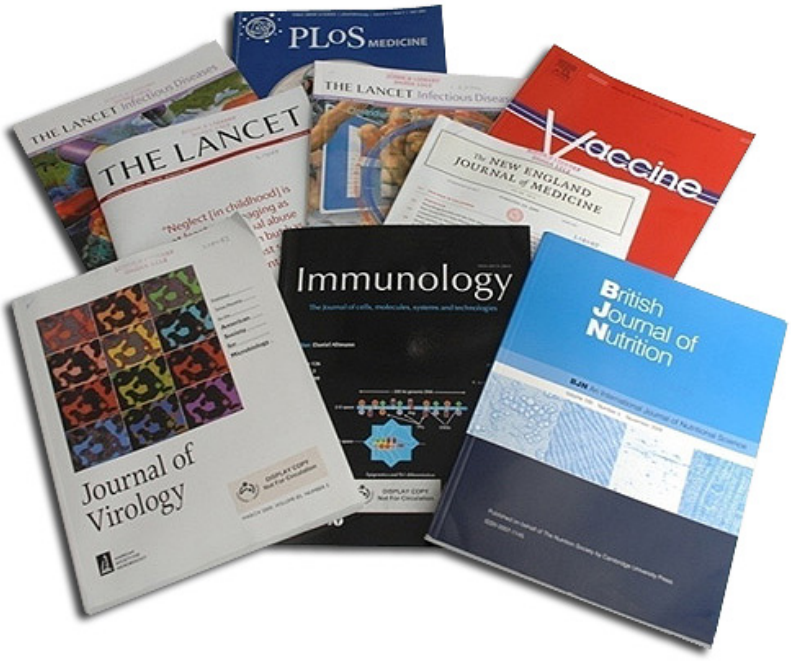
The Pathway to Publishing:

A Guide to Quantitative Writing in the Health Sciences



Steve Luby
Dorothy Southern

Revised
August 2017

Preface

Steve Luby is a medical epidemiologist who has worked for over 20 years conducting public health research in low income countries. This guide grew out of his review of dozens of draft manuscripts from novice scientists in Pakistan in the mid-1990s. To avoid writing the same critique into multiple manuscripts, he developed a short list of 'most common errors' with explanations of how they should be addressed. This allowed him to refer to manuscript errors more quickly by number, and allowed writers to see a more complete description of the problem than might be typed out when they came up again in a manuscript.

Over the years these 'most common errors' multiplied. While working in Bangladesh Steve began collaborating with Dorothy Southern who edited and organized this rather unwieldy list, integrated explanations and examples from a number of different sources, and produced a more systematic guide. As new errors have arisen, they have also been incorporated. Dorothy also worked to broaden the document to describe the mentor-orientated approach to scientific writing that we promoted in the Centre for Communicable Diseases (CCD) at the International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b).

Neither Steve nor Dorothy are now living in Bangladesh, but we both remain involved teaching scientific writing to early career scientists especially those working in low income countries. We have chosen to self-publish the guide so that it can be downloaded at no charge by scientists working in low income countries.

The Pathway to Publishing: A Guide to Quantitative Writing in the Health Sciences focuses on the unique format and data presentation of quantitative studies in the health sciences. It aims to support and encourage scientists who are actively engaged in quantitative research to write effectively, so as to increase the sharing of important scientific results. Since this guide grew out of training public health scientists in Pakistan in Bangladesh, the majority of the examples are from this context, though the principles apply broadly to clear scientific writing.

Bringing scientific work to publication is a group effort. Scientific writing, like the broader scientific enterprise, is a collaboration based on the exchange of ideas. While this guide is primarily focused on providing support to first authors, it also describes the roles and responsibilities of co-authors. Although the specification of these roles were originally articulated to support the management of scientific writing icddr,b in Bangladesh, they remain appropriate principles for the Center for Innovation for Global Health at Stanford University and for other collaborative scientific groups.

We hope you find this guide useful.

Steve Luby, MD
Director of Research
Center for Innovation in Global Health
Stanford University

Dorothy Southern, MPH

Table of Contents

1. Introduction

1.1 The pathway to publishing

Building scientific writing skills is a key skill for researchers. Scientific writing develops critical scientific thinking, helps scientists connect their local results with global understanding and helps scientists identify appropriate next questions to explore. Increased scientific writing capacity means more study results can be shared with the practitioner community and policy makers. More writers mean more work gets published, so all members of the scientific team benefit.

However, there are several barriers to publishing including: lack of focus in framing the research question; inability to explain why the study is important (the ‘so what?’ question); inability to interpret the data and suggest implications for practice or public health policy; unfamiliarity with the requirements of scientific writing formats; and a lack of clarity and conciseness in the use of English language.

The pathway to publishing is a long process that begins with the development of a research idea, and typically requires years to unfold (Figure 1). Often a scientific writers first opportunity as an author will come on a project that was initiated by other scientists. The pathway to publishing process has been diagrammed below to show the relationship between the documents that a researcher might be required to write and the steps along the way to becoming a first author. (Figure 1)

Figure 1: The pathway to publishing

Develop research question(s)
<u>If not funded</u>
Develop a first draft concept note outlining the objectives with broadly summarized methods
After internal review, develop a revised concept note including sample size and budget
After internal review, develop a funding proposal or use the specific donor agency format
<u>If funded</u>
If the funding document lacks sufficient detail, develop a detailed concept note
After review and approval, expand into a study protocol for review by co-authors and institutional review boards
If working with icddr,b in Bangladesh, after co-author review, submit for two external reviews
After responding to all comments, submit to institutional review board(s) (IRB) review
If data will be collected by hand held computer, share data collection tools with the programmers at least six weeks before data collection begins
Implement research activities and collect data
Develop tables shells and then analyze the data to produce completed tables and figures
Develop a high-level outline and share with co-authors and supervisors
After responding to all comments, develop the first draft manuscript
Continue to rewrite again and again, responding to all reviewers comments
Submit for institutional clearance
Submit to appropriate journal
Receive peer reviewers comments and respond appropriately
Re-submit to journal
Congratulations on your first author published manuscript!

1.2 Think before you write approach

To reflect on, think critically about and start writing any type of scientific paper, use the five-step 'Think before you write' approach.

1.2.1 Develop a framing document

The role of a framing document is to assess if proposed results and analysis provide a sufficient basis for a viable manuscript. A single study commonly generates multiple manuscripts. A framing document helps to clarify which results belong in which manuscript. A framing document prevents duplicate publication and provides early feedback to the author to ensure that he/she is on a productive path. Even if there will only be a single manuscript coming out of the study, a framing document helps to clarify the subset of all of the data that the study generated that should be included in a manuscript.

The framing document is primarily a communication to be shared among co-authors familiar with the study. It need not include rationale, detailed methodological explanation, nor Discussion. Think of it as the draft tables and figures for a manuscript with a bit of explanation to clarify framing.

It is, however, important that the framing document be built upon sound data. So first, double-check the quality of your data and your analysis. If you need help, consult a statistician for input. It is a much better learning experience for the author to conduct the statistical analysis with the coaching of a statistician, rather than having the statistician conduct the analysis.

A framing document template is provided in Appendix 4. The framing requires an explicit statement of the objective of the manuscript. A manuscript's objective may be quite well aligned or quite different from the objective of the study. The main results should be specified, if they are either a simple number or are not readily understood from reviewing the tables and figures.

1.2.2 Focus on the high level outline (HLO)

After your senior author and other co-authors have confirmed that the analyses included in your framing document would support a manuscript, the next step is to develop a brief high level outline of the manuscript.

The role of the high level outline is to sketch out the major components of the manuscript that will support the data analysis included in the framing document. This is an outline, that should be no longer than 1500 words (excluding the tables, figures and references). Full sentences are not necessary. A format is provided in Appendix 6.

Keeping the document short helps the author focus on the key elements of the manuscript, and provides early high level input. Because a short document takes less time for authors to produce and less time for co-authors to review, it generates prompt

feedback on key ideas, and so supports a faster path to publication. Using this approach prevents authors investing weeks or months developing full draft manuscripts, that are off target with pages and pages of prose that need to be discarded.

High Level Outline Benefits	
For writers	**For reviewers**
• Bullet points focus on thinking skills, rather than writing skills • Provides framework to guide the thinking process • Allows continuous input and revision	• Content is easy to see and to understand • Short, concise format • Critical importance of results stands out • Easy to change the framing if necessary

1.2.3 Use the 'most common errors'

Use the 'most common errors' listed in '*A Guide to Quantitative Writing in the Health Sciences*' as a method for reviewing and editing the first and all subsequent drafts of a scientific paper. All of the errors listed in the guide have been repeatedly identified in draft scientific papers written by early career writers. These errors range from problems with punctuation, referencing and data presentation to not understanding the difference between association and causality. Examples of the 'most common errors' are provided, along with alternative or better options. Reviewing a paper using the 'most common errors' has several benefits for both the writer and for the reviewer:

The 'most common errors' Benefits	
For writers	**For reviewers**
• Eight categories of errors • Provides more detailed explanations than a reviewer could provide on every point • Real illustrative examples • Systematic learning process	• Covers most errors • Quick easy system • Saves time. No need to repeat explanations • Puts the responsibility on the writer to find the corresponding link to the error and to read and learn about it

1.2.4 Understand authorship and mentoring responsibilities

Scientific writing is a collaborative effort and the framing document and high level outline can provide the opportunity for an initial attempt at identifying the first author and co-authors. Inclusion on an author line is an important indicator of one's contribution to scientific work, and an important professional credential. The roles of both the first author and co-authors should be defined, with clear guidance on what their roles and responsibilities based on the International Committee of Medical Journal Editors (ICMJE). All writers should read the 'Recommendations for the Conduct, Reporting, Editing and Publication of Scholarly Work in Medical Journals (www.icmje.org). The specific responsibilities associated with the various authorship roles are:

First author:

- Conducts the analysis, but may receive substantive input /support from statistical colleagues on complex elements of the analysis
- Constructs the framing document with tables and figures and shares with senior author
- After revision and approval from senior author, shares the framing document with tables and figures with co-authors
- Drafts a <1500 word high level outline
- After revision and approval from senior author, seeks input from co-authors
- Develops multiple high level outline drafts
- Drafts the manuscript
- Follows all the instructions for a draft manuscript as noted in Error A5. (Not using standard draft manuscript form)
- After revision and approval from senior author, seeks input from co-authors
- Develops multiple drafts of manuscript by responding thoroughly and thoughtfully to co-authors' feedback (Avoid Error G2)
- If there is a CDC (Centers for Disease Control and Prevention) co-author, CDC approval is required. The first author:
    - Identifies the first listed CDC affiliate to submit the paper for approval
    - Emails the CDC affiliate the following documents:
        - A MS Word file with the first author and all co-author approvals, with date of approval
        - Completed CDC submission form
        - Draft manuscript to be approved
- Once senior author and co-authors agree, submits the manuscript to a journal
- Circulates submitted draft
- Keeps co-authors informed of all progress on the submission
- Circulates response from editors and comments from reviewers to all co-authors
- Drafts response to reviewers' comments
- Circulates response to reviewers' comments along with a marked up version of the manuscript (to highlight changes) to all co-authors for feedback

Senior author:

- Ensures that the paper is framed to make a meaningful contribution to the scientific literature
- When the first author is an early career scientist, the senior author assumes the role of primary reviewer and assists the first author in:
    - Drafting the author line
    - Selecting an appropriate journal
    - Deciding who should be the corresponding author
    - Identifying external reviewers for journal submission (though first author should generate candidates; see Error G9)
    - Performs the reviews of the initial drafts of the framing document with tables and figures
    - Decides when the framing document with tables and figures is sufficiently developed that it would benefit from review by all co-authors
    - Performs the reviews of the initial drafts of the high level outline
    - Decides when the high level outline is sufficiently developed that it would benefit from review by all co-authors

- Reviews the initial drafts of the draft manuscript
- Decides when the draft manuscript is sufficiently developed that it would benefit from review by all co-authors
- Decides when the draft manuscript is ready for submission to a journal
- Assists the first author in finalizing the author line. For example, if a proposed co-author was included in the initial draft, but never provided any input to the draft manuscript and so does not meet the international criteria for authorship, this co- author would generally be dropped from the author line.
- Carefully reviews the first author's responses to external reviewers' critiques
- Decides when the revised manuscript and responses to external reviewers' critiques are sufficient and the manuscript is ready for re-submission

Second Author:

- The second author is generally the person who made the next largest contribution to the manuscript after the first and senior author, although this designation is sometimes used to denote particularly important institutional collaborators
- The particular role of the second author should be discussed with the senior author. The second author may have additional responsibilities in addition to standard co-author roles including:
    - Drafting sections of the manuscript
    - Performing the role of primary reviewer
    - Functioning as senior author
    - Functioning as the corresponding author

Co-author:

- Provides thorough, substantive review of the high level outline
- Provides thorough, substantive review of the draft manuscript
- Drafts specific sections of the manuscript in one's particular area of expertise and contribution as requested by the first or senior author
- Ensures that the elements of the study that are within his/her area of responsibility and expertise are accurately and appropriately reflected in the manuscript
- Ensures that framing of scientific arguments and references to the literature that are within his/her area of expertise are sound and appropriate
- Assesses whether or not they meet the criteria of co-authorship
- Assesses whether or not they are sufficiently comfortable with the quality of the work, with the integrity with which it was conducted and the conclusions that it reaches that they are willing to accept public responsibility for its content
- Co-authors can opt out of inclusion on the authorship line during any of the drafts, but they should do so before submission to a journal. It is unprofessional to remove one's name after submission because it signals to the journal editor that you were not consulted prior to submission

Getting feedback from the senior author, second author and co-authors is crucial to ensure that a scientific paper clearly describes a valid methodology and communicates convincing results.

1.2.5 Structure the writing and feedback process

In addition to the statistical and writing support recommended above, a feedback system should be put in place to ensure that a scientific paper has the quality to meet the rigor of external reviews and be accepted by a peer-reviewed journal. Getting constructive criticism and focused comments from co-authors helps ensure that a scientific paper clearly describes a valid methodology and communicates convincing results.

First authors should be willing to share any document they are working on: the framing document, the HLO or a draft manuscript. They should expect multiple reviews and revisions, but in a culture of trust and openness. Reviewers should use the 'most common errors' to highlight areas in need of further work, but all feedback provided in the review process must be timely.

A review and feedback schedule needs to be agreed on to ensure the pathway to publishing can be covered in the shortest time possible. Long delays in giving comments and suggestions to improve a scientific paper can de-motivate the writer and delay the dissemination of meaningful research. A suggested time frame for review is:

Structured feedback timeline	
Type of document	**Reviewed within**
Concept note	5 working days
Protocol	5 working days
Framing document	5 working days
Conference abstract	4 working days
Poster	5 working days
High level outline	10 working days
Draft manuscript	10 working days
Reponses to journal editors and reviewers	5 working days

The first author also has a responsibility to continue to dedicate time regularly to the manuscript. Writing well enough so that editors and reviewers agree that your work is a novel and useful contribution to the global scientific literature requires substantial ongoing time commitment from the first author. The biggest difference between people who are authors and people who aspire to be authors but do not achieve this aspiration is that authors dedicate substantial time to writing. A long delay in developing the next draft of a paper means the article loses its developmental momentum and potentially even its relevance to the global scientific discussion. Co-authors usually have multiple on-going projects and will move their attention to other papers. Although a strict schedule for producing revised drafts is difficult to prescribe because substantive critiques may require a deep and critical review of the literature, more in-depth statistical analysis or additional laboratory work or data collection, writers should commit

substantial time each week to keep revising and improving their drafts at short regular intervals.

Summary of the 5 step think before your write process

First authored scientific publications are a prime way to develop scientific reasoning, share an organization's work, and to contribute substantially to global scientific knowledge.

Following this five step 'Think before you write' approach is a win - win situation. Spending initial time developing a framing document and a high level outline saves countless hours in the long run. Responding to the 'most common errors' identified by reviewers dramatically improves the quality of the drafts of any scientific paper in the shortest time possible. Sharing the draft versions of your paper with co-authors on a regular and timely basis will ensure you make steady progress towards publication.

1.3 The scientific writing style

The writing style of quantitative scientific papers is unique. Always use the six 'S's' below to guide your scientific writing:

Structured
Write under the guidance of the high-level outline, knowing where the logic starts from and where it is going.

Sequential
A key characteristic of good scientific writing is reader-centricity. Take the reader by the hand through the sequence of thoughts, step by step, without any leaps or missing links in the development of the ideas. Give the reader information when they need it in a logical sequence that anticipates their questions. This facilitates their ability to interpret and critique the information.

Simple
Use simple words to explain what is meant. Imagine trying to explain the concept to a layperson. Don't use technical or statistical jargon. If you find you about to write or type a word you wouldn't use in every day conversation, stop and simplify.

Short
Use short sentences containing only one idea in each. Split complex sentences. Cut unnecessary information elements and only include those data which relate to the point of your paper. Do not include data just because you collected them. If it is an interesting result, but is not directly related to the focus of the paper, it should not be included in the paper. Remember, 'If it's only nice to know, it ought to go.' If it is a clarifying point, supported by a lot of data analysis, consider including it as supplementary information.

Strong
Use the verb as the center of gravity of your sentence. If the verb is weak, the sentence is weak. For example, instead of, 'We did an interview', write, 'We interviewed'. Use active voice instead of passive. For example, instead of, 'The study was conducted'

write, ‘We conducted the study’. With active voice the subject does the action of the verb, which implies more immediacy and transparency.

Specific
Say clearly and exactly what you want to say. Don’t use qualifiers, which are imprecise and judgmental. Avoid words such as ‘very’, ‘rather’ or ‘much’. Choose your adjectives carefully. Don’t use adjectives that imply subjectivity and/or emotion. For example, ‘It *was a very large outbreak’. What does very mean?* How big is large? Quantitative writing prefers numbers.

2. Most common errors

A. General research and writing practices

A1. Insufficient knowledge of the literature

The first step in developing a scientific document is not writing, but thinking and reading. Good authors are good readers. To write a good paper, you need to develop your own critical thinking, creative thinking, and understanding. You need to have read and critically considered what others have previously reported.

This error can take several forms, such as not having read the relevant literature, not understanding and integrating the work of others into the paper, or ignoring work that threatens or contradicts one’s findings or beliefs. An author needs to understand what has been previously published on the topic in order to frame the research question, and the novel elements of their contribution. If the author lacks sufficient interest in the topic to read about it in detail, he or she is not well positioned to convince readers to be interested in that topic. finally, failure to demonstrate familiarity with the literature and understanding of the topic will jeopardize the credibility of the authors.

Remember, experts in the field will be reviewing your paper. Your initial drafts will be reviewed first by your primary reviewer, then by your co-investigators, co-authors and research group head. When you submit a manuscript to a journal it will be peer reviewed. If you don’t find the most up-to-date relevant information, then a reviewer is likely to do it for you, resulting in embarrassment and/or rejection of your paper.

An author needs to understand and communicate what the state of knowledge in the field is, and describe what your paper adds to what is already known. You are trying to advance the field of knowledge, not just duplicate it. You cannot do this unless you are intimately familiar with what is already known. This should transcend, ‘There is almost no data on this subject in Bangladesh’…the implication being that, anything I say will be an improvement. While limited prior work may be limited, you need to look at similar settings or even dissimilar settings and see what other researchers have found. What are the principle ideas, explanations, and data that are relevant to your particular paper?

If you cannot answer the question, ‘What does this paper add to what is already known about this subject in the literature?’, then you are not ready to write the paper. Expect to spend many days finding relevant articles and reading them critically before you can understand and then communicate clearly what new information or idea your paper adds.

When conducting a literature review, it is, at times, acceptable to put together a concept note or a first draft of a protocol by reviewing abstracts of journal articles. However, to cite information in a paper for submission to a journal you need to have read the complete manuscript, not just the abstract, to understand fully how the information relates to your research. There are two reasons for this. First, on the level of a peer-reviewed publication, the specificity in your statements and the requirements for critical understanding require that you know your colleagues' work at a level of detail that is unavailable from an abstract. Second, there may be something in a manuscript that directly challenges a central idea you are presenting in your paper. If you fail to note it and submit the implications for your paper you will lose credibility in the minds of readers and reviewers.

Finally, the excuse of, 'I couldn't get the paper', is not acceptable in the arena of international scholarship. You can get any paper. Identify what you need and work to secure it. Online resources and collaboration with other institutions and even directly writing authors can secure helpful sources. Different electronic search engines can help you identify different articles: Google Scholar lists the number of times an article is cited; while PubMed lists the most recent articles first.

Examples of the error:	**Alternative, better options:**
× Key studies in the field are not quoted.	✓ Search the literature carefully.
× The studies quoted do not represent the best or the latest studies.	✓ Update literature search, identify "citation classics".
× Studies are misquoted.	✓ Read all cited papers fully, not only the abstracts.

A2. Not referencing statements

Scientific writing demands specificity. All statements that are not common knowledge or do not flow directly from your data need to be referenced. Referencing is a standardized method of acknowledging sources of information and ideas that you have used in your document in a way that uniquely identifies everything readers need to locate each source. Authors must not make general statements about a problem in the absence of quantification, documentation or references.

Example*: It is estimated that by the end of the century, South Asia will surpass Africa to become the region with the greatest number of HIV infected persons*.

Who made such an estimate? On what is this estimate based? This may pass for casual conversation with your colleagues, but in scientific writing the reader needs to know what the precise basis is of everything you are writing. They can then judge whether this specific argument, and ultimately your overall work, is based upon sound research, or not. If it cannot be documented, it must not be said.

Examples of the error:	Alternative, better options:
× Pneumonia is a major public health problem in India.	✓ In 2000, pneumonia was the leading cause of death among children in India. (ref)
× Hand washing is effective against diarrheal diseases.	✓ Community level interventions that promoted hand washing have been associated with reduced incidence of childhood diarrhea. (ref)

Careful referencing is an important strategy to avoid plagiarism. Plagiarism is the appropriation of another person's ideas, words, processes, or results without giving appropriate credit to the original source through referencing. Careful management of references during the research and writing stages of a manuscript or presentation will prevent unintentional plagiarism. In addition, citing up-to-date respected sources will build credibility for your readers.

Learn and use a reference management software. Options include EndNote, Mendeley, Zotero, Mendeley, Papers, JabRef and many others. Reference software helps you track the source of the information and ideas that contribute to your own scientific understanding. Keep a physical or electronic log book during your research. When you identify a good source of information, record the relevant documentation in your notes.

Whether intentional or unintentional, plagiarism is unacceptable. As a scientist your ability to secure funding, to collaborate with other groups on projects, and to have your work published in high profile journals depends on your reputation. Even a single incident of plagiarism can substantially undercut your reputation and so your career.

A3. Weak citations

Scientific reasoning is based upon what can be observed in the world. Authors support scientific arguments by pointing to various observations. An original scientific paper includes new observations and argues that they inform broader understanding. Although it is sometimes appropriate to cite specific arguments, ideas or theoretical models, our most common citations are the observations reported by other scientists. Three common forms of the weak citation error are:

A3a. Citing a secondary source

In this form of the error, the author cites an article that cites the original observation. Standard scientific practice is to cite the primary observation. It is a flagrant error if you cite an article that makes a similar point to the argument you want to make in your article, and the article that you are citing perhaps, in its introduction, cites the primary articles. Avoid this error by simply citing the primary article.

Sometimes it is acceptable to cite meta-analyses or other reviews, but the best practice in most cases is to cite the relevant primary literature even if it requires multiple citations. Citing the primary literature points directly to the empirical basis of the assertion. It specifies where the critical reader should look if s/he is interested in further exploring these data. It also signals to the reader, who may know the literature very well, that you

are also familiar with the relevant literature. If you are citing work that people are not so familiar with, but it is important to your argument, this can be an important pathway for arguing a somewhat different interpretation than the dominant interpretation. This is a process that encourages creative connections, critical thinking and productive scientific argumentation.

A3b. Presenting conclusions rather than data from references

Scientific understanding advances by reasoned interpretation of observation. Indeed, an essential difference between scientific discourse and non-scientific discourse is this reliance on observation as the cornerstone of argument. Thus, if you want to make a persuasive scientific argument you need to present the core data, not just a person's conclusion from that data.

Example: *A baseline evaluation of the quality of sexually transmitted disease case management was conducted in five areas of Madras, in 1992 and it was found that there is an urgent need for health care providers to adopt the syndromic approach to STD treatment*.

In this example, the cited study may well have concluded that the health care providers' performance was so poor in detecting and treating sexually transmitted diseases, that a move to a syndromic approach was the best option. But if this is being presented as evidence that sexually transmitted disease diagnosis and treatment was poor, why should a scientific thinker have to accept the judgment or opinion reached by someone else? Accepting another's judgment without personally evaluating the data upon which that judgment is based is non-scientific reasoning. Non-scientific reasoning is out of place in a scientific manuscript.

Consider the alternative, better option: *In a baseline evaluation of the quality of sexually transmitted disease case management conducted in five areas of Madras in 1992, 74% of persons presenting with symptoms of sexually transmitted diseases were given treatment that differed from World Health Organization guidelines.*

Now, the reader is no longer being asked to accept the interpretation of the author of the original study, or of the author of the present manuscript. He/she has been given the primary observation, the basic unit of reasoning, and so can either accept it as appropriate to the idea being developed or not, but at least can follow the author's reasoning.

A3c. Arguing from authority

An argument from authority asserts that readers should accept a statement as true because of the authority of the person who spoke it. In everyday life we depend upon arguments from authority to help navigate the world. For example, we believe the auto mechanic when he tells us our car will not start because the fuel pump is not working and we believe the attorney we consult who suggests that structuring a contract in a particular way will avoid subsequent legal problems. Arguments from authority are commonly used in many religious traditions and among journalists.

A distinctive feature of scientific reasoning, by contrast, is that it eschews arguments from authority and instead asserts that statements are credible because of the empirical evidence that supports them. Scientists do not believe statements because they were uttered by a prestigious university or government official. Scientific reasoning requires evidence.

Examples of the error:	Alternative, better options:
× Many experts emphasize that shared toilets are the only solution for urban slum residence.	✓ Because of severe constraints on space, shared toilets will continue to be a common option in urban slums the foreseeable future.
× Daniel Kahneman, a Nobel prize winning economist, notes that human decision making is frequently illogical.	✓ Numerous formal assessments find that human decision making is frequently illogical (references).

A4. Endnotes not in standard style

There are many times that a scientist is required to exercise creativity and ingenuity. Writing endnotes is not one of those times. Endnotes for manuscripts have standard formats well detailed in the 'Uniform Requirements for Manuscripts submitted to Biomedical Journals' (www.icmje.org).

There are various software programs that assist in tracking and reporting references including Mendeley, Zotero, Mendeley, Papers, JabRef and many others. Reference software allows writers to format references for various journals with just a few clicks of the mouse. Check the specific format required by the journal you want to submit your manuscript to, and then make sure that you format the references to match those guidelines. Prior to submission be sure to carefully check the specific references, as the software usually makes a few mistakes.

A4a. Varying endnote notation

Different journals use different formats for cited references. There are two basic approaches, either the references are sequentially enumerated in the order that they appear in the narrative, or the references are listed alphabetically at the end of the narrative and within the text one or more of the authors name and the year of publication are noted. Among those journals that prefer sequential numbering, the way they want these numbers displayed varies by journal. Some prescribe that numbers be displayed within square back brackets. Others want numbers in parentheses. Others request superscripts. Some journals want reference numbers to precede periods or commas. Others want them to follow.

If you are drafting a manuscript for journal, look up your target journal's reference format and use it. If you are writing a proposal or other piece of work that does not have a set format, then use a format that is easy for readers to understand. Do not mix formats, that is sometimes using author's last names in parentheses and other times using numbers.

Sometimes copying and pasting from different documents creates this problem. It risks confusing readers and making it difficult for them to connect to your references.

A5. Not using standard draft manuscript form

Most journals have specific instructions for manuscripts submitted to them, usually detailed in their website under 'Instructions to Authors'. However, as a good starting point, the following generic style would be appropriate for a first draft manuscript sent to co-authors for review.

1. Format a title page to include:

    - The title of the article
    - First name, middle initial, and last name of each author (check the journal to see if they have a maximum number of author limit)
    - Each author's institutional affiliation as a superscripted note
    - Targeted journal(s)
    - Main text total word count
    - Abstract total word count
    - Key words

2. Include an abstract in the format and word length of the targeted journal. If the journal choice is uncertain, then include a structured abstract (text separated into sections labelled Background, Methods, Results, and Conclusion) of no more than 250 words.

3. The main text of the article should be in the traditional format of Introduction, Methods, Results, and Discussion. Different disciplines and different journals have different norms regarding the appropriate length of an article. The main text should not exceed the word limit for your target journal. Shorter articles are particularly attractive to most journal editors. If the journal does not suggest a limit, look at the length of articles that they generally publish. A manuscript that is too long risks discouraging reviewers, editors and readers. By contrast, if a report is too short, editors and reviewer can request that more information be included.

4. The manuscript should be double spaced using a common font size 12. This provides more space for comments for reviewers of both the paper and electronic version.

5. The narrative text should be in a single column. Don't try to make it look like a formatted two columned journal article. It makes it harder to review electronically, and it is also not the form it needs to be in for a specific journal submission.

6. Indent the first word of each paragraph one tab width (0.25 – 0.5 inch) or skip a line between paragraphs to signal the reader that this is the start of a new set of ideas. Align text to the left.

7. Insert the acknowledgements after the discussion. Then add references up to the limit permitted by the journal.

8. Tables and/or figures should be placed after the references. There is often a limit of five tables and/or figures.

A6. Repeating information

Editors of scientific manuscripts prefer succinct writing. Don't repeat ideas. Say it well and say it once. A useful strategy to reduce repetition is by carefully considering the logic of your arguments in presenting the ideas so that they build progressively. If a point is so important that you want to ensure that reader see it, then include it in both the body of the paper, and the abstract, which is a summary of the manuscript.

A subtle version of this error is including both proportions of a dichotomous outcome in a results table (see examples).

One situation where a modicum of repetition may be appropriate is in the development of some ideas in the discussion when it is appropriate to link the development of these ideas to specific study results, and/or to issues of study rationale raised in the introduction.

However, in a linked discussion, the important point is not to repeat the words, but rather to make a logical connection between what was raised earlier and the discussion about to take place. Thus, a short recall, without quantitative details, is sufficient. Some journals, including the Lancet, want the first paragraph of the discussion to summarize the main results.

<table>
<tr><th>Examples of the error:</th><th>Alternative, better option:</th></tr>
<tr><td>× "Disease X causes XXX deaths annually worldwide" used in the first paragraph of the introduction and in the first paragraph of the discussion.</td><td rowspan="2">✓ Don't repeat an idea. Say it well and say it once. If you are unsure about where to mention it, review Error B2 that clarifies the respective roles of each section of a manuscript to identify the most suitable place.</td></tr>
<tr><td>× Full repetition of results, with quantified data and statistical tests in the discussion section.</td></tr>
<tr><td>× Sex
o Male 245 (48%)
o Females 273 (52%)</td><td>✓ Males 245 (48%)</td></tr>
<tr><td>× Household pays for electricity
o Yes 3 (10%)
o No/don't know (90%)</td><td>✓ Household pays for electricity 3 (10%)</td></tr>
</table>

A7. Labelling a scientific document as ‘final’

Avoid the word ‘final’ in the title or the description of any scientific document. Scientific thinking is always open to revision. To call a document final implies either dogmatic close-mindedness or naiveté, both characteristics that are inconsistent with a genuine scientific outlook.

Examples of the error:	Alternative, better options:
× Attached is the final version of the protocol	✓ Attached is the version of the protocol approved by the Institutional Review Board
× Here is the final version of the manuscript.	✓ Here is the published version of the manuscript. (Who knows, there may be letters to the editor or subsequent insight that requires further revisions?)

A8. Characterizing an observation as ‘the first’

Scientists take pride in identifying novel observations. Galileo was the first person to see moons around Jupiter. Darwin was the first to both notice the very high variation of bird species on tropical islands and to suggest that this variability was best explained by evolution of species. Watson and Crick were the first to identify the structure of deoxyribonucleic acid (DNA). Part of that task of writing a manuscript is to explain to the readers what is new about the information that is being presented, how this new information changes or refines global scientific understanding. In response, many authors will assert that their scientific findings are ‘the first’. However, there are three problems with describing one's scientific findings as ‘the first’.

1) These assertions can create controversy and ill feeling with scientists writing venomous letters to the editor disputing the claim of primacy. Such ill feelings do not help scientific understanding progress. Indeed, if one of your subsequent papers is then reviewed by one of these scientists who felt slighted by not being appropriately recognized in your earlier work, you risk receiving an unnecessarily devastating review that does not fairly consider the merits or your work. Indeed many journal editors (e.g., those at the Lancet) will not publish claims of first, primarily because they prefer to avoid such non-productive ego driven controversy.

2) Every observation can be described as a first if there are sufficient qualifications. Thus, the assertion of ‘first’ is not, in itself, meaningful. For example, ‘This is the first time that hepatitis E virus has been confirmed using advanced molecular methods in environmental water supplies in Shakira District during the dry season at night using locally trained staff.’ Philosophically, with enough qualifications, every observation is unique, is a ‘first’. Thus, asserting that something is ‘first’ does not communicate why it matters.

3) These assertions distract from useful explanations of how these observations contribute to global scientific understanding. If a health condition has been found in

the other 10 countries where it has been looked for, then saying that this is the first time this has been recognized in Bangladesh tells us more about the interest of Bangladeshi scientists in this condition and the funding available to work in this area than about the health condition itself or the situation in Bangladesh. It does not tell readers why this observation is important.

Like all rules in the guide, this one is not absolute. An occasional claim of first may be defensible and help to clarify to the reader how to interpret the results, but >95% of scientific articles are best written without any claim to 'first'.

