Volume 25 Number 3
September 2016

Medical Writing

Statistics

Also in this issue...

- Where have all the UK entry level pharmaceutical regulatory medical writing jobs gone?

EMWA EUROPEAN MEDICAL WRITERS ASSOCIATION

EUROPEAN MEDICAL WRITERS ASSOCIATION

Medical Writing is the official journal of the European Medical Writers Association (EMWA). It is a quarterly journal that publishes articles on topics relevant to professional medical writers. Members of EMWA receive *Medical Writing* as part of their membership. For more information, contact mew@emwa.org

Submissions:
For instructions to authors, go to the journal section of EMWA's website (www.emwa.org). All manuscripts should be submitted to mew@emwa.org.

Statistics for medical writers 1
Phil Leventhal

President's Message 2

EMWA News 3

Feature Articles

History of biostatistics 8
Rosser Matthews

The illusion of certainty and the certainty of illusion 12
Tom Lang

Never *P* alone 17
Tom Lang

A medical writer's guide to meta-analysis 22
Adam Jacobs

Study design made easy 26
Diogo Bruno

SAMPL guidelines 31
Tom Lang & Douglas Altman

How to interpret and report the results from multivariable analyses 37
Stephen Gillver & Neus Valveny

Biostatistics and medical writing: Synergy in preparing clinical trials documents 43
Scott Miller & Raquel Billiones

Best friends forever: A pattern of collaboration between medical writers and biostatisticians within the Russian CRO 46
Eugenia Radkova & Ivan Dobromyslov

Where have all the UK entry level pharmaceutical regulatory medical writing jobs gone? 50
Elsa Lewis & Jonathan Oliver

Regular Features

News from the EMA 55

Journal Watch 58

In the Bookstores 61
- Essential Statistics for the Pharmaceutical Sciences (Second Edition)

The Webscout 62
- The history of statistics

Good Writing Practice 63
- Syntactic dissonance and impeded immediate comprehension: Coordination nonparallelism

Medical Communications 65
- RMP public summary reloaded: Revision 2 of GVP module V

Getting Your Foot in the Door 69
- Successful internships: a company perspective

Lingua Franca and Beyond 73
- Statistics – are we doing it right?
- Refereeing: a humour-free occupation

Gained in Translation 75
- English to Italian translation: 17 biostatistical terms that we are using in Itanglish that could just be used in proper Italian!

Teaching Medical Writing 77
- 10 tips for organising a successful writing course

Profile 78
- An interview with Professor Peter Jüni on methodology and statistics in scientific manuscripts

Out on Our Own 80
- What's happening at the European level: The European Forum of Independent Professionals (EFIP) and why you should know about it
- The coworking revolution
- Medical translation – a dead-end job or a gateway to opportunity? Part 3

Volume 25 Number 3 | September 2016

Medical Writing

Statistics for Medical Writers

Phil Leventhal, Editor-in-Chief

Until recently, statistics was a subject that I avoided at all costs. I found it difficult to understand and boring, and for years I swore that I would never learn about or be interested in statistics. (By the way, I also swore that I would never become a medical writer.) Since becoming a medical writer just over 10 years ago, I have gradually warmed to statistics and now even find it interesting. However, I still think that statisticians are an alien species that speaks a different language. If you don't believe me, try to get a straight answer from a statistician!

I have come to realise that, to do our jobs well, we medical writers must have at least a basic understanding of statistics and must be able to communicate and collaborate effectively with biostatisticians. This issue of *Medical Writing* provides a wealth of information to help get you there. **J. Rosser Matthews** starts the issue off with an article on the history of biostatistics, which helps provide some context about how and why statistics is used in medicine. Then, in two articles, **Tom Lang** discusses how statistics can be used to mislead the reader and what can be done about it. He and **Douglas Altman** also provide us with an update of the SAMPL (Statistical Analyses and Methods in the Published Literature) guidelines for reporting statistics in medical journal articles. Elsewhere, EMWA's resident statistics expert, **Adam Jacobs**, offers us a guide on understanding and reporting meta-analyses, and *Medical Writing* Co-Editor **Stephen Gilliver** and colleague **Neus Valveny** provide a guide on understanding and reporting multivariable analyses. These articles are complemented by two others, one from **Scott Miller** and **Raquel Billiones** and the other from **Eugenia Radkova** and **Ivan Dobromyslov**, on collaborating and communicating with biostatisticians. In addition, the In the Bookstores, Webscout, Lingua Franca and Beyond, Gained in Translation, and Profile sections add to the wealth of information on statistics in this issue.

Speaking of sections, I would like to announce a new regular section, Getting Your Foot in the Door, led by Section Editor Raquel Billiones. Getting Your Foot in the Door will include articles on how to launch a career in medical writing, and it is part of an initiative that resulted in the first annual Internship Forum at the EMWA 2016 spring conference in Munch, which is described in this first instalment of the section.

To end, I think that we all could use a bit of humour given the Brexit madness, the Trump madness, and a variety of other disturbing current events. So, for a good laugh related to statistics and biostatisticians, take a look at the cartoon "Biostatistics vs. Lab Research" on YouTube (https://www.youtube.com/watch?v=PbODigCZqL8).

Phil

CONTACT

Phillip Leventhal
Editor-in-Chief
editor@emwa.org

President's Message

Dear EMWA Members,

By the time you receive this copy of the journal most of you will be returning from your summer break. I hope you all enjoyed your holiday and returned refreshed and ready for another busy year.

Your new Executive Committee has been working hard since the Spring conference and is already busy with plans for the Autumn conference in Brussels. By the time you read this we will have full details available on the EMWA website. In addition we are looking at improvements in the way we provide documentation for conferences and updating the e-mail system for Executive Committee members. Later in the year we will be sending out a survey to all members to ensure that we are providing the services that you want and need. Please look out for this and provide your views to us.

Over the summer there have been opportunities for learning via our webinar series with a joint webinar with AMWA on CORE Reference and how this can help Medical Writers prepare compliant CSRs in July, and one on how to solve formatting problems in Microsoft Word (something we have all suffered from at one time or another) in August. If you missed either of these they are both available in the Archive section of the website. I urge you to investigate this valuable resource.

EMWA are also involved in local meetings of medical communication specialists. Members presented at a MedComms Networking monthly Brunch Club held in Oxford in June and will be presenting at the First German Medizin-Kommunikations-Forum in Berlin in October.

Finally, for those of you who have been following the political situation here in the UK I would like to reassure you that the decision made by the UK voters in June to leave the European Union will make no difference to EMWA. The UK may have voted to leave the EU but we are still geographically and culturally part of Europe. EMWA is for all writers in Europe (including the UK and Switzerland) and also any writers from outside of Europe who wish to join us.

Best wishes

Alison
alison.rapley@gmail.com

Call for Companies

The 2nd Medical Writing Internship Forum will be held at our May 2017 Conference in Birmingham, UK. Please contact internship@emwa.org for more information.

