Mean (SD), mg/L	Mean ± SD, mg/L	Deviation from target, %
Pig serum	11.4 (2.1)	11.4 ± 2.1	14
Sheep serum	10.7 (1.4)	10.7 ± 1.4	7
Artificial serum	10.3 (0.8)	10.3 ± 0.8	3
Saline	10.1 (0.6)	10.1 ± 0.6	1
Human serum	9.9 (0.6)	9.9 ± 0.6	−1
Cow serum	9.6 (1.4)	9.6 ± 1.4	−4
Horse serum	8.9 (0.7)	8.9 ± 0.7	−11

the experiment performed and the data produced. The stubs are left justified, and their meanings are sufficiently clear from the title that a header is not needed. The units are placed in the headers for the columns because the units apply to all of the data in the column. Two different styles for presenting the mean and SD are shown in columns 2 and 3. In column 2, which follows *Clinical Chemistry* style, the data are aligned on the decimal point. Some journals allow the style shown in column 3, in which the data are aligned on the common element (±). In the last column, the numbers are right justified regardless of whether the numbers are positive or negative. It might help to pretend that a decimal place exists after each whole number and that you are aligning the numbers with this decimal point.

Tables of Lists and 2-Column Tables

Sometimes a table can be a simple list of information, such as the exclusion criteria for a clinical trial, instrument settings, or sequences of primers used in a PCR assay. Although the items in such lists may not have any inherent rank or order, how you decide to order them may affect what the reader infers *(3)*. Table 4 illustrates 2 ways to list the top 10 states for air quality. Table 4A starts with Idaho and ends with Tennessee. As a reader, I would infer that the point of the table is to show that Idaho is the top state for air quality and that Tennessee is 10th among the 50 states, but if the states were ranked alphabetically (Table 4B), I would infer that the point is simply to list the top 10 states, with no implication regarding which state was best in air quality.

With a 2-column table (Table 1), you have the choice of ordering either the stubs or the data fields. Lang refers to this choice as helping the reader see specific information (organizing the table from the outside in) vs helping the reader see patterns (organizing the table from the inside out) *(5)*. In Table 1A, the

Table 4. Two ways to list the top 10 states for air quality.

A. States with the 10 highest air quality indices.
Idaho
North Dakota
Montana
Alaska
Minnesota
Hawaii
New Mexico
Wisconsin
South Dakota
Tennessee
B. States with the 10 highest air quality indices.
Alaska
Hawaii
Idaho
Minnesota
Montana
New Mexico
North Dakota
South Dakota
Tennessee
Wisconsin

country names are listed alphabetically, and the reader is not drawn to any pattern in the data. Rather, the reader must look for the country of interest and then find the corresponding dollar expenditure for that country. This ordering helps the reader find specific information. The layout in Table 1C helps the reader see a pattern. The dollar values in the data field are ranked from highest to lowest, and the reader is drawn to the pattern of the data rather than to the country names. Again, how you organize the table can affect what the reader sees and infers.

The Best Tip of All: Keep Your Tables Small

One trap that authors often fall into is to assume that all acquired data should be placed in a table. In an attempt to be comprehensive in providing information, you may create a table that is so large and complex that the important data—and the message—get lost among column after column (or row after row) of information. The data that belong in a table are the data that are essential for conveying the results you want the reader to see.

So what is the right size for a table? A useful rule provided by Huth *(6)* is that the maximal width should

be 60 characters and spaces in a row for a table running across half a page and 120 characters and spaces for a table running the full width of a page. For a 2-column journal such as *Clinical Chemistry*, a table width of 110 characters and spaces would fit onto a portrait-formatted page. Otherwise, a journal might consider publishing the table in landscape format (sideways), but this formatting is awkward for the reader and just does not look as nice on the page (Tables 5 and 6).

Fortunately, there are several ways to condense the size of a table. One is to reorient the table so that the variables are reversed in the table *(6, 7)*, as shown in Table 7. This table contains the same data as Table 5 but uses half the width. It also fits a portrait layout on the journal page. If the ratio of the number of column headings to row headings is greater than 2:1, you should evaluate whether reorienting the table would allow a better presentation of the data *(6)*.

Another option is to evaluate whether all of the columns or rows are necessary in a table and whether any nonessential data can be removed. In Table 6, for example, are all of the data and text in the first 3 columns necessary? Will the first author and reference number suffice to identify the studies? Is there a strong reason (e.g., the differences in the number of patients are critical to your discussion) to include the number and sex of the patients if the same information can be found in the reference? Isn't the initial value for a variable, such as the leukocyte count, typically considered 100%? Condensing column 1 and removing columns 2 and 3 yield a table (Table 8) that is narrower and looks cleaner on the printed page. Other columns of data that could be removed and explained in a footnote include those that contain a single data point, a column in which all of the values are the same, or a column in which only 1 or 2 of multiple values differ from the others (e.g., 9 positives and only 1 negative). If a column contains a series of *P* values with only 1 or 2 being statistically significant, this column could be replaced with superscripted letters or symbols after the values, with the significant differences and explanations presented in footnotes.