Examples of the error:	Alternative, better options:
× This is the first time that an association between hepatitis C infection and carcinoma of the liver has been demonstrated in Bangladesh.	✓ The link noted between hepatitis C and liver carcinoma in this population in Bangladesh provides further evidence of the importance of hepatitis C as a leading cause of hepatocellular carcinoma globally. It suggests that for a low income country like Bangladesh, preventing the transmission of hepatitis C may be the most cost effective way to prevent liver carcinoma.
× This is the first time that Nipah virus antibodies have been identified in dogs in Bangladesh.	✓ Nipah virus infects a wide range of mammals. Earlier studies in Malaysia identified dogs with evidence of Nipah virus infection, but similar to our findings in Bangladesh, dogs appear to be dead end hosts rather than the reservoir of the infection.

A9. Errors in reasoning

Scientific reasoning is central to interpreting our scientific results and so to sound and persuasive communication with our colleagues. There are many ways that scientific reasoning can go astray. Indeed, one of the main benefits we derive from having co-authors and external reviewers critically review our manuscripts, is that they actively criticize our reasoning and so help us to improve it. Some of the differences in opinion on what constitutes appropriate reasoning are based on different interpretations of what is in the literature. What follows, however, are more formal errors in the structure of argument.

A9a. Casual assertion of causality

Scientists take the idea of causality very seriously. Indeed, much scientific work is centered around developing causal hypotheses that explain the relationship between characteristics and exposures in the world and subsequent outcome. When a scientist concludes that a particular chemical exposure caused illness, this is an argument that is based on careful observation, a biologically plausible mechanism, systematically collected data that demonstrates a statistical association and rejection of alternative explanations including bias and chance (see error A9b).

By contrast, when non-scientists speak they tend to be much less careful in their assertion of causality. Everyday business journalists assert that the stock market went down because, for example, the weather was cold, a large company reported disappointing quarterly results, or investors were concerned about recent political developments. Similarly, politicians will assert, for example, that the reason crime is increased in a population is because there are too few police officers. Sport journalist and our colleagues will assert that the reason the home team lost the soccer match is because they did not take their opponents seriously. Each of these assertions may or may not reflect a genuine causal relationship, but none of the people making the assertion is offering a rigorous scientifically persuasive argument.

Such casual assertions of causality, which might be acceptable in casual conversation political speech or daily journalism is not acceptable in scientific writing. Thus, especially in the Introduction and Discussion sections of the manuscript, when you are addressing issues that may be outside of your immediate disciplinary expertise it is critical for your credibility as a scientist not to assert causality unless there is rigorous evidence to support this assertion.

Examples of the error:	**Alternative, better options:**
× Banning overnight poultry storage at live bird markets have been found to reduce Influenza H9N2 circulation substantially in Hong Kong	✓ After overnight poultry storage at live bird markets in Hong Kong was banned, influenza H9N2 circulation decreased among market poultry.
× Due to higher temperature, the number of non-cholera diarrhea cases also increased among the individuals with lower educational attainment, non-concrete roof and unsanitary toilet user	✓ As temperatures increased the number of non-cholera diarrhea cases also increased among individuals with less education, non-concrete roof and unsanitary toilets
× Development project implementation also faltered, the reasons being: financial constraints that produced cost overruns and procurement delays, foolhardy recruitment of under skilled personnel and ill planned career management, and imprecise delineation of the respective roles of development planning and supporting agencies.	✓ Fewer than 10% of development projects achieved their target objectives. Commentators suggest that the factors that most likely contributed to this underperformance included financial constraints that produced cost overruns and procurement delays, recruitment of under skilled personnel and ill planned career management, and imprecise delineation of the respective roles of development planning and supporting agencies.

A9b. Assuming association is causality

Much of our scientific work involves trying to identify associations between different phenomena. For example, is a particular exposure (drinking raw date palm sap) associated with a particular outcome (developing Nipah virus infection)? When we construct 2 x 2 tables or evaluate if there are different mean values between different groups we are exploring whether there are associations within our data. An important element of our data analysis is to identify important associations within our data.

However, just because we find an association, this does not mean that the exposure caused the outcome. For example, if our analysis shows that people who have a lower income have a higher incidence of tuberculosis compared to people who have a higher income, it would be an error in scientific inference to conclude that low income causes tuberculosis infection. Consider for a moment what mechanism we would be asserting. Does the individual *Mycobacterium* have receptors that only attach to the alveolar cells of persons who have an income less than $100 per month? Does the individual *Mycobacterium* wait to see how much money someone spends a month before deciding whether or not to infect him? In this example, low income is probably not best thought of as a causal, but rather as an indicator of an environment that puts certain people at risk. For example, people who have low incomes more commonly have poor nutrition and this poor nutrition reduces the capacity of the body to defend itself from an infection from *Mycobacterium*. Additionally, people with low income tend to live in more crowded settings where it is easier for respiratory diseases to spread from one person to another. Thus, there is an association between wealth and tuberculosis, but the causal mechanism is a deeper underlying mechanism.

There are a number of other reasons that we might find associations between exposures and outcomes in our data. Three common reasons for associations in our data are bias, chance and confounding. There are entire books written on each of these topics and we encourage you to read them. However, when it comes to interpreting your data, any time you see an association, you need to be asking yourself the following questions: What is underlying this association? Is there bias? Could this have arisen by chance? Is this a marker of confounding?

Scientific writing is most persuasive when it invokes a thoughtful, conservative interpretation of association. When discussing an association in the result section, for example, one should never use language that asserts the relationship is causal. In the results you are only presenting the data and identifying associations.

The argument that an association is causal is an argument that should consider the potential mechanism of action; the possibility that the association is a result of bias, chance or confounding; and results from other studies including different types of evidence that supports a causal mechanism. An assertion of a causal relationship is an argument that should be made in the Discussion section; indeed such an argument is often the major point of the Discussion section.

A9c. Assuming reported behavior reflects actual behavior

Our scientific work often considers human behavior, what people do and what might influence what they are do. Scientific study involving human behavior requires considering how to assess behavior. Usually, the easiest and least expensive approach is simply to ask study respondents how they behave. This can be appropriate and useful, but considerable literature illustrates that compared with actual practice people generally over-report socially desirable behavior and under-report stigmatized behavior. Scientists should not take reported behavior at face value, but consider the likelihood that the reported behavior is not accurately reflecting behavior.[1] These considerations are an important aspect of how we interpret our results and so should be considered in the discussion and the limitations.

Sometimes we use research methods that permit us to directly observe behavior. Although the presence of an observer has been repeatedly demonstrated to alter behavior, observed behavior is often less biased compared with reported behavior. Nevertheless, even scientists who study observed behavior must keep in mind the difference between behavior when an observer is present and the behavior that occurs when people are not being observed.

For example, scientific studies comparing reported handwashing behavior to observed handwashing behavior consistently demonstrates that reported handwashing vastly exceeds observed handwashing.[2-4] Indeed, the differences are so great, that reported handwashing behavior is not a valid proxy measure of handwashing practice. Similarly, the handwashing literature provides strong evidence that the presence of an observer markedly increases handwashing behavior.[5-8]

In scientific narrative when referring to behavior that has been studied by other researchers or when describing your own work, it is important to keep in mind the deep biases associated with reported behavior and so when describing behavior it is useful to clarify whether the behavior was observed or reported.

Examples of the error:	Alternative, better options:
× After the intervention respondents were less likely to defecate in the open.	✓ After the intervention respondents were less likely to report defecating in the open.
× In Bangladesh the rate of exclusive breast feeding in the first six months is 64%.	✓ In the 2011 Bangladesh Demographic and Health Survey 64% of mothers reported exclusively breast feeding their children during the child's first six months.

A9d. Confusing imperfect recall with recall bias

Human memory is imperfect. If you ask people what they ate for lunch 17 days ago, most would be unable to provide an accurate response. We do not remember all of our experiences. This is imperfect recall. Imperfect recall does not necessarily constitute a

bias. Recall bias occurs when different people within the study are likely to remember experiences differently. For example, assume you are conducting a case-control study exploring risk factors for leg fractures. If the injury occurred 2 weeks previously and you ask people what they were doing in the minutes preceding the injury, cases, that is people who had experienced a fracture are much more likely to have carefully considered the events that led up to the fracture and so are likely to recall details of what type of shoes they were wearing where they were, what the visibility and footing was. By contrast, if you ask controls about their precise exposures at the same time of day 2 weeks previously they are much less likely to recall rich details of their experience. Thus, there may be systematic differences in the recall of cases and controls, not because their exposures were different but because their recall of events is different. This is recall bias. Study subjects will always have imperfect recall. If there is no reason to believe that this recall will differentially affect reports of exposures or outcomes it should not be labelled as recall bias.

Examples of the error:	Alternative, better options:
× Since the data on exposures to sick poultry was collected by interview, there is a risk of recall bias.	✓ Although we would expect reports of exposure to sick poultry to be an imperfect measure of exposure, because people in this community do not consider sick poultry to be a risk factor for human illness, we would not expect any bias.

A9e. Confusing absence of recognition with absence

Authors should not blithely assume that all occurrences of a phenomenon of interest are known to science and reported in the scientific literature. Many events of scientific interest are neither recognized nor recorded in the scientific literature.

Examples of the error:	Alternative, better options:
× Mortality in ducks and geese as a result of highly pathogenic avian influenza H5N1 infection had never occurred in Bangladesh.	✓ Mortality in ducks and geese as a result of highly pathogenic avian influenza H5N1 infection had never been confirmed in Bangladesh.
× The last of the two Nipah outbreaks from India was in 2007.	✓ The last recognized outbreak of Nipah in India was confirmed in 2007.

A9f. Asserting seasonality with a single year of data

It is an error in scientific inference to assert that a phenomenon that occurs at different frequencies at different times of a single year of observation is due to seasonality. This is an error because it is assumes a pattern when no repetitive pattern has been observed. With only a single year of data from South Asia, for example, only one rainy season was observed. Cases may have increased during the rainy season because a new strain of the pathogen was introduced into the community, a strain that the community did not have immunity against. The strain may have been introduced during the year of observation during the rainy season, but the following year a new strain might be introduced at a different time of year. We are much less prone to scientific error (and

have much more credibility) if we draw conclusions conservatively from our data. Multiple years of data that show a similar pattern provide a stronger case to assert that the variability in the observation over time is associated with seasonal patterns.

So what should we do if we have one year of data and see more cases in the rainy season than in the dry season? It is reasonable in the Discussion section to note that the cases were more common in the rainy season and that multiple years of data would need to be observed to see if this is a seasonal pattern. It would be an error, however, when referring to a single year of data to describe it as seasonal.

A9g. Drawing conclusions using confirmation bias

Confirmation bias refers to the human proclivity to see patterns in the world that are consistent with previously held beliefs.[9] It is a particularly pernicious bias for scientists, because we strive to bring forward new information and to draw sound conclusions.

Confirmation bias often asserts itself among scientists when we look at our data and see the patterns that we expect, for example if people in the intervention group reported less illness, then the data makes sense to us and we don't dig deeper. By contrast, when we find an association that is unexpected, for example that disease is more common among people who received the intervention, then we carefully re-evaluate the evidence. We check to see if we made a coding error in the analysis or if there was some way the question was framed that might have confused respondents. In short, we invoke a double standard of accepting results that confirm our preconceptions and looking to identify problems with evidence that runs counter to our expectations.

Another common manifestation of confirmation bias in science is interpretation of borderline p-values. If the association is not consistent with the unifying theory that the author is proposing, it is either not mentioned or dismissed as not significant. By contrast the p- value of 0.10 that supports the author's tidy interpretation is often framed as a "borderline result that supports this interpretation."

Confirmation bias is so deeply rooted in our human capacity to see patterns in information and the incentives that scientists have to find interesting associations that it is difficult to avoid. One of the benefits of peer review is that reviewers may not share the authors' preconceptions and so offer alternative interpretations of the data.

As an author, consider the risk of confirmation bias in your interpretation. Seriously consider the strengths and weaknesses of alternative interpretations. Consider the limitations in your data and available data in supporting the most likely interpretation. A conclusion that is based on evidence while also conceding weaknesses and alternative interpretations is more persuasive to a scientific audience.

Examples of the error:	Alternative, better options:
× The evidence supports that pesticides contributed to the elevated lead levels among mothers.	✓ The evidence that pesticides contaminated with lead were associated with elevated blood levels is mixed. We found a strong association with reported use of a particular brand of pesticide and blood lead levels, but when we later collected some samples of this pesticide, those samples did not contain lead. It is possible that lead arsenate intermittently contaminates commercial pesticides, but further study will be needed to assess this.
× We found no association between child nutritional status and risk of infection.	✓ Both well-nourished and poorly nourished children were at risk of infection. Indeed, we found no association between child anthropometric measures and risk of infection, though the number of observations were small so we had limited power for this assessment.

A10. Constructing a multivariate model using only statistical criteria

Scientists are commonly interested in explicating causal pathways, that is understanding how various factors interact to produce a particular outcome. Much of our research efforts are aimed at explicating these pathways. When we explore statistical associations between exposure and outcome we are usually striving to understand if there is an underlying causal connection.

Real world causal pathways are usually complex. Multiple factors generally need to be present (e.g. there is a pathogen in the environment, there is a person who is exposed to the environment, the person is susceptible to the infection). In addition, causal pathways typically have sequences where one exposure must precede another in order for the effect to occur. For example, the pathogen must be present in the environment before the person enters the environment. We are much more likely to add insight to global scientific understanding of underlying causal pathways if we seriously reflect on the likely underlying causal mechanism and then construct our investigations and our data analyses to query these pathways.

All too commonly analyst simply dump all of their exposure variables into a multivariate model and use backward elimination to identify those exposures that are most strongly associated with the outcome and then offer this as a final model. This approach provides no consideration for the potential that two variables may be measuring the same underlying characteristics. It also invokes an implicit causal structure that all of the exposures occur simultaneously and without interacting with each other to generate the outcome. This is a naïve and unlikely map of the way processes unfold in the world.[10]

A better approach is to develop a causal model that explicates how the scientist believes the various factors are likely to co-produce the outcome and use this conceptualization to decide which factors to test in the model. There is considerable scholarship on directed acyclic graphs which provide graphical support to help illustrate proposed causal paths and the impact of confounding and temporal sequencing.[11,12] The researcher's proposed causal model can be included as a figure in the paper. This way,

readers can follow the hypothesized causal map and so the judgments used in building a multivariate model.

This is a very different approach than large machine learning efforts that aim not to detect causal relationships but only to find associations and then use those associations to predict subsequent activity. This type of prediction algorithm development has been remarkably successful at identifying patterns in marketing data. In some settings this widespread search for association in large data sets have been used to identify unexpected associations that may be worth further exploration. This approach remains uncommon among scientists who generally strive to elicit causal understanding. The statistical approach employed should align with the analyst's aspiration.

Examples of the error:	**Alternative, better options:**
× Tobacco use and male sex are highly correlated (1/34 female respondents reported regular tobacco use as compared to 11/16 males); therefore, although both characteristics meet the specified criteria for inclusion in the final model, only male sex is included.	✓ Tobacco use and male sex are highly correlated (1/34 female respondents reported regular tobacco use as compared to 11/16 males); because tobacco use is known to affect taste (the primary outcome) it was included in the model and sex was dropped.
× We used univariate logistic regression to select predictor variables significant at the $p<0.2$ level for inclusion in the full model. We used sequential backward elimination of variables with the weakest association to reach the final model of variables all with $p<0.05$.	✓ Exposures were grouped in four blocks by following the conceptual model: (1) attitude, (2) knowledge, (3) school facilities and programs, and (4) practices. We performed bivariate analysis between exposures and outcome to calculate crude association. We further considered only those exposures associated with outcomes with a $p<0.2$. We then conducted multivariable analysis among the exposures within each block including confounders identified in the conceptual model. We retained exposure within each block associated with outcome at the $p<0.05$ level. We then built an overall multivariate model by using exposure variables from each block that were associated with school absence at the $p<0.05$ level and which captured most of the measurement.

B. Content of quantitative papers

B1. Improper focus or format of title and abstract

The title and the abstract are the most visible parts of your manuscript. Because most people rely on electronic search engines to find articles, it is particularly important to catch the reader's attention by drafting a title and abstract that is as concise, accurate, and readable as possible, and to include key words that potential readers of the paper are likely to use during a literature search. When writing a title be as descriptive as possible and use specific rather than general terms.

Check the specific 'Instructions to Authors' for the journal you plan to submit your manuscript to and note the permissible length of the abstract and whether they are looking for a structured or unstructured abstract. As most readers will only read the abstract, it is important that you craft your abstract so that it includes all of the essential information within this word limit.

The abstract must stand alone. It must tell the reader why the topic is important, what the researchers did, what they found out (the most important results and data from the study) and how these findings make a contribution to knowledge. Do not cite references or use abbreviations. In an unstructured abstract, methods and results can be merged to a certain extent. A structured abstract should include the following separate sections:

- Background: Explains the rationale for conducting the study, that is, why is this study question important? The last sentence in the background should state the objective of the abstract / manuscript. If space limitations are severe, and there is only sufficient space for a single sentence of background, the one sentence should be a statement of the objective.
- Methods: Summarizes how the study was carried out. It describes the study population and explains the key techniques used to generate the primary results reported in the article.
- Results: Presents the main findings of the study as specific quantitative results
- Conclusion: A brief interpretation of the findings, why these results matter, what their broader implications are

B2. Confusing the role of Introduction, Methods, Results, and Discussion

The standard structure that most journals prefer for a quantitative scientific paper typically includes the Introduction, Methods, Results And Discussion (IMRAD). The IMRAD structure is explicitly recommended in the 'Uniform Requirements for Manuscripts submitted to Biomedical Journals' (www.icmje.org).The content of each of these sections is ruled by conventions that help readers quickly understand the article. The Introduction explains why the research question addressed in the manuscript is important, the Methods describe how the study was conducted, the Results present the findings, and the Discussion builds upon the results to draw conclusions.

These conventions allow the reader to quickly look for the information they are interested in if they choose to read selectively (a common practice). See Appendix 6 for more clarification about what to include in each section.

Examples of the error:	Alternative, better options:
× Too many details in the Introduction section.	✓ Bypass burden of disease and other general considerations and use a direct sentence that drives the reader towards the research question or problem statement.
× Too many details in the Methods section.	✓ Focus on key considerations needed to understand what was done. Do not spell out methods for which you do not present results.
× Too many details in the Results section.	✓ Narrow down on a set of results that are key for the conclusion.
× Too many details in the conclusion section of the discussion.	✓ Use two short sentences: one to give the big picture related to how your results help us understand a broader topic; then one to state what implications your results have for public health actions or policy.

B3. Not writing the Methods section in chronological order

The Methods section typically involves explaining a number of interrelated activities. A common error is a disorganized series of sentences that jumps back and forth between various activities. This risks confusing the reader. The order that is generally easiest for a reader/reviewer to understand is chronological order. The first part of the Methods section for a public health paper is commonly a brief description of the study site and population to explain the context. Then, the method section explains in detail the study activities that were performed in sequential chronological order. In a protocol, the methods are written in future tense as these are planned activities. In a manuscript, the methods section is always in past tense, to tell the reader exactly what the researcher did.

Example of the error:	Alternative, better option:
× We will also obtain age and socio-economic status data over the phone and demonstrate distribution of typhoid fever mortality in different age groups and income groups. We will select these participants from the laboratory records. We will chose those who are blood culture positive. This will serve as our secondary study objective.	✓ Break down the "methods" section considering the suggested subheadings in Appendix 2, i.e.: o Study site and study population o Study design o Key definitions (e.g., case definitions) o Sampling methods o Data collection tools and processes o Laboratory analysis o Sample size assumptions and calculation o Data analysis plan, including statement of the primary outcome o Ethical considerations

B4. Not emphasizing steps taken to protect human subjects

When describing how ethical considerations were addressed by the study team, a writer can misplace the emphasis by first citing that it was approved by a specific human subjects review committee, and then explaining how the participants' rights were protected, and if there was any benefit or risk to them. This structure mistakenly implies that the cornerstone of ethical practice is approval by a review committee.

Instead, lead off this section by describing key activities undertaken by the study team to conduct an ethical study. Only the last sentence, somewhat as an afterthought, should confirm that all of these procedures, which we developed and carefully and systematically implemented, were reviewed and approved by an appropriate committee. The idea is that the study team acting as moral agents; they are neither delegating the ethical conduct of the study to an external group, nor simply seeking the permission of some institutional authority.

Example of the error:	**Alternative, better option:**
× Our study protocol was approved by the ethical review committee of Aga Khan University. Before collecting data we obtained written informed consent form each adult study participant in the household.	✓ We obtained written informed consent from the adult study participants in each household. The study protocol was reviewed and approved by the ethical review committee of Aga Khan University.

B5. Listing interpretations, but not defending one in the Discussion

The role of the Discussion section is to explain what the results mean. Sometimes it is tempting to list all the possible interpretations and 'let the reader choose' what is the most reasonable. This is an abrogation of the responsibility of the author. As the person who analyzed the data and knows the study, you are in the best situation to explain what the most likely interpretation is and defend it. This is not to say that other important potential interpretations shouldn't be mentioned, but rather that you as the author should clearly state what you believe the data means and why. For example, the reader who looks at the following text has no idea which of these interpretations is the most plausible:

'*The difference between the commuting rate and the injury rate may be because men are more likely than women to exhibit risky behaviour, particularly not waiting for the bus to stop, hanging on side and climbing on the roof, and running to catch the bus. It could also be explained by a different gender mix on buses during the observation period in these high risk areas than at other times, or perhaps there are fewer males injured by buses, but this is more than compensated by a disproportionate number of males injured from motorcycles.*'

B6. Not fully explaining limitations

The objective of a section on limitations is not to list all aspects of the study that could be done differently with infinite money and flawless data collection tools in a perfect world. Instead, this section identifies limitations in the inferences that can be drawn from the study. There are four rules for discussing study limitations:

1. State only the most serious limitations. Don't list every possible problem. Although a thesis advisor may be interested in them, a journal reader is not.
2. Explain the limitation, don't just label it. Instead of writing, 'One of our limitations is selection bias', discuss how you enrolled subjects and how this may result in an unrepresentative study estimate.
3. Be as precise about such limitations as possible, e.g., what were the confidence intervals, and level of detection or discrimination allowed by your sample size.
4. Discuss how you interpret the data in the light of this potential problem, e.g., 'It is unlikely that this procedure substantially affected our results, because...'

Example of the error:	**Alternative, better option:**
× Our study was limited by focusing on only one sub-district and so the results might not be generalizable.	✓ Our study focused on only a single subdistrict and so is not representative of the whole country, but the level of economic development, the percentage of the population engaged in agriculture and the seasonal availability of water is fairly typical of the country.
× Our study was limited by its small sample size	✓ The association of illness with date palm sap consumption was unlikely to be due to chance, but because the outbreak only affected 12 people we had limited statistical power to identify other potentially important exposures with smaller effect sizes.

B7. Writing generic recommendations

Only make recommendations that your data can support. They should be applicable to the specific context. For example, avoid suggesting interventions in low income countries that require a level of national income and government capacity equivalent to that of Western Europe (Error B11a).

Generally, recommendations should not simply call for 'more research'. Such generic calls appear self-serving and do not guide the field. By contrast, it is very useful to reflect on what was learned through your study and identify for the global scientific community (including funding agencies) the one or two important research questions that should next be addressed. Don't provide a laundry list of everything you think should be done. Usually you should make no more than two practical recommendations.

Recommendations have to be carried out by someone or some agency. Useful recommendations give clear statements about who the actor is, what they should do,

and when. Within public health and other applied sciences, scientists are often asked to actively assist in translating scientific knowledge to practical advice for non-researchers. A mechanism to achieve this is through knowledge translation briefs, or one-page summaries of key messages and evidence-based recommendations for action derived from the research results. Aimed at the right institutions and interest groups, evidence-based information and recommendations can inform national policy and programs to address important problems.

B8. Presenting new data in the Discussion

The role of the Discussion is to tell the reader what the authors believe the results mean. It is a violation of the standard IMRAD (Introduction, Methods, Results, Discussion) format to present new data in the Discussion section to support an argument you are trying to make. If the data are important enough to be referenced in the Discussion, then these data should be presented in the Results.

B9. Reporting the number of enrolled subjects in the Methods

For studies of human subjects, the Methods section should describe the enrollment criteria for study subjects and how the investigators trained the study workers to apply these criteria to the community where the study was implemented. The first line of the results should describe how many people were approached, how many agreed to participate, and how many were ultimately enrolled and had all samples collected. Think of this first sentence as describing to the reader how you reached the study population, the underlying denominator for the rest of the paper.

The number of enrolled subjects is a topic for the Results section rather than Methods because there is often some difference between what was planned and what was ultimately implemented. The sample size section of the methods describe the planned sample size. The first line of the results describes the sample size that was ultimately realized.

Examples of the error:	Alternative, better options:
× In the Methods Section: "Altogether 330 questionnaires were taken for study."	✓ In the Methods Section: "Study workers visited the study hospital each morning, approached each inpatient who met the enrollment criteria, and invited them to join the study." In the first sentence of the Results Section: "Study workers ultimately approached 349 patients meeting the eligibility criteria; 19 refused and 330 completed an interview.

B10. Specifying the contents of a questionnaire

Journals generally limit the number of words in a manuscript. This both saves on paper for printed journals and also helps to preserve the time and attention of readers so that they can focus on the most important elements of the manuscript. With all scientists having more articles to read than they have time, succinct writing improves the influence of your article.

Listing the various content area queried within a questionnaire used in the study is neither a good use of precious space nor of the reader's attention. Results from a questionnaire that are relevant to the issues raised in the manuscript will be presented in the results. Readers can infer that this reported information was asked about in the questionnaires. Information that was asked about the questionnaire that is not presented in the manuscript need not be included in the methods.

By contrast, the physicals samples that were collected or the standardized observations that the research team made should be specified.

	Examples of the error:		Alternative, better options:
×	Between February and March 2014, researchers administered a 45 minute exposure questionnaire among case and control households. The questionnaire collected information about household size, education level, occupation, and age as well as potential exposures related to food consumption, jewelry, kitchenware, housing materials, and farming practices	✓	Between February and March 2014, researchers administered a 45 minute exposure questionnaire among case and control households.
×	Field workers collected data using a standardized questionnaire and performed spot checks on hand and domestic hygiene and collected food samples. **The questionnaires included questions on household socio-demographic factors, household assets, drinking water source, sanitation facilities, food storage duration, food reheating history, and food serving practices**. After the interview was conducted, field workers performed spot checks on food and hand hygiene practices including container types used for cooking and food storage, container cover status, animal presence in the food storage area, feces in the household compound and food storage area, and cleanliness of utensils, mother and child hands.	✓	After conducting interviews using standardized questionnaires, field workers performed spot checks on food and hand hygiene practices including container types used for cooking and food storage, container cover status, animal presence in the food storage area, feces in the household compound and food storage area, and cleanliness of utensils and mothers' and children's hands.

B11. Naïve theories of change

The underlying motivation for public health research is to generate knowledge that can be used to improve health (in contrast to pure academic research that generates knowledge that is interesting but is not primarily justified by its impact on the world). The Introduction of a public health manuscript explains why the question addressed by the manuscript is important, and the Discussion explains the implications of this knowledge. Authors of public health research are expected to explain where their research should lead. This often involves making recommendations that are outside the set of issues where an individual researcher has been professionally trained. When a narrowly trained researcher asserts how to bring about change, the suggestions risk being naïve and therefore not useful. By beginning with a more realistic model of how changes occur, a model of change that is informed with some understanding of history, political science, economics and sociology, then you can make your scientific work more impactful.

Invoking naïve theories of change create who problems. First, they are lost opportunities. Your published manuscript presents an opportunity to make a credible suggestion to an interested audience on the way forward, to have an impact on public health, and you provide no useful guidance. Second, naiveté undermines the credibility of your voice, and so of the work, and even of your reputation. Readers conclude, “this is a narrow scientist who does not understand the world.”

B11a. Recommending a massive increase in funding

When we evaluate a public health problem in the context of a low income country, and compare how a similar problem is addressed in a high income country context, it seems reasonable to ask that local government authorities take the same steps to resolve the problem. The difficulty with this practical sounding advice is that low-income country government authorities do not have the funds available to them that authorities in high income countries have.

Of course you are concerned about the specific public health problem that is the focus of your paper. However, if everyone working on their area of interest always requests the government to provide more money to replicate what high income countries do, this becomes an impossible agenda for the government to fulfil. Indeed, from the perspective of government decision makers, every sector, including transportation, infrastructure, education, economic development, energy and health, wants more money. While we may passionately believe that allocating more money to the specific problem that is the focus of our research within the health sector would create a better society, in general, this is not a particularly useful suggestion. The demands on government funds so exceed the available funds, that your recommendation is only one among a never ending chorus of similar requests.

If we cannot make a particular government sector richer, what should we do? As the expert on the topic of the paper you are writing, reflect upon and propose practical suggestions that are cost effective, or even better, that cost no money or cost less money than is currently being spent to address the issue. Such recommendations are much more likely to be implemented. Identifying practical solutions to problems, or at

least pointing out where we can begin to develop practical solutions, is a centrally important way that scientists can improve public health.

B11b. Ignoring incentives and barriers

Public health studies commonly assess knowledge among residents of low income countries at risk for a particular health condition. Quite unsurprisingly, such studies generally find that these populations have imperfect knowledge about the health condition under study including ignorance regarding the exposures that increase risk. Many scientific authors then call for an intervention to improve the knowledge of the population, to tell them what they should do.

Such recommendations are naïve because they assume that ignorance is the primary determinant of unhealthy behavior. However, there is abundant evidence both in everyday life and in the scientific literature that knowledge is rarely the primary determinant of behavior. Are people obese because they do not know that eating excessive calorie dense food leads them to gain weight? Do people who smoke cigarettes because they believe they are using a healthy natural product? Do impoverished households in Bangladesh not serve their children fish more frequently because they are unaware that fish is nutritious?

In general, it is much more productive to consider the incentives and barriers that people have to perform actions rather than their knowledge. People consume excessive calories for a variety of reasons including the pleasure of eating, emotional connections to food and acquired habits. People smoke cigarettes because of addiction to nicotine and enjoyment of the smoking ritual. Poor Bangladeshi households do not eat much fish because they do not have the money to pay for it.

Occasionally, improving knowledge can help to facilitate behavior change, but most knowledge interventions fail to improve health. Thus, when a scientist recommends improving knowledge of the at risk population as the primary intervention, it suggests to readers that the author is not well read (Error A1) and is unaware of the strong and consistent evidence that such knowledge interventions nearly always fail to change behavior. We are much more likely to contribute towards improving health, by examining more thoroughly and critically the likely determinants of behavior and then suggesting prevention efforts directed at these determinants.

Example of the error:	**Alternative, better option:**
× Half of the duck flock owner reported disposing of dead ducks by throwing them into an adjacent water bodies. Duck owners should be taught that the Food and Agricultural Organization recommends burying carcasses on site to control avian influenza transmission.	✓ Biosecurity interventions that cost-effectively improve duck survival and egg production are much more likely to be adopted. We recommend further research to develop and evaluate interventions that simultaneously improve duck raisers' profitability and bio-security.

B11c. Assuming weak states can implement

When working on public health problems, we often consider regulatory approaches to constrain unhealthy practices, for example requiring factories not to discharge toxic pollution into the environment or to have people who provide housing, food, water, education and healthcare meet certain standards. Most people expect that some agency within government should maintain and enforce such regulations. Indeed, most people think of such regulations and their enforcement as one of the primary roles of government. Government officials often describe this as being a central part of their role including passing many such regulations.

However, many governments have limited capacity to enforce such regulations. For example in rural Bangladesh, or Pakistan, or Kenya, or Malawi do drivers on rural roads routinely obey the speed limits? Do industrial factories routinely treat their emissions so that they do not pollute air or water? In cities, are building codes enforced? Are criminals who commit serious crimes routinely identified tried, convicted and punished?

Bangladesh and most low income countries are what political scientists refer to as 'weak states'. These states have limited capacity across a range of functions, including limited capacity to enforce regulations. Weak states are unable to enforce regulations both because of a lack of technical capacity as well as pervasive incentives that undermine enforcement. Government agencies in weak states lack bureaucratic autonomy.[13] This means that enforcement actions are strongly influenced by political actors. Indeed, some political scientists describe many governments of low income countries as 'predatory states'. They argue that such governments exploit their position to extract resources from the citizens, without providing the basic functions of government. In weak states it's relatively easy to pass laws, so it may appear that there is substantial progress, but there is very limited capacity to enforce such laws, so there is no discernible difference in a situation before and after a law is passed.

If authors are working in a weak state, but then suggest solutions that presuppose a strong state, for example Singapore or the United Kingdom, perceptive readers conclude that the authors do not understand the context they are working in. This does not improve your credibility as a scientist. Moreover an approach that requires a strong state will not be effective in a weak state, so the suggestion is not a useful. It does not help move towards a healthier situation. Public health problems generally result from multiple determinants that create an unhealthy situation. Suggesting practical low-cost approaches that can actually improve the situation within the constraints of a low income context within a weak state is difficult, but this is why deep creativity, hard work, wide reading, conversation with colleagues and iterative efforts are required tasks of effective public health researchers.

Example of the error:	Alternative, better option:
× The Government of Bangladesh should adopt the manufacturing standards promoted by the European Union to ensure a healthy environment.	✓ In Bangladesh the current equilibrium of low-priced bricks with high externalities to environment and health results from the structure of the underlying incentives of stakeholders. Transitioning these long-standing equilibria to alternative equilibria that produce less damage to health and the environment will require new approaches.