EMWA News

Editorial

A lot was going on from April to June this year. Certainly the main event was our spring conference in Munich. Aside the established workshops, Freelance Business Forum, Expert Seminar Series, and updates on Special Interest Groups – it included an outstanding Symposium Day, the launch of the Internship Forum, and a poster exhibition.

Mainly, but not only, based on the very interesting symposium theme, EMWA was asked to present at the Brunch Club meeting of the MedComms Networking group (http://www.medcommsnetworking.com). Also, EMWA has been presented during talks at a careers fair at 'The Organisation for Professionals in Regulatory Affairs' (TOPRA) in London, at the Max-Planck Institute in Munich and EMWA attended the CTrials conference in Tel Aviv.

Beatrix Doerr
beatrix.doerr@gmail.com

Spreading the word – EMWA's presence at the CTrials conference

In a new departure, EMWA took a stand at the CTrials conference in Tel Aviv in April 2016. This conference is organised by The Israeli Association for the Advancement of the Biomedical Research Community and covers hot topics in the field of clinical trials. Two members of the EMWA Executive Committee, Barbara Grossman, a fluent Hebrew speaker, and Diarmuid De Faoite voluntarily manned the stand during the conference. Approximately 400 people attended the CTrials conference and more than half of them visited the EMWA stand to find out more about what EMWA has to offer. A targeted follow-up email was sent to all those who registered their interest.

EMWA was also invited to give a 20 minute talk to the conference attendees and Diarmuid De Faoite gave a well-received presentation on Important Documents in Clinical Research. Of course, EMWA already has members in Israel and we are indebted to Sharon Furman-Assaf and Miriam Aghassi-Ippen for their help in making this event such a success.

The EMWA Executive Committee will carefully assess the impact of this initiative with a view to further expanding the organisation's scope of actions.

Diarmuid De Faoite
emwaweb@gmail.com

SECTION EDITOR

Beatrix Dörr
beatrix.doerr@googlemail.com

Elsa Lewis from Lioness Writing Ltd reports on presenting at the TOPRA careers fair on 17 April 2016

More than 100 students and young professionals in Regulatory Affairs in the pharmaceutical industry attended the inaugural TOPRA (The Organisation for Professionals in Regulatory Affairs) careers fair called 'Regulatory Careers Live' at the Royal Pharmaceutical Society in central London.

Presenters included representatives from pharmaceutical companies, regulatory agencies, and contract research organisations. Elsa presented 'What colour is your paraglider' as an interactive introduction to Medical Writing within Regulatory Affairs and for the wider industry. Within this presentation, EMWA was introduced as an organisation for Medical Writers. During the networking sessions there was enthusiasm from participants to learn more about EMWA and careers in Medical Writing.

The second TOPRA career fair is planned for 2017.

Elsa Lewis
Lioness Writing Ltd
elsa.lewis@hotmail.co.uk

Munich: A report on EMWA's record-breaking conference

What a record-breaking conference it was! Some 419 participants enjoyed a short stay in picturesque Munich, the main city of Bavaria, Germany. The delegates were mostly drawn from Germany and the UK, but some also came from as far away as Argentina, Australia, China, India, South Korea, Singapore, Japan, the US, Lebanon and Israel! They chose what to attend from a total of 50 workshops – 34 at foundation and 16 at advanced level.

At the opening session, Beatrix Dörr, EMWA's PR Officer, gave a great insight into the region with a talk titled 'Servus Bavaria: The Land of Beer, Crazy Kings and Medical Writers'. She was followed by an invited speaker, Stefanie Weber from the Audi Accident Research Unit, who gave a stimulating talk on how it is possible to learn from road accidents by integrating technical, medical and psychological perspectives.

EMWA is always working hard to make the conference experience as rich as possible. New medical writers were particularly well served this year with two new features that look set to become staples at future conferences. The first-ever internship forum attracted over 50 participants and was a matching exercise par excellence. Medical writers seeking internships had the unique opportunity to present themselves to companies open to taking on internees. Allied to this, a new seminar by Philip Leventhal, the Editor-in-Chief of EMWA's *Medical Writing* journal, imparted many valuable tips in his talk, Getting Your Foot in the Door: How to Build Experience to get a First Medical Writing Job. There was also a stimulating poster session in the exhibition area during the duration of the conference.

For more experienced medical writers there was also an array of offerings to avail of. The six Expert Seminars presented as part of EMWA's second Expert Seminar Series (ESS) were suited to senior and experienced medical writers. International experts held lectures with either a panel or participant discussion or demonstration on topics including automated authoring systems, building medical writing teams in the Far East and India, and how transparency and disclosure initiatives will impact clinical document structures.

The Symposium Day, entitled *Scientific and Medical Communication Today* focused on the evolving field of medical communications, focusing on the importance of medical writers as medical communicators.

The Pharmacovigilance Special Interest Group (PVSIG) held its first session with presentations from Industry and Regulators on the latest aspects of Pharmacovigilance. The CORE Reference team also held an open session. Since the CORE Reference launched open access on 03 May 2016, resources are available at the dedicated website: www. core-reference.org. EMWA also launched the Regulatory Public Disclosure SIG (RPD SIG), as a natural follow-up to CORE Reference at the conference.

At the Annual Meeting we said goodbye to outgoing EMWA President Sam Hamilton who drove many initiatives in the course of her time on the Executive Committee. Alison Rapley is the new EMWA President, supported by Abe Shevack as Vice-President. Education Officer Barbara Grossman also stepped down from her role and will be sorely missed. Marian Hodges will step into Barbara's shoes on the EMWA Executive Committee. Slavka Baronikova was re-elected as Conference Director for another two years and we congratulate her on her success.

Of course, EMWA conferences are also about networking, meeting old friends and making new ones. Apart from the coffee and lunch breaks, the organised events are a great way for delegates to mingle. All of the social events were fully booked. The outdoor events such as the walking tour and bike ride were all a great success, despite the inclement weather. Over 150 people signed up for the Bavarian spring dinner and dance which showcased many elements of Bavarian culture.

Don't miss out on the next EMWA experience! The 43rd EMWA Conference in Brussels, Belgium, will be held from 3-5 November 2016 at The Sheraton Brussels Hotel.

Diarmuid De Faoite
emwaweb@gmail.com
Slavka Baronikova
slavka.baronikova@gmail.com

The 4th EMWA Symposium "Scientific and Medical Communication Today"

The 4th EMWA Symposium focused on the ever-changing field of medical communications and the importance of medical writers as medical communicators.

After the welcome from Symposium moderators and introductory polling questions to characterise the audience, Prof. Nico Pitrelli (Scuola Internazionale Superiore di Studi Avanzati – Sissa – Trieste, Italy) set the ground with a wide and thoughtful view on the evolving environment of scientific communications facing the challenges to the professional functions as well as working practices of scientific and medical writers. During the second engaging and inspirational presentation by Chris Colaço (Initiate Training & Development, Switzerland), focus was oriented on the importance of medical writers' reputation, branding, brand promise and value. Moving from paradigmatic examples of brand, Chris explained the principles and significance of building a medical writer's brand.