The use of abbreviations instead of longer names also can reduce the width of a table substantially. Journals allow more leeway in the use of abbreviations in tables because the definitions can be provided in footnotes. As an example of this approach, one could try to abbreviate the types of acute myelogenous leukemia in Table 5 (e.g., UL, undifferentiated leukemia; MML, myelomonocytic leukemia; EL, erythroleukemia). Try this, and see how much this table could be condensed by doing so.

If you have data that could be removed from a large table but are of consequence for the reader to best understand your message, consider placing the most

Table 5. Age-related 5-year survival for forms of acute myelogenous leukemia.

Age, years	Undifferentiated leukemia, %	Myeloblastic leukemia, %	Promyelocytic leukemia, %	Myelomonocytic leukemia, %	Monocytic leukemia, %	Erythroleukemia, %	Microkaryoblastic leukemia, %	Megakaryoblastic leukemia, %
<21	91	80	85	81	82	73	62	52
21–40	89	83	79	77	68	61	57	41
41–60	74	62	68	59	40	37	31	24
>60	51	48	39	34	28	21	16	9

Table 6. Previous studies of leukocyte reduction during kelvac therapy in patients with chronic myelogenous leukemia.

Study	No. of patients	Leukocyte count, %[a]						
		Day 0	Day 7	Day 14	Day 21	Day 28	Day 56	Day 84
Wilkins and Potter, Ref[b] 11	M11;F11	100	97	—	84	—	—	70
Pillsbury et al., Ref 12	M10;F18	100	100	81	—	76	—	64
Annesley et al., Ref 18	M27;F20	100	89	76	—	63	—	62
Kronnenberg and Stenmeyerson, Ref 20	M9;F7	100	103	95	—	88	69	—
Flowers and Peterson, Ref 25	M20;F23	100	101	96	93	89	86	98
Flloyd et al., Ref 26	M27;F23	100	95	—	—	91	—	79
Robinson et al., Ref 27	M19;F20	100	—	100	—	96	—	94
Nowicki and Phillips, Ref 32	M15;F16	100	—	92	—	82	74	—

[a] Percentage of initial value at start of treatment.
[b] Ref, reference; M, male; F, female.

important data showing the results in a printed table and placing the secondary data in a supplemental or online table. In the current era of electronic publishing, most journals make use of these supplemental electronic files as a way to optimize the use of an article's allotted page space, yet such files still allow authors to provide secondary details and data in an easily accessed format.

Sometimes the best option is to split a large table into 2 separate tables. Authors tend to avoid this in an attempt to count 1 large, complex table as a single table to stay within the journal's limit on tables and figures. This trick more often backfires than helps, however. It takes a peer reviewer more time to try to understand a large table, it takes more time and money for an editor or copy editor to reformat the table, and it takes more time for the reader to understand. None of this added effort creates a good impression. My advice is to make the effort up front to split the table.

Table 7. Age-related 5-year survival for forms of acute myelogenous leukemia (AML).

AML type	Age			
	<21 Years	21–40 Years	41–60 Years	>60 Years
Undifferentiated, %	91	89	74	51
Myeloblastic, %	80	83	62	48
Promyelocytic, %	85	79	68	39
Myelomonocytic, %	51	48	39	34
Monocytic, %	82	68	40	28
Erythroleukemia, %	73	61	37	21
Microkaryoblastic, %	62	57	31	16
Megakaryoblastic, %	52	41	24	9

Learning Exercise

Look at Table 9. This table can be improved in several ways to make it more clear and informative. Compare your suggested changes with those provided at the end of this article.

Final Thoughts

When you write a paper, you know that any typographical and grammatical errors will be corrected during the

Table 8. Previous studies of leukocyte reduction during kelvac therapy in patients with chronic myelogenous leukemia.

Study (reference)	Leukocyte count, %[a]					
	Day 7	Day 14	Day 21	Day 28	Day 56	Day 84
Wilkins *(11)*	97	—	84	—	—	70
Pillsbury *(12)*	100	81	—	76	—	64
Annesley *(18)*	89	76	—	63	—	62
Kronnenberg *(20)*	103	95	—	88	69	—
Flowers *(25)*	101	96	93	89	86	98
Flloyd *(26)*	95	—	—	91	—	79
Robinson *(27)*	—	100	—	96	—	94
Nowicki *(32)*	—	92	—	82	74	—

[a] Percentage of initial value at start of treatment.

Table 9. Effect of tacrolimus or sirolimus on everolimus measurement.