B12. An insufficiently focused Introduction

In a standard scientific manuscript the role of the Introduction is very specific: The Introduction is not a mini review of interesting themes within the broader field of your study question. The Introduction is an argument crafted to persuade the reader of the importance of the study question. After outlining the Introduction, review each assertion and ensure that it directly contributes to a logical, coherent argument that supports the claim that this study question is important. Remove any other points.

Sometimes understanding the study question requires an explanation of the context of the study, or how the present analysis fits within other analyses that have already been published. When this kind of explanation is required so that the reader can understand the relevance and framing of the study question, these elements should also be included.

B13. Failure to clarify key sample size assumptions

Estimating a reasonable sample size for a study requires that the researcher predict what his/her results will be, and then apply the laws of probability to calculate the number of observations that would be reasonably expected to demonstrate a difference of this magnitude with a low probability that the difference was only due to chance. The most common version of this error, which appears in draft concept notes and protocols, is the failure to specify a predicted outcome, or the failure to explain why the predicted outcome asserted by the scientist is reasonable.

Scientists do not conduct studies when they already know what the results will be. The argument, ‘I don't know what the outcome is; that is why I am conducting this study’, is not an acceptable reason for the absence of a defensible argument for sample size. If it were an acceptable argument, it would apply to all studies. Estimating a sample size is an exercise similar to developing a budget for an activity. We cannot foresee all expenses, but we make a judgment based on prior experience to estimate the costs. Similarly, when calculating sample size we make an estimate of what we think we will find, and explain why we think so. Perhaps there will be studies from other regions that have looked at this phenomenon or a similar phenomenon. You may argue that unless a problem is of a certain magnitude, then either it is not important enough or we accept that we won't have sufficient power to see it. A funding agency will look at the sample size estimate, and ask if the money they are investing is likely to achieve the study

objectives. They do not want to overpay, but they want reassurance that their money will not be wasted because the sample size was too small to reach the objectives.

A common variation on this error occurs when the primary study outcome is prevalence. The scientist predicts that the outcome will be 50% because they read in a statistics textbook that estimates near 50% require the largest sample size and so they want to be maximally conservative. This is unreasonable because calculating sample size requires both the outcome variable and a reasonable level of precision. If the estimated prevalence is 50%, then a study that estimates this prevalence ±5% may be reasonable. By contrast if the estimated prevalence is 3 per 10,000 then an estimated prevalence of 50% ±5% would be provide a sample size estimate that is far too low.

There is no simple statistical rule that will allow a scientist to assert a sample size by a mechanical process that bypasses estimating an outcome and making a reasoned argument for this judgment. When writing a manuscript, the methods section should clarify the assumptions that the scientists originally made of the study outcomes.

Examples of the error:	Alternative, better options:
× We calculated a sample size of 400 based on 80% power and 95% confidence.	✓ We assume, based on studies of indoor air pollution from cooking (Alam NE 2004, Jones FJ 1997), that children living in a village located within one mile of a brick kilns will be at 30% increased risk of pneumonia compared with children who live in villages > 5 kilometers distant from brick kilns. If we assume an incidence of pneumonia in this community will be 45 per 100 child years of observation (SE Arifeen 2007) then a sample size of 400 will provide 80% power to detect a difference in groups of 30% at 95% confidence.
× We assume that 50% of the poultry workers (~380) will experience at least one episode of symptomatic illness during the study period.	✓ An earlier study found that 44% of adults in an urban community in Dhaka developed a symptomatic episode of influenza like illness between March and September (MA Azziz 2006). We assume that 44% of poultry workers will experience at least one episode of influenza like illness during six months of observations.

B14. A high level outline that is not high level

The objective of a high level outline is to sketch out the major components of the manuscript that will support the data analysis included in the framing document (See 1.2.2). The phrase “high level” means that the document outlines the major issues for the manuscript, not all of the details or even all of the components that will be included in the manuscript. The narrative should be no longer than 1500 words; 1200 words is even better.

If your narrative outline is longer than 1500 words, respect the time of your co-authors, and edit it to focus on key ideas prior to sharing it. Bullet points are fine. This is not the time for refining grammar and English language scientific prose.

B15. Specifying software used for routine data analysis

There are specific elements that contribute to a study that results in a scientific manuscript, but the manuscript need not, indeed cannot, specify all of these elements. For example it is not necessary to mention the brand and version of word processing software that was used to craft the study protocol. It is not necessary to specify the e-mail program that the principal investigator used to communicate with co-investigators or the operating system that was used on the data server. Similarly, if the statistical analysis is routine, the name of the software program used for data analysis need not be specified. Routine analysis includes calculations of means, medians, standard deviations, interquartile ranges, prevalence, incidence, odds ratios, prevalence ratios risk ratios and their accompanying 95% confidence intervals, simple linear regression, multiple linear regression and multiple logistic regression.

The underlying guiding principle for writing the methods section is that the methods should be presented in sufficient detail so that other investigators could replicate the study. If the statistical calculations are routine, they could be conducted on any available statistical platform, but if they are unusual using a non-standard approach that perhaps required special programming in R or a module that is available only in a particular software package, but is not widely available, then it is appropriate to specify the software and procedure that was used. If not, don't squander the readers limited attention with this irrelevant detail.

In the interest of improving validity and reproducibility, more and more journals are requiring scientists to make the primary data and their analytical code publicly available. There are several platforms including the Open Science Framework (https://osf.io/) that permit this. When posting the analytical program used to analyze the data, it is important to characterize the analytical software and the version number within the posting, though it need not be mentioned in the manuscript.

Examples of the error:	Alternative, better options:
× We performed descriptive statistics using STATA version 14 software.	✓ ~~We performed descriptive statistics using STATA version 14 software.~~
× We conducted all of our analysis using R version 3.1.0.	✓ Our pre-specified analysis plan is available (https://osf.io/6u7cn/)

B16. Presenting rationale in the last sentence of the Introduction

In a standard public health or biomedical manuscript the last sentence of the Introduction is a succinct statement of the objective of the manuscript. All of the rest of the introduction is basically an argument on why the objective is important. When an author inserts a sentence or two after the statement of objectives and concludes with a rationale, this confuses the reader because the basic narrative form has been violated. It makes it difficult for someone who is scanning a paper to quickly identify the objective.

Examples of the error:	Alternative, better options:
× This study aimed to identify national-level menstrual hygiene management knowledge and practices among adolescent school girls and facilities provided by their schools. We examine the association of menstrual hygiene management knowledge, practice and school facilities with absence from school during menstruation. Findings from this study can guide Bangladesh government policy on female education and inform future initiatives to increase female student attendance and school performance.	✓ This study aimed to identify national-level menstrual hygiene management knowledge and practices among adolescent school girls and facilities provided by their schools. We examine the association of menstrual hygiene management knowledge, practice and school facilities with absence from school during menstruation. ~~Findings from this study can guide Bangladesh government policy on female education and inform future initiatives to increase female student attendance and school performance~~.

The rationale is an important aspect of the introduction (see Error B12). It is simply out of place at the end of the introduction.

One exception to this rule is articles in economics journals. In economics journals, the last sentence of the introduction is an explanation of the organization of the paper.

C. Mechanics of writing

C1. Using non-standard abbreviations

One of the great barriers to communication is overuse of TLAs. What happens is that you work in a specific area and you are quite comfortable with a TLA. You make it up, or hear others in your project or area use it and pretty soon you are using it. Now when you have a chance you start writing, instead of recording words you spout TLAs throughout your manuscript. A TLA is a three letter abbreviation. It is annoying to read a passage that is written in code.

While acronyms mean something to those who use them every day, as soon as a document is shared with outsiders, they become an obstacle to understanding. Writers have a tendency to assume that everyone understands them. This is untrue. It is best to avoid all acronyms, all the time. Using the replace feature of any word processor, you can remove them from your text. This means more people can understand your writing, including, for example, journal editors and journalists who are not topic experts in your area and staff who work for policy makers. An article that can be understood without decoding will be understood by more people. It will have a greater influence on global understanding.

It demonstrates respect for readers to avoid insider abbreviations. Using words removes the burden from readers requiring them to refer back to the first use of the abbreviation to decode meaning.

The few exceptions to this rule pertain to acronyms that are so standard that the general population would understand them (e.g., HIV). However, even for these, the acronym should be spelt out the first time it is used in the manuscript. The Editor of the American Journal of Public Health states this succinctly, “We frown on all acronyms but those in universal use.” The editors of The American Journal of Tropical Medicine and Hygiene maintain, “Abbreviations are commonly overused, compromising the clarity of manuscripts. Authors are advised to keep abbreviations to a minimum, using them when they are clearer than long terms (e.g. PCR, DNA), but avoiding them when possible when they are non-standard and idiosyncratic.” The ‘Uniform Requirements for Manuscripts submitted to Biomedical Journals’ (www.icmje.org) recommends, “Avoid abbreviations in the title and the abstract.”

Examples of the error:	**Alternative, better options:**
× The NTCP has not been evaluated.	✓ The National Tuberculosis Control Program (NTCP) has not been evaluated.
× The CSF is scheduled to begin at 12 noon every Monday.	✓ The Centre for Scientific Forum (CSF) is scheduled to begin at 12 noon every Monday.

C2. Using non-standard spaces

This error is particularly common among authors who draft their manuscripts using both right and left justification (Error C8). Perhaps all the squeezing and spreading of spaces

required by bilateral justification makes it difficult for the author to see the error. It remains distracting to the reader, and is a reason to align all text to the left.
Non-standard spacing includes:

1) The absence or too many spaces before or after parentheses.

Example: *To evaluate compliance with current World Health Organization(WHO) guidelines of post-exposure rabies treatment(PET), we interviewed all animal bite victims*. *One-hundred-nine(76%) bites were category III and 33(23%) were category II.*

This is incorrect. There should be a space after 'Organization' and before '(WHO)'. Similarly there should be a space after 'treatment' and before '(PET)'. These should be a space after 'nine' and before '(76%)'. There should be a space after '33' and before '(23%)'.

2) The absence of spaces following a comma.

Example: *On average the workers completed five household interviews,three child assessment,and one structured observation per day.*

This is incorrect. There should be a space after the word interviews, and after the word assessment.

3) Inserting more than one space between words.

Example: *Approximately six million people annually undergo post-exposure treatments worldwide, most in Third World states as a consequence of failure of canine rabies control* programmes *or strategies.*

This is incorrect. There should be only one space after the word *'undergo' and* only one space after the word *'as'.* Non-standard spacing makes a document quite distracting to read, an irritant that you want to avoid with reviewers and editors. There should also be one space between sentences, not two.

4) Inserting a space within a numeral > 1,000

Example: Field workers collected samples from 12, 456 patients.

This is incorrect. There should be no space after the comma. The numeral should be written as 12,456.

If this error has been pointed out anywhere in your document, then search your entire document and ensure that there are no non-standard spaces. This is an easy error to check for and correct on any word processor. Use the 'Find and Replace' feature. Search for two spaces and replace them with one. If you click on the replace all button, then this removes all of the double spaces in the document. You may have to repeat this process a couple of times if you also have some triple or larger series of spaces within your document.

Example of the error:	Alternative, better option:
× Iodine deficiency disorders, including goiter, have been reported in northern areas for many years(5). In 1908, a survey estimated that 80% of the population had visible goiters(6).	✓ Iodine deficiency disorders, including goiter, have been reported in northern areas for many years (5). In 1908, a survey estimated that 80% of the population had visible goiters (6).

C3. Improper spelling

Improper spelling is distracting and unnecessary with the advent of spell checking. Be sure to thoroughly spell check any document you ask others to review. In Microsoft Word either click on the ABC icon or under Tools use the Spelling option.

Note whether your target journal is published in the United Kingdom or the United States. Set the English language appropriate for the journal as the spell check dictionary for the narrative elements of the manuscripts. Do not change the spelling in references, nor in the proper names of institutions.

Example of the error:	Alternative, better option:
× Mixture of American and British English.	✓ Harmonize spelling in article. See 'Instructions to Authors' for guidance.
× Centres for Disease Control and Prevention	✓ Centers for Disease Control and Prevention

C4. Capitalization problems

C4a. USING ALL CAPITAL LETTERS

LOOK AT AN ARTICLE IN YOUR JOURNAL ARTICLE. IS THE TITLE OF THE ARTICLE WRITTEN IN ALL CAPITAL LETTERS? ARE THE TITLES OF THE TABLES AND FIGURES IN ALL CAPITALS? ARE THE WORDS THAT ARE COLUMN AND ROW HEADINGS IN ALL CAPITALS? THE REASON THAT PORTIONS OF JOURNAL ARTICLES ARE NOT WRITTEN IN ALL CAPITALS IS THAT READING TEXT THAT IS WRITTEN IN ALL CAPITAL LETTERS IS ANNOYING. INDEED, RESEARCH HAS DEMONSTRATED THAT PEOPLE READ ALL CAPITAL LETTERS MORE SLOWLY THAN THEY READ STANDARD SENTENCE CASE. THUS, PREPARE YOUR DRAFT IN ACCORDANCE WITH THE STANDARDS OF THE LITERATURE.

To learn more navigate to Google scholar (http://scholar.google.com/). Input the search terms "Reading speed all capitals" and review the nearly 100 year history of research demonstrating the reduced readability of all capital lettering.

Take a lesson from the clarity of scientific findings. Avoid all capitals. If you want to emphasize a divider or a heading, use a larger font or **bold.**

C4b. Capitalizing non-proper nouns

Although you may commonly use an acronym, IEC, to refer to information, education, and communication, that does not make these words proper nouns requiring capitalization. A proper noun refers to a specific person or place. Barak Obama, or the Director General of Health are proper nouns requiring capitalization, but not acquired immune deficiency syndrome (AIDS).

Example of the error:	Alternative, better option:
× In low-income countries, Information, Education and Communication (IEC) should focus on high-risk sexual behaviour.	✓ In low-income countries, information, education and communication should focus on high-risk sexual behaviour.

C5. Failure to spell out an isolated numeral < 10

The International Committee of Medical Journal Editors (www.icmje.org) used to suggest that numbers < 10 should be spelled out in the text ('four' instead of 4). However, in their April 2010 guidance, they no longer make this recommendation. Different journals have different rules on this. Unless journal copy editors recommend otherwise, we recommend you present numerals if you have a direct comparison or multiple numbers in a sentence, some less than ten and some more than ten, but write out numbers if they stand alone.

Example of the error:	Alternative, better option:
× The field team identified 6 community residents with fever and mental status changes.	✓ The field team identified six community residents with fever and mental status changes.
× Following the intervention, five of the 45 health centres were observed to have adequate practices.	✓ Following the intervention, 5 of the 45 health centres were observed to have adequate practices.

C6. Starting a sentence with a numeral

Example: 43 (56%) individuals tested positive to more than one dengue serotype. 24 of them were reactive to type 1 and 2.

Historically many journal and copy editors have considered this incorrect, and not permitted it. However Bob Fontaine, the resident advisor of China's Field Epidemiology program, argues that we should present numbers so they can be easily assimilated and compared. Trying to compare a number that is spelled out in English to a number that is numerically presented in the same sentence is an unnecessary chore - much like trying to read material that is in all capitals.

If you look in leading scientific journals, e.g., *Lancet* and *Science,* you can find examples of articles with numerals beginning a sentence and numerals less than 10 presented numerically.

What should a writer do? The first goal of a writer is to provide clarity and quick understanding. If it is reasonable to initiate a sentence with a number, then do so. If editors (e.g. editors of Centers for Disease Control and Prevention publications) do not permit it, then alternative strategies include:

- Write out the numeral in words.
- Recast the sentence so that it doesn't begin with a numeral, but be careful not to make the sentence too awkward.
- String sentences together with semicolons because the next word following a semicolon does not need to be capitalized, thus numerals are OK.

Examples of the 'error':	**Alternative options:**
× 50 respondents did not complete the survey.	✓ Fifty respondents did not complete the survey.
× 24 study participants (45%) correctly recalled the health education message that they had received.	✓ Of the respondents, 24 study participants (45%) correctly recalled the health education message that they had received.
× 43 (56%) individuals tested positive to more than one dengue serotype. 24 of them were reactive to type 1 and 2.	✓ Forty-three individuals (56%) tested positive to more than one dengue serotype; 24 were reactive to type 1 and 2.

C7. Not indenting paragraphs

To make it clearer to your readers how your paper is organized into different ideas and/or sections, it is important to indicate when one paragraph ends and when another begins. The standard format is to indent the first word of each paragraph one tab width (0.25 – 0.5 inch). An alternative form is to skip a line between paragraphs. If you do skip a line between paragraphs, it is still most appropriate to indent the first word, but is acceptable if you just skip a line. Using either of these formats sends a clear signal to the reader that this is a new paragraph with new information.

C8. Not aligning text to the left

Setting your word processor so that it aligns text to both the left and right margin (justify), distorts the space between letters and makes it more difficult for the reader to read the text. Although it creates a clean look along the left and right side of the page, it makes it difficult to identify spacing errors. Leave such text alignment to the journal that will finally format your article. For drafts that you send for review you want to make these as easy on your co-authors and reviewers as possible. Align all text to the left.

C9. Problems with parentheses

In general, parenthetical phrasing should be avoided in the narrative portion of a manuscript. The major exceptions are to report data or to cite a source that is not appropriately included as an end-note. If you find yourself wanting to use parenthetical structure, consider that you may have written your ideas with sufficient clarity.

1.) Using parentheses to clarify language.

Incorrect example: *Personal harm (physical injury) of a friend was reported by 10%.*
Alternative: *Ten percent of students reported that a friend was physically injured.*

2.) Putting numbers and percentages in parentheses.

Incorrect examples:
The majority (n=64, 92%) of women reported associated symptoms.
The majority (64, 92%) of women reported associated symptoms.

Correct example: *The majority (64,[92%]) of women reported associated symptoms.*
If you want to include both the number and percentage in a narrative results section, use square brackets around the percentage.

C10. Not recognizing when an abbreviation has become a name

Institutions often begin with one name, but as they evolve the original name no longer describes the institution and so the name changes. For example AT&T used to be the American Telephone and Telegraph Company. BRAC used to be the Bangladesh Rural Assistance Committee. Sometimes institutions have an official name (Leland Stanford Junior University) but a different name that the institution actually uses as its regular name and brand (Stanford University).

The acronym for the International Centre for Diarrhoeal Disease Research, Bangladesh is a communications nightmare. It is not simple. It is not easy to understand. It does not accurately describe what the institution does. Beginning in 2010, the institution has re-branded, and like AT&T and BRAC, now wants its former acronym to be a name.

Note that with this re-branding the institution's name it is not capitalized. icddr,b should be written in lower case when it stands alone, or is included in a sentence, even if it is at the beginning of a sentence. This format will make it consistent across all publications. It is also intended to stop readers unpacking it, or spelling it out. Note there is no space between the comma and the letter b. Additionally, avoid the acronym of a specific program or other sub-division next to icddr,b.

Journal editors or reviewers will often assume that icddr,b is an abbreviation (as it was in the past) and request that it be in all capital letters and spelled out. When responding to this request, it may be helpful to use examples of other acronyms that have become names, for example IBM and AT&T.

Examples of the error:	Alternative, better options:
× ICDDR, B	✓ icddr,b
× icddr, b	✓ icddr,b
× PEI/icddr,b	✓ Program for Emerging Infections, icddr,b
× Leland Junior Stanford University	✓ Stanford University

C11. Misplaced commas in large numbers

The standard placement of commas in numbers greater than 999 in international communication is with a comma after every 3 digits and no spaces between digits or between the comma and the digits. The comma is optional, but it can be particularly helpful to readers to understand numbers especially when they exceed 5 digits. The placement of commas and the use of spaces is often different in the Asian subcontinent, but for scientific writing, or anytime you are writing for an international audience, large numbers should be recorded in standard international form.

Examples of the error:	Alternative, better options:
× 7, 51,842	✓ 751,842
× 51, 00,000 doses of vaccine	✓ 5,100,000 doses of vaccine

C12. Varying fonts within the narrative

The font used for narrative text of the manuscript should be a consistent size and style. If the first paragraph is Times New Roman 12 point, then so should each of the subsequent paragraphs. Sometimes during copying, pasting or other editing variable font sizes or types are introduced. Consistency avoids distracting the reader.

Examples of the error:	Alternative, better options:
× We randomly divided consented households into two groups. Fieldworkers visited Group A each Sunday and Wednesday and visited Group B on Tuesdays and Thursdays.	✓ We randomly divided consented households into two groups. Fieldworkers visited Group A each Sunday and Wednesday and visited Group B on Tuesdays and Thursdays.
× We randomly divided consented households into two groups. Fieldworkers visited Group A each Sunday and Wednesday and visited Group B on Tuesdays and Thursdays.	✓ We randomly divided consented households into two groups. Fieldworkers visited Group A each Sunday and Wednesday and visited Group B on Tuesdays and Thursdays.

C13. Using bulleted lists rather than sentences

A list of phrases or words with preceding bullets works well for outlining, for quickly communicating a list on a website summary or (if not overused) on slides that accompany an oral presentation. Although we are quite accustomed to communicating ideas in this format, this is not the standard technique for communicating in a scientific manuscripts. Scientific manuscripts use sentences that flow together in paragraphs. There is quite a long history of written English language that uses sentences and

paragraphs. Indeed, the complexity and nuance that characterizes scientific ideas makes this traditional format work quite well. Moreover, it is what editors and readers expect.

Examples of the error:	Alternative, better options:
× The field team also conducted spot checks to observe the following: o latrine status (hygienic or unhygienic) o presence of open feces (both human and animal) within courtyard o raw food remnants (food waste produced during food preparation) or leftover food within the courtyard o food storage practices	✓ The field team also conducted spot checks within the household compound to observe the type and cleanliness of the latrine(s), presence of animal or human feces and food waste within the courtyard, and food storage practices.
× The defining features of all of these permutations of a District-Based Approach to intervening in the water and sanitation sector are: o A commitment by an intervening organization to work for a period of time longer than that needed for single projects within a specific sub-national administrative district, such as a municipality or county (but smaller than a state or equivalent). o Sets a goal of achieving universal access to water and sanitation services within that district. o Explicitly working with the local government. o Develop capacity in the local government/public sector for planning and maintaining water and sanitation services. o Align with the country's national water and sanitation policy, while engaging, to varying degrees, with national government.	✓ The defining theoretical features of the District-Based approach appear to include: an institutional commitment to work long-term in a specific sub-national administrative district, a goal of achieving universal access to water and sanitation services within that district, explicitly working with the local government, and alignment with the country's national water policy while engaging, to varying degrees, with national government.

C14. Uninformative document names

The scientific document that you create will be shared with co-authors many of whom are likely to be co-authors on many other scientific documents. It helps your co-authors and reviewers keep track of your work if you to create names for your document that are specific. The most clear document names include a description of the document and a version number. It also can be helpful to include your name.

Examples of the error:	Alternative, better options:
× Manuscript.docx	✓ Simple Soap Man v13.docx
× Review response.doc	✓ Response to Ecohealth Reviewers v3.doc
× Concept note.docx	✓ Detecting_lead _in_spices_Concept_note_Jenna v2.docx

D. Grammatical structures and stylistic strategies

D1. Using present rather than past tense

When your work is published it becomes a historical document. Years, even decades, later, people can look back at what you did at that time in that place, and what you learned. The present tense might sound OK to your ear as you are writing your first draft and the project is still ongoing, but after one or two years elapses before your manuscript appears in print, and another couple of years before a reader pulls it out of a MEDLINE search, the present tense will not be correct. Editors will insist on the past tense, so from the beginning draft it in the past tense.

Present tense can only be used in the introduction or the discussion to report established facts, e.g., 'Tuberculosis is a leading cause of death among adults in low income countries.'

Examples of the error:	**Alternative, better options:**
× We enroll every fourth house as part of our study.	✓ We enrolled every fourth house as part of our study.
× Data derived from the Thatta Health System Research Project are used for the study.	✓ We used data derived from the Thatta Health System Research Project for the study.

D2. Failure to use definite and indefinite articles

What is an article? An article modifies a noun. English has two articles: **the** and **a/an**. Neither Bengali, the language of Bangladesh and West Bengal, nor Urdu, the most common language spoken in Pakistan use definite or indefinite articles. Speakers whose first language does not use articles, do not have an intuitive pattern to apply to English.

The is a definite article. It is used to refer to specific or particular nouns. For example, if I say, "Let's read **the** book.", I mean a *specific* book.

A/an are indefinite articles. Indefinite articles modify non-specific or non-particular nouns. For example: If I say, "Let's read **a** book", I mean *any* book, rather than a specific book. If I say, "I would like to go see **an** art exhibit.", I don't have a *specific* art exhibit in mind. There are many art exhibits, and we could be talking about *any* art exhibit. The indefinite article **a** is used when the next word begins with a consonant (e.g. a paper, a writer). The indefinite article **an** is used when the next word begins with a vowel (e.g. an article, an author).

To find out more about definite and indefinite articles go to www.owl.english.purdue.edu

A specific error commonly made by scientific writers for whom English is a second language and whose first language does not use articles, is use of the word "majority" without a preceding definite article. Whenever you use the word "majority" in your scientific writing, ensure that an article precedes it.

Examples of the error:	Alternative, better options:
× Majority of cases (83%) took advice, while very few (17%) did not consult anybody.	✓ The majority of cases (83%) took advice, while very few (17%) did not consult anybody.
× We reviewed the hospital log book to determine in which sub-districts majority of patients resided.	✓ We reviewed the hospital log book to determine in which sub-districts the majority of patients resided.
× Majority of respondents thought the new design was an improvement.	✓ A majority of respondents thought the new design was an improvement.

D3. Excessive use of passive voice

In general, writing should be composed in the active voice because of the sense of immediacy and concision conveyed when the subject of the sentence carries out the action. Fewer words are usually required for the active voice, it is more efficient, and it takes the reader from point A to point B in a 'straight line'. It communicates who the actor was and so provides greater detail and precision. Active voice is closer to normal conversational speech and usually reads easier and with greater clarity. There is nothing inherently wrong with the passive voice, but if you can communicate the same idea in the active mode, you should do so. Your text is likely to have more impact. In other areas of writing, for example business writing and journalism, active voice is almost universally preferred.

Although passive voice is used in many scientific articles, especially in the methods section, active voice is increasingly common. Passive voice is imprecise. It allows you to write without using personal pronouns or identifying the specific people who are the subjects of sentences. Although some writers use it convey the appearance of an objective, fact-based discourse, not limited to or biased by individual perspectives or personal interests, it risks conveying that the authors are not willing to take responsibility for the specified actions. If you are willing to use the word 'we', your manuscript will be more readable.

Active example: The *study team administered a questionnaire.*

With active voice the subject does the action of the verb. The study team is the subject. The subject performed the action, administered the questionnaire.

Passive example: *A questionnaire was administered*.

In passive voice the subject is acted upon. It does not actively perform the verb. The subject is passive. The questionnaire did not do the action of the verb. The questionnaire did not administer. It was acted upon by the verb. It was administered.

<u>When to use passive voice:</u>

The passive voice exists for a reason and using it is not automatically the wrong choice. The passive is particularly useful (even recommended) in two situations:

1.) When it is more important to draw our attention to the person or thing acted upon.

Correct passive example: The results of the study will be published in the next issue of the journal.

Instead of writing: *The editor of the journal will publish the results of the study in the next issue.*

2.) When the actor in the situation is not important: Passive voice is especially helpful in scientific or technical writing or lab reports, where the process or principle being described is of ultimate importance.

Correct passive example: *The first coat of primer paint was applied immediately after the acid rinse.*

Instead of writing: *I applied the first coat of primer paint immediately after the acid rinse.*

Examples of the error:	**Alternative, better options:**
× A non-inferiority analysis was done.	✓ We conducted a non-inferiority analysis.
× A sample was selected.	✓ We selected a sample.
× Questionnaires were administered to the household head.	✓ Field workers administered the questionnaire to the household head.

D4. Improper use of "we"

A major advantage of using active voice is that it specifies who did which action. It is important that this attribution of action be correct. A manuscript's authors collectively write the manuscript. When the manuscript uses the word "we" this refers to the authors. Work that is conducted by field workers or other members of the team who are not on the author line, should not be attributed to the authors.

Examples of the error:	**Alternative, better options:**
× We revisited households three and six months after receiving the filter to assess usage.	✓ Fieldworkers revisited households three and six months after receiving the filter to assess usage.
× We interviewed households at baseline and weekly from August 2005 – September 2006.	✓ Trained enumerators interviewed households at baseline and weekly from August 2005 – September 2006.

D5. Writing from a psychological perspective

Science assumes that the external world, the world outside of our minds, is real. Scientific articles describe observations of this external world, and attempts to integrate them into larger theoretical understanding. What interests or surprises people varies and is more likely due to their own background, their affection for their own hypotheses or transient fads than from valid induction from scientific observations. Thus, when you write emails to your family or articles for the popular press, you can include subjective considerations, e.g., interests, surprises, shock. However, when you are writing a scientific manuscript, you should focus on the ideas relevant to the issues examined in your study, and the consistency of ideas and theories with available evidence.

Examples of the error:	**Alternative, better options:**
× We were surprised to find that people admitted to using alcohol in a country where its use is restricted.	✓ The proportion of people reporting using alcohol was substantial despite the prohibition in place in the country.
× The incremental cost of adding Haemophilus influenza type B vaccine to the existing immunization schedules in low income countries may not be as high as imagined.	✓ Adding Haemophilus influenza type B vaccine to the existing immunization schedules in low income countries would lead to an incremental cost ranging between XX% and XX% of the national immunization budget. (ref)

D6. Using sub-headings in the discussion section

For most articles presenting original research in most journals the Discussion section (unlike the Methods section) is not subdivided. In standard manuscript format, a section explaining limitations, and a section drawing conclusions are included in the Discussion section as outlined in Appendix 6. These sections should not have a separate header labeled "limitations", “recommendations” or “conclusions" unless the journal you are preparing the article for has a specific requirement for such a section.

D7. Misplaced modifiers

A misplaced modifier is a word or phrase that is meant to modify one object in a sentence, but its placement in the sentence implies that it modifies a different object. Sometimes, the reader can figure out what the author meant; other times the meaning is ambiguous. Even if the reader can figure out the meaning, it is sloppy grammar that risks distracting readers.

Examples of the error:	**Alternative, better options:**
× Then field staff selected four girls from each school for interview who had reached menarche.	✓ Then field staff selected four girls who had reached menarche from each school to interview.
× Since 2006, surveillance physicians maintained a registry of patients admitted to three Nipah surveillance hospitals— Rajshahi, Rangpur and Faridpur Medical College Hospitals — meeting the encephalitis case definition: fever or history of fever with axillary temperature >38.5°C (101.3°F) with altered mental status, new onset of seizures, or new neurological deficit	✓ Since 2006, surveillance physicians at the three Nipah surveillance hospitals — Rajshahi, Rangpur and Faridpur Medical College Hospitals — maintained a registry of admitted patients who met the encephalitis case definition: fever or history of fever with axillary temperature >38.5°C (101.3°F) with altered mental status, new onset of seizures, or new neurological deficit
× Interventions to reduce the risk of pig-related diseases can compromise the social and economic situation of pig raisers in predominately Muslim countries who may already be stigmatized.	✓ Interventions to reduce the risk of pig-related diseases in predominately Muslim countries can compromise the social and economic situation of pig raisers who may already be stigmatized.

E. Achieving clarity and conciseness

E1. Labelling rather than explaining

We love our technical terms. We've studied them; we learn them and now while writing a manuscript we finally have a chance to use them! Right? Well, actually not.

Labelling is shorthand for the full development of an idea, but people often have a different idea of exactly what that shorthand actually means. Different people, especially with different disciplinary backgrounds, use the same term differently and read the same term with different interpretation. This makes using these terms a barrier to clear communication.

Strive to explain exactly what you did. Do not label it. The more specific you are about exactly what you did, the easier it is for someone else to understand it. If a methods section reads, "For the hospital catchment area survey, we selected 20 unions, using a probability proportional to size sampling approach." A reader may wonder, What is a probability proportional to size sampling approach? How did you apply this concept to your site? The Methods Section should report methods in sufficient detail so that other investigators could repeat them. So, skip the label, and instead use the space to describe the steps you took to identify and enroll the population.

The three most common labelling issues in papers concern study design, sampling methods and limitations.

Examples of the error:	**Alternative, better options:**
× The population of the catchment area was projected for 2008 on the basis of the 2001 Bangladesh census using population estimation by component method.	✓ We began with the 2001 Bangladesh census of sub-district populations, applied national estimates of crude birth rate, net external migration and national crude death rate (ref) to estimate the population in the sub-districts at the time of the assessment.
× Confounding by wealth is a potential limitation.	✓ Households who had windows that provided cross ventilation, may have been wealthier and possessed other characteristics that improved their children's health and that we were unable to completely control for in the analysis.

E2. Using weak opening phrases for sentences

You should try to use phrases and transitions that move along and develop the central theme of the paper. However, most of the phrases below only reflect the psychological state of either the reader or the writer. Strive to write from the perspective of the ideas you are developing. You are better off having no transition than using such vacuous phrases as the examples below:

Examples of the error:	Alternative, better options:
× It was found out that...	✓ *Delete*
× One important observation from the findings of this study was that...	✓ *Delete*
× We conclude from our data…	✓ *Delete*
× Moreover, our survey showed that...	✓ *Delete*
× Therefore, this will not be an overstatement that...	✓ *Delete*
× It is known that...	✓ *Delete*
× It can be seen from the above table that...	✓ *Describe*
× The explanation could be that...	✓ *Explain*

E3. Using adjectives and qualifiers

Adjectives are words that modify a noun. Adjectives often imply substantial subjective and emotional content, both of which should be minimized in conventional scientific writing. For example, what is 'important' or 'large' to one person, may not be 'important' or 'large' to another.