After the first coffee break, different perspectives on what scientific communications means have been discussed. Jan Geissler (European Patients' Academy on Therapeutic Innovation, EUPATI, Germany) presented the patients view: the need to see patients as centre of any healthcare service and communicate to them appropriately. As the new EU Clinical Trials Regulation (http:www.ema.europa.eu) requires that clinical trial communications will also include lay person, medical writers have to be able to prepare these documents and communicate scientific data to patients as well. An example of effective communication of medical data to patients and lay public was reported by Fabienne Huebener (inword.de, Germany) that narrated a story that emotively involved the audience and inspired writers on the difference between 'writing' and 'communicating' medicine. The morning was closed by the EMWA past-president Laurence Auffret (CINETIQUE Translations, UK) that highlighted the concept of effective translation. This cannot be ensured by the application of translation's standards but needs to be targeted to cultural environment.

The role of regulatory authorities and their initiatives on communication and transparency were presented by Juan Garcia Burgos (European Medicines Agency, UK), highlighting the importance and benefits linked with their effective development and use. Hartwig Buettner (Eli Lilly, Germany) shared industry's expectations and issues through some examples on the importance of high quality disclosures and their link with the status of drug development. Hartwig highlighted industry appreciation of medical writers as a key figure in communication about addressing unmet medical and patient's needs by clinical development. This session was closed by Chris Winchester (Oxford PharmaGenesis, UK) experience from medical communication agency point of view, highlighting that planning and high quality delivery are the constants that ensure successful collaboration with the medical communications agency for achieving high quality scientific communications.

Past, present and future trends for communicating scientific and clinical research were the natural conclusion of the day. Andrea Bucceri (Dove Press, UK) described new technologies and methods of communicating scientific data and facilitating access to information highlighting their crucial role in the present and future scenario. Jan Seal-Roberts (Adis, Springer Healthcare, UK) predicted the possible development of scientific articles and their management in the next 5, 10 and 20 years according with the evolving reading habits of healthcare professionals and technical evolution. The day had its natural conclusion with the presentation on extending the impact and reach of science publications by Martin Delahunty (Springer Nature Publishing Group, UK). Martin exacerbated the central role of scientific journals for future high-quality research disclosure in an environment where open access to content and data will extend the reach and impact of publishing beyond the traditional research communities to anyone who has an interest, need or wants to advance better medical practice and health outcomes.

Each session ended with a Q&A session where presenters answered questions from the audience.

Description of the contents presented are available also at https://www.gkm-therapieforschung.de/emwa16/. All presentations are available at EMWA website and some of the presenters will write an article on their presentation to be published in the 2016 December issue of EMWA's *Medical Writing* journal.

Andrea Rossi
molezzano.1@libero.it
Slavka Baronikova
slavka.baronikova@gmail.com

EMWA at the Max Planck Institute of Psychiatry

On behalf of EMWA Christopher Marshallsay and Beatrix Doerr joined invited speakers from a broad spectrum of areas to introduce the career option "medical writing" at the 2nd Career Day of the Max Planck Institute of Psychiatry held on May 11th, 2016. They presented an introduction to the medical writing profession which included the different type of medical writers, what medical writers do, the 'typical' medical writer profile, career progression in medical writing, and the pros and cons of the role. They also introduced the new EMWA Internship Forum and the advantages of EMWA membership. The audience, largely BSc, MSc and PhD students, posed many questions and – as so often – were not aware of the role. One excited attendee reported "that's me", obviously a budding medical writer, subsequently attended the open sessions at the EMWA Spring Conference and was thrilled to be able to talk with medical writers and to learn about the profession, the opportunities it offers, and how best to apply.

Christopher Marshallsay
christopher.marshallsay@grunenthal.com

RPD SIG – Launch meeting May 2016

The Regulatory Public Disclosure Special Interest Group (RPD SIG) launched at the EMWA Spring conference and held an introductory session on 12th May, alongside the launch of CORE Reference. The session was very well attended and was received enthusiastically by the audience.

Tracy Farrow, Senior Director of Medical Writing at PPD, introduced the session by taking the audience through some background to the current regulatory public disclosure environment and why it is important and of interest to medical writers. She described the objectives of the EMWA RDP SIG in providing a forum for the discussion and sharing of information, best practices, and ideas with EMWA members, and named the proposed advisory panel who will be supporting this important SIG by freely sharing their knowledge and expertise.

Dr Christopher Marshallsay, Head Medical Writing and Public Disclosure at Grünenthal, described the new RPD SIG website and the available resources that include a glossary of terms, a library of key references and background reading as well as a question and answer log. He finished off discussing the next steps and request that volunteers share ideas and experience. This will be a key component of RPD SIG.

The session was brought to a close with an opportunity for the audience to ask questions and pleasingly many of the conversations continued into the refreshment area after the session. The interest in the RPD SIG and the fluid nature of topic in general should generate an informative and interesting session at the Spring conference in 2017 when the RPD will have its first formal session as part of the Symposium Day.

Tracy Farrow
Tracy.Farrow@ppdi.com
Christopher Marshallsay
Christopher.Marshallsay@grunenthal.com

PV SIG – Meeting update May 2016

The pharmacovigilance special interest group (PV SIG) held its first session at the EMWA Spring conference. The session title was 'Are we ready for the patient's voice through social media in the benefit-risk assessment of drugs?' and despite being held late on Friday afternoon, was very well attended.

The participants were given the latest update on the MHRA's WebRADR initiative, which is collecting and collating adverse event data. The MHRA's Special Projects Officer for Vigilance and Risk Management of Medicines, Dr Alicia Ptaszynska-Neophytou, outlined the project and explained the problems involved in dealing with the huge amount of data available and the approaches that the MHRA are taking. Alicia described what had been learnt in the 21 months since WebRADR was launched, and the progress that the WebRADR consortium are making, along with their plans moving forwards.

Dr Ulrich Vogel, Head of Strategic Data Analysis and Global Pharmacovigilance at Boehringer Ingelheim then described the collection of data from patient support programmes (PSPs) – an alternative source of Pharmacovigilance data that many medical writers are less familiar with, but that has gained importance in the periodic safety update report (PSUR) assessment of some products. Ulrich explained what kind of information writers could expect from a PSP database, and how these data could be analysed and described. Ulrich explained his company's approach and the challenges faced when dealing with this kind of information.

To round of a very interesting and informative session, there was an excellent discussion, and both presenters took a variety of questions, covering topics from how to address quality versus quantity, audit ramifications and why 'death' had been classified as a 'non serious event'! The session was enjoyed by all and we are looking forward to the next one in Spring 2017.

Lisa Chamberlain James
lisa@trilogywriting.com

CORE Reference news, June 2016

Sam Hamilton wrote a Guest Blog at BioMed Central's invitation in late May 2016: 'Safeguarding the privacy of clinical trial patients': http://blogs.biomedcentral.com/on-medicine/2016/05/27/safeguarding-privacy-clinical-trial-patients/.