Specimen	Measured concentration	Bias, %	*P*[a]
Blood + 10.0 μg/L everolimus	9.9 μg/L	−1	
Blood + 10.0 μg/L everolimus + 10.0 μg/L tacrolimus	10.5 μg/L	5	0.052
Blood + 10.0 μg/L everolimus + 10.0 μg/L sirolimus	14.3 μg/L	43	<0.001

[a] *P* value compared with everolimus alone. $P < 0.05$ considered significant.

final editing process. Nonetheless, you always perform a spelling and grammar check, because you know that confusing or poorly composed sentences do not make a good impression during the review process. The same holds true for tables. Although the editors and printer will make sure that your table is formatted properly before publication, the time to make the best impression is during the review process. A table that not only is easy to read but also emphasizes the point you are trying to make will get you started on a solid footing.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures of Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:*

Employment or Leadership: T.M. Annesley, *Clinical Chemistry*, AACC.

Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

References

1. Hamilton CW. On the table. Chest 2009;135:1087–9.
2. Iverson C, Christiansen S, Flanagin A, Fontanarosa PB, Glass RM, Gregoline B, et al. AMA manual of style: a guide for authors and editors. 10th ed. New York: Oxford University Press; 2007.
3. Browner WS. Publishing and presenting clinical research. Philadelphia: Lippincott Williams & Wilkins; 2006.
4. Annesley TM. Put your best figure forward: line graphs and scattergrams. Clin Chem 2010;56:1229–33.
5. Lang TA. How to write, publish, and present in the health sciences. Philadelphia: ACP Press; 2010.
6. Huth EJ. Writing and publishing in medicine. 3rd ed. Baltimore: Lippincott Williams & Wilkins; 1999.
7. Zeiger M. Essentials of writing biomedical research papers. New York: McGraw-Hill; 2000.

Answers to Learning Exercise

1. The title would be more informative if it included the analytical technique or assay used.
2. An em dash may be added to the last column.
3. Because the unit of measure is the same for each concentration measured, it could be presented once after the corresponding column heading.
4. The numbers, including the *P* values, should be aligned on the decimal point.
5. The information in the stubs for the second and third rows wraps to a second line. The data in the column entries should be aligned with the top line of the corresponding stub.
6. Because only 1 *P* value is statistically significant, one could remove the last column and provide the same information in a footnote.

Clinical Chemistry 56:11
1671–1674 (2010)

The Discussion Section: Your Closing Argument

Thomas M. Annesley*

In the judicial system in many countries, a jury decides the final outcome in a court case. The proceedings begin with a lawyer providing an opening statement telling the jury what he or she plans to present. Through a variety of chosen tactics and methods, the lawyer then presents the various pieces of evidence, all of which lead up to the closing argument. A poor closing argument can hurt even the best case. A great closing argument can convince the jury that the evidence is sound and the lawyer's interpretation of it has merit. In the original *Law & Order* TV show, which incorporated both the investigation of a crime and the courtroom proceedings, the closing arguments were often the most compelling and defining moments in the show.

The Discussion section in a scientific paper and the closing argument in a courtroom have similarities. For many readers, the most important information is not what your results show but what your results mean. The purpose of the Discussion section is to explain what your results mean and what contribution your paper makes to the field of study. The Discussion section is your closing argument. Numerous scientists have told me that when reading a paper they first look at the Abstract to get an overview of the topic and the purported findings. If the topic appears to be of interest, they then skip to the Discussion section. If the Discussion is neither stimulating nor convincing about the meaning and importance of the findings, it does not really matter how the experiments were performed or what results were reported. A poor Discussion detracts from a scientific paper. A good Discussion adds a strong finish to a scientific paper. It brings meaning to your study. My goal with this article is to help you understand the characteristics of a good Discussion section.

Invert the Cone

In a previous article on the Introduction section of a scientific paper *(1)*, I discussed how this section could be envisioned as having the shape of a cone or funnel. The information in the Introduction flows from broad to narrow. The first paragraph provides general background material on the topic, and the last paragraph focuses on the specific question(s) being asked in the study.

In contrast, the Discussion can be envisioned as an inverted cone or funnel, from which the flow of information goes from narrow (top) to broad (bottom). This analogy helps to emphasize the need for the first paragraph of the Discussion to be very specific and focused. This goal is accomplished by getting right to the point, which is to answer the questions(s) presented in the Introduction. As Zeiger states *(2)*, "Thus, the answer to the question is the culmination of the paper. It deserves the most prominent position in the Discussion—the beginning." The hypothetical example below shows how the Discussion picks up where the Introduction leaves off.

End of the Introduction:

We therefore investigated whether β-selectin, vascular lipoprotein-binding molecule (VLM), and interleukin-6γ (IL-6γ) play a role in the vascular inflammation associated with atherosclerotic disease or are just markers that reflect vascular inflammation. Using a herpes simplex virus type 2 (HSV2) infection protocol to stimulate continuous production in mice, we investigated the effects of β-selectin, VLM, and IL-6γ production on the development of atherosclerotic lesions.