Qualifiers are words that modify an adjective, but do not carry a specific meaning, such as 'very'. The addition of a qualifier adds to the subjectivity, as in 'very important'. It is better to try to choose the best adjective, and provide justification of its use, and not to use a qualifier.

Examples of the error:	Alternative, better options:
× The outbreak caused very high mortality.	✓ 56% of people infected in the outbreak died.
× This very large outbreak.	✓ This outbreak affected 300 school children.
× The incidence was much higher in children < 5.	✓ The incidence among children < 5 had was six times higher than older children.

E4. Overusing studies or authors as sentence subjects

In general, when referring to other scientific work, the subject of the sentence should not be the study, or the study's author, but the core ideas or results that connect to your manuscript. Ideas and observations referenced from other studies are central to scientific reasoning. The use of a study or a study's author as the subject of a sentence risks distracting the reader from the substance that links to the author's own study. The structure of your sentences should reflect this prioritization of ideas and results over individuals and authors. underlying structure and hierarchy, while the ideas you present flow one into another logically and persuasively.

Example of the error:	Alternative, better option:
× A study by Yoruba in Tanzania suggested that 78% of the clients who presented to traditional healers were females, 95% of who were illiterate and of a low socio-economic group (ref).	✓ Demographic parameters are important because they may influence health seeking behaviour; a study in Tanzania, for example, indicates that educated mothers are more likely to discourage traditional healing practices (ref).
× Curtis et al have championed structured observation as the preferred approach to measuring handwashing (ref).	✓ Using structured observation to assess handwashing behavior has consistently identified lower frequencies of handwashing with soap compared with reported behaviors.

E5. Using non-descriptive numeric or alphabetical labels

Study teams commonly develop some study specific vocabulary (e.g. Group 1 and Group 2, Phase 1 and Phase 2). The study team becomes so familiar with these labels that denote differences that are meaningful to the team that they use these labels in everyday conversation within the study team. It is not surprising, then that when team members start writing about the study, they use these same labels.

However, such labels are inappropriate for a scientific document. Such non-descriptive numeric or alphabetic labels requires your readers to learn your private code, which is useless information not applicable to any other manuscript they will ever read. You want to make your paper as easy to understand as possible. Use descriptive labels for each group.

Examples of the error:	Alternative, better options:
× At baseline, group 1 participants were somewhat less likely to own a television than group 2.	✓ At baseline, participants enrolled from Tongi were less likely to own a television than participants enrolled from Narshindi.
× Group 1 consisted of formal health care providers and Group 2 consisted of informal providers.	✓ The formal health providers had a higher education level than the informal health providers.
× Category A symptoms included cough and difficulty breathing, while category B symptoms included diarrhoea and vomiting.	✓ Respiratory symptoms included cough and difficulty breathing. Gastrointestinal symptoms included diarrhoea and vomiting.

E6. Using respectively

Avoid the respectively structure. It forces reader to go backwards and re-read to mentally connect the pieces. It is an extra effort and breaks the reader's flow of understanding your message. You want to make it easy for them to read from the beginning to the end.

Examples of the error:	Alternative, better options:
× Of the Plasmodium positive children, 17 (4%) and 9 (2%) were positive for *P. falciparum*, and *P. vivax* respectively.	✓ Of the smear positive children 17 (4%) had *P. falciparum* and 9 (2%) had *P. vivax.*
× Attack rates for any post-operative infection between the suspected outbreak period January - December 1996 and for comparison period June - December 1995 were 14% (10/72) and 6% (2/31) respectively.	✓ The attack rate for any post-operative infection between the suspected outbreak period January to December 1996 was 14% (10/72) compared with 6% (2/31) between June and December 1995.

E7. Using the word etcetera

Scientific writing is characterized by precision. 'Etcetera' is not specific. This imprecision suggests that the author's ideas have not been fully formulated or have not been fully thought through. 'Etcetera' should never appear in a scientific concept paper, protocol or manuscript.

Example of the error:	Alternative, better option:
× Medical costs in the hospital included admission fees, bed rent, diagnostic tests, medicine, consultation fees, etc. Non-medical costs included travel, food, tips etc.	✓ Medical costs in the hospital included admission fees, bed rent, diagnostic tests, medicine and consultation fees. Non-medical costs included travel, food, and tips.

E8. Using Bangla as an English word

*Bangla** is not an English word. The English language word for the language spoken in Bangladesh is Bengali (not italicized). When writing about questionnaires in Latin America, scientists do not use the Spanish word for the Spanish language (*español*). They do not write that the questionnaires were translated into *español*. Instead, they write that the questionnaires were translated into Spanish. Similarly, when writing in English about work in Bangladesh, we should describe the local language as Bengali.

* Note that words from other languages used in an English scientific report should be *italicized.*

E9. Using local words, expressions or monetary figures

Most scientific manuscripts are designed to be a form of international communication. If the writer uses words and expressions that are specific to the country where the work was conducted, this information might not be communicated clearly to the reader. The

information might not mean anything, or it might mean something entirely different to readers in other countries. For example, to a reader from North America a 'block' will suggest a group of houses located between four streets in a city and not an administrative division. A *gacchi* will not be recognized as a date palm sap harvester. A *Taka* will not have much significance outside of Bangladesh and most readers outside of the subcontinent will not know the exchange rate between a local currency and their own. If you want your scientific manuscript to be more broadly understood report the information in terms of internationally recognized definitions. For monetary information, report the figure in a major international currency (US dollar, British pound or Euro). At the very least include an appropriate conversion (the one prevailing at the time data was collected) between the local currency and an international currency, so that persons reading it can figure out how much money that is by local and international standards.

Examples of the error:	**Alternative, better options:**
× We conducted a case control study in two *upazilas* in Rajshahi district.	✓ We conducted a case control study in two sub-districts (*upazilas*) in Rajshahi district.
× The cost per fully treated patient was 500 *taka*.	✓ Provide equivalent in US$ and mention in the Methods section the exchange rate that was used.

E10. Using the term 'developing country'

The term 'developing country' is non-standard, imprecise and inaccurate. All countries are developing. Japan is a different country in 2017 than it was in 2000. It has higher income and a greater number of internet connections. It is developing. Japan will look different in 2030 than it does today. It will develop further. Although the term historically connotes industrial development, there is no standard definition of what constitutes a developing country.

By contrast, the World Bank has clear standards for characterizing low income countries. There is an accepted definition for country classification and using the criteria of gross national income is meaningful. For more information see www.worldbank.org under Data and Statistics. Since 2015 Bangladesh has been classified as a lower middle income country.

E11. Using the term 'socio-economic status' as a synonym for wealth

When referring to income or poverty/wealth among persons, households or communities, many writers mistakenly use the term socio-economic status. If the available measurements are strictly measurements of wealth or income, e.g., household assets, then use terms that refer to this more narrow concept precisely, e.g., wealth, income, or poverty level. Socio-economic status and wealth are not synonyms. The concept of socio-economic status captures more than just wealth. It refers to income, education, and profession, and also includes the idea of social class. Restrict the use of the term socio-economic status only when the available data supports this broader conceptualization.

E12. Using a technical term in its non-technical sense

Several technical scientific terms also have a somewhat less specific meaning when used in general speech. To avoid confusing the reader avoid using technical terms in their non-technical sense.

E12a. Using the term 'random' in its non-technical sense

The term 'random' has a specific technical meaning within science. Random selection, for example, implies that the entire population is enumerated and that a process, such as a lottery or a random number generator, is used to select individuals from among the entire population. In a scientific manuscript the word 'random' should only be used within this specific context. In common speech the word 'random' is often used as a synonym for 'haphazard'. For example, "I was walking down the street and selected a restaurant for lunch at random." To a scientist, this was not random selection of a restaurant. Rather the choice of lunch location was based on convenience.

Example of the error:	Alternative, better option:
× In-depth interviews were conducted with 10 randomly selected key informants working there.	✓ We conducted in-depth interviews among 10 key informants we identified working in these communities.

E12b. Using the term 'reliable' in its non-technical sense

The term 'reliable' has a specific technical scientific meaning that is somewhat different than its meaning in more common speech. Within science 'reliability' refers to whether the repeated measurements of the same phenomenon are similar. A blood test is reliable if it provides the same result on repeated testing of the same sample. The synonym for 'reliability' in this technical sense is 'repeatability'. To avoid confusing your scientific reader, the words 'reliable' and 'reliability' should only be used in their strict technical sense in any scientific document.

Example of the error:	Alternative, better option:
× The self-reported data may not be reliable.	✓ The self-reported data may not be valid.
× The direct observations were conducted to cross check the responses and ensure reliability of the data collected in the self-administered survey.	✓ We cross checked the findings from the self-administered survey by comparing them with results from direct observation.

E12c. Using the term 'significant' in its non-technical sense

The term 'significance' has a specific technical meaning in quantitative scientific writing. Specifically, it refers to statistical associations that are less likely than would be expected by chance. Conventionally these are associations with a probability of occurring by chance of less than 5%. Many thoughtful commentators on scientific writing are critical of the narrow dichotomous thinking that divides all results into 'significant' or 'not significant' (see Error F1a). Despite these criticism of its overuse, when scientific readers see the term 'significant' in a scientific manuscript they will assume the author is referring to statistical significance. Therefore, do not use the term in a different context, because you risk confusing the reader. A confused reader is less likely to maintain interest in your article and so this lessens your contribution to global scientific communication.

Example of the error:	Alternative, better option:
× A significant number of respondents could not identify common signs of H5N1 in poultry (Table 2).	✓ Most respondents could not identify common signs of H5N1 in poultry (Table 2).
× Backyard poultry can be a significant source of high quality protein for rural low-income families.	✓ Backyard poultry can be an important source of high quality protein for rural low-income families.

E12d. Using the term 'valid' in its non-technical sense

The term 'valid' has two related technical meaning in quantitative scientific writing. When used to describe a measurement, it implies that the measurement reflects the underlying phenomenon of interest and is not an artifact of the instrument being used for measurement nor other cause of inaccuracy.

When used to describe a scientific inference the term valid implies that the inference is sound given the results and the way the data were collected. The term is used more loosely in general communication. To avoid confusing readers in scientific manuscripts, only use the term in its technical sense.

Example of the error:	Alternative, better option:
× Preventing nosocomial transmission of tuberculosis is especially valid in Bangladesh because of its high tuberculosis burden.	✓ Preventing nosocomial transmission of tuberculosis is important in Bangladesh because of its high tuberculosis burden.
× The similarity of results from the repeat assessment of the samples suggests that the assay is valid.	✓ The similarity of results from the repeat assessment of the samples suggests that the assay is reliable.

E12e. Using the term 'incidence' incorrectly

Epidemiologists define incidence as the number of new cases of illness that occur in a specified population in a specified time. For example, the incidence of hepatitis B in the population was 23 cases per 10,000 persons per year. The numerator for incidence is a count of new cases (or new events). The denominator is person-time, that is a measure

that captures both population size and time. Because time is in the denominator, incidence is always a rate. Thus, the second word of the phrase 'incidence rate' is redundant.

Prevalence, by contrast, is the number of cases in a population. It includes both new cases and old cases. For example, there may be 400 cases of hepatitis B in the same population of 10,000 people. Most of these cases are old cases. The prevalence of hepatitis B in the population is 4%.

Reporting incidence as an unqualified percentage is incorrect, because it does not communicate the time frame that the new cases occurred.

Example of the error:	**Alternative, better option:**
× We followed a cohort of live poultry market workers in Bangladesh to determine the seroprevalence and incidence rate of seroconversion of antibodies to H5N1 virus.	✓ We followed a cohort of live poultry market workers in Bangladesh to determine the seroprevalence and incidence of seroconversion of antibodies to H5N1 virus.
× The incidence of diabetes among Marin County residents, 5%, is the lowest in the state.	✓ The prevalence of diabetes among Marin County residents, 5%, is the lowest in the state.

E13. Using the verb 'documented'

The word 'document' is a noun. English often turns nouns into verbs, but not always with good results. To 'document' means to make a document, that is to write something down. So if I write down on a piece of paper the phrase, 'the earth is flat', then, strictly speaking, I have documented that the earth is flat. Creating a document is unrelated to the validity of an assertion. Therefore, we should not use this verb to communicate scientific validity of a statement.

Example of the error:	**Alternative, better option:**
× Studies in Bangladesh, India and Malaysia also documented neutralizing antibodies against Nipah virus in *Pteropus* bats.	✓ Studies in Bangladesh, India and Malaysia also identified neutralizing antibodies against Nipah virus in *Pteropus* bats.

E14. Framing an argument in terms of need

Quite often arguments in draft scientific papers are framed in terms of needs. The underlying message is that we 'need' to do something. Usually the authors are asking the reader, the government or society more generally to care about the issue in the same way that the authors care about the issue and follow the specific advice of the authors.

In a scientific manuscript it is reasonable to talk about a need for water, oxygen, and food for survival, but it is less appropriate to assert a need for health-care reform or a need for social change. The problem with this language is that it disguises the goals and

aspirations of the authors in terms of a need, when the issue of what constitutes a legitimate need is an open question for individuals, for society and for science.

Scientific writing is most persuasive when it demonstrates the connection between a set of conditions and consequences. Rather than framing arguments in terms of needs, the same ideas should be described as steps that are required to achieve a particular outcome. Importantly, the outcome should be specifically stated.

Examples of the error:	Alternative, better options:
× There is a need to standardize and expedite the assignment of causes of death, thereby enhancing a timely process of appropriate decision-making.	✓ If the assignment of causes of death could be standardized, appropriate decision-making based on these data could be expedited.
× A low-cost, accurate approach to characterize handwashing behaviour is needed.	✓ A low-cost, accurate approach to characterize handwashing behaviour would improve the assessment of handwashing promotion programs.

E15. Using the term 'illiterate' as a synonym for 'no formal education'

We frequently see studies that asked respondents about their years of formal education and then the findings state, 'The respondents were illiterate'. Although we often use the word 'illiterate' as a synonym for 'no formal education', these terms are not synonymous. Generally, literacy is evaluated by asking people if they can read or write, and is validated by specific literacy tests. People may have attended school for some years and still not be able to read or write. What we really are reporting is that because they completed so little formal education they probably cannot read. The term illiterate is also commonly used with a condescending tone, and so risks communicating a lack of respect for one's study subjects.

Examples of the error:	Alternative, better options:
× The age range of programme beneficiaries was 18–65 years old and over 25% who took part in activities were illiterate.	✓ The age range of programme beneficiaries was 18–65 years old and over 25% who took part in activities had less than 4 years of schooling.
× Educated mothers were 2.3 times more likely to wash hands at key times than illiterate mothers.	✓ Educated mothers were 2.3 times more likely to wash hands at key times than those with no schooling.

E16. Using the word 'challenging' as a synonym for 'difficult'

We often use the word difficult to describe public health problems or solutions. The word difficult means that the problem or solution is not simple or easy. However, when substituting the word challenging for difficult, the implication is that by engaging in this issue we are somehow tested, and that something about ourselves, our capacity to take on new issues and to grow to address these issues, is revealed. When a situation is

difficult, motivational coaches encourage us to see this difficulty as a personal challenge, so that we can strive to overcome it.

This implicit motivational jargon is out of place in scientific writing that values precise description. The substitution of challenging as a synonym for difficult is so overused, that it sounds insincere. It is the kind of language typical of hucksters selling products on late night infomercials. If the situation is difficult, then call it difficult. If you want to challenge a group, in an editorial or in the discussion section, then do so explicitly. (If you disagree vehemently with this advice, we recommend that you consider it a challenge to write without using the word challenging.)

Examples of the error:	**Alternative, better options:**
× We will explore challenges in implementation, as well as find out what factors motivate children to participate.	✓ We will explore difficulties in implementation, as well as find out what factors motivate children to participate.
× In these impoverished contexts, changing child feeding behavior is challenging.	✓ Poverty is a major barrier to improving child feeding behavior
× These modest findings highlight the challenges of maintaining high quality implementation of interventions at scale.	✓ These modest findings highlight the difficulties of maintaining high quality implementation of interventions at scale.

E17. Describing a laboratory test result as positive

Scientific communication is characterized by specificity and nuance. It avoids unqualified generalizations. Scientific thinking eschews narrow dichotomies, such as stating that an intervention was a success or failure. Instead, a scientific approach is more likely to identify aspects that achieved objectives, and aspects that did not.

Scientific writing should bring this framework to our description of laboratory results. No laboratory test is 100% sensitive and 100% specific. A laboratory test provides additional information that scientists can interpret. When describing laboratory results, use sufficient specificity so that readers can interpret the meaning without having to jump back to the methods section to review which laboratory tests were conducted and how they were interpreted.

Examples of the error:	**Alternative, better options:**
× Out of 23 samples tested for different respiratory viruses, 21 were positive for respiratory syncytial virus.	✓ Out of 23 samples tested for different respiratory viruses, 21 had detectable RNA for respiratory syncytial virus.
× From the surveillance database, we identified 209 influenza positive patients during May to October, 2010.	✓ From the surveillance database, we identified 209 laboratory confirmed influenza patients during May to October, 2010.
× Among the 123 people tested six were positive for Nipah.	✓ Among the 123 people tested, six had IgM antibodies against Nipah virus.

E18. Using increase or decrease in the absence of a time trend

The words increase or decrease imply a change in quantity over time. They should not be used when comparing two groups during the same time interval.

Example of the error:	Alternative, better option:
× Children under age 5 years had an increased risk of infection compared with school aged children.	✓ Children under age 5 years had a higher risk of infection compared with school aged children.
× Children in the nutrition intervention group had a decreased prevalence of anemia compared with controls.	✓ Children in the nutrition intervention group were less likely to have anemia than controls.

The words increase and decrease can be used appropriately when evolution over time has occurred. For example, the incidence of anemia decreased between 2003 and 2015.

F. Recording scientific data

F1. Using statistics in place of the study question to frame results

We become so enamored with the output of our statistical programs and our statistical understanding that sometimes our writing reads like the output of our statistical analysis program. You know you are making this mistake when words like ‘association’, ‘analysis’, or ‘relationship’ are the subject of a sentence.

The point of analysis of health data is not mathematical findings, but what these results mean in terms of the lives and health of people. The statistical analysis is a means to an end and the results should be expressed and communicated with other health professionals in terms of the research question.

Examples of the error:	Alternative, better options:
× Father's literacy was associated with child working as helper in specific skilled services (p=.007).	✓ Children whose fathers were educated were more likely to work in skilled jobs than children of uneducated fathers. (xx% vs., yy%, p =.007).
× In simple regression analysis, education and pregnancy status give highly significant relationship, while language and counselled by give significant relationship on screening.	✓ Women who were educated, who spoke Hindi, and who benefited from counselling from a physician, were more likely to consent to the screening test.
× The analysis of association among the independent variables showed that there is an association between the main exposure variable (Distgrp2) and the costgrp and between costgrp and the duration of disease (Durdgrp2).	✓ People who lived farther from health facilities spent more money per visit to the health care facility.

F1a. Framing narrative results around p-values

A p-value assesses the probability that results as extreme as observed in the analyzed groups could have arisen by chance enrollment of a non-representative study population. This is an aspect of the associations within their data that scientific authors should assess, but a p-value < 0.05 does not prove an association is causal. It does not provide insight on whether the association is due to bias. It does not assess whether the association is due to confounding.

A low p-value conflates whether an association between exposure and outcome have a large effect (which may have quite important impacts on the scientific or public health implications of the results) or whether there is a small, or even trivial effect in a large number of observations. (If the sample size is 1 million all of the P-values will be <0.001.)

As the authors of a formal assessment of the use of p-values in biomedical literature noted “P values do not provide a direct estimate of how likely a result is true or of how likely the null hypothesis is (“there is no effect”) is true. Moreover, they do not convey

whether a result is clinically or biologically significant. P values depend not only on the data but also on the statistical method used, the assumptions made, and the appropriateness of these assumptions."[14]

In short, P-values are silent on most important dimensions of assessing valid scientific inference. Presentations of results should not be framed around p-values. Indeed, framing results around p-values communicates to the reader that the author has a naïve approach to data interpretation. Instead frame results around effect sizes and presenting work in an order so that readers can consider issues of confounding, bias, and dose effect. Present p-values like a footnote, not as a central finding. Think of "statistical significance" as, only an issue of second order concern, i.e. if there is a difference that is potentially meaningful and interesting, it provides a test of whether this difference is likely due to chance selection of a non-representative study population.

Framing a scientific narrative around p-values also encourages a naïve dichotomous conceptuality, i.e. that a factor is either present or absent. Science is characterized less by these sort of absolute binary frames, and more about measuring degrees of difference.

The editors of the *International Journal of Epidemiology* explain their perspective on this issue. "We actively discourage use of the term 'statistically significant' or just 'significant' and statements in method sections such as 'findings at p<0.05 were considered significant'. Where used, we ask authors to provide effect estimates with confidence intervals and exact P values, and to refrain from the use of the term 'significant' in either the results or discussion section of their papers. Our justification of this position is given in: Sterne J, Davey-Smith G. "Sifting the evidence - What's wrong with significance tests?" *BMJ* 2001: 322:226-231.""

Examples of the error:	**Alternative, better options:**
× When we looked at the contamination of each toy ball separately, two toys did not reach statistical significance for fecal coliform contamination.	✓ When we compared fecal coliform contamination between groups for each toy ball separately, toys were consistently less contaminated in the cleaner households compared to the less clean households. However, the comparison between groups of fecal coliform contamination of toys 2 and 4 did not reach statistical significance (Table 2).
× Compared with persons who contracted Nipah infection from another person, Nipah cases who drank raw date palm sap were more likely to develop convulsion (log rank p value <0.001), altered mental status (log rank p value <0.001) and die (log rank p value <0.001).	✓ Compared with persons who contracted Nipah infection from another person, Nipah cases who drank raw date palm sap were three times more likely to develop convulsions, 50% more likely to develop altered mental status and 58% more likely to die (log rank p values all <0.001).

F2. Not presenting the core data

It is crucial that scientific readers be able to evaluate your data. They don't want to just read your conclusions, they want to look at the data and draw their own conclusions. This is the essence of science--reflective consideration of empiric observations. Your manuscript should present the data in a way that allows the reader to form an independent opinion as to whether the data were analyzed properly and interpreted prudently. As a matter of transparency, the reader should be able to re-do the key calculations. Thus, basic frequencies, rates or means comparing groups on your central findings, are crucial.

A common variant of this error occurs when comparison between groups is presented. In its most extreme form the measure of association are omitted entirely. Only a p-value is presented (See Error F1a).

Examples of the error:	**Alternative, better options:**
× Most subjects (62%) were not aware of ….	✓ Of 113 subjects, 70 (62%) were not aware of…. *[Always show numerators and denominators in the calculation of proportions]*.
× There was a significant difference in the proportion of case-patients and control-subjects who reported eating the potato salad (p=0.0001).	✓ Of the 42 case-patients, 30 (71%) reported eating the potato salad compared with 19 of the 120 control-subjects (16%, Odds ratio=13.3 p<0.01).
× Proportions only in the tables	✓ Always provide numerators and denominators.

F3. Using too many decimal places

When the results of a study are presented with an excessive number of decimals, communication between the writer and the reader is impaired. The extra digits distract the reader from the message and usually add no significant meaning. Another reason to avoid presenting too many decimal places is they imply a precision that the data generally lack.

This error is most commonly seen with percentages. Data are presented as percentages, e.g., 39%, rather than as frequencies e.g., 321/815, so that it is easier to remember and compare one group or scenario to another. Although ten thousand decimal places is a more precise report of the percentage, it is also burdensome to the reader. For example, if 13 of 17 enrolled study subjects have a particular characteristic, this can be reported as 76%, 76.5%, 76.47%, 76.461%, 76.46706....in fact, with a powerful enough calculating programme you could report thousands or millions of decimal places.

However, after reporting percentages to one or two decimal places, the numbers are no longer easy to remember and compare. Active readers who want to understand the meaning of your scientific writing will often compare reported numbers to each other. It is much easier for readers to compare numbers and to perform mental arithmetic on rounded numbers. Thus, wherever possible, note percentages without decimal places.

Only if the percentage is less than 10, and the figures beyond the decimal point have public health significance, then it might be reasonable to include them.

Similarly, when people report relative risk or confidence intervals they are often reported to two decimal places. For example, the statement that people who ate goat curry were three times more likely to become ill than persons who did not (Relative risk of 3.24, 95% Confidence Interval CI=0.74-12.99, p value=.143). Can your investigation reliably estimate the relative risk and the confidence interval to 2 decimal places? Almost certainly not! If the study cannot support such precision, then you should not imply that level of precision by reporting the extra decimal places.

One rule of thumb for confidence intervals for odds ratio is that they should not have more than two meaningful figures. Whether or not these figures are decimals or not depends upon where the odds ratio fit on a log scale. Remember that the odds ratios for 'protective exposures' and 'risk factors' are symmetrical around the number one on a log scale. Thus, reporting an odds ratio of 243 represents the same amount of precision as an odds ratio of 24.3, an odds ratio of 2.43 and an odds ratio of 0.243. Thus, try to round up (add or subtract digits) so that you always display two meaningful figures, e.g., 24, 2.4, or .24.

Examples of the error:	**Alternative, better options:**
× The prevalence of active trachoma was 21.01% (95% confidence interval: 6.23-36.77%).	✓ The prevalence of active trachoma was 21% (95% confidence interval: 6.2-37%).
× People who ate goat curry were three times more likely to become ill than persons who did not (Relative risk of 3.24, 95% confidence interval CI=0.74-12.99 p value=0.143).	✓ People who ate goat curry were three times more likely to become ill than persons who did not (Relative risk of 3.2, 95% confidence interval CI=0.74-13, p value=0.15).

F4. Using too few decimal places

In the enthusiasm to avoid using too many decimal places, occasionally authors present too few. In most contexts you want to communicate two digits of numerical information. (25% is two digits. $1.2 million is two digits). As noted above in reporting a percentage greater than 10, adding a third digit, a decimal place, is generally distracting and uninformative. However, if you are reporting an odds ratio or other relevant small number then it is important to communicate two digits of information (2.1 or 0.63), even if one or more of these digits are decimal places. Count digits, not decimal places!

Examples of the error:	**Alternative, better options:**
× Children whose mother completed primary education were less likely to be hospitalized for diarrhea (odds ratio 0.6, 95% confidence interval 0.4, 0.8)	✓ Children whose mother completed primary education were less likely to be hospitalized for diarrhea (odds ratio 0.57, 95% confidence interval 0.42, 0.77)
× Ambulatory case-patients spent a median of US$2 (IQR=$1–4) in the public hospitals.	✓ Ambulatory case-patients spent a median of US$1.8 (IQR=$1.1–3.6) in the public hospitals.

F5. Using incomplete headings for tables and figures

In a biomedical manuscript the figures and tables should stand alone. A reader should be able to look at the table or figure, read the title, and understand it. Readers should not have to refer to the narrative methods or results to understand the table or the figures. Thus a typical heading will need to include person, place, and time characteristics. The number of study subjects and statistical methodology should be communicated. Use footnotes to explain apparent discrepancies or other issues in the table/figure. However, for oral presentations, brief titles for tables and figures are fine.

Examples of the error:	**Alternative, better options:**
× Figure 1: Epicurve of the measles outbreak.	✓ Figure 1: Cases of measles by date of onset, Chennai city, Tamil Nadu, November 2004.
× Table 2: Risk factors associated with illness, univariate analysis.	✓ Table 2: Characteristics of meningitis case-patients and control subjects, Kano city, Nigeria, March 1996.

F6. Imbalance between table and narrative presentation of the results

a) Too little narrative

Just as tables, figures and graphs should stand on their own and not require accompanying text, the narrative section of the results should stand alone. A reader should be able to read only the narrative text, not look at any of the figures or tables, and come away with a clear understanding of the important findings from the analysis. This error most commonly takes the form of several well-constructed tables being presented in the results section with only a sentence or two in the narrative results section pointing to each table. The results section should not repeat all the data that is in a table, but rather should focus the reader on the highlights. Look at several quality journal articles related to your research question and note the balance between what is presented in the narrative text and what is presented in the tables. Strive for a similar balance.

b) Too much narrative

The other side of this error is when the narrative goes on and on, often through several paragraphs citing innumerable, often minor, comparisons within the table that do not address the core issue of the manuscript. One of the responsibilities of the analyst is to reduce data so it is more easily understandable to the reader. *The Lancet* does not permit authors to mention any numbers in the narrative that are already presented in the table. The idea is that the narrative is used to highlight the core ideas or patterns that can be seen from the data presented in the table. Most scientific writing need not invoke *The Lancet's* standard of no repetition of data, but the role of the narrative in the results section of scientific writing should be more summary and perspective, and less repetition of data that is more easily seen and compared in a well-constructed table.

Example of the error:	Alternative, better option:
× Of all the food items, only the vanilla ice cream was associated with illness (Table X).	✓ The risk of illness was estimated according to consumption of each of the eight menu items that were served at the lunch (Table X). Eating vanilla ice cream was the only exposure that was significantly associated with illness (relative risk: 8.6, p=0.001) and that accounted for the majority of cases (population attributable fraction: 86%).

F7. Pointing too explicitly to tables and figures

In your results section if the words 'Table 1' or 'Figure 2' are the subject of a sentence, you have likely committed this error. The whole paper should be organized around the central ideas you want to communicate and that you want the reader to focus on. Thus, lead with your findings, and compose your language around those findings and related ideas, rather than around structures, i.e., pages, tables, or figures.

Examples of the error:	Alternative, better options:
× Table 1 describes the forms in which areca nut was used.	✓ Sweetened varieties of areca nut were the most popular (Table 1).
× Figure 2 presents the age, sex and geographic distribution of our sample across the four study districts.	✓ The age, and sex of distribution of the samples was similar across the four study districts (Figure 2).

F8. Using inappropriate figures

Edward Tufte in his excellent book, "*The Visual Display of Quantitative Information*" argues that figures for scientific manuscripts should be evaluated using a data to ink ratio, e.g., the amount of data that can be presented with the least amount of ink. Excessive ink in figures mean they include unnecessary axes, gridlines, borders, 3-D effects and other elements that do not add anything, and make the figures less understandable.

Space is always at a premium for journal editors, who weigh this issue more from the perspective of data to space ratio. Both pie charts, and simple frequencies presented as bar charts, are inefficient. It is reasonable to assume that the reader of a scientific manuscript understands the difference between 20% and 40% and so does not need it illustrated by comparing relative widths of a pie or relative heights of a bar. A simple table can efficiently present proportions.

Thus, use figures to achieve key communication objectives. Figures are best used in two situations:

1.) When they permit presenting a large amount of data in a format that is revealing of the underlying characteristics of the distribution. For example, scatter-plots that show trends.

2) When they communicate in a more effective and efficient visual format than could be done with a narrative description or a table, e.g., a figure that presents multiple components of a phenomenon, such as different age trends by sex.

F9. Using the wrong symbol to designate degree

Wrong example: 4 ^{0}C or 4 oC

To make the degree symbol use the insert symbol feature of Word, select a circle (i.e., not the letter 'o' or the number zero) and then make the circle superscript.

Correct example: 4°C.
Recent versions of MS Word, include a degree symbol. Go to Insert, then Symbol to find the figure.

F10. Using non-standard footnote symbols in tables

Footnotes contribute important explanations to data presented in tables. They are useful to clarify analytic approach, groups being compared, statistical significance and other explanatory information. The International Committee of Medical Journal Editors (www.icmje.org) specifies the symbols and their sequence for footnotes.

*, †, ‡, §, ||, ¶, **, ††, ‡‡, §§, ||||, ¶¶, etc.

Do not use other symbols or other sequences unless the journal recommends them (e.g., PLoS Med uses a, b, c, d, . . .).

You can find these symbols using the insert symbol feature of Microsoft Word. Note that these symbols should be in superscript.

F11. Comparing to a varying baseline

We often analyze data where observations are grouped into multiple levels of exposure. In the example below we have categorized observed handwashing behaviour into mutually exclusive categories:

Handwashing after defecation	Group A		Group B		Odds ratio	
	Number	%	Number	%	varying baseline	reference group
No handwashing	75	12%	150	19%	0.6	--
Washed one hand with water alone	150	23%	150	19%	1.3	2.0
Washed both hands with water alone	125	19%	150	19%	1.0	1.7
Washed one hand with soap	150	23%	100	13%	2.1	3.0
Washed both hands with soap	150	23%	200	25%	0.9	1.5
Total	650		750			

The common error is to compare the prevalence of each level of the variable in group A to the prevalence of the same level of the variable in group B. Thus if we compare the prevalence of washing both hands with water alone, the prevalence is the same (19%) in group A and group B, so we could say that people in group A and B are equally likely to wash both hands with water alone, which is equivalent to an odds ratio of 1.0. The problem with this comparison is that the people who are not washing both hands with

water alone are quite a heterogeneous group. Some of them are practicing less intense handwashing (not washing their hands at all or only washing one hand) and others are practicing more intense handwashing. Indeed, even if we have an elevated odds ratio with such a comparison it is difficult to interpret, because we don't know if this elevation results from a difference in more intense or less intense handwashing. The standard approach to resolve this dilemma is to arrange the exposure level into a mutually exclusive hierarchy. Set the lowest level of exposure as the baseline and then consider the 2 x 2 table comparing each level of exposure to the baseline. Using this approach illustrated in the final column, we can conclude that compared with Group B, Group A is more likely to wash one hand with water rather than not washing at all.

F12. Generic data tables that lack a clear message

There is no single standard format to present data in tables. Tables are an integral element of the broad scientific argument that you compose through your manuscript. Tables should be organized based on the communication objective of the article. Thus, the first step in drafting a table is to identify the communication objective for the table. Examples might be to describe the baseline characteristics of the population, to compare the outcome of a group who received an intervention with the outcome in a non-intervention group, or to compare the characteristics and exposures of persons who became ill with persons who remained well.