This clear and nontechnical article shows patients, doctors and researchers how this important topic relates to them. The blog is expected to receive 2000-3000 hits a day, and should drive up traffic to http://www.core-reference.org as well as the technical publication: http://dx.doi.org/10.1186/s41073-016-0009-4

Perhaps more importantly, the Blog helps those outside the sphere of regulatory medical writing understand that CORE Reference is a **freely available resource for the reporting of human medicinal trials**. Increasing awareness in the pharmaceutical research sector of the availability of CORE Reference means that just one month after its release, CORE Reference downloads reached 1,000 and this number is increasing exponentially, as a look at the download counter on the website will tell you. **Principal Investigator-led clinical trial units in universities, hospitals and medical charities should also know that this free resource is available for them.** Please use Sam's Blog to help spread the word, and also outside your professional circles. Let's encourage patients and the public generally to be better informed.

The open comment period on CORE Reference ended on 14 June 2016. Comments and responses are shared via http://www.core-reference.org 'Comments and Responses' page.

EMWA and AMWA workshops are planned at forthcoming conferences from autumn 2016.

Finally, despite the encouraging numbers downloading CORE Reference, we need you to tell us about its adoption and use via the dedicated page on http://www.core-reference.org. We know from your personal emails that support is widespread, but we need your public declaration of support.

We would also like you to tell us if you have submitted CSR(s) redacted for public disclosure to regulators, and share any feedback that you may have received via the Contact page.

Dr. Sam Hamilton
sam@samhamiltonmwservices.co.uk

We continue to develop the resources to help you use CORE Reference effectively:

Resource	Where and When	Presenter
Webinar: CORE Reference: A Medical Writer's Guide to Preparing CSRs in an Evolving Context	AMWA (online) 27 July 2016 13.00-13.45 EDT	Aaron Bernstein
Presentation: CORE Reference: A Medical Writer's Guide to Preparing CSRs in an Evolving Context	11-12 July 2016 3rd ExL Clinical Medical Writing Forum Sonesta Hotel Philadelphia, Philadelphia, PA, USA	Aaron Bernstein

See page 54 for more details

Call for Companies

The 2nd Medical Writing Internship Forum will be held at our May 2017 Conference in Birmingham, UK. Please contact internship@emwa.org for more information.

History of biostatistics

J. Rosser Matthews
University of Maryland, College Park, MD, USA

Correspondence to:
J. Rosser Matthews
Professional Writing Program
English Department
1220 Tawes Hall
University of Maryland
College Park, MD, USA
+1 (301) 405-3762
jrmatt3@umd.edu

Abstract
The history of biostatistics could be viewed as an ongoing dialectic between continuity and change. Although statistical methods are used in current clinical studies, there is still ambivalence towards its application when medical practitioners treat individual patients. This article illustrates this dialectic by highlighting selected historical episodes and methodological innovations – such as debates about inoculation and blood-letting, as well as how randomisation was introduced into clinical trial design. These historical episodes are a catalyst to consider assistance of non-practitioners of medicine such as statisticians and medical writers.

Methodologically, clinical trials and epidemiological studies are united by a population-based focus; they privilege the *group* (i.e., population) over the clinically unique individual. Over time, this population-based thinking has remained constant; however, the specific statistical techniques to measure and assess group characteristics have evolved. Consequently, the history of biostatistics could be viewed as an ongoing dialectic between continuity and change. The continuity derives from focusing on the group rather than the clinically distinct individual. The change derives from developments in statistical theory that have led to more sophisticated analyses. In this article, I will illustrate this dialectic by discussing examples from antiquity to the emergence of the clinical trial in the mid-20th century.

Ancient sources: Hippocratic writings and the bible

Although the Hippocrates writers (active in the 5th century BCE) did not employ statistical methods, one treatise does stand out as a pioneering example of an environmental epidemiological study– the treatise *On Airs, Waters, and Places* (c. 400 BCE).[1] Relying on a view of disease as based on an imbalance in bodily fluids– known as *humours* – the work emphasised how climatic changes throughout the seasons of the year contributed to the spread of different types of diseases.[1] While basically qualitative, the work is historically significant because it looked beyond the individual to suggest a role for larger geographic and environmental factors. Furthermore, it relied on naturalistic explanations rather than invoking various deities to account for illness and therefore anticipated a modern scientific outlook.

Another ancient forerunner of contemporary clinical trials is discussed in the Bible's *Book of Daniel*. King Nebuchadnezzar of Babylon wanted all of his subjects to eat a diet of only meat and wine. However, Daniel and some of the other Jewish children wanted to eat a diet of legumes and water. The King permitted them this diet for 10

days – after which it was determined that they were indeed healthier. Consequently, they were allowed to continue on this diet.[2] Although not having the "apparatus" of a modern clinical trial (e.g., statistical tests to determine p-values, confidence intervals etc.), it does illustrate the use of a comparison to test the efficacy of a dietary intervention.

Eighteenth century developments

In the 18th century, one prominent example of using statistical methods to resolve therapeutic debates centred on the practice of smallpox inoculation. This involved inserting actual smallpox pustules under an individual's skin in the hope of creating a mild (i.e., non-disfiguring) case of the disease that would induce later immunity. Since this actually put patients at risk of contracting a potentially fatal form of the disease, this became the subject of much controversy.

Some argued against this procedure based on the Hippocratic injunction "first, do not harm." However, many writers justified the procedure based on arguments that today would be called "risk-benefit analysis." For example, the London physician John Arburthnot (1665-1735) published an anonymous pamphlet in 1722, in which he examined the *London Bills of Mortality* from earlier years and estimated that the chance of dying from naturally-occurring smallpox was 1:10. He then asserted (without evidence) that the chance of dying from inoculation-induced smallpox was 1:100. This ten-fold reduction made him conclude that inoculation made sense: "A Practice which brings the Mortality of the Small Pox from one in ten to one in a hundred, if it obtain'd universally would save the City of London at least 1,500 People Yearly; and the same Odds wou'd be a sufficient prudential Motive to any private Person to proceed upon."[3] In 1760, a more mathematically sophisticated version of this type of analysis took place in a debate between the Swiss mathematician Daniel Bernoulli (1700-1782) and the French mathematician Jean d'Alembert (1717-1783). Bernoulli drew on probability mathematics to contrast life expectancies for inoculated and non-inoculated individuals; also, he calculated the benefits of inoculation broken down by age. D'Alembert challenged Bernoulli's assumptions and said that Bernoulli's model had not accurately captured the psychology of human decision making – would an individual accept the risk of death now (from inoculation) for an expected "pay-off" of additional years of life when one was old and feeble?[3]

While the debates about inoculation relied on mortality statistics, the individual that is more often credited with designing a controlled clinical trial (i.e., intentionally dividing the participants into two or more comparable groups to test hypotheses) is James Lind (1716-1794). In 1757, Lind (a ship's surgeon) had to deal with an outbreak of scurvy. He selected 12 of the sailors and divided them into six groups of twos. All were given the same diet – except for a key different ingredient for each of the distinct six groups. For the two sailors who received oranges and limes as supplement, there was one complete and one near recovery; none of the other five groups improved as much. Despite some obvious structural similarities to the Biblical account, Lind is today regarded as the (modern) "father" of the controlled clinical trial.[2]