Beginning of the Discussion:

In this study, ***we investigated whether β-selectin, VLM, and IL-6γ play a role in the vascular inflammation associated with atherosclerotic disease or are just markers that reflect vascular inflammation.*** *Our results show that* ***in mice, IL-6γ*** *(a)* ***appears to play a role in vascular inflammation*** *and (b)* ***increases the development of atherosclerotic lesions****.*

Three points about the beginning of the Discussion are worth emphasizing here. First, because the Introduction and Discussion sections are separated by other sections of the paper, it is acceptable to provide in the Discussion an introductory sentence that restates the question or purpose of the study. One sentence is usually enough. Second, the restatement at the beginning of the Discussion must match the statement of purpose in the Introduction. In the example above, the authors found it convenient to use the same sentence

University of Michigan Health System, Ann Arbor, MI.
* Address correspondence to the author at: University of Michigan Health System, Rm. UH2G332, 1500 East Medical Center Dr., Ann Arbor, MI 48109-5054. Fax 734-763-4095; e-mail annesley@umich.edu.
Received August 18, 2010; accepted August 23, 2010.
Previously published online at DOI: 10.1373/clinchem.2010.155358

from the Introduction to restate their purpose in the Discussion (boldface text). Third, it is important to answer the question as it was asked in the Introduction, with the same words and key terms. In the example, the authors indirectly asked 2 questions in the Introduction: whether any of 3 compounds played a role in vascular inflammation and whether these compounds had any effects on the development of atherosclerotic lesions. The second sentence of their Discussion thus contains 2 answers (boldface text) that clearly match the original questions in both wording and key terms.

After answering a specific question with a specific answer, you need to describe how the answer is supported by your results:

Our first finding that IL-6γ appears to play a role in vascular inflammation is supported by our protein expression experiments. Twenty-four weeks after injection with cDNA-transfected viral units, only β-selectin was present in sera from mice injected with HSV2–β-selectin, and only VLM was in sera from mice injected with HSV2-VLM. The injection of mice with HSV2–IL-6γ, however, yielded not only high serum IL-6γ concentrations as expected but also high concentrations of VLM and β-selectin, both of which are known to increase with vascular inflammation. Our second finding that IL-6γ appears to contribute to the formation of atherosclerotic lesions is supported by our observation that the mean areas of lesions in mice injected with HSV2–IL-6γ were nearly 3-fold larger than the mean areas of lesions in control mice.

The scope of the Discussion should then be broadened by describing how your results and your interpretation of the results are supported by, consistent with, or related to the results (evidence) from other published studies. If your results support the work of others, you can also use this approach to discuss your results:

Evidence of a possible linkage between serum IL-6γ concentrations and the formation of atherosclerotic lesions comes from the study by Proctor and Schlesser. These authors showed that dietary polyphenols in fruits and vegetables decrease the size and number of atherosclerotic lesions. A secondary finding in their study, which was not explored further, was the observation that polyphenols also reduce serum IL-6γ concentrations. Additional evidence comes from recent studies showing that HMG-CoA reductase inhibitors rapidly reduce both serum C-reactive protein and IL-6γ concentrations, followed by a reduction in arterial plaque density.

Or:

Previous studies have found an association between serum IL-6γ concentrations and the formation of atherosclerotic lesions. Proctor and Schlesser showed that dietary polyphenols in fruits and vegetables decrease the size and number of atherosclerotic lesions. A secondary finding in their study, which was not explored further, was the observation that polyphenols also reduce serum IL-6γ concentrations. A second study by the Canadian All Cause Mortality Coalition showed that HMG-CoA reductase inhibitors rapidly reduce both serum C-reactive protein and IL-6γ concentrations, followed by a reduction in arterial plaque density. The results of our study not only confirm an association between IL-6γ and the formation of atherosclerotic lesions but also show that overproduction of IL-6γ promotes the formation of atherosclerotic lesions.

Toward the end of the Discussion, the "big picture" should now be considered. It is important to describe the contribution your study makes to the field and how your findings can be applied to existing and future studies. For example:

Our demonstration that increased production of IL-6γ is associated with both vascular inflammation and a significant increase in the size of atherosclerotic lesions indicates the existence of multiple pathways that can promote vascular inflammation. It should be possible to design IL-6γ antagonists as therapeutic agents for those individuals who have high serum IL-6γ concentrations. Structural analog antagonists have already been described for IL-4, which plays a role in the allergic response. Another similar advance that supports the potential benefit of IL-6γ reduction is the development of tumor necrosis factor α antagonists that have been effective in treating rheumatoid arthritis.