Having identified the communication objective of the table, you then construct the table so that the message comes through clearly. The patterns in the data which you are striving to illustrate should be obvious at a glance, or at least should be obvious once they have been pointed out by the narrative description in the results section of the manuscript.[15] Just like narrative scientific writing, expect that you will have to develop and revise tables through several drafts.

F13. Table layout that impairs comparisons*

An advantage of presenting data in tables, rather than in a narrative paragraph, is that by clearly aligning numbers different groups and different characteristics can be readily compared. Numbers are easier to compare reading down columns than across rows especially for larger numbers of items. Such comparisons are often the central communication objective of a table. To facilitate comparison avoid:

- Columns that are too wide. This makes it difficult to compare data between columns. One common form of this error is to set the width of the table column based on the length of the column heading, rather than on optimizing column width to permit comparison of data.
- Ordering data haphazardly. Rather than presenting characteristics in the table in alphabetical order, or in the order they were asked in the questionnaire, consider the easiest way for the reader to understand the information. Ordering characteristics from smallest to largest or largest to smallest is an intuitive approach that helps the reader to quickly and easily understand.
- Poorly aligned data that impedes comparison. Align data and decimals so that a vertical list is readily comparable.

Hard to compare	Easier	Still Easier
23 42 34 109 87 42 27 98 114 75	23	23
	42	27
	34	34
	109	42
	87	42
	42	75
	27	87
	98	98
	114	109
	75	114

*These examples and much of the text was contributed by Robert Fontaine with help from Ehrenberg ASC, J. R. Statist Soc. A, (1977), 140, Part 3, pp. 277-297.)

Use the table layout effectively to help the viewer -- place numbers for comparison close together

Year	Both Sexes	Male	Female
1973	600	500	99
1970	670	580	87
1968	550	460	89
1966	330	260	71

Draw columns and rows close together

Year	Both Sexes	Male	Female
1973	600	500	99
1970	670	580	87
1968	550	460	89
1966	330	260	71

Move and minimize intervening numbers

	Rate per 1000 (SE)					
Year	Male		Female		All	
1993	83	(2.3)	78	(2.2)	80	(1.9)
1994	62	(2.5)	66	(2.7)	63	(1.8)
1995	58	(2.1)	54	(2.0)	56	(1.7)
1996	55	(2.0)	45	(2.0)	51	(1.7)

Remove intervening numbers entirely if consequence minimal

	Rate per 1000[a]		
Year	M	F	All
1993	83	78	80
1994	62	66	63
1995	58	54	56
1996	55	45	51

a. Standard errors for all rates less than 5% of rate.

Organize data by magnitude

Exposure	1000 Cases	Rate	Rate Ratio	p
A	11	2.9	1.3	> 0.10
B	6	9.9	4.3	< 0.001
C	34	5.4	2.3	> 0.1
None	27	2.3	1.0	Ref*

a = p-value
b = reference exposure category

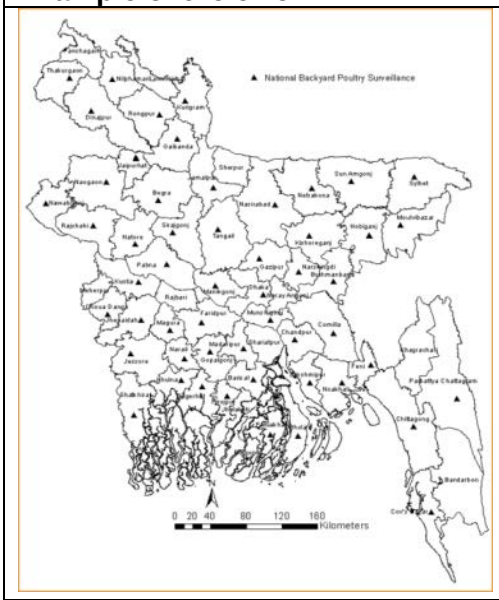
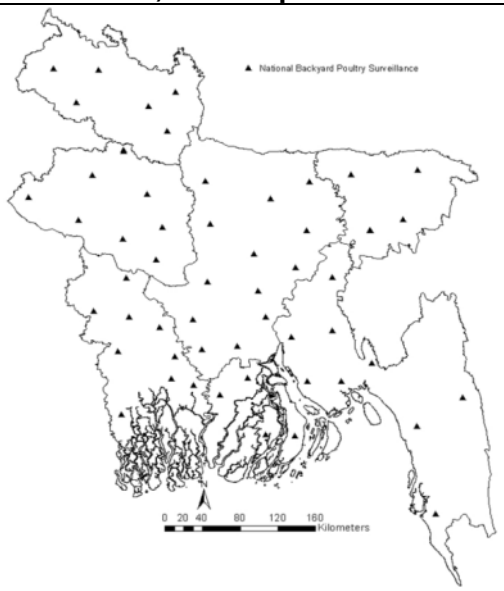
Organize data by magnitude

Exposure	1000 Cases	Rate	Rate Ratio	p[a]
B	6	9.9	4.3	< 0.01
C	34	5.4	2.3	< 0.05
A	11	2.9	1.3	> 0.001
None	27	2.3	1.0	Ref[b]

a. = p-value
b. = reference exposure category

F14. Maps with irrelevant details

When a map is included in the manuscript, its role is to communicate specific geographical information, for example the location of the study, spatial relationships among cases, or the spatial distribution of exposures. Inserting a map constructed by someone else that is filled with details that are irrelevant to the communication role for the map, e.g. district divisions, rivers, or railroad lines, distracts readers from the message. Draw your own map or begin with a generic map and add the elements that are essential to the message.

Example of the error:	Alternative, better option:
	

F15. Numbering figures or tables out of sequence

Readers expect and journals require tables and figures to be numbered in the order that they are referred to in the narrative text of the paper (i.e. Table 1, Table 2, Table 3, Figure 1, Figure 2, Figure 3). In addition, each table and figure needs to be cited in the narrative text (otherwise readers and editors will assume it is not important and can be dropped).

The most common form of this error is when authors mention an element of complicated data analysis in the methods section and refer to a later table or figure in the manuscript. Usually, the best approach in this situation is to describe the statistical method without pointing to the results table or figure. The problem with citing the advanced table or figure as Table 1 or Figure 1 is that it will confuse readers to have this more complicated analysis presented before the more basic results that build toward the more complicated analysis.

The other common form of this error is renumbering the tables or figures, but not updating these numbers in the narrative text.

G. Approaching publication

G1. Failure to respond to reviewers' comments

Among the most consequential errors that researchers commit is ignoring advice given in the previous reviewed draft. As the first author, you have the right to decide what goes in it and what does not. Nevertheless, you should respond to every issue raised by a reviewer or co-author. It is acceptable to reject the advice offered by a reviewer. Indeed, it is important to reject inappropriate or unsound advice. In a scientific environment, reviewers fully expect that some of their advice will be rejected. However, if you choose to reject the advice of a reviewer or a co-author, you need to defend that decision when you submit the next draft.

To address every point raised by a reviewer, either change the manuscript accordingly, or explain in a separate note the issues you chose not to change and defend why you chose not to change them. If you simply ignore the advice you will just get the same comments from the reviewer again. The paper does not develop further, and both reviewer and author feel like they are wasting time. Often the situation is a problem with written communication. The reviewer doesn't understand something that the author has done. This can be an important clue that you need to add something to your writing to make it understandable. At other times some language needs to be changed to clarify the point. The key is to respond to every issue raised by a reviewer. Be prepared to write and rewrite before and after submission to a journal.

a) Responding to internal primary reviewers and co-authors: How-to tips

Remember it normally takes 10 working days to get all the reviewers comments. Indeed, it is a good practice when circulating a draft manuscript to request input by a specific date. 10 working days is a reasonable timeline. If you provide less time than this, you risk communicating a lack of respect for the time of your co-authors. Similarly, when you are a co-author it is a responsibility to provide input within a reasonable time-frame.

Read all reviewers' comments carefully before starting to revise to get an overall picture of how others interpreted your paper. Oftentimes it is useful to read the comments all through once to get a general idea of the criticism (and feel the pain that not every reader loved every decision you made). Then, after a day or two go back through each of the comments carefully. Often, taken together there are a number of major changes you will want to make to your manuscript. We recommend implementing those and then returning, to the line by line critique.

Sometimes reviewers 'double-up' on a manuscript and add multiple comments, or sometimes comments are all on individual copies. How can you manage this? Make hard copies of all comments and after reading them thoroughly start from the beginning making changes on a newly named file (Abbreviated Title, Draft 2 Oct 12). Or use multiple monitors, one with newly named file and the other with all copies opened, ready to pull up and incorporate into the new draft.

Remember, not all comments may be useful or even correct. You, as first author, need to make the decision about what comments to accept and what to revise. If there is a major comment that you do not agree with, you should explain why by either inserting a

comment (using track changes), or by stating the reason in the accompanying email or in attached document.

b) Responding to editors' and external reviewers' comments: How-to tips

After you submit your manuscript to a journal, the editor will make a decision on whether the article is of interest to the journal or not. Many articles are rejected by the editor after his/her personal review or other in-house review. If after internal review, the editor is interested in the manuscript, then he/she will send it out for an expert peer review. Each review will be a critique that includes an overall evaluation and a list of specific items that need improving. Based on the reviews, the editor's letter will put your paper into one of three categories:

- The manuscript is accepted, pending specified changes.
- The manuscript requires revision and then the editor will review it again.
- The manuscript is rejected.

First, take time to read the all the reviews carefully and completely. Understand, in a holistic way, what the reviewers consider the most important weaknesses of the paper. Then begin revising. You will need to resubmit:

- A cover letter that summarizes the changes you made in your manuscript.
- A separate response to each itemized comment.
- Two versions of the manuscript: a marked up version that reflects all the changes you've made, and a clean version.

In the cover letter addressed to the editor, you will briefly describe the changes you have made, both those that were prompted by the reviewers and others that you have added during your review.

Make a copy of the itemized comments, and then draft a document that details the response to each of the comments raised by the reviewer. If a comment is acceptable and responding strengthens the paper, make the change in the manuscript using track changes, and then describe this change under the comment. It is often helpful to include a direct quote of added or revised text. If you do not believe that a change suggested by a reviewer, will improve the manuscript, be polite and professional in tone (even if the reviewer is not), while defending your rationale thoroughly.

G2. Incomplete response to reviews

The task of responding to comments is not to provide a minimalist justification why you wrote what you wrote. Instead the task is to demonstrate to coauthors, editors and reviewers that you fully understand the critique and the implication of the critique for your paper. If the reviewer raises a meaningful issue, you need to respond to that critique and revise the manuscript so that other readers do not face similar questions and confusion. Indeed, this is a great benefit of having your work undergo peer review. We should not lament that "the reviewer did not understand our work." or that the reviewer did not see that the current text already addressed their question. If the reviewer did not understand, we should take this as a signal that our message was not written clearly enough to be readily understood, and consider what changes we can make to the paper so that future readers will not suffer the same misunderstanding.

Make clear in the document you draft responding to the reviewers' comments what changes you made in the manuscript as a response to the comment. If you only respond to the reviewers' criticism, but don't change the manuscript many future readers will likely have the same unanswered question or criticism. If you change the manuscript, but don't make it clear in the cover letter that you made these changes, then the editor has to go point by point and try to figure out what you changed and what you didn't change. This is a painstaking, annoying and frustrating task. If you want your manuscript to be accepted, avoid annoying and frustrating the editor. Demonstrate to the editor that you have thoroughly considered and responded to each of these issues. Make it easy for the editor to accept your work.

It is completely acceptable, indeed expected, to disagree at times with some points made by a reviewer or co-author, but such disagreement must be framed within the context of a full understanding of their critique. The editor will review this response carefully, and may ask the reviewer(s) to look again at the manuscript and your responses.

G3. Invalid authorship line

Inclusion on an author line is an important indicator of one's contribution to scientific work, and an important professional credential. However, the authorship line can sometimes be controversial, so it is important to understand who should be included and who should not. All writers should read the 'Recommendations for the Conduct, Reporting, Editing and Publication of Scholarly Work in Medical Journals', a document developed in 2013 by the International Committee of Medical Journal Editors (ICMJE) available at www.icmje.org. Essentially, authorship credit should be based on four criteria, with authors meeting each criteria:

- Substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data
- Drafting the article or revising it critically for important intellectual content
- Final approval of the version to be published
- Agreement to be accountable for all aspects of the work and ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

If you follow these guidelines, your choices can be defended in any academic setting. Clarify in your own mind who clearly fulfils the criteria for authorship. Have a separate discussion with your supervisor if you believe that any other person needs to be included, for example, a government colleague who is critical to the government acting on the manuscript recommendations or an institutional collaborator who is essential to support ongoing scientific collaboration. Know your institutional or program criteria.

Generally, the first author is usually the one who participated significantly in the research by:

- Being involved in the conception and design of the research and collecting the data
- Interpreting the results
- Writing the first drafts of the paper
- Responding to co-authors and supervisors comments

- Submitting the manuscript to the journal and responding to the editors and peer reviewers suggestions

In the best-case scenario all co-authors should discuss and agree on the responsibilities and contributions early on, preferably during the development of the protocol when the roles of the investigators are specified. Practically, however, which specific analyses will ultimately support a manuscript, and so how many manuscripts will be appropriate and how each should be framed are usually impossible to anticipate before the data are analyzed. In addition, the composition of the scientific team and interest and availability of potential authors is often different by the time the data are available compared with the original plan, and so authorship typically needs to be revisited.

A tool to help you decide who should be listed as an author on a paper, and the ordinal ranking of authors listed on a paper, is the authorship ranking scorecard. We recommend that you use this authorship scorecard to share your ideas of authorship with your primary reviewer when you develop your framing document. (See Appendix 8)

G4. Missing acknowledgement section

Many research organizations and academic institutions have a specific policy, template and language for acknowledging the financial or material help from the agency or government who funded your research. Check your institution's policy. Confirm the donor's grant number by reviewing the contract. Many funding organizations, including US government organizations have preferred language for acknowledging their support. US government donors also often require some statement that the conclusions of the article are the authors' own and should not be construed as official government policy. Clarify from your US government contacts the specific language which they prefer.

People who contributed to the study, but do not fulfil the criteria for co-authorship, should be listed in the acknowledgment section. These may include:

- Community members of the study site
- Data collectors
- Laboratory support
- Administrative support
- Statistical assistance
- Writing assistance

Look at examples of the acknowledgement section from the journal you are planning to submit to. Usually the wording is straightforward. Don't be too informal in your language. Journals commonly require that anyone listed by name in the acknowledgment section must agree to have their name listed. If you want to acknowledge a person by name, send an email requesting permission to list his/her contribution in the acknowledgments section. If he/she responds affirmatively simply save the email in case a question is raised by the journal editor.

See the 'Uniform Requirements for Manuscripts submitted to Biomedical Journals' (www.icmje.org) for additional guidance.

G5. Choosing an inappropriate journal

It is rarely obvious which journal is best for your article. Many early career scientists request their senior author to recommend the target journal. This approach undercuts the opportunity to learn how to choose a journal. Instead, early career scientists should consider candidate journals and then propose and defend a prioritized list of journals to their senior author. By considering feedback from their senior author and ultimately their own experience in attempting to publish in various journals, authors can develop and hone their judgment regarding optimal journal choice.

Choosing a journal depends on who the audience is in relation to your research question. Before you start writing, start exploring some journals by reviewing previous issues. Have they published similar studies? Look at the references from an up-to-date manuscript you have found during your literature search. Do you see any pattern in terms of where this type of paper is being published? When you have identified several journals that have published similar topics, read and critique an article.

Another consideration is the journal's impact factor. The impact factor is a measure of the frequency that the 'average article' published in a given scholarly journal has been cited in a particular year or period. This metric reflects the importance of communication in scientific work. As science is a social activity, articles that are noted and cited by other researchers are influencing the field. This factor is often used to measure or describe the importance of a particular journal to its field. The Institute for Scientific Information (ISI) ranks, evaluates, and compares journals within subject categories and annually publishes the results in Journal Citation Reports.

The formula to determine impact factor 2016 for a journal would be calculated as follows:

A = the number of times articles published in 2014-15 were cited in indexed journals during 2016
B = the number of articles, reviews, proceedings or notes published in 2014-15

Impact factor 2016 = A/B

Impact factors can have a controversial influence on the way published scientific research is perceived and evaluated and the following criticisms have been made of the system:

- Journal impact factors depend on the research field: high impact factors are more likely in journals covering large areas of basic research and less likely in more subject-specific journals.
- Although Journal Citation Reports includes some non-English journals, the index is heavily skewed toward English-language journals, leaving out important international sources.
- Researchers may be more likely to pursue fashionable topics that have a higher likelihood of being published in a high-impact journal than to follow important avenues that may not be the as popular.
- Review articles are often highly cited, but they are a quite different contribution than highly cited original work.

Because there are so many journals today, and because most scholars look for articles using electronic search engines, the impact factor of the journal may be less important

now than it was a generation ago. Many very highly cited articles are published in journals that do not have a particularly high average impact factor. You want to select a journal whose editors will be interested in your work and who are able to identify good peer reviewers. Often a specialty journal with a somewhat lower impact factor is the best place to reach readers interested in your topic and where journal editors can find high-quality reviewers.

Good reviewers identify important issues for further development in your manuscript. Good reviewers improve your manuscript. Better manuscripts have more influence. If you have results that you and your supervisor believe represent broad international interest, it is reasonable to submit it to a more competitive high impact journal. Recognize however that these high impact journals, for example the *Lancet*, *Science* or *Nature,* reject 97% or more of all submitted manuscripts. Each manuscript submission takes time, time that could be deployed in writing your next manuscript. Therefore, spending time to reach for a high impact journal for a special manuscript may be a good idea, but it is generally prudent to submit to journals where the type of work that you are submitting is common published. For help with finding appropriate journals, explore the website JANE. (See Appendix 9)

G6. Not following a specific journal's details of style

All journals periodically publish their style rules in a hard copy edition, or these style rules are always available on the journals' website under 'Instructions for Authors' or 'Requirements for Manuscripts'. Go online and read the individual journal's instructions and follow them exactly before you submit your manuscript.

G7. Not using a checklist to review your paper before submission

After your manuscript is published it will be read, critically appraised, and hopefully will contribute to systematic reviews, inform specific public health guidelines, and influence overall public health practice. Before you submit your paper to a journal, you should consider if you have provided enough details so that the work can be used for these additional purposes. Some peer-reviewed journals require authors to follow a pertinent guideline. A comprehensive list of the available reporting guidelines appropriate to different study types, including systematic reviews, meta-analyses, and economic evaluations, is available at the EQUATOR Network library for health research reporting at www.equator-network.org/resource-centre/library-of-health-research-reporting/

Two checklists have been included that might help to prevent inadequate reporting of both observational studies and randomized controlled trials. The STROBE and CONSORT statements both provide an evidence-based, minimum set of recommendations for reporting these types of research studies. Use these checklists to review your paper to make sure all information is included, and also to critically review other scientific research papers. (See Appendix 10 & Appendix 11)

G8. Exceeding the journal word limit

Exceeding your target journal's word limit for manuscript length, especially for an initial submission, increases the risk that the editor will reject the paper without sending it for external review. The most common form of this error is an author circulating a draft

manuscript that is over the journal word limit and then asking co-authors to edit the draft for them.

It is an art to write succinctly, an art that is worth cultivating because readers' attention is a scarce resource, and holding readers' attention with your scientific writing is essential for your ideas to influence global scientific discourse.

An initial draft circulated to co-authors may be a little long, but do not circulate a late stage draft of the manuscript where either the abstract or the body of the manuscript exceeds the specifications of the target journal.

When your manuscript is less than 15%- 20% over limit and you've had one or more rounds of input from co-authors, dedicate several hours to reviewing every single sentence and asking yourself, "How can I communicate these ideas clearly with fewer words?" Smile every time you reduce a couple of words, and cheer when you realize you can drop a whole sentence by reorganizing your arguments and dropping some repetition. If you specifically focus on succinct language, you can often markedly reduce word count without eliminating ideas. Focusing on writing succinctly increases the clarity of your scientific reasoning. This laborious task is a first author responsibility.

A version of this error is circulating a draft manuscript with an abstract that is longer than permitted by your target journal. This invites your reviewers to waste their precious time. Such invitations discourage people from dedicating their time to review your work. Instead, exert the effort so that your abstract is a genuine draft abstract formatted for your target journal so your co-authors' suggestions can be focused, efficient and useful.

G9. Asking your senior author to recommend reviewers

Many journals request that authors recommend reviewers at the time of manuscript submission. This assists editors, because authors are in a good position to identify people who are expert in the area of their submitted work. If an early career author asks a senior author for a list of potential reviewers, then he/she undermines the opportunity to learn how to select reviewers.

A good reviewer is someone who would be interested in your work and has published work that is closely enough related that he or she would have an informed opinion. A good place to begin is considering the authors of the references cited within your manuscript. Also conduct some brief literature searches and review abstracts to identify other potential candidates. When considering subject matter to search, consider not only the central subject of your manuscript, but also related subjects or authors who have reported work using a similar method.

More senior scientists will have more requests for reviews, and so will likely decline to review a larger proportion of review requests. Scientists who have very recently published in a related subject area may be particularly interested in providing a review.

Draw up a list of reviewers, provide a reason for selecting each reviewer and then ask for input from your senior author. This way, you will both generate a reasonable list of reviewers, and have gained experience to help you select reviewers for future articles.

G10. Responding to journal reviewers using the first person singular

In group authored papers, the manuscript is the product of the work of the group. All authors agree to publically defend what is written. Similarly, the response to reviewers is not only what the author who drafted the response is saying, it is a statement from all authors. Once you have responded to external reviews, you should provide all co-authors a 1 week opportunity to review those comments and make any suggestions. (Early career author's should first have their senior author review the response to reviewers before circulating to all authors.) Because the responses to reviewers reflect the combined responses of all authors, the first person singular “I” should not be used in the response document.

Examples of the error:	**Alternative, better options:**
× I have revised the related text to provide the details of the selection process of the informants.	✓ We have revised the related text to provide the details of the selection process of the informants.
× I have tried my best to address all of your major and minor comments.	✓ We have tried to address all of the comments.

G11. Retaining comments in subsequent drafts

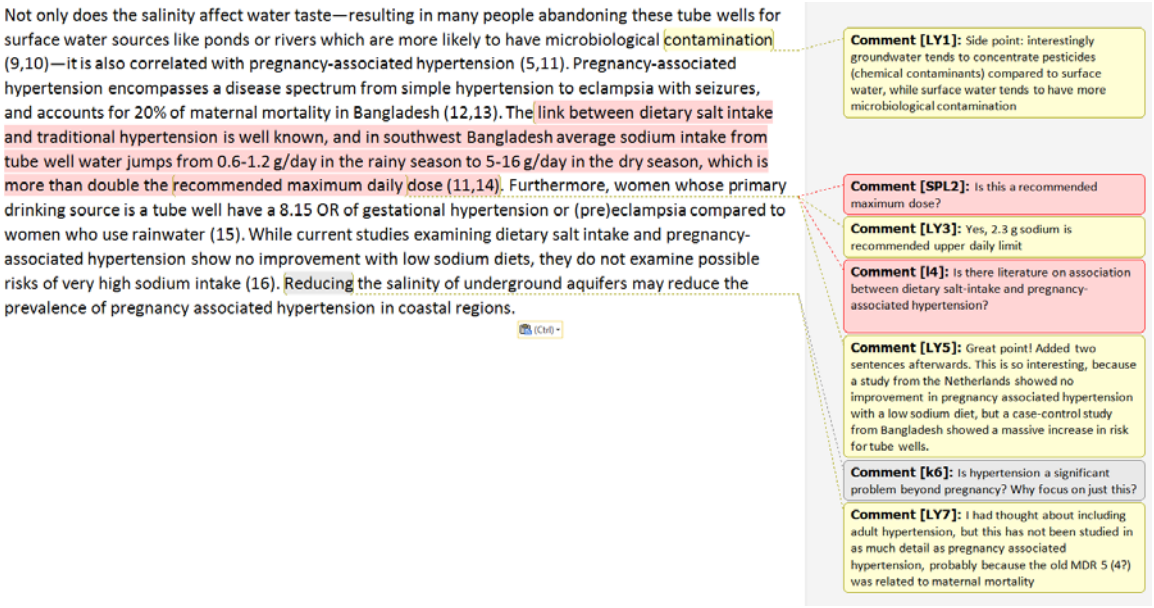
Many co-authors make comments on draft manuscripts using the comment feature of word processors. These can provide very useful input to the author. Often the authors are tempted to respond to comments by continuing the conversation within the comment bubble. The result is the next draft embodies 2 conversations. First is the narrative text. Second, is a side conversation among co-authors. Complex drafts that include a lot of historical commentary from multiple reviewers are burdensome and distracting to review.

The goal in drafting a scientific manuscript is a narrative text that is clear and stands on its own. Readers of the published manuscript will not have access to all of the side commentary. The task of a scientific author is to write clearly and strive to address the primary concerns of most readers. Responding in comment form risks Error G2.

Retaining a couple of comments that are addressing central issues where there is some appropriate conversation can be helpful, but these should be minimized so that the focus remains on creating a clear text that stands on its own.

If there are a number of comments from co-authors that would benefit from explaining why you did not take certain suggestions (i.e. avoiding Error G1) this is often better communicated by a separate response document. Each co-author can see that their issues were considered, but you keep your main document self-explanatory. Alternatively, you can circulate a clean and marked version. The marked version can have detailed responses to comments and show track changes, but the clean document is the working document that presents the draft close to how a new reader would see it.

Example of the Error:



Alternative, Better option:

Not only does the salinity affect water taste—resulting in many people abandoning these tube wells for surface water sources like ponds or rivers which are more likely to have microbiological contamination (9,10)—it is also correlated with pregnancy-associated hypertension (5,11). Pregnancy-associated hypertension encompasses a disease spectrum from simple hypertension to eclampsia with seizures, and accounts for 20% of maternal mortality in Bangladesh (12,13). The link between dietary salt intake and traditional hypertension is well known, and in southwest Bangladesh average sodium intake from tube well water jumps from 0.6-1.2 g/day in the rainy season to 5-16 g/day in the dry season, which is more than double the recommended maximum daily dose (11,14). Furthermore, women whose primary drinking source is a tube well have a 8.15 increased odds ratio of gestational hypertension or (pre)eclampsia compared to pregnant women who use rainwater (15). While current studies examining dietary salt intake and pregnancy-associated hypertension show no improvement with low sodium diets, they do not examine the risks of very high sodium intake (16). Reducing the salinity of underground aquifers may reduce the prevalence of pregnancy associated hypertension in coastal regions.

G12. Not finding a description of the error code

Writers sometimes receive an error code from a manuscript reviewer, but cannot find the error code in their copy of “A Guide to Quantitative Writing in the Health Sciences.” Although it is possible that the reviewer mistyped the error code, more commonly the reviewer is using a more up to date version of the guide, than is the author. Before complaining that you cannot find the error code in your version of the guide, ensure that you are working from the latest version of the guide available at: http:/globalhealth.stanford.edu/resources.html

G13. Requesting an unprofessionally short turnaround time

Asking others to be a co-author is requesting that they assume a substantial responsibility. By affiliating their name with the article, co-authors are accepting accountability for the work. They are publically connecting their reputation to the quality and the veracity of the scientific work, its analysis and its interpretation. Assuming this responsibility requires careful review of draft manuscript and ensuring that important issues are resolved prior to submission.

Co-authors are busy. Knowledge workers characteristically have more demands on their time, than they have time in a day. This means when you request that they give time to your article, you should be sensitive to how much of a request this is, and so provide a reasonable time for co-authors to respond (See Section 1.2.3 Timely reviews).

In the absence of exceptional extenuating circumstances, asking for a review within a few days communicates a lack of professionalism, and a lack of respect to co-authors. It is not a recipe for productive long term collaboration.

Examples of the error:	Alternative, better options:
× Attached is the final version of our paper. Please send me your consent to be a co-author by tomorrow so we can proceed with journal submission.	✓ Attached is the most recent version of our paper. I have attempted to address all of the concerns raised by co-authors. I am anxious to proceed with submission. Please look over the draft, and if you concur please send a statement that you agree to be listed as a co-author and agree with its submission for publication. Of course, any additional suggestions to improve the paper, would be welcome. Please respond within 2 weeks. (use specific date).

G14. Sending blank forms for co-authors to complete

Journals often require signed forms from co-authors reflecting their contribution to the manuscript, their willingness to be included as a co-author and declarations of potential conflicts of interest. These forms typically require the name of the manuscript and other details that are the same for all co-authors and then a number of co-authors specific issues. Both as a courtesy to your colleagues as well as to boost team efficiency, when sending out these forms, the first author should complete as much of the form as possible, so that, for example, each co-author doesn't need to go back through their files and find out what the exact title of the manuscript is.

G15. Not providing co-authors a copy of the submitted manuscript

Co-authored manuscripts reflect the collective work of the co-authors. When submitting a manuscript to a journal, most journal websites generate a PDF version of what is actually submitted or permit the submitting author to generate such a document. A copy of this document should be provided to all co-authors so that each has the most up-to-date version of the group's collective work. This way if questions come up about the article or the analysis prior to publication, co-authors still have access to the best collective understanding.

G16. Not keeping co-authors informed of journal discussions

Co-authored manuscripts reflect the collective work of the co-authors. When editors and reviewers raise concerns, this discussion is relevant to the whole co-author team. The best practice is to circulate comments as soon as they are received so that all co-authors can consider them. Next, the first author should respond to each of the critiques and make appropriate changes in the manuscript. Often there are several iterations of responses to revisions between the first author and the senior author. Once the senior author is satisfied, then the first author should send around the responses and manuscript changes to all co-authors for their input. It is best to give co-authors one week to review this. Journals often set deadlines for when responses and revisions need to be returned, so it is best to begin working on these revisions promptly to allow the opportunity for all co-authors to weigh in and improve the collective responses and so the final manuscript.

The exception to this approach is when the editor has asked for only minor changes in style or correction of a couple minor errors. Then it is more efficient to simply respond to the journal and send a copy to all co-authors of the responses and the revised submission.

H. Slide and poster presentations

H1. Bullets on the wall

These are slides that present a detailed outline of the talk as bullet points that are projected on the screen / poster board. In the days before slides and screen protectors, speakers commonly used an outline as a prompt to help them remember the key points of their talk. A written outline of the ideas that you want to cover in a talk remains a useful aid to a complete and coherent presentation, especially if you are speaking without slides. However, projecting a detailed outline of your talk on the wall, and then talking through the points bullet by bullet, or even worse, reading them directly to the audience, is a misuse of oral presentation format and a huge turnoff to the audience.

Do you like attending oral presentations where bullets are projected on the wall and the speaker reads them to you? When a Fortune 500 company has a new product to advertise, do they use a bulleted list to communicate its attributes to potential customers? Of course not. We are drawn to engaging speakers and engaging presentations. One of the roles of a scientist is to communicate her/his findings and ideas so that a broader audience considers them, so it affects the audience's understanding and impacts serious discussions.

A verbal presentation is an opportunity to engage the full range of your interpersonal skills to communicate your ideas with your audience. For centuries people have made compelling oral presentations without visual aids. The slides that support an oral presentation should be constructed to reinforce your communication objectives, so it helps the audience understand the ideas you are presenting. Bullets after bullets after bullets bore an audience. This is a recipe for losing the audience attention and a failure to achieve your communication objectives.

Background

- Respiratory viruses can cause pandemics and epidemics
- Emergence of Severe Respiratory Distress Syndrome (SARS) led WHO revise, adopt and implement the IHR (2005) to detect emerging pathogens
- Strong surveillance system is the cornerstone of pandemic preparedness and response
- Early detection of unusual clusters and human to human transmission is the most important function of a surveillance
- Individual disease cluster investigations may not be fruitful unless the causal mechanism is single and the relative risk is high
- Investigation of clustering of a given disease detect space-time aggregation of cases which is caused by environmental agents

Background cont'd

- Proactive identification system can enable public health officials to identify problems earlier
- Bangladesh, most populous country in the world with widespread H5N1 outbreaks in poultry is at particularly high risk for emergence of new strains with pandemic potential
- In 2007, ICDDR,B set up national hospital based influenza surveillance in collaboration with IEDCR i.e. the Government of Bangladesh and Center for Disease Control and Prevention
- In 2009, ICDDR,B embedded cluster investigation to identify new strains and viral etiology of clusters of severe respiratory infection

Figure H1a. Opening slides for an influenza surveillance talk with too many bullets.

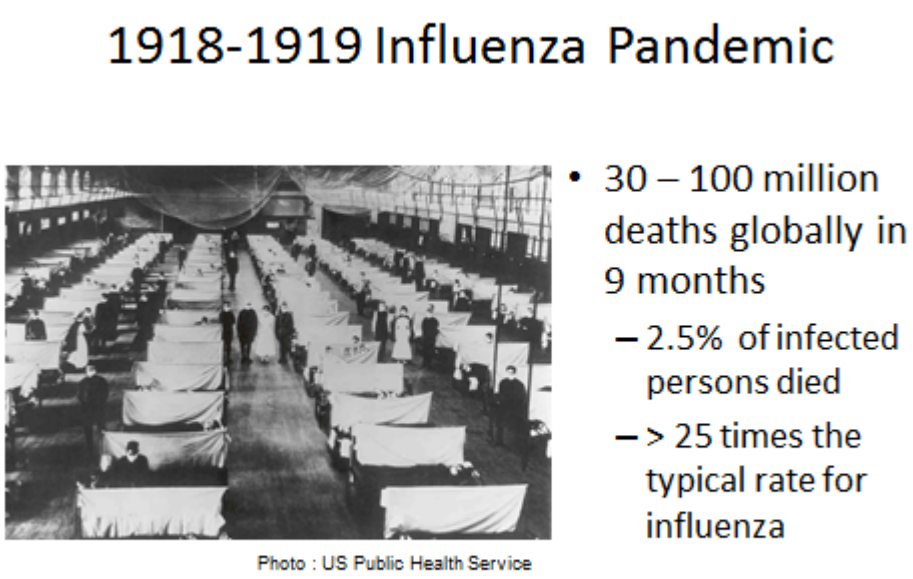


Figure H1b. An alternative opening slide for an influenza surveillance talk that communicates to the audience why this is a compelling issue.