Nineteenth century developments

In several areas of 19th century scientific endeavour, statistical reasoning was introduced – and the field of medicine was no exception. In the 1830s, one of the most prominent advocates for applying the "numerical method" to medicine was the French clinician Pierre-Charles-Alexandre Louis (1787-1872) (Figure 1). By collecting data on patients admitted to hospitals, Louis argued that the practice of bloodletting was actually doing more harm than good. In his 1835 treatise *Recherches sur les effets de la saignée*, Louis pointed out that 18 patients died out of the 47 who had been bled (approximately 3:7) whereas only nine died out of the 36 patients not bled – producing a lower mortality rate of approximately 1:4.[4]

Louis justified his approach by claiming that the difference between numbers and words (such as "more or less" and "rarely or frequently") is "the difference of truth and error; of a thing clear and truly scientific on the one hand, and of something vague and worthless on the other." Furthermore, Louis prophesied that, with the widespread introduction of numerical reasoning, "we shall hear no more of medical tact, of a kind of divining power of physicians."[4] In language that foreshadows contemporary discussions of "evidence-based medicine," Louis was basically saying that the key to transform medicine into a science was to rely on population-based thinking rather than individual expertise.

Some of Louis' contemporaries criticised his approach for failing to acknowledge that the physician had to treat the individual as a patient rather than a statistical construct. For instance, the physician Benigno Risueño d'Amador (1802-1849) used an analogy to maritime insurance. Although past experience might tell you that 100 vessels would perish for each 1,000 that embarked, these population-based regularities could not tell you which specific ships

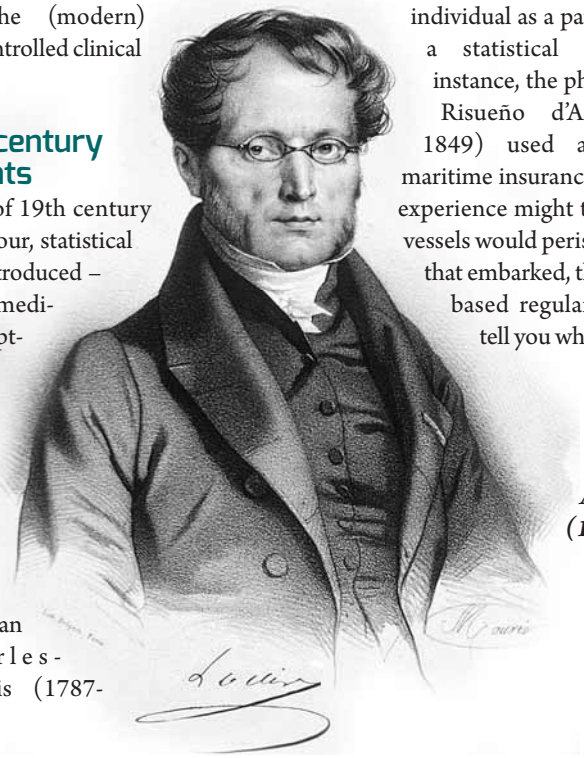


Figure 1. Pierre-Charles-Alexandre Louis (1787-1872) was a pioneer of the "numerical method" in medicine.

would be destroyed. Analogously, Risueño d'Amador argued that the calculus of the mathematicians "cannot be used to forecast a determined event, but only to establish the probability of a certain numerical proportion between two classes of possible events. But it is precisely this fact which makes it completely useless in medicine."[5] Drawing a different analogy, the physician François Double (1776-1842) claimed that relying on the numerical method would reduce the physician to "a shoemaker who after having measured the feet of a thousand persisted in fitting everyone on the basis of the imaginary model."[6]

These types of criticisms discussed statistical reasoning in the context of medical ethics: should the physician be concerned primarily with advancing scientific knowledge (through collecting empirical data), or with treating the individual in need of medical care? At the same time, however, a more mathematically sophisticated critique of Louis' work was developed by the French physician Jules Gavarret (1809-1890) (Figure 2), who had been trained as an engineer before becoming a physician and therefore understood probability mathematics. Gavarret published a treatise in 1840 entitled *Principes généraux de statistique médicale* in which he pointed out that Louis's averages could vary between what he called "limits of oscillation" if multiple samples were taken from the same population. For instance, Louis had observed 140 cases of typhoid fever with 52 deaths and 88 recoveries, or a mortality of 37%. Relying on probabilistic considerations, however, Gavarret posited that the results could vary by 11.55%, or between 26% and 49% in every 140 cases observed.[7] In modern day parlance, Gavarret was reporting the "confidence interval" associated with Louis' result.

To modern eyes, Gavarret seems remarkably prescient; however, there was no receptive audience for this marrying of statistics to probability mathematics in mid-19th century medicine. While his treatise was commented on throughout the 19th century (with varying degrees of mathematical sophistication), no "school" of followers committed to Gavarret's specific mathematical approach emerged. As a result, the meaning of statistical evidence remained contentious throughout the 19th century. For example, the famous surgeon, Joseph Lister (1827-1912), argued for his particular method of antiseptic surgery based on statistical studies; however, his critics had alternative theories of how to make surgery safer, citing other statistical studies that claimed to establish the superiority of their alternative theoretical approaches.[8]

The creation of the modern clinical trial

The move to standardise and "mathematise" statistics came with the creation, at University College London, of the Biometric School in 1893 and the Biometric Laboratory a decade later.[9] Heading the School and Laboratory was the pioneering statistician Karl Pearson (1857-1936) (Figure 3) who developed many modern statistical techniques to study biological variation – such as curve-fitting and goodness-of-fit tests, as well as methods for measuring correlation.[9] While Pearson's interest in developing statistics derived from a desire to make explicit the statistical implications of Darwin's theory of natural selection, he also advocated the extension of these methods into medicine. To that end, he often contributed to the *British Medical Journal*, the *Lancet*, and *The Royal Society of Medicine* as attempts to "educate" the medical profession on the proper methods of statistical reasoning.[10]

One physician who would actively embrace Pearson's recommendations was Major Greenwood (1880-1949). Greenwood studied under Pearson in 1904-1905 at the same time that he received his licence to practice medicine. At the beginning of 1910, Greenwood would be awarded a full-time position as a medical statistician at the Lister Institute of Preventive Medicine. Like Pearson, Greenwood would proselytise for statistical methods by debating with physicians. One of his most noteworthy encounters involved an exchange in the *Lancet* in 1912-1913 with the bacteriologist Sir Almroth Wright (1861-1947) over Wright's use of vaccines to combat pneumococcal infection among South African mine workers. Centring on the issue of the accuracy of a blood test, the debate evolved into a more generalised discussion over which forms of scientific evidence were more credible.[11]

By forging a career in academic science, Greenwood would help lay the foundations for a mathematically-informed understanding of epidemiology. In 1927, he would become the first professor of epidemiology and vital statistics at the London School of Hygiene and Tropical Medicine (LSHTM);[12] in 1928, he would be elected to the Royal Society. Also, Greenwood would train many students, of which the most prominent would be Sir Austin Bradford Hill (1897-1991).