Close the Discussion with 1 or 2 sentences that provide a take-home message for the reader. This take-home message can restate the answer one last time and/or indicate the importance of the work by stating implications, applications, or recommendations *(2)*. It is important, however, not to repeat items already discussed. Some journals include an actual Summary or Conclusion section at the end of published articles, where these summary points belong. Whether such a message is considered a summary or a conclusion, the worst thing you can do is end with a weak statement, such as "further work is needed to solve this problem," or "we plan on conducting future experiments," or "we have already begun experiments to test our new theory." The final sentences should provide a strong finish:

In summary, our study shows that IL-6γ induces the production of known inflammatory markers and appears to cause an increase in the size of atherosclerotic lesions in mice. Since IL-6γ binds to a different receptor family than C-reactive protein in both mice and humans, there are now at least 2 mechanisms that must be considered when developing new strategies to reduce the incidence and severity of atherosclerotic disease.

Be Fair and Balanced

There is a major news network in the US that has used the trademark slogan "fair and balanced"—fair by being impartial and free from bias, and balanced by presenting all sides of a story. Your Discussion section should also be fair and balanced. There are important points that should be considered to help achieve fairness and balance *(2–4)*.

First, be sure to give credit where it is deserved. If the methods or results from other studies added an important element to your study design or if the work of others supports your findings, state that in the Discussion. Conversely, if you believe that your work supports the findings of others or improves upon what others have done, give yourself the same credit. The key here is to be factual and not boastful about what you have done:

Of the 4 published procedures for nucleic acid insertion, we chose the one described by Wallenburg and Hughes because their procedure yields the highest percentage of cDNA-transfected viruses. Other researchers have successfully used the same procedure to generate viral vectors for the in vivo production of ferritin and transcobalamin, 2 smaller proteins with molecular weights in the same range as our 3 proteins. We were able to improve the yield of transfected viruses 2-fold by adding 0.01% glycerol to the trypsin-EDTA solution.

Second, if you encounter any unexpected results or find that your results (or answer) disagree with other studies, be transparent about these differences, and try to explain them rather than pretending they do not exist:

Smith et al. previously reported that in vitro exposure of cultured smooth muscle cells to IL-6γ does not elicit the release of β-selectin. We were able, however, to elicit the production of β-selectin when mice were injected with HSV2–IL-6γ. It is known that the vascular endothelium must sense both a hemodynamic pressure change and a modulation in signal from receptor–protein binding before microvascular changes occur. This combination can occur only in vivo. Therefore, the difference between our findings and theirs might be due to the in vivo nature of our experiments.

Third, use the Discussion section to acknowledge any limitations of your study and any alternative explanations for your findings. Acknowledging limitations up front makes you look better because you considered, even in retrospect, how the study could have been done better or differently. Recognizing alternative explanations shows that you have a good breadth of knowledge of the field and the factors that might have come into play throughout your experiments. Most importantly, acknowledging any limitations of your study or any alternative explanations preempts a peer reviewer's discovery of them and the opportunity to point them out. If you can explain how the conclusions drawn from the results are probably not affected, do so:

One limitation of our study is that our experiments have thus far been conducted only on mice; however, the results for many studies of atherosclerosis, such as those examining the effects of cholesterol-lowering drug therapy, that were originally performed on mice have subsequently been extended to humans. Such evidence suggests that IL-6γ might have the same effect on the formation of atherosclerotic lesions in humans. Another limitation is that the continuous production of IL-6γ in our cotransfected HSV2 model may not reflect the rate of production or the serum concentrations of IL-6γ that would be required to promote atherosclerotic lesions. Our experiments, however, were designed to examine a cause-and-effect scenario rather than a relative response.

Although our results show that IL-6γ appears to play a role in vascular inflammation and the development of atherosclerotic lesions, its contribution could be primary or secondary in nature. Our initial evidence, when combined with evidence from other studies, supports a primary effect; however, we did not construct a serum or tissue metabolomic profile to identify compounds that might have been up-regulated by increased IL-6γ concentrations. Thus, we cannot discount the possibility that IL-6γ acts in combination with another compound to promote vascular inflammation or that IL-6γ induces the production of another compound that itself is the active agent in vascular inflammation.

Use Transition Words and Phrases

In my previous article on the Introduction section of a scientific paper *(1)*, I talked about how the story becomes clearer if transition words and phrases are used. Transition words and phrases allow the author to emphasize important points and to help the reader recognize a switch from one topic to another. They also serve the same purpose in the Discussion section. Examples from the hypothetical Discussion above include:

- *Our results show that . . .*
- *Our first finding that . . .*
- *Our second finding that . . .*
- *Evidence of . . .*
- *Additional evidence comes from . . .*
- *Our demonstration that . . .*
- *Previous studies have found . . .*
- *The results of our study not only . . .*
- *Therefore, . . .*
- *However, . . .*
- *Thus, . . .*

Learning Exercise

Answer the following questions:

1. What is the purpose of the Discussion?
2. How are the formats of the Discussion and Introduction different?
3. Name 3 types of information that should be included in the Discussion.
4. What is a good way to end the Discussion?