H2. Chart junk

Edward Tufte in his classic book, *The Visual Display of Quantitative Information*, defines chart junk as, visual elements in charts and graphs that are not necessary to comprehend the information represented on the graph, or that distract the viewer from this information. The worst promoters of chart junk are institutions that want all slides to have a common look that advertises the institution. These objectives run counter to clear communication. Clear communication will better promote a scientist and their institution's reputation compared with tacky backgrounds that obstruct and detract. Clear, large and simple is the most effective pathway to clear visual communication. If your institution insists on a stylized template, we recommend using it only on the opening and closing slides.

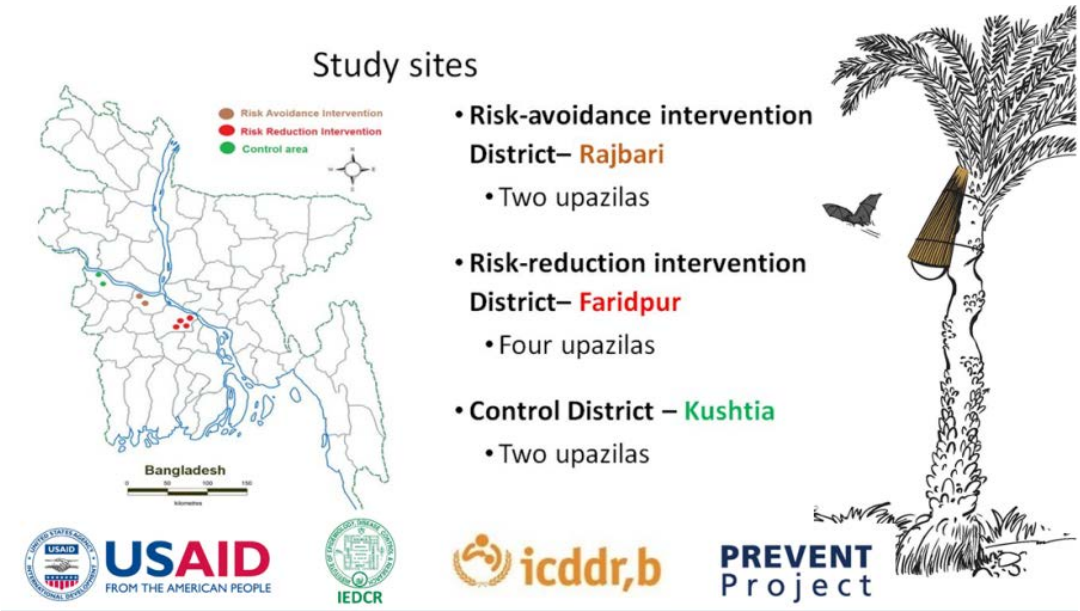


Figure H2a. A slide from a presentation using a template requested from the study funder designed to give credit to funders and a uniform look to the presentation.

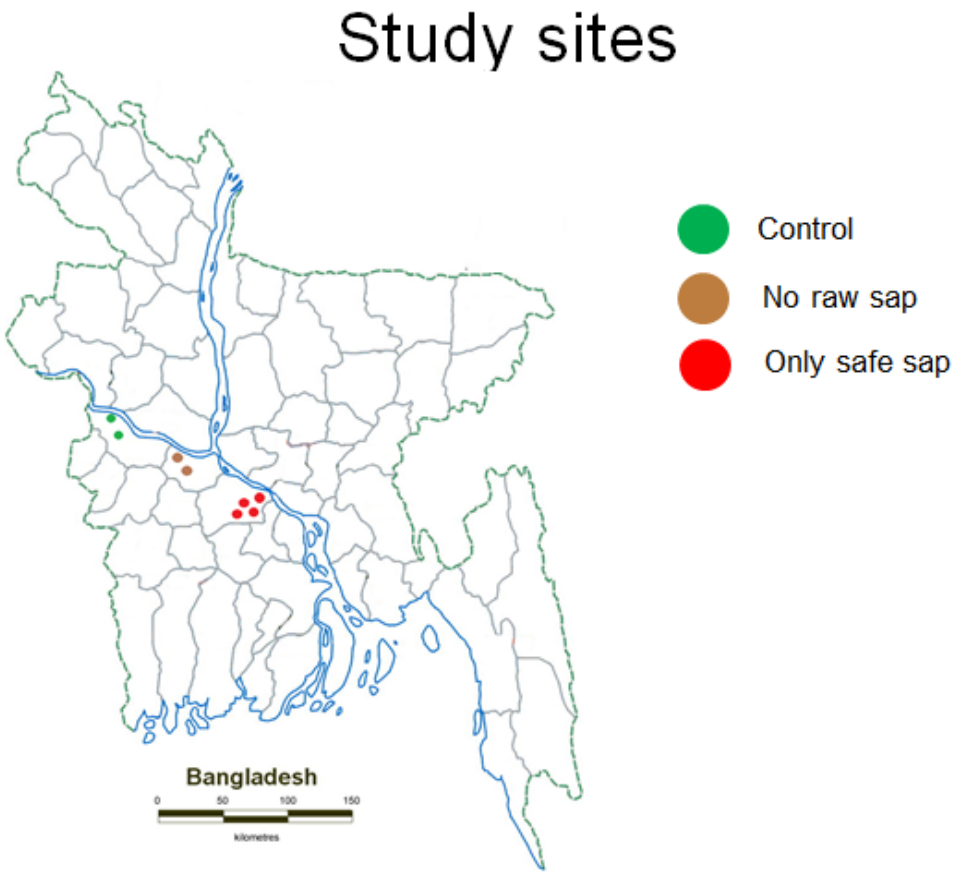


Figure H2b. A cleaner presentation of the slide with chart junk and extraneous information removed to permit attention to the key communication objectives.

H3. Copying a manuscript figure instead of developing a custom figure

Constructing high quality slides to support an oral presentation requires considerable thought, creativity and time. It might save time to use figures developed by others in your own presentation. Especially if you are reporting information from someone else, it is quite tempting to copy directly from their manuscripts or, if you have access to their slides, directly from his/her slides. The drawback to this approach is that visual presentations used for one speaker in one context or as part of a manuscript, often have a somewhat different role in your own presentation. Copying and pasting somebody else's work (even if appropriately attributed) is often not the best way to achieve your communication objective.

Each slide should be integrated with the narrative and communication objectives of your presentation and should be designed to help the audience succinctly understand your ideas. A visual presentation is quite different from reading a manuscript. Figures or tables in the manuscript can include more detail, because the reader can take the time to work carefully through the details. By contrast, the pace of an oral presentation is quicker and so the supporting information should be presented more simply in a clear format that audience can quickly grasp. If you find yourself saying "I apologize for the messiness of the slide but I want to focus on this one issue . . ." or "This is hard to read, but . . . " this is a message to yourself that the slide needs to be revised. Remove the messiness. Clearly communicate the one issue to the audience and jettison the apology.

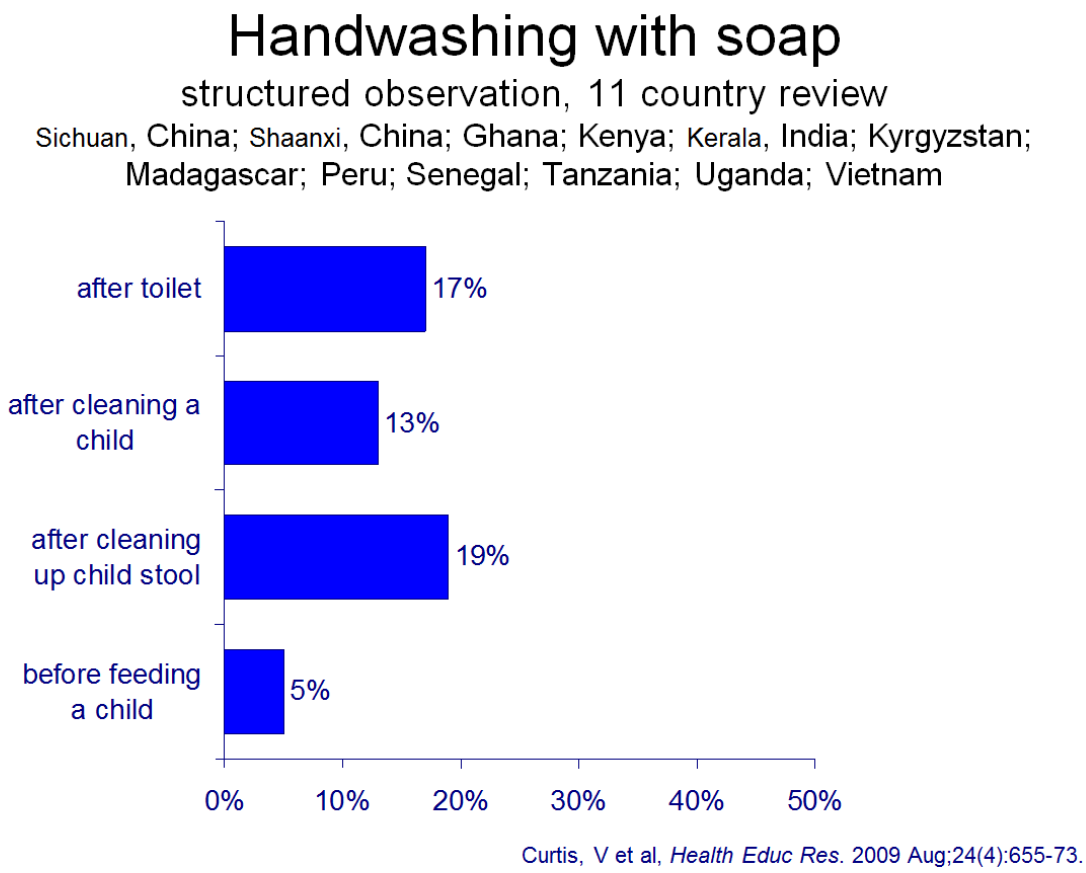
Handwashing with soap (HWWS)

structured observation, 11 country review

Country	*n*	HWWS after toilet (%)	HWWS after cleaning child (%)	HWWS after cleaning up child stools (%)	HWWS before feeding index child (%)
Ghana	500	3	2	—	1
Kerala, India	350	42	—	25	—
Madagascar	40	4	—	—	12
Kyrgyzstan	65	18	0	—	—
Senegal	450	23	18	—	—
Peru	500	14	—	—	6
Sichuan, China	78	13	—	16	6
Shaanxi, China	64	12	—	—	16
Tanzania	30	13	13[a]	13[a]	4
Uganda	500	14	19	11	6
Vietnam	720	—	14	23	5
Kenya[b]	802	29	35	38	13
Average		17	13	19	5

Curtis, V et al, *Health Educ Res*. 2009 Aug;24(4):655-73.

H3a. Slide developed by lifting a table from a manuscript.



H3b. Custom graphic derived from the table to communicate key messages to an audience. Note the elimination of most of the numbers, the removal of the confusing nonstandard abbreviation, yet noting the countries that were actually included.

H4. Photos with an unnatural aspect ratio

Digital photography allows us to insert engaging photographs into our presentations. Often, to make the text fit more neatly with the photograph we adjust the size of the photograph, but sometimes inadvertently also affect the aspect ratio. The aspect ratio is the ratio of the width to the height. If the ratio of the height to width is changed, the photograph appears distorted. This is particularly common when using PowerPoint, and resizing the image by clicking and dragging. Below is the same photograph, with 3 different aspect ratios.

H4A: The photographic subjects have been squeezed, that is the horizontal aspect ratio is too small compared with the vertical

H4B: Here the photograph has been stretched horizontally

HC: This is the photograph as taken by the camera

Changing the aspect ratio distorts the picture and makes readers wonder whether the photographic subjects are oddly disproportioned. To make a photograph fit within a space, consider careful cropping and selecting the right size, but don't change the aspect ratio. You may also need a photograph with a different orientation. When combining text and photographs on a PowerPoint slide, vertically oriented photographs generally use the space better and are easier seen from the back of the room. Encourage your field team to compose photographic subjects that work well with a vertical orientation.

One way to avoid distorted aspect ratios, is to use the insert function on MS Word or MS PowerPoint to directly insert the file, rather than using copy and paste. You can then adjust the size of the photograph by right clicking on the photograph, select size and position, ensure that the "lock aspect ratio" box is checked and then change the size of the photograph by incrementing the height or width using the arrow keys.

H5. Too many photographs on a single slide

Context is characteristically critical for communicating public health scientific results. Many people in the audience will never have visited communities similar to where your study was conducted nor understand the local practices and conditions. Photographs can communicate to an audience the situation that gave rise to the issue of public health interest and the people who are at risk through visual pathways that complement spoken description and written text.

A common saying asserts that a picture is worth 1000 words. Especially in oral presentation when timing is strictly limited, an extra thousand words to communicate your study is a huge asset, but we would slightly modify the saying, i.e. one good picture is worth 1000 words. A good picture illustrates your point, and is easily seen by your audience. A plethora of pictures risks being distracting, because they are too small to see by the half of your audience who are sitting in the back half of the room. Moreover, multiple pictures mean multiple messages, and so the audience may be focusing on

trying to figure out what is in each of the tiny pictures rather than listening to the substance of your verbal presentation.

- **Data collection**:
  - Respondent was a mother or caregiver of child under 5 yrs of age
  - Pre tested questionnaire included:
    - Face-to-face interviews
    - Spot checks
    - Hand washing demonstrations
  - Data were collected in Smart phones/PDA
- **Data Analysis:**
  - Descriptive statistics
  - General regression models
  - Accounted for clustering

H5A Cluttered difficult to see slide

Study Participants

- Mother or caregiver of child under 5 yr of age

Data collection

- Pre tested questionnaire included:
  - Face-to-face interviews
  - Spot checks
  - Hand washing demonstrations

H5B The photograph is large enough that the audience can see the field worker visually inspecting the child's hand

H6. Field workers as the dominant subject of photographs

We cannot usually afford to include professional photographers on our field teams to capture images of the context where we work. Consequently, we depend upon fieldworkers or other members of the study team to take pictures that can be used to communicate context to our audience. Fieldworkers, however, are often particularly interested in pictures of the field team. Although this is occasionally a useful complement to a verbal presentation, photographs that illustrate the conditions as experienced by the target population are generally much more useful. We recommend specifying to the photographers on your team the photographic subjects what you are particularly interested in. Verbal presentations are often given to audiences who have never been in the country nor seen the conditions where the work was conducted, so photographs that provide an evocative illustration of these contexts are particularly usually useful to improve audience understanding.

H6a. Photograph of a water treatment device affixed to a hand pump surrounded by study personnel and men in the compound. This staged photograph displays involved workers, the device, and some information on context, but does show the device being used, or even include women who are the primary caretakers of household water.

H6b. This photograph shows women working with a compromised water supply near an open drain. It illustrates the cramped surrounding and the proximity of supply water to contamination.

H7. Using bullets without hanging indents

Bullets help to format text so that it is clear there are a series of points. They improve readability of narrative. It is easiest to see the difference between points when a hanging indent is used on subsequent lines so that the separation between ideas is clear. In addition a slightly larger spacing between points in contrast to lines within points further makes this separation easier to see and read.

Example H7a: Bullets without hanging indent (The common error):

● Antibiotic use has the potential to contribute to antibiotic resistance.
● Empiric prescription rates for mild respiratory illness range from 40-60% in developed countries.
● We reviewed icddr,b-IEDCR's collaborative hospital-based influenza surveillance data collected from May 2007 to August 2014 to assess antibiotic prescriptions for mild respiratory illness.

Example H7b: Bullets with hanging indent:

● Antibiotic use has the potential to contribute to antibiotic resistance.
● Empiric prescription rates for mild respiratory illness range from 40-60% in developed countries.
● We reviewed icddr,b-IEDCR's collaborative hospital-based influenza surveillance data collected from May 2007 to August 2014 to assess antibiotic prescriptions for mild respiratory illness.

Example H7c: Bullets with hanging indent, single space within points, with 1.2 spaces between lines:

● Antibiotic use has the potential to contribute to antibiotic resistance.

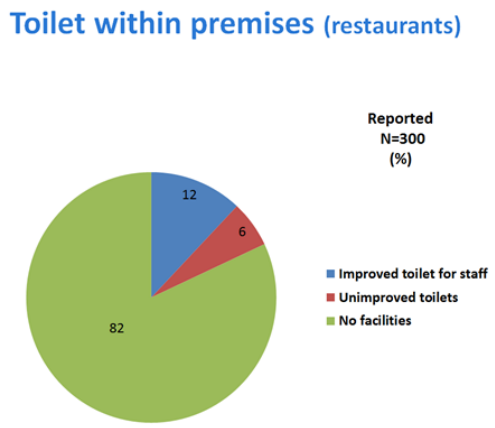
● Empiric prescription rates for mild respiratory illness range from 40-60% in developed countries.

● We reviewed icddr,b-IEDCR's collaborative hospital-based influenza surveillance data collected from May 2007 to August 2014 to assess antibiotic prescriptions for mild respiratory illness.

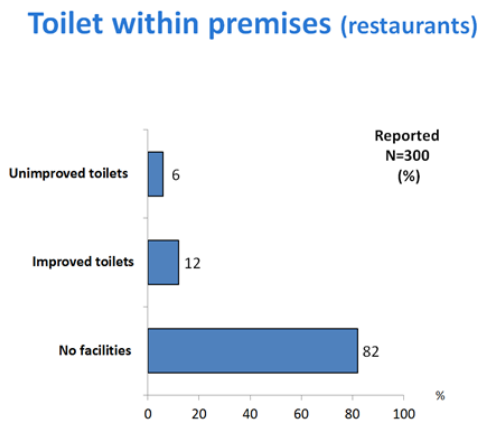
H8. Using a pie chart

For a scientific presentation simple pie charts are best avoided. It is safe to assume that a scientific audience understands percentage without having it illustrated, i.e. they don't need an illustration to appreciate that 25% is one quarter of a pie.

Pie charts made using the default features of PowerPoint are particularly bad. In the PowerPoint pie chart, the reader has to jump back and forth between the pie and the legend to sort out what the particular proportion represents. This requirement that the reader decodes, adds another cognitive task that detracts from simple communication. It invites the audience to focus attention on decoding your graphic at the expense of listening to what you are saying. If there is a compelling reason for a pie chart, use labeling that avoids a legend.

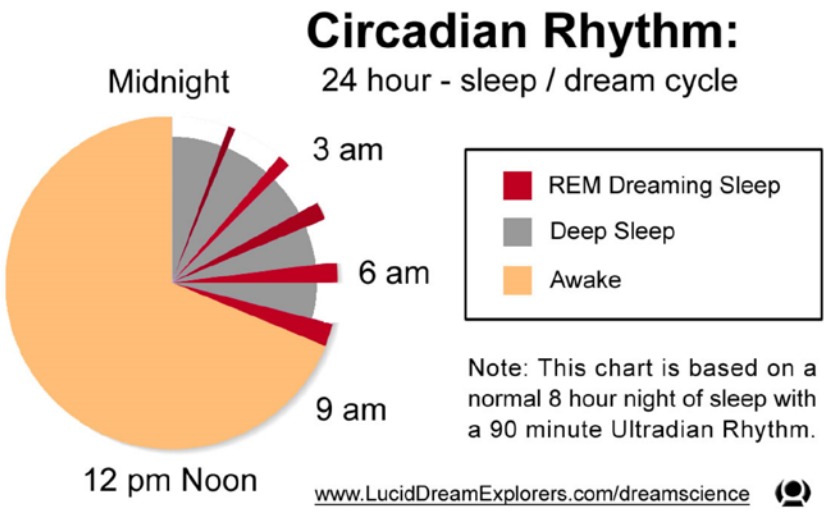


H8A. Default Pie chart from PowerPoint. It is both under-informative and requires decoding.

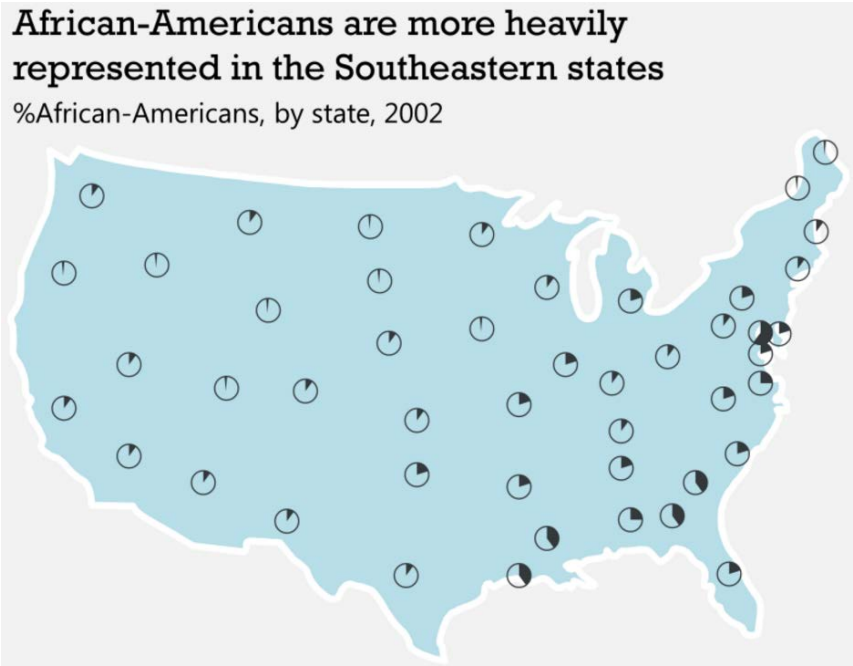


H8B. Easier to interpret visualization of data from H8A. The labels are right next to the numbers. No decoding required.

An exception to the avoid pie chart rule, is when a comparison between 2 groups or a breakdown of a subgroup of a pie provides a useful illustration that engages the audience's visual understanding to interpret patterns in the data.



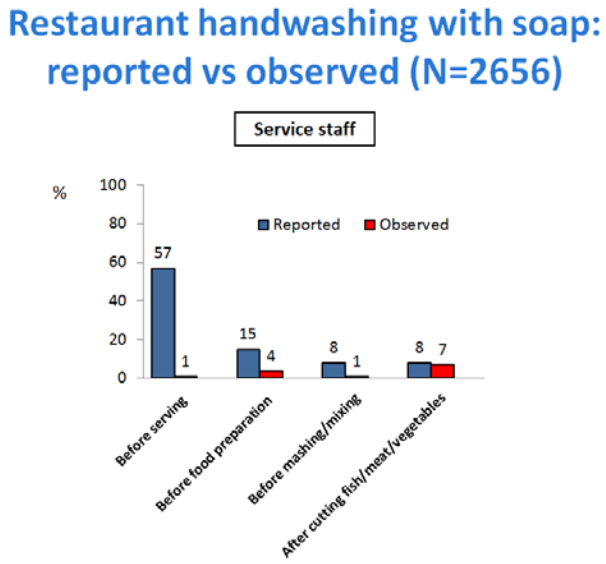
H8C. An illustrative pie chart which effectively embeds additional meaning, and communicates effectively



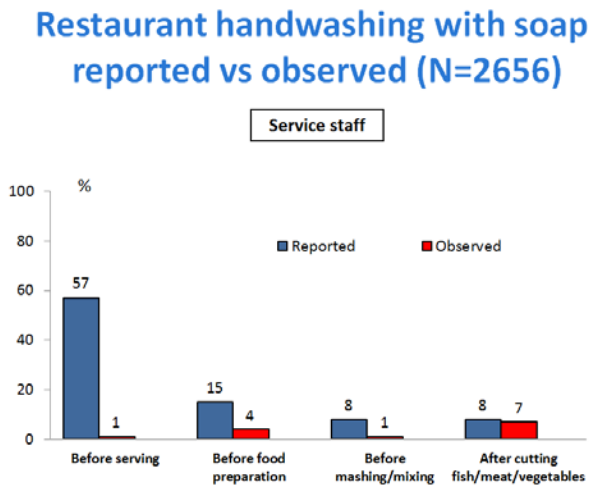
H8D. A comparative pie chart, that supports a visual understanding of a distribution. From (http://speakingppt.com/2013/03/18/why-tufte-is-flat-out-wrong-about-pie-charts/)

H9. Using vertical bars when horizontal bars would communicate better

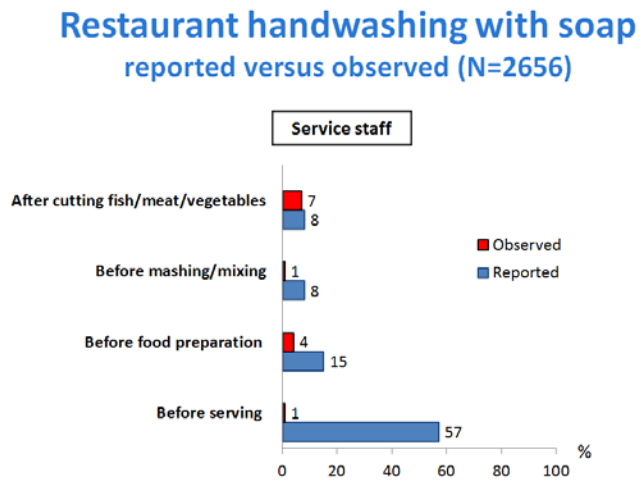
Vertical bar charts are commonly used default formats in PowerPoint, but they are often not the best way to present data. If a useful description of the variable being presented is long, it is difficult to read in the constrained space or in an odd angle at the bottom of a slide. A horizontal bar allows more space and larger font to facilitate quick communication.



H9A. Vertical bar chart with long labels. Note that the titles do not align intuitively with the bars. Our eyes are not accustomed to reading across odd angles.



H9B. Vertical bar chart with multi-line descriptions. These are often small and difficult to read.



H9C. Simpler, easier to read horizontal bar chart.

PowerPoint is quirky. In many versions of PowerPoint, the order of appearance of the horizontal bars is directly counter-intuitive, i.e. when you construct the data table, the first variable you enter displays at the bottom of the chart, and the bottom variable is at the top. You can simply reverse the order in the data table to have it present according to what aligns best with your communication objectives.

H10. Including a final “Thank you” slide

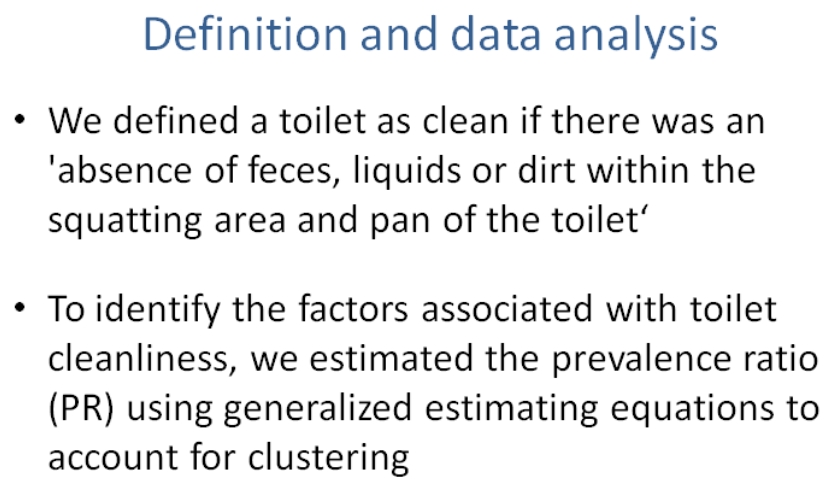
Having your final slide say “Thank You” (presumably to the audience for their attention) often accompanied by an illustration that is irrelevant to the theme of your talk is common in some contexts. Such slides are less common in an international scientific forum. Indeed, they often appear out of place. The gratuitous graphics distract from the major communication message of your talk. Drop such slides. Your final slide should either be acknowledgements, or conclusions.

H10A. A final “Thank You” slide best left out of the presentation.

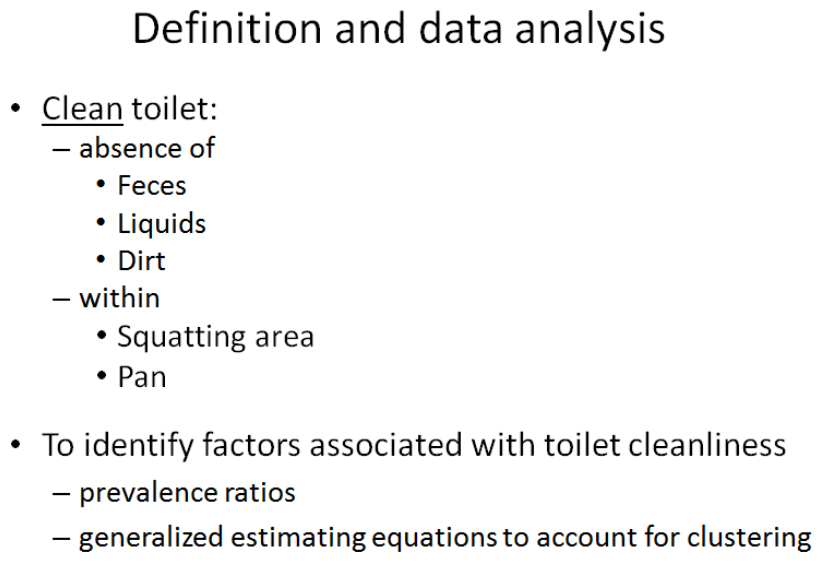
H11. Using sentences for bullet points

Bullet points should be terse summaries that help the audience follow your key points. They should not be full sentences nor paragraphs that you read. Full sentences and paragraphs are appropriate for scientific writing, but it is mind numbingly boring to have full sentence after full sentence projected with the speaker reading the sentences to the audience. The average audience member can read such sentences 3 – 5 times faster than the presenter can speak them, so this is not an efficient method to communicate. It is a misuse of a verbal presentation opportunity.

Posters are meant to be read, and so somewhat longer lines of text can be used than in a verbal presentation, but ideas that break down into sections should still be presented as brief bullet points so people can quickly grasp the structure of the ideas.



H11A. Sentences making minimal use of visual organization of ideas



H11B. Ideas organized as bullets. This would also accommodate a nice picture of a clean toilet which would further enhance communication.

- In courtyard meetings, participants received messages to wash hands with soap after defecation, after cleaning a child's anus and before preparing food. HW before child feeding was not included in the promotional messages

H11C. Paragraph like bullet from a draft poster.

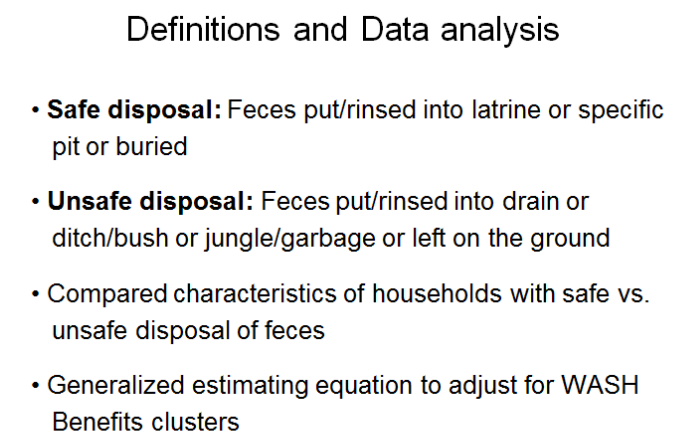
- In courtyard meetings, participants received messages to wash hands with soap:
  - after defecation
  - after cleaning a child's anus
  - before preparing food.
- Handwashing before child feeding was not included in the promotional messages

H11D. Information recast as quick to read organized bullets

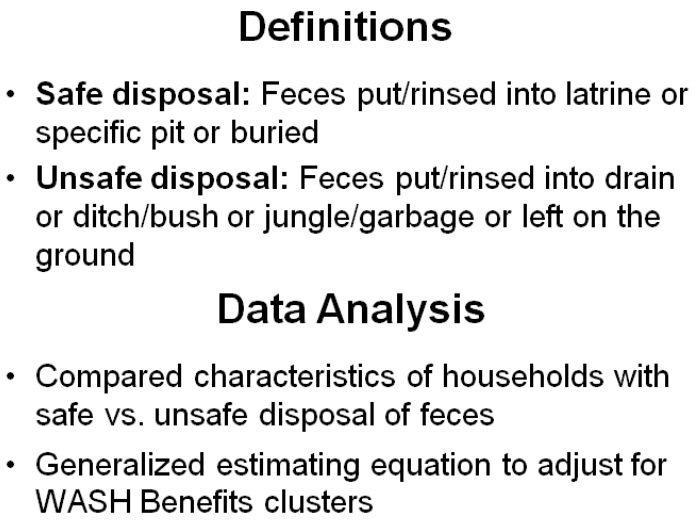
H12. Too much space between bullets

Oftentimes PowerPoint inserts substantial space between lines of text. This can occur, both as too much space between lines within a bullet, as well as too much space between bullets. All of this white space reduces the amount of space for communication and forces smaller font sizes that becomes difficult or impossible to read, especially from the back of the room.

These spacing issues can be addressed by using the paragraph features of PowerPoint. Set the line spacing to single, and make spacing Before and After small (e.g. <6 pt.). Another strategy to modify space between bullets is to insert a line with a single letter of text. Color the text the same color as the background and adjust the font size to something small that optimizes spacing.