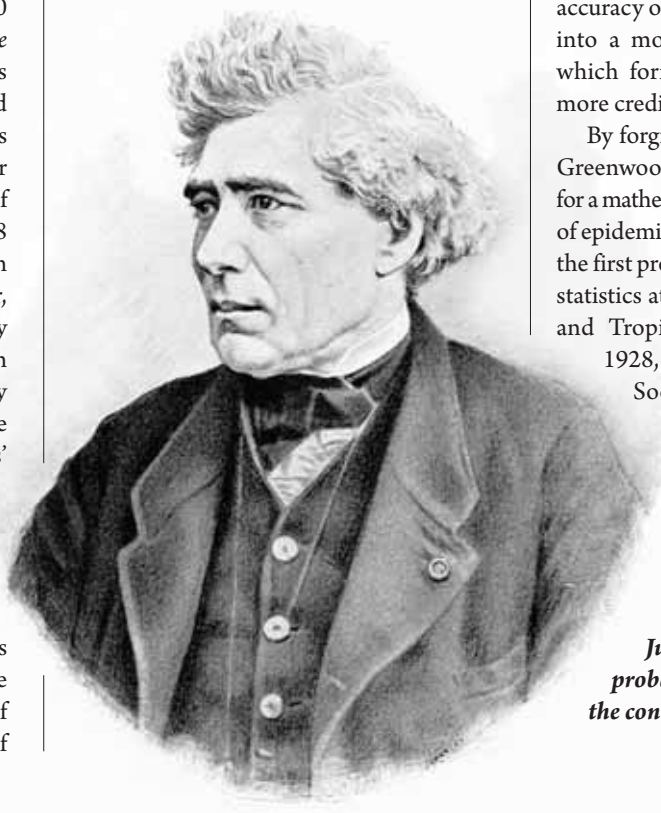


Figure 2. Jules Gavarret (1809-1890) used probability mathematics by applying the concept of the confidence interval to medical statistics.

Figure 3. Karl Pearson (1857-1936) developed curve-fitting methods and measures of correlation.

Bradford Hill was the third son of the physiologist Sir Leonard Hill (1866-1952) and had planned on following his father's medical profession. However, he contracted tuberculosis during World War I, and eventually earned an economics degree by correspondence. Hill gravitated towards statistics attending Pearson's lectures. In 1933, he would be appointed to a Readership at the LSHTM; upon Greenwood's retirement in 1945, Hill would succeed him as the head of the Statistics and Epidemiology Unit.[9] Like Pearson and Greenwood, Hill sought to educate the medical profession on the proper use of statistics. In 1937, he wrote a series of articles explaining statistical methods for the *Lancet*; subsequently, they would be republished as *Principles of Medical Statistics* and go through multiple editions and translations. In 1946, Hill would design a famous clinical trial to test the efficacy of streptomycin in treating tuberculosis – a methodologically noteworthy trial because it used a series of random sampling numbers to assign patients to the control (bed rest) or experimental (streptomycin) group. This trial is often characterised as the first clinical trial to use a randomisation scheme effectively. In 1965, Hill would articulate what have come to be known as the "Bradford Hill Criteria." These criteria can be used to determine whether an empirically observed association (e.g., between cigarette smoking and cancer) might be suggestive of an underlying causal relationship.

Today, the clinical trial is held as the gold standard for certain knowledge, and statistically-based epidemiological studies are widely reported in the news. However, as this brief historical sketch has illustrated, the current ascendency of these population-based thinking masks a larger ambivalence towards statistical methods within the medical profession. Even as statistical methods have been used to justify notable therapeutic breakthroughs, the population-based thinking on which they are predicated still runs counter to the individualistic focus of clinical practice. Perhaps, this historical legacy is one of the reasons that clinical trials often require the services of "outside" experts – such as statisticians and professional medical writers.

References

1. The Internet Classics Archive. On Airs, Waters, and Places by Hippocrates. 2016. [cited 2016 June 30]. Available from: http://classics.mit.edu/Hippocrates/airwatpl.html.
2. Collier R. Legumes, lemons and streptomycin: A short history of the clinical trial. CMAJ. 2009;180(1): 23-4.
3. Rusnock AA. Vital accounts: Quantifying health and population in eighteenth-century England and France. Cambridge: Cambridge University Press; 2002.
4. Louis PCA. Researchers on the effects of bloodletting in some inflammatory diseases, and on the Influence of tartarized antimony and vesication in pneumonitis. Putnam CG, translator. Boston: Hilliard, Gray; 1836.
5. Risueño d'Amador B. Memoire sur le calcul des probabilites applique a la medicine. Montpellier: L.F. Gaste; 1838.
6. Bulletin de l'Académie Royale de Médecine. 1836/37; 1: 702.
7. Gavarret J. Principes généraux de statistique médicale. Paris: Libraires de la Faculté de Médecine de Paris; 1840.
8. Tröhler U. Statistics and the British controversy about the effects of Joseph Lister's system of antisepsis for surgery, 1867-1890. JLL Bulletin: Commentaries on the history of treatment evaluation. 2016 [cited 2016 June 30]. Available from: http://www.jameslindlibrary.org/articles/statistics-and-the-british-controversy-about-the-effects-of-joseph-listers-system-of-antisepsis-for-surgery-1867-1890/.
9. Hardy A, Magnello ME. Statistical methods in epidemiology: Karl Pearson, Ronald Ross, Major Greenwood and Austin Bradford Hill, 1900-1945. Soziologie-Präventivmedizin. 2002; 47: 80–9.
10. Magnello E. The introduction of mathematical statistics into medical research: The roles of Karl Pearson, Major Greenwood, and Austin Bradford Hill. In: Magnello E, Hardy A, editors. The road to medical statistics. Amsterdam-New York: Rodopi; 2002. p. 95-123.
11. Matthews JR. Almroth Wright, vaccine therapy, and British biometrics: Disciplinary expertise versus statistical objectivity, In: Magnello E, Hardy A, editors. The road to medical statistics. Amsterdam-New York: Rodopi; 2002. p. 125-147.
12. Farewell V, Johnson T. Major Greenwood (1880–1949): A biographical and bibliographical study. Stat Med. 2016;35:645–70.

Author information

J. Rosser Matthews received his undergraduate degrees in mathematics and philosophy, and then earned a Ph.D. in the History of Science and Medicine. He is the author of the 1995 book *Quantification and the Quest for Medical Certainty*, which traced the history of statistical reasoning in 19th and 20th century medicine.

The illusion of certainty and the certainty of illusion: A case study of misunderstandings in scientific articles

Tom Lang
Tom Lang Communications and Training International, Kirkland, WA, USA

Correspondence to:
Tom Lang
10003 NE 115th Lane
Kirkland, WA 98933
USA
tomlangcom@aol.com
+1 425 242 1370

This article was reprinted from: Lang T. The illusion of certainty and the certainty of illusion: a case study of misunderstandings in scientific articles. *AMWA J.* 2013;28(3):105-9.