Final Thoughts

There is a well-known saying, "You don't get a second chance to make a first impression." This saying certainly holds true if you consider the importance of the title and abstract of a scientific paper; however, for scientific papers there should also be a saying, "You don't get a second chance to make a final impression." The Discussion is your opportunity to make a good final impression. If you apply the information presented in this article, you will be on your way to doing just that.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures of Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:*

Employment or Leadership: T.M. Annesley, *Clinical Chemistry*, AACC.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: T.M. Annesley, *Clinical Chemistry*, AACC.
Research Funding: None declared.
Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

References

1. Annesley TM. "It was a cold and rainy night": Set the scene with a good introduction. Clin Chem 2010;56:708–13.
2. Zeiger M. Essentials of writing biomedical research papers. New York: McGraw-Hill; 2000. p 176–219.
3. Browner WS. Publishing and presenting clinical research. Philadelphia: Lippincott Williams & Wilkins; 2008. p 115–28.
4. Hess DR. How to write an effective discussion. Resp Care 2004;49:1238–41.

Answers to Learning Exercise

1. The purpose of the Discussion section is to explain what your results mean and what contribution your paper makes to the field of study.
2. The Introduction presents information from broad to narrow (from the larger picture to the specific question). The Discussion presents information from narrow to broad (from the answer to a specific question to the larger picture).
3. The Discussion section should include:

   The answer to the question

   How the answer is supported by the results

   How the results are supported by other studies

   How the results support other studies

   How the results differ from those from other studies

   Any limitations to the study

   Any alternative explanations for the results
4. Restate the answer one last time, and/or indicate the importance of the work by stating implications, applications, or recommendations.

Clinical Chemistry 57:1
14–17 (2011)

Giving Credit: Citations and References

Thomas M. Annesley[1*]

You are nearing the end of the process of writing your scientific paper. You have carefully written a concise introduction, provided a detailed description of your methods, reported your results clearly, and discussed the meaning of your results. You may even have the perfect title and abstract ready to go. But the need to keep your focus and attention to detail remains, because there is still an area where you can stumble and hurt your efforts: the citations and references. Lest you consider references to be a minor component of a paper, consider the fact that the Council of Science Editors devotes 86 pages in their style manual to the proper use of references *(1)* and the *AMA Manual of Style (2)* includes 41 pages covering references.

During the writing process you compiled a file (a stack of photocopied articles or an electronic database) of previously published papers that directly or indirectly contributed to your study. Therefore, it is important to give credit to (cite) the ideas, methods, and results of others. It is also important to tell readers where they can access documentation of this work (references). A citation (typically a number or the author name and year) inserted in the text identifies material that should be attributed to or associated with previously published work. A corresponding reference documents the original source of the material. Citations and references can be a source of information for readers, but they can also become a source of frustration if not selected and used wisely. So let's go over some basics of the use of citations and references.

Organizing Citations and References during the Writing Process

The ideal time to organize (or perhaps reorganize) the materials you might cite is when you begin writing a paper. Organizing potential references at this stage is useful for several reasons. First, the process allows you to identify where in your paper a previously published article is most relevant and should be cited. For example, articles that help you define your topic for the reader or help the reader understand the knowledge gaps that need to be filled might be categorized as important for the Introduction. Prior publications that contain details of methods you applied would be cited in the Methods/Experimental section, and articles that support your results—or help you interpret your results—would be considered relevant for the Discussion section. Second, organizing and reviewing potential references allows you to see the larger picture of the types and scope of articles you have assembled. This process can help you recognize whether your references appear to emphasize one aspect of your study at the expense of another. This process also helps you remove early in the writing process any references that ultimately did not contribute to the study. Third, organizing your references as you begin writing helps you get an early count of the number of citations that are accruing in your paper. Many journals limit the number of references, so early recognition that you are likely to approach this limit can save you grief later. Perhaps a review article or a couple of more-recent papers could substitute for a larger number of general papers that describe your research topic. The one nice thing about the Internet era is that readers can more readily access the older publications cited in newer references or review articles. Thus, you do not need to be ultracomprehensive in citing the literature.

The 2 most common citation and reference formats used today are the citation-sequence or consecutive-numbering system (the "Vancouver system") and the author name–publication year system (the "Harvard system"). During the writing process, it is important to follow the format used by the particular journal you have in mind. This information can usually be found in the Information/Instructions for Authors. Even if the selected journal uses the consecutive-numbering system, many authors find it helpful to use the author name–publication year system in early drafts of a paper *(3)*. The reason is that the numbers assigned to references likely will change if you add or delete a reference or if you modify the sequence of citations in the text. If you insert an author name and year of publication in the text instead of a number, any subsequent changes to citations or references are easy to match. When you prepare the final version of the

[1] University of Michigan Health System, Ann Arbor, MI.
* Address correspondence to the author at: University of Michigan Health System, Rm. UH2G332, 1500 East Medical Center Dr., Ann Arbor, MI 48109-5054. Fax 734-763-4095; e-mail annesley@umich.edu.
Received October 14, 2010; accepted October 19, 2010.
Previously published online at DOI: 10.1373/clinchem.2010.158048

paper, you can assign numbers to both the citations in the text and the corresponding references.