H12A. Lots of white space, not well used that limits font size



H12B. Reorganization of slide redistributes white space to better group and communicate ideas. Animation features could be used so that the top of the slide appear first and the Data Analysis section appears when the presenter clicks

Conclusion

- Restaurant staff and street food vendors in Bangladesh do not usually wash their hands with soap during food handling
- Limited facilities contributes to a lack of hand and food hygiene
- Impractical for street food vendors to carry and store water
- Cost of soap is also a barrier

27

- Equated handwashing with hands contacting water
- Respondents perceived that customers are satisfied if they get tasty food

H12C. So much space between the bullets that the list stretches across 2 slides.

28

Conclusion

- Restaurant staff and street food vendors in Bangladesh do not usually wash their hands with soap during food handling
- Limited facilities contributes to a lack of hand and food hygiene
- Impractical for street food vendors to carry and store water
- Cost of soap is also a barrier
- Equated handwashing with hands contacting water
- Respondents perceived that customers are satisfied if they get tasty food

H12D. Same bullets with reasonable spacing between fit on a single slide

H13. Failure to separate ideas in a multi-lined title

When typing a sentence, after producing sufficient text to fill a line, the next word appears on the next line. This works fine for sentences, but is sub-optimal for titles. Titles are an integral element of the visual presentation of your ideas. By thoughtfully dividing the title into natural parts the audience can more quickly understand your message.

H13A. Multi-line title running to the end of the line

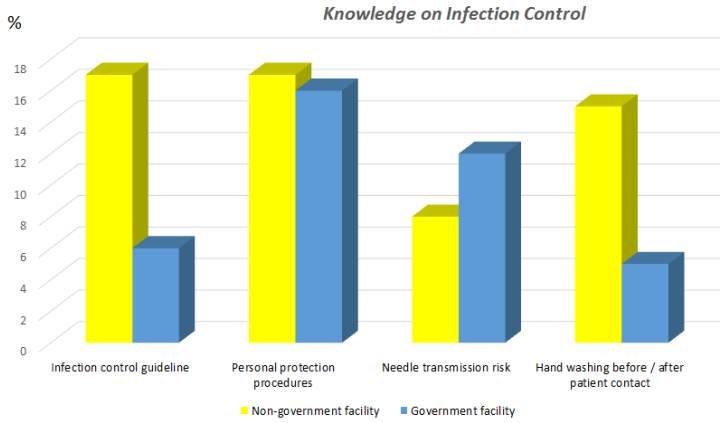
H13B. Better title split by ideas

H13C. Default splitting of title

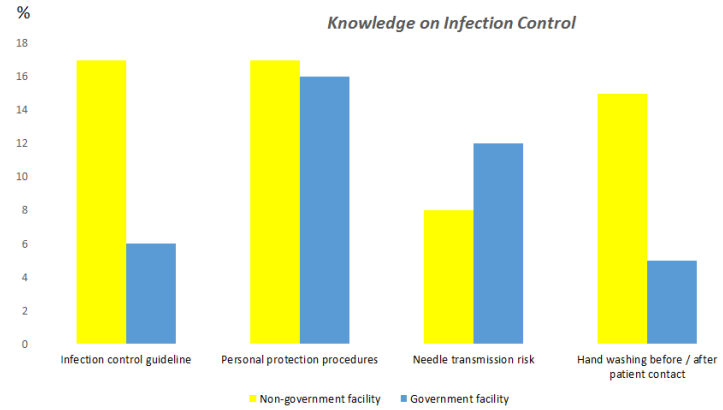
H13D. Improved title with ideas grouped together.

H14. Using 3 dimensional chart features as decorations

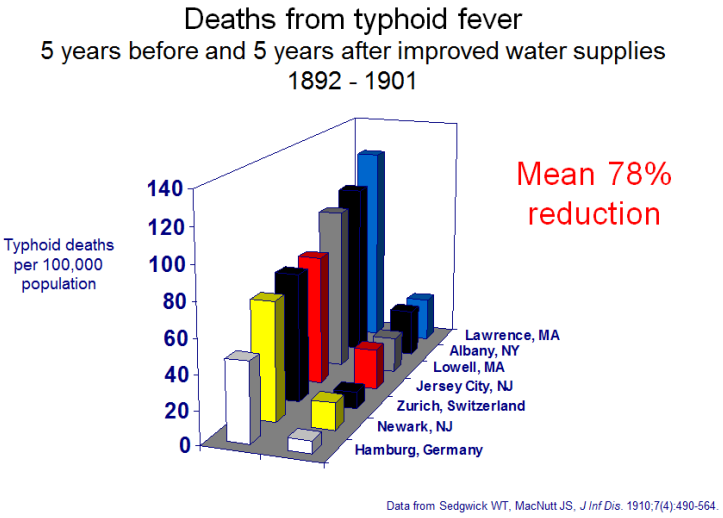
Charts and figures are used to connect to the visual centers of human perception and so improve communication of your quantitative results. Adding three dimensions to these charts adds complexity. This complexity should only be invoked if it is used in the communication of the data. Otherwise, this three-dimensional imagery is chart junk (Error H2) that risk distracting the audience. Strive for minimalist elegant images that communicate without distraction.



H14A. 3 dimensions used as un-informative chart junk

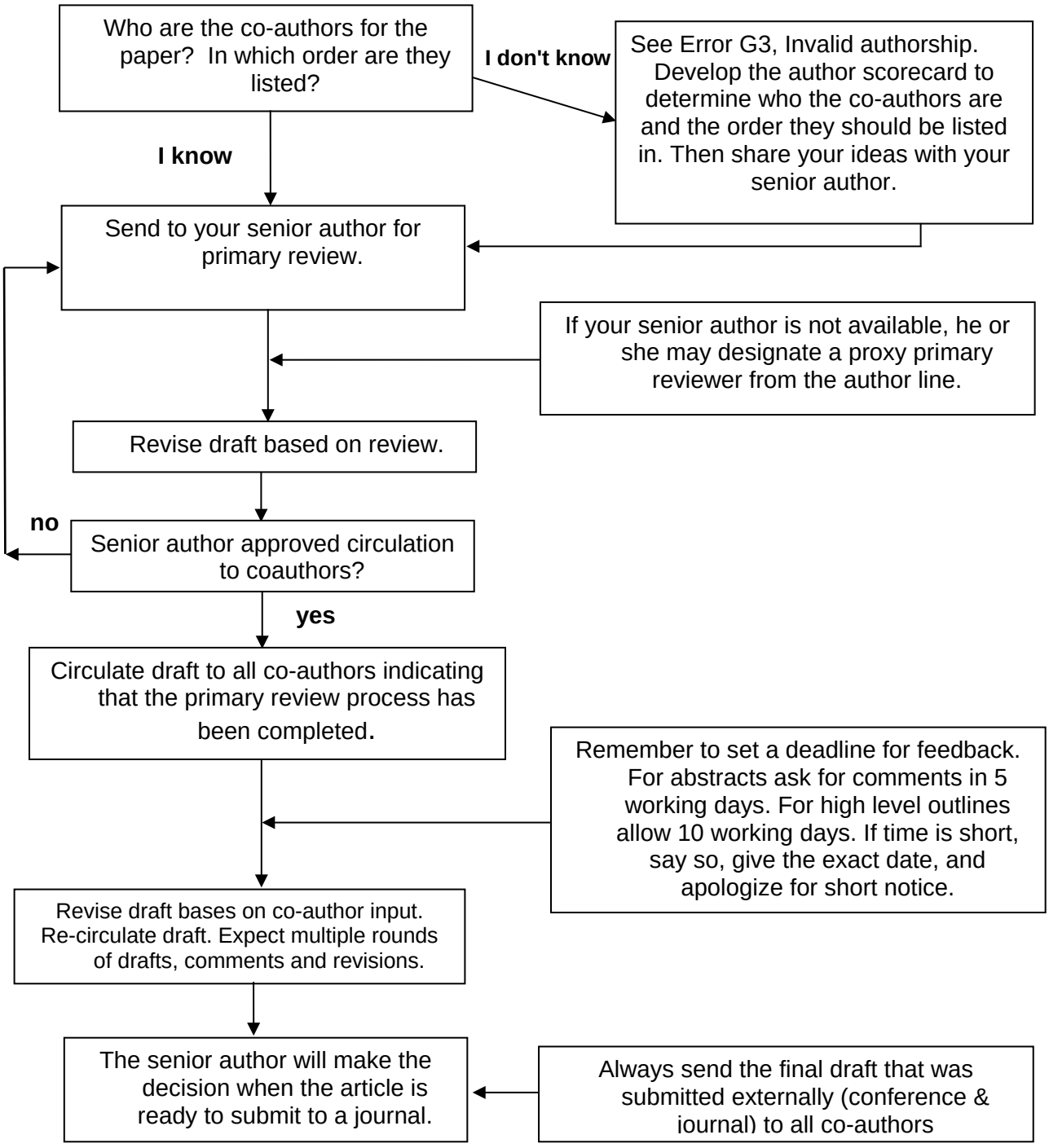


H14B. Simpler cleaner chart



H14C. 3-dimensional features used to support data communication

Appendix 1: Flowchart for reviewing scientific documents



Appendix 2: Concept note outline

1) Title of the proposed study

2) Objective(s)
   a) What will the study funder receive if they invest in this study?

3) Background
   a) Current state of knowledge on specific study question
      i) Not a general review, but tightly focused on study question
      ii) Cite key literature
   b) Specify the gap in current knowledge
   c) Describe the relevance of the study question. Why should readers/funders care?

4) Methods
   a) Study site and study population
   b) Study design
   c) Key definitions (e.g., case definitions)
   d) Sampling methods
   e) Data collection tools and processes
   f) Laboratory analysis
   g) Sample size assumptions and calculation
   h) Data analysis plan, including statement of the primary outcome
   i) Ethical considerations

5) Timeline
   a) Gantt chart

6) Budget
   a) Only major items (personnel, transportation, laboratory tests, materials)
   b) Based on sample size
   c) Help researchers understand the cost implications of methods
   d) Help decision makers understand the resources required

Appendix 3: Critical questions for protocol development

Thinking Critically

1. What is your over-all research question?
2. What is the hypothesis that you want to test?
3. What is the aim(s) of your study?
4. What do you already know about the subject?
5. What don't you know about the subject? (the gap in knowledge)
6. Why is this research important? What kind of answers will the study provide?

Research Design and Methods

7. What is the identified target group?
8. What type of study design did you choose to test your hypothesis?
9. What is your sample size?
10. How did you estimate your sample size?
11. What is the statistical power of your study?
12. How did you select your study unit of population (explain sampling method)?
13. How will you collect your data?

Data Analysis

14. What variables are you going to study?
    a. Outcome variables
    b. Exposure variables
15. How are you going to measure these variables?
    a. For categorical variables, what are the category definitions?
16. How will you analyze your data to test your hypothesis?

Ethics

17. How will you provide ethical assurance for protection of human/animal rights?

Logistics

18. How long will the study take? What is your time line?
19. How much will this cost?
20. When will the results become available, how will you disseminate them?

Appendix 4: Framing document

Title of paper:

Objective(s) of the paper:

Main result(s)

1.

2.

3.

Tables, figures or graphs that support your main results:

(Example only....you might have 5 tables, or any combination)

Table 1:

Table 2:

Figure 1:

Graph 1:

Table 3:

Appendix 5: Conference/scientific meeting abstracts

Domestic and international conferences often publish a 'Call for Abstracts' to identify oral presentations and posters on relevant subjects that can be featured in that meeting. Before you think of applying, read all of the information about the conference carefully. Ensure that the potential audience is the right fit to showcase your particular results. Conference presentations are excellent opportunities to collect feedback on your work. Such feedback can help in the development of your manuscript. You want to choose a conference where the attendees will be interested in your work and so likely to provide thoughtful feedback.

Usually the conference will give specific guidelines on the length of the abstract and how to submit on line. Read all the instructions carefully before you start developing your abstract. You can think of your abstract as a mini-version of your study that includes four sections: background, methods, results, and conclusion. Do not need include references. You can use numerals instead of words to save characters and space. But make sure to include your main results, i.e. the specific numbers, especially for primary outcome measures.

To develop an abstract, follow these steps in sequence:

Step 1: Results

- Use your framing document to identify main results.
- Include raw data including percentages, confidence intervals (CI), odds ratios (OR), p-values, or whatever statistical analysis is important to showcase your results.

Step 2: Conclusion

- Write a broad statement interpreting your results and how they link to your objective and what they mean for public health.
- Write a practical recommendation and/or next steps for research.

Step 3: Methods

- For each result, check that you have included a corresponding method.

Step 4: Introduction

- Background: Provide concise information directly related to your objective and results.
- Last sentence should be a clear statement of your objective.

Review examples of accepted abstracts from prior years of the targeted conference. These are generally available on line.

Appendix 6: Quantitative manuscript high level outline (HLO)

Use sub-titles that match your study;
Limit to ≤1500 words excluding tables, figures and references

Introduction

- **Context**
  - Introduce the subject to provide context for the objective.
- **Information gap**
  - What don't we know that this manuscript will address?
- **Relevance**
  - Why is this knowledge important (the 'so what?' question).
- **Objective of the manuscript**

Methods

- **Study site and population.**
  - Outline the setting in which the study was carried out, e.g., urban vs. rural.
  - Mention the study participants, e.g., women, or children under five years of age.
- **Design and sampling**
  - Describe the study design / approach
  - Provide key operational definitions
  - Outline sampling methods
- **Data collection?**
  - Outline data collection approaches
  - Outline any special laboratory materials, equipment, or reagents.
- **Data analysis**
  - Outline primary approach

Results

- One bullet point to summarize each table
- One bullet point to summarize each figure

Discussion

- **Summary interpretation of overall results**
  - Link to objectives and rationale
  - Avoid repeating the results (no statistics).
- **List the primary conclusions that you can logically and defensibly draw from the results.**
  - Outline that supports this conclusion.
  - If a statistical association represents one of the core conclusions and you believe the association reflects an underlying causal relationship, then outline evidence to support that this association is likely causal. Also outline alternative potential interpretations and evidence that supports that.
- **Limitations**
  - Focus on the impact that these limitations have on the conclusions we can draw the study.
  - Outline how you interpret the data in light of these limitations.
- **Conclusions**
  - Outline the big picture: How do your results help us understand a broader topic?
  - What implications do your results have for public health or related policies?
- **Recommendations**

- What are the key next steps that are practical and applicable to the context?
- What specific research question should next be pursued?

- **References**
  - Need not be complete, but helps clarify the key issues in the Introduction and Discussion.
  - Permits the author to offer an interpretation based on the literature of key issues and provides the co-authors the opportunity for input on this framing

- **Tables and figures**

Appendix 7: Example of quantitative manuscript HLO

Title: Difficulties in Maintaining Improved Handwashing Behaviour, Karachi, Pakistan[16]

Introduction
- Handwashing with soap can reduce diarrhoea and respiratory illness *(Refs)*
- Handwashing promotion that requires repeated household visits is prohibitively expensive on a large scale *(Refs)*
- In 2003, we conducted a cluster randomized control trial in low-income squatter settlements in Karachi, Pakistan
- Field workers promoted improved handwashing by providing households with free soap and weekly visits over a 9 month period up to December 2003
- We conducted a follow-up study 18 months later to determine how long selected households sustained improved handwashing practices

Methods

Study Setting
- Adjoining multi-ethnic squatter settlements in central Karachi
- Field work was conducted by Health Oriented Preventive Education (HOPE), a local non-governmental organization

Study Design
- In the 2003 cluster randomized control trial, 47 clusters of households were selected and randomly assigned 5 intervention groups:
    - 9 clusters received soap and encouragement
    - 10 clusters received soap, handwashing promotion and flocculent disinfectant
    - 9 were controls that received no intervention
- In the 2005 follow-up cohort study, field workers, who had not participated in the 2003 study, attempted to revisit households assigned to either of the intervention clusters that included soap and handwashing promotion or to the control group (Figure 1)

Data Collection
- Field workers conducted a re-enrolment survey using a standard questionnaire and performed spot checks of facilities for handwashing
    - They asked the mother or caregiver of the household:
        - To demonstrate usual handwashing practices
        - If any children in the household had diarrhea (three or more loose stools within 24 hours) in the preceding week, and, if so, for how many days
        - If mother or caregiver had diarrhea
        - How much hand soap was purchased in the preceding week

Data Analysis
- We compared characteristics of re-enrolled households by originally assigned intervention groups with the control group using generalized estimating equation
- We calculated respondents' longitudinal prevalence of diarrhoea
- To assess the relationship between soap consumption and diarrhoea, we used the number of bars of soap purchased during the week divided by the number of persons in the households as the independent variable, and the longitudinal prevalence of diarrhoea in the subsequent week as a dependent variable in a generalized estimating equation model
- For all generalized estimating equation models, we used an exchangeable correlation structure applied to neighborhoods to account for clustering derived from spatial proximity

Results

Descriptive

- A total of 577 households were enrolled: 69% (560) were re-enrolled from the original study's 810 households; 17 were households that split and set up new households in the same study area
- The 560 re-enrolled households were similar to the 250 households that declined re-enrolment by household size, water supply, reported income, and amount spent on soap and water (Table 1)
- Households that re-enrolled were more likely to have been assigned to the handwashing promotion with soap intervention during the original study and were more likely to own a refrigerator and television (Table 1)

Handwashing behaviour

- At re-enrolment, intervention and control households were just as likely to have soap in the house and reported similar spending on hand soap (Table 2)
- Households originally assigned to handwashing promotion, but with no water treatment, were more likely to have a handwashing station with soap and water (79%) than control households (53%, $P = 0.001$), or households that received both handwashing promotion and water treatment (64% $P = 0.05$)
- During the 63 week follow-up, intervention households purchased a similar quantity of soap and used a similar amount of soap per capita per week compared with control households (Table 2; Figure 2)

Diarrhea prevalence

- During the first 5 months of follow-up, households from the different intervention groups reported different prevalences of diarrhea. In the subsequent 8 months, the prevalence was similar across the groups (Figure 3)
- The overall longitudinal prevalence of diarrhoea was 15–16% lower in the intervention households. After accounting for clustering, neither the longitudinal prevalence among all ages, nor any of the age specific diarrheal prevalences were significantly different between intervention and control households (Table 3)
- There was no association between weekly per capita soap consumption and longitudinal prevalence of household diarrhoea in the following week ($P = 0.38$)

Discussion

- These findings illustrate important barriers to improving handwashing behaviors globally. Households that received the handwashing intervention:
  - Acquired the habit of washing hands properly and maintained it for several months.
  - Had a better place to wash hands
  - Experienced a substantial reduction in diarrhoea
- When soap was no longer provided free, and regular encouragement to wash hands stopped, their behaviour reverted to less soap consumption and a disease experience that was no different than households that received no intervention
- These results are similar to findings from a follow up of a randomized controlled trial of household water treatment that found high levels of product use during the study period accompanied by a marked reduction in diarrhoea, but no sustained regular use
  - Only four evaluations of long term sustainability of handwashing promotion were identified *(Refs)*
- In the Karachi study the lack of a sustained improvement in handwashing behaviour suggests that specific methods used for short term efficacy, e.g., free soap, did not produce long term behaviour change

    - This is consistent with behaviour change specialists who note that maintaining a changed behaviour is fundamentally different from acquiring a new behaviour: Maintenance has different determinants and requires different interventions (Refs)
- In the first 6 months there was some difference in diarrhea prevalence, but later there was none, suggesting the declining impact of the intervention over time, that might have been lessened with occasional refresher visits
- The amount of soap purchased by households was used as an indirect measure of handwashing, taking into account that soap is used for many household purposes and is sold in different sizes
    - We hypothesized if handwashing increased, then soap purchases would increase
    - No difference in amount of soap or an increase in spending on soap suggests no sustained change behaviour by this intensive intervention

Limitations

- Limited power to detect a difference in the longitudinal prevalence of diarrhoea between the intervention and control arm
- Of the originally enrolled households, 29% did not participate in the follow-up evaluation.

Conclusion

- Improved handwashing behaviour is not guaranteed to be maintained when the activities promoting that behaviour are withdrawn

Recommendation

- Like other behaviour change interventions, maintaining effective handwashing behaviour requires focused efforts and research on optimal strategies

Tables and Figures

Appendix 8:Authorship Scorecard

A Worksheet for Authorship of Scientific Articles Author(s): Robert H. Schmidt
Source: Bulletin of the Ecological Society of America, Vol. 68, No. 1 (Mar., 1987), pp. 8-10
(Included with permission of publisher and author)

A WORKSHEET FOR AUTHORSHIP OF SCIENTIFIC ARTICLES

Inclusion as an author in a scientific publication is important to many ecologists for reasons of prestige and advancement. Publications are a key factor in deciding on promotions for many ecologists at universities (Jackson and Prados 1983, Croll 1984). The order of listed authors on a paper is assumed to be an indication of the relative contribution of each of the included authors.

Day (1983:15–19), Croll (1984), Kennedy (1985), and Jackson (1986) reviewed contemporary difficulties with decision-making in assigning authorship. Dickson et al. (1978) proposed guidelines for determining inclusion and ranking in authorship of a scientific publication. They divided research investigations into five areas: conception (including funding), design, data collection, data analysis, and manuscript preparation, and recommended that authors need to make, at a minimum, a significant contribution in manuscript preparation and in at least one other area. Authorship order was determined by a ranking of the number of areas in which significant contributions were made.

This paper details a method for assisting in (1) deciding who is to be listed as an author on a paper, and (2) the ordinal ranking of authors listed on a paper. Of course, the best procedure for dealing with potential problems in assigning authorship is to deal with the issue at the beginning of a study.

8

Table 1. Format and example of a worksheet for determining the relative contributions of participants in a research project. Values listed are percent relative contributions. In this exapmle, a natural cutoff for authorship status would be between Technicians C and D. Authorship ranking should be Leader A, Leader B, and Technician C. The number in parentheses is a multiplier (see text for details).

Investigator	Conception (1.0)	Design (1.0)	Data collection (1.0)	Data analysis (1.0)	Writing (1.0)	Total
Leader A	50	90	0	70	40	250
Leader B	50	10	20	0	30	110
Technician C	0	0	40	30	30	100
Technician D	0	0	40	0	0	40
Column totals	100	100	100	100	100	500

However, even preassigned roles can have complications, especially when personnel on a project change, or when responsibilities are transferred. In addition, people often underestimate the inputs required, especially time, for the various contributions, making initial agreements, in retrospect, seem unfair. The trend toward multiauthored papers may indicate how research is becoming increasingly interdisciplinary. In these situations a method for defining authorship roles becomes useful. This simple technique should be a useful decision-making aid, especially for projects with many researchers involved.

A general framework for a decision-making worksheet, with an example, is given in Table 1. For each of the five parts of the research investigation (as defined by Dickson), the relative contribution of each participant is assessed. For each part, total contributions should equal 100%. When all contributions have been assigned, row values are added, resulting in a "score" of between 1 and 500. The relative contribution of all participants can then be assessed, and a natural break between subsets of scores on the lower end of contributions can be used as a cutoff to delineate inclusion as an author. Scores can then be ranked for order of authorship.

This technique has a number of assumptions. First, it assumes that each of the five parts of a research investigation are weighted equally. In some situations, this may not be the case. For example, a study may require minimal funding, the infrastructure of a principal investigator's laboratory may be essential to a successful project, or the data set may be collected over several years. This situation is easily dealt with by weighting the unbalanced part with a multiplier. For example, all values in the "data collection" column can be multiplied by 1.2, if data collection is judged to have been 20% more important than the other areas.

Secondly, this technique assumes that all contributions can be judged fairly and accurately. This may not always be the case; indeed, it may be that this technique would only be necessary for papers where it is difficult to assess contributions. Two points are suggested for resolving this. It must first be recognized that each contribution score is usually an estimate, and, as such, has some corresponding error associated with it. Therefore, the difference of only a few points between participant's scores is probably not sufficient to rate relative contributions, and other methods must be utilized to determine authorship ranking (perhaps even the flip of a coin). As the second point, a consensus-type survey system, such as the Delphi system (Schuster et al. 1985), may be useful as an in-house tool for resolving difficult authorship assignment problems, although it is recognized that assigning authorship is rarely a democratic process.

How are contributions assessed? One method that could be used is the actual time (hours, days, years) put into each of the five parts of the research investigation. A key problem here is the importance of experience. For example, how would you compare a two-hour contribution to a project's design from a person with 30 years of experience with a two-hour contribution from a person with little or no experience? Another method, admittedly subjective, is an assessment of the "importance" (relating to intellection) of contributions in each area. Again, a consensus-type survey can be helpful in arriving at an

acceptable and agreeable assessment. The development of some criteria for better assessment of contributions is needed. Time should be minimized, while intellectual contribution should be maximized, yet it is easy to visualize a project in which time is a real measure of effort.

Finally, there is a situation which involves teams of workers involved in one of the five parts. A realistic example would be having many workers assisting in data collection. Although the team's contribution may be large (perhaps 100% of the data collection), the relative contribution of each team member is small. The "points" given to this team may then be assigned to the team coordinator or leader. There is some question whether a "technician" should ever be a coauthor, especially if his or her sole responsibility is data collection or data collection and analysis, when the analysis is limited to performing routine operations rather than interpretation (Dickson et al. 1978).

It must be repeated that this system for determining authorship of scientific articles should not replace consultation among authors. However, it should be useful in delineating relative individual contributions when there are many, and it can help project coordinators or senior authors identify personnel who have contributed in a significant way to a study's conclusion. Authorship is a symbol that means taking responsibility for the contents of the paper (Jackson 1986). If the responsibility is there, inclusion as a co-author is appropriate. This worksheet should be helpful in defining this responsibility.

ACKNOWLEDGMENTS

For helpful comments on this essay, I thank C. Shugart, W. Howard, J. Aloi, R. Case, T. Tomasi, D. Anderson, R. Johnson, J. Tully, P. Moyle, and T. Salmon.

LITERATURE CITED

Croll, R. P. 1984. The noncontributing author: an issue of credit and responsibility. Perspectives in Biology and Medicine **27**:401–407.

Day, R. A. 1983. How to write and publish a scientific paper. ISI, Philadelphia, Pennsylvania, USA.

Dickson, J. D., R. N. Conner, and K. T. Adair. 1978. Guidelines for authorship of scientific articles. Wildlife Society Bulletin **6**: 260–261.

Jackson, C. I. 1986. Honor in science. Sigma Xi, New Haven, Connecticut, USA.

Jackson, C. I., and J. W. Prados. 1983. Honor in science. American Scientist **71**: 462–464.

Kennedy, D. 1985. On academic authorship. American Council of Learned Societies, Office of Scholarly Communications and Technology, Scholarly Communication Reprint **4**:1–5.

Schuster, E. G., S. S. Frissell, E. E. Baker, and R. S. Loveless, Jr. 1985. The Delphi method: application to elk habitat quality. United States Forest Service Intermountain Research Station Research Paper **INT-353.**

Robert H. Schmidt
University of California
Hopland Field Station
4070 University Road
Hopland, CA 95449

Appendix 9:JANE (Journal/Author Name Estimator)

This information and more and is available on http://biosemantics.org/jane/

Summary:
With an exponentially growing number of articles being published every year, scientists can use some help in determining which journal is most appropriate for publishing their results, and which other scientists can be called upon to review their work. Jane is a freely available web-based application that, on the basis of a sample text (e.g., the title and abstract of a manuscript), can suggest journals and experts who have published similar articles.[17]

How does Jane work?
First, Jane searches for the 50 articles that are most similar to your input*. For each of these articles, a similarity score between that article and your input is calculated. The similarity scores of all the articles belonging to a certain journal or author are summed to calculate the confidence score for that journal or author. The results are ranked by confidence score. For more information, you can read .

How often is the data behind Jane updated?
We are currently updating the data once every month.

Which journals are included in Jane?
Basically, all journals included in Medline are included in Jane. However, in order to show only active journals, we do not show journals for which no entry was found in Medline in the last year.

Which authors are included in Jane?
All authors that have published one or more articles in the last 10 years that have been included in Medline, are included in Jane.

Which papers are included in Jane?
All records in Medline have been included that 1) contained an abstract, 2) were published in the last 10 years, 3) did not belong to one of these categories: comment, editorial, news, historical article, congresses, biography, newspaper article, practice guideline, interview, bibliography, legal cases, lectures, consensus development conference, addresses, clinical conference, patient education handout, directory, technical report, festschrift, retraction of publication, retracted publication, duplicate publication, scientific integrity review, published erratum, periodical index, dictionary, legislation or government publication.

* For the computer geeks: we use the open source search engine . Queries using keywords are parsed with the Query Parser class, titles and abstracts are parsed using the MoreLikeThis parser class.

Appendix 10: STROBE Statement

The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) recommendations are aimed at improving the quality of reporting of observational studies. The STROBE Statement provides guidance to authors about how to improve the reporting of cohort, case-control, and cross-sectional studies. It facilitates critical appraisal and interpretation of studies by reviewers, journal editors and readers through the use of a checklist of 22 items, which relate to the title, abstract, introduction, methods, results and discussion sections of the article. Eighteen items are common to cohort studies, case control studies and cross-sectional studies and four are specific to each of the three study designs.

The STROBE checklist is best used in conjunction with an Explanation and Elaboration article that discusses each checklist item and gives methodological background and published examples of transparent reporting.[18] More information about STROBE is available at www.strobe-statement.org.

Manuscript Section	Item Number	Recommendations
TITLE and ABSTRACT	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found
INTRODUCTION		
Background/ rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any pre-specified hypotheses
METHODS		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(a) Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case-control study—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants (b) Cohort study—For matched studies, give matching criteria and number of exposed and unexposed Case-control study—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential

Manuscript Section	Item Number	Recommendations
		confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) Cohort study—If applicable, explain how loss to follow-up was addressed Case-control study—If applicable, explain how matching of cases and controls was addressed Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses
RESULTS		
Participants	13*	(a) Report the numbers of individuals at each stage of the study—e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) Cohort study—Summarize follow-up time (e.g., average and total amount)
Outcome data	15*	Cohort study—Report numbers of outcome events or summary measures over time

Manuscript Section	Item Number	Recommendations
		Case-control study—Report numbers in each exposure category, or summary measures of exposure Cross-sectional study—Report numbers of outcome events or summary measures
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses
DISCUSSION		
Key results	18	Summarize key results with reference to study objectives
Limitations	19	a) Discuss limitations of the study, taking into account sources of potential bias or imprecision. b) Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
OTHER INFORMATION		
Funding Acknowledgement	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based
*Give such information separately for cases and controls in case-control studies, and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.		

Appendix 11: CONSORT Statement

Investigators and editors developed the CONSORT (CONsolidated Standards of Reporting Trials) Statement to help authors improve reporting of two-parallel design Randomised Control Trials by using a checklist. The most up-to-date revision of the CONSORT Statement is CONSORT 2010, which is shown below. The checklist items pertain to the content of the Title, Abstract, Introduction, Methods, Results, Discussion, and Other information. The checklist includes the 25 items selected because empirical evidence indicates that not reporting the information is associated with biased estimates of treatment effect, or because the information is essential to judge the reliability or relevance of the findings. The CONSORT group has developed additional guidance for multiple types of trials including cluster randomized trials and non-inferiority trials. To download these documents and get more information on the CONSORT group go to www.consort-statement.org.