Abstract
Critical thinking is necessary to edit a scientific article. However, in addition to questions about the language, we can also question the assumptions, documentation, and implications of the research, in a process I call "analytical editing." A text with unverified assumptions, missing documentation, and unconsidered implications can lead readers into believing that they understand an article when they do not, creating the "illusion of certainty." Here, I present an example of the analyses needed to understand a single sentence; a case study, if you will, of analytical editing. A close look at the sentence raises several important questions about meaning, measurement, statistical analyses, how data are presented, and how results are interpreted. Analytical editing, in conjunction with traditional substantive editing, allows editors to increase their professionalism and value-added to clients.

The single biggest problem in communication is the illusion that it has taken place.
George Bernard Shaw[1]

Science is based on writing. Only writing allows science to be recorded, evaluated, and reproduced and enables it to be systematic, cumulative, and public; the characteristics that distinguish it from authority, intuition, and tradition as a way of establishing "truth."

Publication–the final stage of research–depends on writing, as does evidence-based medicine, which is *literature-based* medicine.[2]

Given the importance of writing in understanding and advancing science, one would think that physicians and researchers would be provided full support in preparing publications. However, at least in clinical medicine, such support is often inadequate. Researchers are not expected to do their own literature searches and so are given access to librarians. They are not expected to do their own data analysis and so are given access to statisticians. They are not expected to render their own graphs and drawings and so are given access to medical illustrators. But for some reason, we expect them to do their own writing–to communicate technical information accurately, completely, clearly, and economically, in words and images–without specific training, and often without the support of professional medical writers and editors. Thus, we shouldn't be surprised that a large portion of the scientific literature is not immediately, accurately, and completely understandable.

One of the most important lessons I have learned in almost 40 years of editing is that the certainty we believe we have about understanding even a simple, straightforward sentence is often illusory. The sense of certainty is so strong that we don't even think to question the meaning. Only on closer examination does the illusion become apparent. Further, such sentences are found in most scientific articles, which is to say, these illusions are also a certainty in the scientific literature.

I encountered a good example of a sentence in which the actual meaning differs remarkably from its apparent one. In this article, I pose some questions that need to be answered if this sentence is to be understood correctly. These questions are part of what I call "analytical editing," or editing to assure that research designs and activities are documented appropriately and explained adequately.[2] Analytical editing seeks to meet the needs of evidence-based medicine by making sure the evidence itself is completely and clearly reported. Analytical editing does not require us know medicine. It does require that we know how medical research is conducted–or at least what questions to ask about the research–as well as the standards to which this research should be documented. A task often left to peer reviewers, analytical editing can be done by trained editors and, in conjunction with traditional substantive editing, allows editors to increase their professionalism and value-added to clients.

The Example

This sentence was in the results section of a poorly written abstract: *One group of patients was significantly less depressed than the other*. The sentence seemed straightforward, but the more I analysed it, the more questions I had.

The questions

Question 1: What is the context of the sentence?

The sentence was the second in the results section of the abstract. Taken by itself, the sentence could have been a description of the patients at baseline, an incidental finding that might confound the results, or a result itself. Given the context of the article–a study of a new antidepressant–it was probably the result of the study.

Meaning is a product of message and context, in the same way that the meaning of a picture is a product of image and background (Figure 1). Change the context, and the same message has a different meaning. *The wall was built to scale* means something different to an architect than it does to a climber. For this reason, the context of every scientific article needs to be clear to rule out other interpretations made possible by different contexts. One function of a good introduction is to put the research in the proper context.

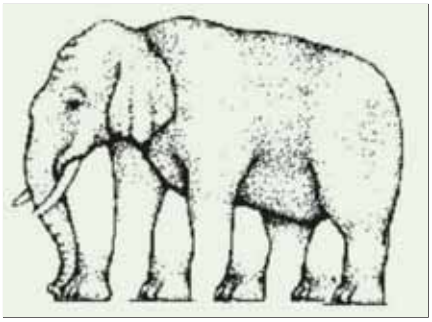


Figure 1. The "figure-ground" effect that becomes apparent from trying to make sense of this image is similar to what happens when we interpret a written message in different contexts. The context determines the meaning to some extent.

Question 2: Who was studied?

The article stated that the participants were women outpatients with moderate-to-severe depression being treated at a university hospital. The two groups mentioned in the sentence were the treatment and control groups of the study, something the sentence could have said, "Patients in the *treatment* group were significantly less depressed than were patients in the *control* group." We also need to know the patients' age, diagnosis, how they were selected for the study (the sampling method and eligibility criteria), other health conditions, and so on.

How the sample size was determined also needs to be explained. Especially in randomised trials, sample size should be determined with a power calculation. Basically, a power calculation tells investigators how many patients they need to enroll in a study to have, say, an 80% chance of detecting a difference of a given size *if such a difference actually exists in the population from which the sample was taken*. Investigators rarely get a chance to study an entire population. Instead, they have to study of a sample of that population. However, there is a chance that the sample won't include patients that express the difference of interest, a problem called "sampling error." The power calculation estimates the size of the sample likely to be large enough to include patients that express the difference at a degree of uncertainty acceptable to investigators.

In "underpowered" studies–studies that did not enroll enough patients to detect the desired difference–the lack of a statistically significant difference doesn't mean the groups are similar, it means the study was inconclusive: "absence of proof is not proof

… the lack of a statistically significant difference doesn't mean the groups are similar, it means the study was inconclusive: "absence of proof is not proof of absence."

of absence." The difference of interest is usually the smallest considered to be clinically important, so we have to determine this difference and whether the study enrolled enough patients to have a reasonable chance (often 80% or 90%) of detecting it.

Question 3: What was studied?

Depression can be treated in several ways, so the treatment needs to be described in detail. If the treatment is a drug (as it was in this example), we need to know the generic name, manufacturer, dosage, route of administration, and perhaps the indications, possible side effects, and the degree to which each group took the medication as planned. The rate of protocol adherence is usually higher in in-patient studies than in outpatient studies, for example.

Question 4: How was depression measured?

All study variables must be defined in objective, measurable terms. In this case, we need to know how depression was measured. Was the diagnosis based on a physician's judgment, a self-report questionnaire, or some other way? The text said that "All patients completed the Beck Depression Inventory before and after treatment." The Beck Depression Inventory is a common, validated instrument for measuring depression. This information was encouraging. Many authors do not say how they measured their variables, often because "my readers will know." Right.

Question 5: What type of comparison is being made?

In a study with two groups in which both pre- and post-treatment values are measured, two comparisons are possible. The within-group comparison looks at the changes between pre- and post-test values for each group, whereas the between-group comparison looks at the differences between groups at the beginning or end of the study. In a study like this one, both comparisons are likely. However, the sentence in question says that one group was less depressed than the other, so we have to ask whether the statement refers to a *between-group comparison*–at the end of the study, mean depression scores in one group were lower than the mean of those of the other (and presumably the baseline scores were similar)–or a *within-group comparison* – the change in depression scores during the study was greater in one group than in the other (and the baseline scores were not necessarily similar).