Fortunately, many word-processing programs have built-in functionality for inserting citation placeholders, creating and editing citations and references, and creating bibliographies. There are also several citation/reference-management software programs that can be purchased (e.g., Thomson Reuters' EndNote, ProCite, and Reference Manager; ProQuest's RefWorks) or downloaded for free from the Internet (e.g., Zotero). Most of the available programs will format citations and references in a variety of standardized formats, including the citation-sequence and author name–publication year systems, and will even reformat citations and references from one style to another. If you use a citation-management software program, however, be aware that as the author, you are still responsible for checking the accuracy of your references.

Accuracy and Value

All references included in a paper should be accurate and add to its value. Therefore, the selection of references should follow some fundamental guidelines. Accuracy requires 2 elements. First, make sure that you have read and verified every article or document that you plan to use as a reference. You must verify that the article (*a*) indeed contains the information you are citing and (*b*) is the original source of this information. I have looked up references included in papers and frequently found that they were incorrect in some manner. I have also encountered instances in which a cited paper was actually not the source of the details of the original method or study but instead referred to an earlier publication that was the original source of the information. The only circumstance in which this practice might be excusable is when the original source is very old and difficult for the reader to access and the citing publication does a good job of describing this information or the details of the original method. The second element of accuracy is the correct order and spelling of author names, the correct journal name, and the correct year, volume, and page numbers. Besides being insulting to the authors of the paper you are citing, errors make it more difficult for readers, peer reviewers, and editors to find the cited article. Remember that citations and references serve purposes beyond simply giving credit to others. References provide other investigators with the sources of your methods *(4)*. References add support to the interpretations and conclusions drawn from your results *(5)*. References help editors pick our peer reviewers. References help peer reviewers evaluate your work more effectively and efficiently. Thus, a misspelled author name, an incorrect journal name, or an incorrect volume number can make it difficult for others to access the correct article. In the current age of electronic publishing, the references in many online journals have a direct link to PubMed or the cited journal that makes it easy to access the selected article. An error may cause the reference not to link properly, thus requiring a manual search for the article and leading to wasted time.

References have value only if they contain all of the information and facts to which the author had access. The issue of value is why most journals do not allow authors to cite submitted papers or unpublished results, and why many journals discourage citing abstracts and personal communications. Submitted papers may not be accepted, in which case they are of no value to anyone. Even if a paper is accepted at a later date, the final published version may differ in content from the version being cited. The same problem arises with unpublished results.

Although citations to abstracts might credit researchers who first reported an idea, abstracts have value issues similar to those described above. Readers can usually find abstracts that are published in a regular issue of a journal tracked by PubMed; however, abstracts published in a separate volume of meeting proceedings may be accessible to only a very limited audience. Additionally, because of their short length, abstracts contain limited or no information about many aspects of a study. Unless the readers saw the poster or presentation slides during the actual scientific meeting, they will have no idea whether the abstract matches the final content of the presented material. Therefore, although citing an abstract may be the only available option, be aware of the problems it can cause.

Unless it is vital to the message you are trying to convey, avoid citing personal communications. This type of citation has the same limitations as the citation problems discussed above. Other than what the author states, readers have no access to the actual communication that took place between the author and the person being cited or to the context in which the communication took place. There are generally no supporting data or results. If you decide to refer to a personal communication, it is imperative that the source being cited provide the journal with written permission and confirmation of the accuracy of the cited statements.

Use of Citations and References in a Paper

Maximal clarity should be the goal when deciding where to use citations and references in a paper. In the citation-sequence system, which the International Committee of Medical Journal Editors *(6)* prescribes and which *Clinical Chemistry* follows, references are listed and numbered in the order in which they are

cited in the text. References cited only in figure legends and tables should be numbered according to the point in the paper where the figure or table is first mentioned. This format initially may seem inconsistent because authors are typically told to place figures and tables at the end of a submitted paper, but it becomes clearer if you consider that (*a*) journals insert figures and tables in the final print version near where they are first mentioned in the text and (*b*) readers look at a figure or table when so directed to see what the author is describing and therefore should see citations that directly relate to the figure or table.