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	
Introduction			
Background & Objectives	2a	Scientific background and explanation of rationale	
	2b	Specific objectives or hypotheses	
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	
Participants	4a	Eligibility criteria for participants	
	4b	Settings and locations where the data were collected	
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	
	6b	Any changes to trial outcomes after the trial commenced, with reasons	
Sample size	7a	How sample size was determined	
	7b	When applicable, explanation of any interim analyses and stopping guidelines	

Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	
	11b	If relevant, description of the similarity of interventions	
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	
	13b	For each group, losses and exclusions after randomisation, together with reasons	
Recruitment	14a	Dates defining the periods of recruitment and follow-up	
	14b	Why the trial ended or was stopped	
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	

Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	
Other information			
Registration	23	Registration number and name of trial registry	
Protocol	24	Where the full trial protocol can be accessed, if available	
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	

Appendix 12: List of common errors

A. General research and writing practices

A1. Insufficient knowledge of the literature
A2. Not referencing statements
A3. Weak citations
  A3a. Citing a secondary source
  A3b. Presenting conclusions rather than data from references
  A3c. Arguing from authority
A4. Endnotes not in standard style
  A4a. Arguing from authority
A5. Not using standard draft manuscript form
A6. Repeating information
A7. Labelling a scientific document as 'final'
A8. Characterizing an observation as 'the first'
A9. Errors in reasoning
  A9a. Casual assertion of causality
  A9b. Assuming association is causality
  A9c. Assuming reported behavior reflects actual behavior
  A9d. Confusing imperfect recall with recall bias
  A9e. Confusing absence of recognition with absence
  A9f. Asserting seasonality with a single year of data
  A9g. Drawing conclusions using confirmation bias

B. Content of quantitative papers

B1. Improper focus or format of title and abstract
B2. Confusing the role of Introduction, Methods, Results, and Discussion
B3. Not writing the methods section in chronological order
B4. Not emphasizing steps taken to protect human subjects
B5. Listing interpretations, but not defending one in the discussion
B6. Not fully explaining limitations
B7. Writing generic recommendations
B8. Presenting new data in the discussion
B9. Reporting the number of enrolled subjects in the Methods
B10. Specifying the contents of a questionnaire
B11. Naïve theories of change
  B11a. Recommending a massive increase in funding
  B11b. Ignoring incentives and barriers
  B11c. Assuming weak states can implement
B12. An insufficiently focused introduction
B13. Failure to clarify key sample size assumptions
B14. A high level outline that is not high level
B15. Specifying software used for routine data analysis
B16. Presenting rationale in the last sentence of the introduction

C. Mechanics of writing

C1. Using non-standard abbreviations
C2. Using non-standard spaces
C3. Improper spelling
C4. Capitalization problems
  C4a. USING ALL CAPITAL LETTERS
  C4b. Capitalizing non-proper nouns
C5. Failure to spell out a numeral <10
C6. Starting a sentence with a numeral
C7. Not indenting paragraphs
C8. Not aligning text to the left
C9. Problems with parentheses
C10. Not recognizing when an abbreviation has become a name
C11. Misplaced commas in large numbers
C12. Varying fonts within the narrative
C13. Using bulleted lists rather than sentences
C14. Uninformative document names

D. Grammatical structures and stylistic strategies

D1. Using present rather than past tense
D2. Failure to use definite and indefinite articles
D3. Excessive use of passive voice
D4. Improper use of 'we'
D5. Writing from a psychological perspective
D6. Using sub-headings in the discussion section
D7. Misplaced modifiers

E. Achieving clarity and conciseness

E1. Labeling rather than explaining
E2. Using weak opening phrases for sentences
E3. Using adjectives and qualifiers
E4. Over using studies or authors as sentence subjects
E5. Using non-descriptive numeric or alphabetical labels
E6. Using respectively
E7. Using the word etcetera
E8. Using Bangla as an English word
E9. Using local words, expressions or monetary figures
E10. Using the term 'developing country'
E11. Using the term 'socio-economic status' as a synonym for wealth
E12. Using technical terms in their non-technical sense
  E12a. Using the term 'random' in its non-technical sense
  E12b. Using the term 'reliable' in its non-technical sense
  E12c. Using the term 'significant' in its non-technical sense
  E12d. Using the term 'valid; in its non-technical sense
  E12e. Using the term 'incidence' incorrectly
E13. Using the verb "documented"
E14. Framing an argument in terms of need
E15. Using the term 'illiterate' as a synonym for 'no formal education'
E16. Using challenging as a synonym for difficult
E17. Describing a laboratory test result as positive
E18. Using increase or decrease in the absence of a time trend

F. Recording scientific data

F1. Stating results in statistical terms rather than on study question
  F1a. Framing narrative results around p-values
F2. Not presenting the core data
F3. Using too many decimal places
F4. Using too few decimal places
F5. Using incomplete headings for tables and figures
F6. Imbalance between table and narrative presentation of the results
F7. Pointing too explicitly to tables and figures
F8. Using inappropriate figures
F9. Using the wrong symbol to designate degree
F10. Using non-standard footnote symbols in tables
F11. Comparing to a varying baseline
F12. Generic data tables that lack a clear message
F13. Table layout that impairs comparisons
F14. Maps with irrelevant details
F15. Numbering tables or figures out of sequence

G. Approaching publication

G1. Failure to respond to reviewers' comments
G2. Incomplete response to external reviews
G3. Invalid authorship line
G4. Missing acknowledgement section
G5. Choosing an inappropriate journal
G6. Not following a specific journal's details of style
G7. Not using a checklist to review your paper before submission
G8. Exceeding the journal word limit
G9. Asking your senior author to recommend reviewers
G10. Responding to journal reviewers using the first person singular
G11. Retaining comments in subsequent drafts
G12. Not finding a description of the error code

G13. Requesting an unprofessionally short turnaround time
G14. Sending black forms for co-authors to complete
G15. Not providing co-authors a copy of the submitted manuscript
G16. Not keeping co-authors informed of journal discussions

H. Slide presentations

H1. Bullets on the wall
H2. Chart junk
H3. Copying a manuscript figure instead of developing a custom figure
H4. Photos with an unnatural aspect ratio
H5. Too many photographs on a single slide
H6. Field workers as the dominant subject of photographs
H7. Using bullets without hanging indents
H8. Using a pie chart
H9. Using vertical bars when horizontal bars would better communicate
H10. Including a final “Thank you” slide
H11. Using sentences for bullet points
H12. Too much space between bullets
H13. Failure to separate ideas in a multi-lined title
H14. Using 3 dimensional chart features as decorations

Appendix 13: Concept note example

Temporal Variability of Chlorine Demand in Dhaka, Bangladesh

By Fred Goddard

Study Question

What is the temporal variability in chlorine consumption by inorganic and organic materials in water, both daily and seasonally, in the piped water supply system of Dhaka, Bangladesh? How does this affect the type of chlorination injection systems required to continuously provide water with a minimum chlorine residual of 0.2 mg/l, congruent with WHO standards, for safe drinking water?

Objectives

The goal of this study is to better understand how the effectiveness of chlorination methods at the point of distribution, point of collection or point of use are affected by temporal variability in chlorine demand (chlorine that has been added to water and is consumed by organic and inorganic matter in the water) by:

1. Assessing the daily patterns and variability of chlorine demand.
2. Assessing the impact of rainfall events and temperature on chlorine demand by compiling chlorine demand data in Dhaka's three weather seasons: summer, monsoon and winter.
3. Generate hypotheses for major contributors to elevated chlorine demand, such as power outages or increased water residence time, that could help manage spikes in chlorine demand across water points.

Rationale

Lack of access to safe drinking water is estimated to cause 23% of deaths by diarrhoeal diseases amongst children under the age of five in South Asia (C Boschi-Pinto 2009). Dhaka is one of the most densely populated cities in South Asia, with over 50,000 people per square mile, and approximately one-third of its residents live in slums (G Angeles 2009). A 2010 survey conducted by the Lotus Water team at the International Centre for Diarrhoeal Disease Research, Bangladesh found that 80% of randomly selected samples of 127 slum water points in Dhaka were contaminated with *E.coli*. The primary reasons for unsafe water in Dhaka, which is mostly pumped from central groundwater pumping stations, are considered to be the leaky, intermittently pressurized water distribution systems, a common issue not only in Dhaka but throughout Asia with over half of its water supply shown to be intermittent (van den Berg 2011). In addition, high temperatures and severe weather events, particularly during the monsoon season, typically affect the quality of the drinking water negatively (Mirza 2007). The challenges faced in Dhaka due to the nature of the water supply infrastructure, the limited capacity and resources available to the municipality, as well as the difficult weather conditions are challenges encountered by many other urban areas in South Asia (UN 1987).

Chlorination is widely considered to be one of the more cost effective water disinfection methods to provide water that is safe for consumption (Water Quality & Health Council 2003). In Dhaka, it is being implemented at the point of distribution via chlorine injection pumps as well as at the point of use with chlorine tablets. In addition, the Lotus Water team (www.lotuswater.org) is

piloting an automatic chlorination device that disinfects the water at the point of collection, for example at shared handpumps and shared water points. Regardless at what stage the water is disinfected, it is desirable to not only have sufficient amount of chlorine in the water to kill pathogens but to also have a residual amount of chlorine, specified as a minimum of 0.2mg/L by the World Health Organization, to ensure the water is safe in storage and stop it from becoming re-contaminated. The total chlorine residual in a sample of water is influenced by the dose of chlorine added and the chlorine demand of the water, which is the chlorine that reacts first with inorganic and organic materials in the water and is thus not available for disinfection (CDC 2014).

However, it is necessary to dose accurately and not significantly exceed this minimum, because the taste and odor of chlorinated water becomes unacceptable at high dosing, particularly in settings where the populations are not regularly exposed to chlorinated drinking water (Flanagan 2013). We recently conducted a pilot study in Dhaka to determine a threshold of chlorine concentration at which the water becomes unacceptable to drink to local communities. Our preliminary data suggests that this threshold lies between at 0.8mg/L and 1mg/L. Subsequently, chlorine demand presents a particular challenge to this context, because there is a fine line in dosing enough chlorine to exceed the minimum of 0.2mg/l but stay below the threshold of 0.8-1mg/L. As a result, to be able to ensure that an adequate and acceptable dose of chlorine is added, so that the water is continuously fully disinfected and safe in storage as well as being of acceptable taste and odor, it is important to understand the nature of chlorine demand in Dhaka's water supply system.

In August 2014 we conducted a study pilot to compare the spatial distribution of chlorine demand between 18 water samples collected in four different slums in central Dhaka. This pilot allowed us to develop a method to best collect water samples in Dhaka and detect chlorine demand in the sampled raw water. To find the chlorine demand, it is the chlorine residual 30 minutes after the manual addition of chlorine (WHO 1996) in each water sample that is measured, which is the remaining chlorine that is available for disinfection in storage. The chlorine residual from identical doses of chlorine between chlorine demand free water, for example distilled water, is compared to a raw water sample that has been collected from at a water point in Dhaka. This is further outlined in the Data Collection section.

Chlorine demand can vary over time in Dhaka because of water use patterns and inconsistent pumping regimes caused by electricity cuts, causing fluctuating pressures and residence times in the system. Low pressure and high residence times are conducive to an increased infiltration of organic and inorganic matter in to the leaky piped water system, leading to reduced water quality and hence increased chlorine demand. In addition, weather conditions that vary daily but especially seasonally, such as temperature and rainfall events, have an impact on water quality in South Asia (Abdul Hussain Shar 2008) and Bangladesh (MOEF 2001). Microbiological quality of the water is negatively impacted by the addition of organic and inorganic matter, and changes in water quality affect chlorine demand.

Finally, it is not only important to gain a better understanding of temporal chlorine demand patterns, but also to generate hypotheses for the primary causes of spikes in chlorine demand, to be able to better predict elevations in chlorine demand without having to frequently take measurements. Chlorine demand is not only affected by differences in water quality caused by local conditions, but also by changes in water quality caused by low pressure and high residence times commonly found in an intermittent water supply system (E Kumpel 2013). This study will

help gain a better understanding of the causes of chlorine demand, as well as its temporal variability, which will guide future decisions on what chlorine injection technologies are appropriate and effective.

Outcomes & Exposures

Primary Outcome

Total chlorine demand in water samples collected in Dhaka is the primary outcome variable measured by this study. It is a continuous variable that will be measured with our previously developed chlorine demand detection method, which will be described further in the data collection section.

Secondary Outcome

Our hypotheses for spikes in chlorine demand will be tested and additional hypotheses will be generated through qualitative research outlined in the Study Design section.

Primary Exposures

The primary exposures that can influence different water quality parameters, which in turn affect chlorine consumption, are external events such as pump outages due to power cuts, high residence times due to low water demand or rain events during the monsoon season. Other exposures have so far not been considered for this study, but will be identified during the qualitative research component described under Objective 3 in the Study Design section. Exposures will not be measured but rather identified by examining existing records for the relevant exposures and comparing those to chlorine demand measurements taken in the field. For example, DWASA (Dhaka Water and Sewerage Authority) have records on pump outages and weather data can be pulled from weather reports.

Secondary Exposures

During our study pilot of 18 water points in Dhaka, we aimed to confirm what we hypothesized to be the primary drivers for chlorine consumption in Dhaka's piped water supply. We tested for iron, manganese, turbidity and flowrate. Our results showed that there was no statistically significant correlation between any of these exposures and levels of chlorine demand. As a result of our findings from the study pilot, this study is not designed to conduct further work to identify the exposures in the water quality for chlorine consumption in Dhaka's water. Approximately 30% of the resources for the study pilot were used to measure these exposures and it is not considered to be the most effective and relevant use of funds for the study in temporal variability

Study Design

To fulfill objectives 1-3 the most appropriate approach identified is a combination of two study designs. A non-experimental exposed cohort study will fulfill objectives 1 and 2 by taking chlorine demand measurements at water points and analyzing them for variability. We can assume that all water points are exposed to some organic or inorganic matter, leading to chlorine demand. To fulfill objective 3 a case review study design will be employed. This will be accomplished by using chorine demand data that has been gathered to fulfill objectives 1 and 2 and comparing that to hypotheses generated for exposures found in the qualitative research component of this study. This qualitative research will be conducted retrospectively by questioning water point users on the water quality as well as accessing publicly available data on hypothesized causality for spikes in chlorine demand.

Objective 1: To fulfill this objective, water samples shall be collected two times daily with consideration to DWASA's water use data. For example, the first sample shall be collected after the longest period of low water demand, typically early in the morning. The second sample shall be collected after the longest period of peak demand, typically in the late afternoon/early evening. Should water not be available at the time of collection, the water samples shall be collected as soon as steady-state conditions are reached once the pumps come back on line.

Objective 2: There are three seasons in Bangladesh. A hot and humid summer from March to June, a milder and wet monsoon season from June to October and a cool and dry winter from October to March. Water samples as outlined in Objective 1 shall be collected during all three seasons. The summer data shall be collected during the first and last week in April, the monsoon data in the first and last week in July and the winter data in the first and last week in January. The data collected in each of these months for each water point and for each time period (morning and afternoon/evening) shall be compared to weather data for each day that samples have been collected.

Objective 3: To fulfill the final objective, the quantitative data collected shall be compared to qualitative data collected retrospectively on reasons outlined for poor water quality by local users as well as the previously hypothesized causality (high residence times, pump system failures and monsoon rains). This qualitative data will be collected using two methods. During each water sample collection compound members will be asked about any changes in water quality since the previous collection using a questionnaire. For example, if a data collector visits a water point on a Tuesday morning for sample collection, he/she will inquire to water point users about any notable changes in water quality or other events relating to the water (e.g. no water available, low pressure) since Monday evening's visit. In addition, data will be gathered from available local information on our hypothesized causality for chlorine demand. For example, DWASA reports on pump failures, external weather reports on heavy rainfall events and correspondence with the DWASA operators responsible for pump stations shall be compared to the relevant time periods of chlorine demand data collected.

Analysis

Objective 1: The primary analysis will begin by finding the ranges in chlorine demand for every water point on a daily basis in a given season. It is the distribution of these ranges we are particularly interested in, because it is distribution rather than the absolute values of chlorine demand that will present the greatest challenge to effective continuous chlorination. If the absolute value of chlorine demand is comparitatively high but there is no significant differnce over time, chlorine injection technologies can account for this by continuously adding a higher dose of chlorine, providing the remains acceptable to drink for taste and odor. If there is high variability, current technologies might not be able to continuously ensure a chlorine residual in the water without a difficult to implement complex treatment algorithm. By finding the ranges in chlorine demand we can understand to what extent we might be able to add a sufficient amount of chlorine to the water. To investigate the ranges for every water point, the maximum value for chlorine demand is subtracted by the minimum value for each day.

Objective 2: For objective 2 we would primarly compare chlorine demand measurements immediately before and after major weather events, whether these spread across several days or are just on a single day. Chlorine demand data will be collected over two different sets of days for each season, during the first and last week in January, April, and July respectively. In

addition, we will compare the distribution of ranges between the three seasons to paint a more general picture of differences arising from overall varying conditions.

Objective 3: For the final objective, the measured data for each water point will be analyzed in comparison to data collected through our qualitative research on perceived causality for low water quality by water point users as well as major external events (such as weather or system failures) to generate hypotheses for spikes in chlorine demand. The questionnaire used for interviews of water point users shall be analyzed using a combination of a priori (such as "low pressure" and "no water available") codes and emergent codes. Answers to this questionnaire shall be coded by the person running the chlorine demand study with extensive knowledge on causality for low water quality and how this relates to chlorine demand, since some of the reasoning for changes in water quality provided by the users might not be reasons supported by scientific evidence.

Study Sample

Target Population
Urban water points in low and middle-income countries used for household purposes, such as drinking and cooking, supplied by a piped distribution system.

Study Population
Shared water points in the slums of Dhaka, Bangladesh supplied by the local municipal piped distribution system.

Sampling Method
The sampling for this study shall be conducted with a multi-stage sampling method, where a set of water points shall be chosen from different clusters for purposive sampling. Each cluster represents a different set of water points that are supplied by the same central municipal groundwater pumping station. The chosen clusters must have daily water supply data for their respective pumping stations available. Water points in each cluster shall be chosen systematically but must meet the criteria outlined below:

- Shared water point
- Connected to the DWASA piped distribution system
- Distribution system shall be supplied by one of DWASA's deep groundwater tube wells
- Water point shall be accessible at all times during the day, from early in the morning to late in the evening
- Owner of water point must agree to two four day period of sample collection during the first and last week in January, April and July respectively.

Sample Size
The sample size was determined with a focus on gathering enough samples to fulfill objective 1. However, this sample size will have to be big enough to allow for changes in weather conditions in a given season to fulfill objective 2. Since objective 3 is a hypothesis generating rather than a hypothesis testing exercise, the sample size will not be influenced by this objective. The primary outcome variable measured to fulfill objectives 1 and 2 is the level of chlorine demand in samples drawn from shared water points connected to the municipal piped water system in

Dhaka. The sample size will have to allow for an analysis that shows if there are differences in chlorine demand in a given day and what those differences are. We hypothesize that samples collected in the morning before daily water use has established itself will have higher chlorine demand than samples in the afternoon and evening during peak water use. So the underlying question is: **How many water points do we need to access twice a day to show the anticipated intra daily differences in chlorine demand?**

No published findings were identified that show variations in chlorine demand in water distribution systems comparable to Dhaka. The difference in the means and standard deviations for the morning versus evening sample utilized for this sample size calculation is an estimate formulated from our previous experience with water quality research in Dhaka and our chlorine demand pilot study. Sample 1 outlined in table 1 is the morning sample and sample 2 the evening sample.

Table 1 - Iterations of sample size calculations

Sample 1 mean	Sample 1 Standard Deviation	Sample 2 mean	Sample 2 Standard Deviation	Confidence Interval	Power	Sample Ratio	Design Effect	Sample size*
0.8	0.25	0.4	0.1	95%	80%	1:1	1.8	**16**
0.8	0.25	0.4	0.1	95%	90%	1:1	1.8	**18**
0.8	0.25	0.5	0.15	95%	90%	1:1	1.8	**36**
0.8	0.35	0.5	0.25	95%	80%	1:1	1.8	**60**
0.8	0.35	0.5	0.25	95%	90%	1:1	1.8	**78**

* Sample size is given as 'Total Sample Size', so a sample size of 16 would equal eight water points where two water samples are collected daily

The first two iterations (even at a non-standard power of 0.9) showed a smaller sample size than anticipated because of the high difference between morning and evening samples (0.4ppm) and low standard deviations that were assumed. The next iterations were performed with a smaller difference between the two sample means at 0.3ppm. The two sample means, 0.5ppm and 0.8ppm are in line with our chlorine demand pilot where we found a mean of 0.64ppm chlorine demand for water points that were tested at various times during the day. The standard deviations where chosen to allow for a distribution with chlorine demand values closer to zero as well as values above 1ppm. As comparison, the standard deviation during our 18 water point study pilot was 0.27ppm, in line with our standard deviations for the final two sample calculations. A design effect of 1.8 was chosen to account for clustering – water points are divided in to clusters, where each cluster represents a set of water points that is supplied by the same groundwater pumping station and the same distribution network. Power of 0.9 is not considered to be necessary for this type of study, so the penultimate iteration was chosen to determine the sample size for this study.

This leads to a sample size of 60, which amounts to 30 water points. Using eight water points across four clusters will fulfill this sample size and total 32 water points for sample collection. Sample collectors will rotate the order in which water points are accessed for sample collection, i.e. by accessing water points in order 12345678 on day one, 23456781 on day two, 34567812 on day three and so on. This will require eight days of sampling – four days at the beginning of each month and four days at the end of each month - to ensure samples have been collected in all orders.

Data Collection

The primary measurements required for this study are chlorine consumption in water samples collected at Dhaka's water points.

Strategy

Water samples will have to be collected in the field and be processed in the lab the same day. Eight 0.5L water samples will fit in to a standard cooler (water can be stored at 4°C for up to 24 hours (WHO 1996)). One cooler can be carried through a cluster by the data collector to collect all samples and can be transported back to the office via local public transportation. Water shall be sampled twice a day at each water point. The morning sampling for each cluster shall begin two hours before domestic water demand establishes itself, for example from 5-7am, since our previous work on collecting water samples in Dhaka suggested a maximum of 15 minutes for each sample, which will total to two hours of sampling. The sampling at all four clusters shall begin at a time relevant to their respective water supply data. Each cluster will require a different data collector, as much of the sampling will be carried out simultaneously. After the first sample collection the sample collector shall fill out the brief questionnaire outlined below:

- Compound Name
- Location
- Site Nr.
- DWASA tubewell
- Connection legal or illegal
- Type of connection (Flex. Pipe, Direct to main line, Tank connected to main line)

Collection at the same water points shall be repeated in the late afternoon/evening for all 32 water points across the four clusters during two hours of peak demand. As outlined previously, sample collectors shall rotate the order in which water points are accessed for sample collection on a daily basis. After every sample collection the sample collector shall fill out a questionnaire prompting the users of the tested water point on any issues with the water since the last visit:

- Have you experienced anything different with your water, such as taste, odor, color or flowrate, since our last visit on [day] at [time]?
- If yes, can you please describe what these differences were?
- Where do you think these differences may come from?

In addition, each water sample shall be labeled as outlined below:

- Name of Collector
- Site Nr.
- Date
- Time

After the eight samples have been collected they shall be delivered to the lab for testing immediately.

Chlorine Detection

Our previous work on chlorine demand in Dhaka has allowed us to develop a method to best detect chlorine demand in Dhaka's raw water in the laboratory. To find the total chlorine consumed by organic and inorganic matter, it is the chlorine residual after manual addition of chlorine in each water sample we are analyzing, which is the remaining chlorine that is available for disinfection in storage. The chlorine residual from identical doses of chlorine between

chlorine demand free water, for example distilled water, is compared to a raw water sample that has been collected from a water point in Dhaka. For example, if you introduce 1.5mg of sodium hypochlorite in the form of locally available liquid bleach in to a 0.5L sample of distilled water, you would expect to have a chlorine concentration of 3ppm (or 3mg/L) in the chlorine demand free water. However, if you introduce the same amount of sodium hypochlorite in to raw water, you might get a chlorine concentration of 2ppm. This would mean that the inorganic and organic matter present in the water have consumed 1ppm of chlorine, which in turn means the raw water has a chlorine demand of 1ppm. A more detailed protocol for this sample collection in the field and measurements taken in the lab is outlined in Appendix A.

Human Subjects

Institutional Review Boards will review the protocol for this study for human subjects consideration at Stanford University and at the International Centre for Diarrhoeal Disease Research, Bangladesh. While this study does not influence its human subjects directly through an intervention, data collectors will have to enter slum compounds and access water points shared by households twice a day during the eight days allocated for each season, thus impacting people's daily routines and needs. In addition, data collectors will be briefly interviewing a compound member during each visit as outlined in the Data Collection section.

Informed consent will be collected from water point owners, typically compound landlords (compounds in Dhaka are typically shared between 5 and 50 households in Dhaka). Water point owners must consent to two sample collections over four day periods during the first and last week in January, April and July respectively. Informed consent will also be collected from compound members before the short interview at sample collection. Water point owners and interviewees will not be compensated.

Collaboration

This study will be conducted in collaboration between Stanford University and the International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B). It will build on a multi-year relationship between the two institutions, while leveraging the resources of the Lotus Water project team. Lotus Water was established four years ago as a partnership between ICDDR,B and Stanford. The project is developing community based water disinfection devices for shared water points to provide safe water to Dhaka's slums. We have faculty and students based at Stanford and an eight-person strong field team based in Dhaka. This study will build on the capacity and experience of our team in water quality and chlorination research in Dhaka's slums. Our field assistants in Dhaka were trained on data collection during our chlorine demand study pilot and we have two field research assistants to help manage our field team. Our principal investigators, research associate and graduate students frequently travel to Dhaka for other project work.

Timeline & Budget

The anticipated duration for this study is 23 months and it will cost $66,500. A timeline is outlined in Appendix B and the budget in Appendix C. The budget and timeline were developed taking in to consideration that our team has already built the capacity at ICDDR,B in Dhaka. This study will run in parallel to our other project work and as a result the gaps in data collection that are shown on the timeline are not a concern for the graduate students or the field assistants that will help with data collection.

Limitations

There are several limitations and weaknesses to the proposed study design that are worth noting:

- **Purposive sampling**: From the sample size generated for this study, we will not have the power to draw a purely representative sample for temporal variability in chlorine demand in Dhaka, but we will be able to gain an understanding for the extent of a challenge temporal variability in chlorine demand may present to chlorine injection systems.
- **Clustering**: A multi-stage sampling method was chosen for this study with four clusters that each represent a set of water points supplied by the same groundwater pumping station and distribution system. Even with the design effect added to the sample size calculations, there will be clustering in the chlorine demand data, because the water points are being supplied from the same source and are supplied by the same piped system. However, the clustering of water points is necessary to reduce travel time and allow sample collectors to gather data in the two-hour window described previously.
- **Causality:** The qualitative research for this study is designed as a hypotheses generating exercise. As a result, it will not be possible to conclude causal relationships between the hypotheses generated and spikes in chlorine demand.

References

Hussain Shar, Yasmeen Faiz Kazi and Irshad Hussain Soomro. "Impact of Seasonal Variation on Bacteriological Quality of Drinking Water." *Bangladesh Journal of Microbiology*, 2008: 69-72.

C Boschi-Pinto, L Velebit, K Shibuya. "Estimating child mortality due to diarrhoea in developing countries." *Bulletin of the World Health Organization* 86 (2009): 710-717.

CDC. *Centers for Disease Control and Prevention: The Safe Water System.* 2014. http://www.cdc.gov/safewater/chlorine-residual-testing.html (accessed January 2015).

Dean, A., K. Sullivan, and M Soe. *OpenEpi: Open Source Epidemiologic Statistics for Public Health.* 9 22, 2014. www.OpenEpi.com (accessed 2 23, 2015).

E Kumpel, K Nelson. "Comparing microbial water quality in an intermittent and continuous piped water supply ." *Water Resarch* 37 (2013): 5176-5188.

Flanagan, S. Meng, X. and Zheng Y. "Increasing acceptance of chlorination for household water treatment: observations from Bangladesh." *Waterlines* 32, no. 2 (2013): 125-134.

G Angeles, P Lance, J Bardon-O'Fallon, N Islam, A Mahbub, N I Nazem. "The 2005 census and mapping of slums in Bangladesh: design, select results and application." *International Journal of Health Geographics* 32 (2009): 1-19.

Mirza, M. *Climate Change, Adaptation and Adaptive Governance in Water Sector in South Asia.* Toronto: Adaptation and Impacts Research Division - Environment Canada, 2007, 1-19.

MOEF. *Bangladesh: The State of the Environment.* Dhaka: Ministry of Environment and Forest Bangladesh, 2001.

UN. *Report of the World Commission on Environment and Development: Our Common Future.* New York City: WCED, 1987.

van den Berg, C. and Danilenko, A. *The IBENET Water Supply and Sanitation Performance Blue Book.* Washington, DC: World Bank, 2011.

Water Quality & Health Council. "Drinking Water Chlorination: A Review of Disinfection Practices and Issues." 2003.

WHO. *Guidelines for Drinking-Water Quality.* Geneva: World Health Organization, 1996.

WHO. *Water Quality Monitoring - A Practical Guide to the Design and Implementation of Freshwater Quality Studies and Monitoring Programmes.* Nairobi/Geneva: UNEP/WHO, 1996.

Timeline

		Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	
	Task																								**Personnel**
1	Protocol Development*	■	■																						PI, RA, GS
2	Stanford University approval			■																					
3	ICDDR,B approval			■																					
7	Survey water points that meet criteria				■																				GS, FRA, FA
8	Winter sample collection and measurements					■																			GS, FRA, FA
9	Preliminary winter quantative data clean up and analysis					■																			GS
10	Winter qualitative data analysis (Objective 3)					■	■																		FRA, GS
11	Summer sample collection and measurements								■																GS, FRA, FA
12	Preliminary summer quantative data clean up and analysis								■																GS
13	Summer qualitative data analysis (Objective 3)								■	■															FRA, GS
14	Monsoon sample collection and measurements											■													GS, FRA, FA
15	Preliminary monsoon quantative data clean up and analysis											■													GS
16	Monsoon qualitative data analysis (Objective 3)											■	■												FRA, GS
17	Full quantative data analysis (Objective 1 & 2)												■	■											GS
18	Manuscript - First draft													■	■	■									GS
19	Manuscript review by co-authors (first draft)																■	■							PI, RA
20	Manuscript - Second draft																		■						GS
21	Manuscript review by co-authors (second draft)																			■					PI, RA
22	Manuscript submission																				■				GS
23	Manuscript review																					■	■		
24	Manuscript revision and re-submission																							■	PI, RA, GS

* Assuming I have funding by September - Timeline would be shifted depending on when funding comes in (Data collection may begin in a different season)

Personnel Key:	Time (months)	% Time
PI = Principe Investigator	6	4
RA = Research Associate	6	8
GS = Graduate Student	15	31.7
FRA = Field Research Assistant	7	30
FA = Field Assistant	4	60

Total % time for GS during month of sample collection	80
Total % time for GS during time in Dhaka	40

Budget

Stanford University Budget

Personell	Quantity	Cost ($/month)	Time (Months)	Percent time (%)	Total cost
Research Associate	1	7500	6	8	3600
Graduate Student	1	6250	15	32	30000
				Total	**33600.00**

Travel	Quantity	Unit Cost	Total Cost	
International flights	1	2000	2000	Trip from Dec-Aug for all work in Dhaka.
Visa	1	160	160	
		Total (40%)	**864**	Study will only take on 40% of this cost, because that is how much time student will spend working on the project while in Dhaka (based on timeline)

Transport	Study Days	Cost ($/day)	Total Cost
Student	96	2.50	240
		Total	**240**

Lodging	Months	Cost ($/month)	Total Cost	
Dhaka student apartment	12	500	6000	
Student services fee	12	41.67	500	
		Total (0%)	**0**	(Lodging is included in graduate student stipend)

Per diem	Days	Cost ($/day)	Total Cost	
Daily expenses	225	15	3375	
		Total (0%)	**0**	(Per diem is included in graduate student stipend)

Overhead	Items charged to overhead	Cost	Overhead* %	Total cost
Stanford	Salaries, Travel, Transport	34704.00	60.5	20995.92
			Total	**20995.92**

*For federal funding sources

Stanford cost ($) **55700**

ICDDR,B Budget

Personell	Quantity	Cost ($/month)	Time (Months)	Percent time (%)	Total cost
Field Research Assisstant	1	461	7	30	968.10
Field Assisstant	3	225.00	4	60	1620.00
				Total	**2588**

Transport	Study Days	Cost ($/day)	Total Cost
FA & FRA	186	2	372
		Total	**372**

Equipment	Quantity	Unit cost ($)	Total Cost
Lamotte Colorimeter	1	400	400
DPD liquid reagents	4	6	24
Sample collection bottles	24	1.80	43.16
Glass testing bottles	24	2.04	48.86
Liquid bleach	3	2.50	7.50
		Total	**524**

Overhead	Cost	Overhead %	Total cost
Salaries, Transport	2960.10	30	888.03
		Total	**888**

ICDDRB cost ($) **4372**

Total Cost $60072

References

1. Bernard HR, Killworth P, Kronenfeld D, Sailer L. The problem of informant accuracy: The validity of retrospective data. Annual Review of Anthropology 1984;13:495-517.
2. Manun'Ebo M, Cousens S, Haggerty P, Kalengaie M, Ashworth A, Kirkwood B. Measuring hygiene practices: a comparison of questionnaires with direct observations in rural Zaire. Trop Med Int Health 1997;2:1015-21.
3. Curtis V, Cousens S, Mertens T, Traore E, Kanki B, Diallo I. Structured observations of hygiene behaviours in Burkina Faso: validity, variability, and utility. Bulletin of the World Health Organization 1993;71:23-32.
4. Halder AK, Tronchet C, Akhter S, Bhuiya A, Johnston R, Luby SP. Observed hand cleanliness and other measures of handwashing behavior in rural Bangladesh. BMC Public Health 2010;10:545.
5. Pedersen DM, Keithly S, Brady K. Effects of an observer on conformity to handwashing norm. Percept Mot Skills 1986;62:169-70.
6. Munger K, Harris SJ. Effects of an observer on handwashing in a public restroom. Perceptual and Motor Skills 1989;69:733-4.
7. Drankiewicz D, Dundes L. Handwashing among female college students. American journal of infection control 2003;31:67-71.
8. Ram PK, Halder AK, Granger SP, et al. Is structured observation a valid technique to measure handwashing behavior? Use of acceleration sensors embedded in soap to assess reactivity to structured observation. The American Journal of Tropical Medicine and Hygiene 2010;83:1070-6.
9. Nickerson RS. Confirmation bias: A ubiquitous phenomenon in many guises. Review of General Psychology 1998;2:175.
10. VanderWeele T, Staudt N. Causal diagrams for empirical legal research: a methodology for identifying causation, avoiding bias and interpreting results. Law, Probability & Risk 2011;10:329-54.
11. Robins JM. Data, design, and background knowledge in etiologic inference. Epidemiology 2001;12:313-20.
12. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. Epidemiology 1999:37-48.
13. Carpenter DP. The forging of bureaucratic autonomy: Reputations, networks, and policy innovation in executive agencies, 1862-1928: Princeton University Press; 2001.
14. Chavalarias D, Wallach JD, Li AH, Ioannidis JP. Evolution of Reporting P Values in the Biomedical Literature, 1990-2015. JAMA 2016;315:1141-8.
15. Ehrenberg ASC. Rudiments of numeracy. Journal of the Royal Statistical Society Series A (General) 1977:277-97.
16. Luby SP, Agboatwalla M, Bowen A, Kenah E, Sharker Y, Hoekstra RM. Difficulties in maintaining improved handwashing behavior, Karachi, Pakistan. The American Journal of Tropical Medicine and Hygiene 2009;81:140-5.
17. Schuemie M, Kors J. Jane: suggesting journals, finding experts. Bioinformatics 2008;24:727-8.
18. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. PLoS Medicine 2007;4:e297.