Citations should be inserted right after a fact is introduced in a sentence. Thus, a citation might occur in the middle of a sentence (Examples 1 and 2). Wherever you place it, make sure to insert the appropriate citation after the corresponding fact (Example 3). Unless a sentence ends with a fact (in which case the citation follows), do not pool all the citations at the end of a sentence.[2]

Example 1: *Because of the high reported incidence of infections following hip replacement, we added a 2-week course of ampicillin and sulfadrexin (1).*

Modified Example 1: *Because of the high reported incidence of infections following hip replacement (1), we added a 2-week course of ampicillin and sulfadrexin.*

In Example 1, the authors appear to be citing the fact that they added a 2-week course of antibiotics, which could have been a protocol they used in another published study. In actuality, they were citing the fact that the known high rate of infections (a fact needing a reference) prompted them to use antibiotics. Thus, the modified version inserts the citation at the end of the fact, not at the end of the sentence.

Example 2: *Fetal hemoglobin is replaced by adult hemoglobin (1) during the first 6 months of life.*

Modified Example 2: *Fetal hemoglobin is replaced by adult hemoglobin during the first 6 months of life (1).*

In Example 2, the complete fact is not just that fetal hemoglobin is replaced by adult hemoglobin but also that this process occurs during the first 6 months of life. The example refers to a complete, rather than a partial, fact. Therefore, the citation is now inserted at the end of the fact in the modified example, not in the middle.

Example 3: *Although carcinoembryonic antigen is a good prognostic marker for colon cancer, it can also be found in cancer of the pancreas, breast, ovary, or lung (1–8).*

Modified Example 3: *Although carcinoembryonic antigen is a good prognostic marker for colon cancer (1–3), it can also be found in cancer of the pancreas (4), breast (5,6), ovary (7), or lung* (8).

In Example 3, multiple facts occur in the same sentence, and it is important to make it clear to the reader which references correspond to which fact (type of cancer). The modified example accomplishes this goal.

If more than one reference is used to support a fact, list the references in chronological order. In modified Example 3, there are 3 cited references to support the fact that carcinoembryonic antigen is a good prognostic marker for colon cancer. In this case, the oldest reference would be listed as reference 1, and reference 3 would be the most recently published reference. If 2 references are from the same year, list the references alphabetically by the last name of the first author.

If you are submitting your paper to a journal that uses the author name–publication year (Harvard) system, a similar chronological hierarchy holds true. If more than one reference supports a fact, the oldest reference is cited first in the text (e.g.: Smith, 2003; Hopewell, 2005; Corrigan, 2006). If there are 2 references with the same first author, cite the oldest reference first in the text (e.g.: Hopewell, 2003; Hopewell, 2005). If more than one reference has the same publication year and the same first author, differentiate the references by alphabetical letters after the year of publication (e.g.: Hopewell, 2003a; Hopewell, 2003b).

Check Before You Submit

Before you submit your paper, make sure that every citation has a corresponding reference and that every reference is cited in the appropriate spot in the paper. Also make sure that no reference has been included twice in your list of references. Check that every reference is in the proper format for the selected journal and that you have not exceeded the allowed number of references. Make sure that you did not include any references in the Abstract. Take a hard look at any reference citations within the Results section. It is likely that such sentences or ideas more appropriately belong in the Discussion section.

Final Thoughts

Citations play an important role throughout a scientific paper because they occur in nearly every section of the paper, including figures and tables. Similarly, the references at the end of a scientific paper play an important role because they direct readers to resources that can help them understand the study, reproduce the results, and critically evaluate what contribution the study makes. Citations and references that are clear, are accurate, and add value may not get you bonus points, but citations and references that are unclear,

[2] The reference numbers used in the examples do not correspond to any actual references at the end of the article.

inaccurate, or unhelpful will hurt a paper's chances of acceptance. Never underestimate the power of a reference.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures or Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:*

Employment or Leadership: T.M. Annesley, *Clinical Chemistry*, AACC.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

References

1. Council of Science Editors, Style Manual Committee. Scientific style and format: the CSE manual for authors, editors, and publishers. 7th ed. Reston (VA): Council of Science Editors; 2006. Chapter 29, References; p 490–575.
2. Iverson C, Christiansen S, Flanagin A, Fontanarosa PB, Glass RM, Gregoline B, et al. AMA manual of style: a guide for authors and editors. 10th ed. New York: Oxford University Press; 2007. Chapter 3, References; p 39–79.
3. Huth EJ. Writing and publishing in medicine. 3rd ed. Baltimore: Lippincott Williams & Wilkins; 1999. Chapter 12, The first draft: text; p 121–30. Information cited from p 126–7.
4. Browner WS. Publishing and presenting clinical research. Baltimore: Lippincott Williams & Wilkins; 2006. Chapter 9, References and electronic publishing; p 129–35. Information cited from p 129–32.
5. Lang TA. How to write, publish, and present in the health sciences: a guide for clinicians and laboratory researchers. Philadelphia: ACP Press; 2010. Chapter 7, How to write a journal article reporting original research; p 139–74. Information cited from 163–8.
6. International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals; writing and editing for biomedical publications. http://www.icmje.org/urm_main.html (Accessed October 2010).