Question 6: How large was the difference between groups?

The authors reported that "The mean depression score of the treatment group was 38% lower than that of the control group." Fine, but results expressed only as percentages are *always* suspect. Numerators and denominators should always be available for all percentages.[3]

There is an old laboratory joke about how 33% of the rats lived, 33% died, and the last one got away. It is also usually true that a 50% reduction from 2 to 1 is not the same as a reduction from 2,000 to 1,000. Hence, the need to provide numerators and denominators when reporting and interpreting percentages

Mean values can also be a problem. If Bill Gates walks into a room, the average income of people in the room skyrockets, but nobody makes any more money. Here, it is possible that the lower mean depression scores represent not an overall decrease in the severity of depression but rather an effect caused by a few patients who responded unusually well to treatment (Figure 2).

Question 7: What does the author mean by "significantly"?

In medical writing, *significant* should be reserved for its statistical meaning, but the term is still often used to mean *markedly* or *substantially*.[2-4] An accompanying *P* value or a 95% confidence interval usually indicates that the term is used for its statistical meaning, but not always. In the present example, *significant* was used in its statistical sense.

The most common reporting error in medical articles is confusing statistical significance with clinical importance.[2,3] Relying on *P* values to interpret results is often easier than considering whether a result is clinically important. However, even when used appropriately, *P* values themselves must be reported correctly. We need to know the actual *P* value (*P*=0.03, not *P*<0.05); the alpha level (usually 0.05) that defines the threshold of statistical significance; the statistical test used to calculate the *P* value; whether the assumptions of the test have been met by the data (eg, whether the data are independent or paired); whether the test was 1- or 2-tailed; and the statistical software program used in the

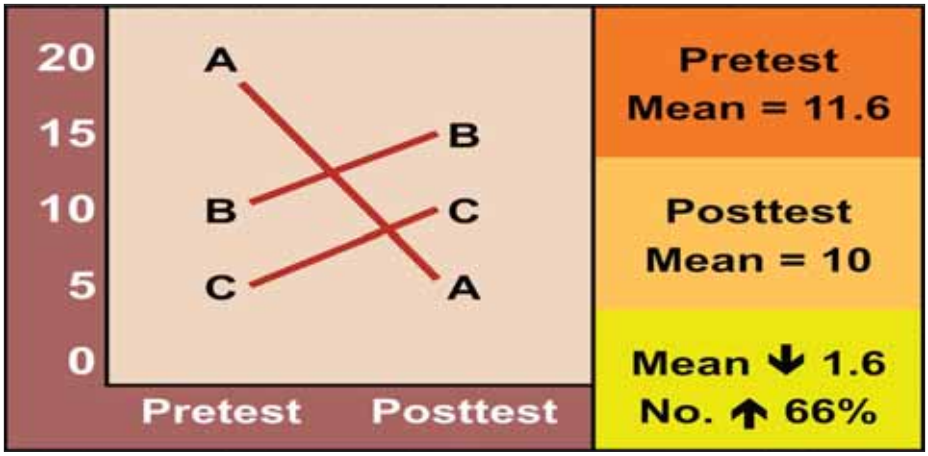


Figure 2. The problem of reporting a change in group means or the number of patients in whom change occurred. Here, the large change in patient A has had a disproportionate affect on the mean of all three patients. Thus, the data can be reported either as the fact that the mean decreased by 1.6 points, from 11.6 to 10 (14%), or that 67% of the patients had increased values. (Of course, the 67% is only 2 of 3, but it's still 67% . . .)

analysis (to establish its validity).[3]

Returning to the manuscript at hand, had the authors said something like, "One group was less depressed than the other (*P*=0.02)," we would have known that "significant" was used in its statistical meaning.

Question 8: How precise is this estimate of the difference?

The results of most biomedical studies are, in fact, estimates, and estimates require a measure of precision.[3] In medicine, this measure is usually the 95% confidence interval. I think of the interval as being the range in which the mean difference is expected to occur in 95 of 100 similar studies and in which the difference would be outside the range in the remaining 5 of the 100.[3]

Confidence intervals are useful because they keep the interpretation focused on the effect size and therefore on the medicine, not the *P* value.[3] Confidence intervals that contain both clinically important and clinically unimportant values ("heterogeneous" intervals) suggest that, even if the difference in means is statistically significant for the current trial, the estimate is not probably not precise enough to conclude that the treatment will likely be effective in 95 of 100 similar trials. In other words, the result is clinically inconclusive.

Typically, a larger sample size gives a more precise estimate (a narrower confidence interval). What is important is not the width of the confidence interval but its "homogeneity." When the confidence interval contains only clinically important values, or only clinically unimportant values, then we have a more definitive answer to the research question.

Ideally, the authors would have written something like: "The difference between means was 3 points (95% confidence interval, 1.5 to 4.5 points)." But they didn't.

Question 9: What is the measurement scale for depression?

The Beck Depression Inventory is a scale that runs from 0 to 63 (Figure 3). Scores of 0 to 9 indicate no or minimal depression; 10 to 18, mild depression; 19 to 29, moderate depression, and 30 to 63, severe depression.[5]

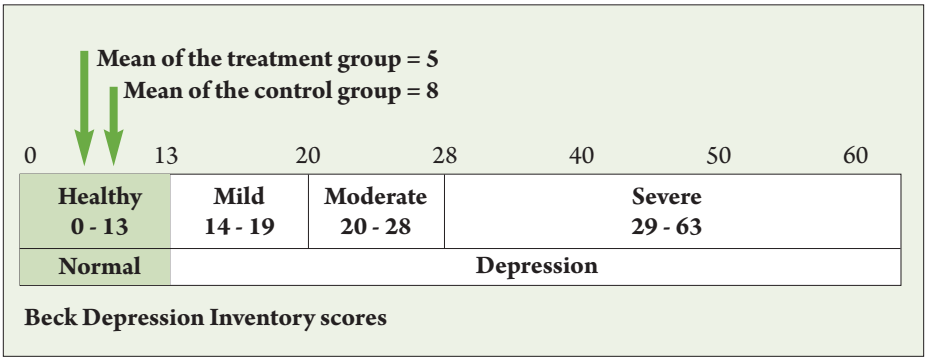


Figure 3. The Beck Depression Inventory is a common, validated instrument for measuring depression. To understand the measurement, however, we must answer several questions: 1. Is the scale linear? That is, does a 3-point difference at one end of the scale mean the same thing as a 3-point difference at the other end? 2. What is the smallest difference in scores that is clinically meaningful? 3. Are their any threshold scores

So, the 3-point difference between means, and its 95% confidence interval, has to be interpreted accordingly.

When we know the scale, we can also infer something about the baseline values. Remember, the text said that "All patients completed the Beck Depression Inventory before and after treatment." It is reasonable to conclude, then, that all women had Beck scores of at least 20 at baseline, and we hope the text will confirm this fact. The results are reported as the means of the post-treatment Beck scores, but it would be nice to know the mean baseline values of both groups. In some studies, if mean baseline values are close to normal, even the best treatment may show little effect because the range over which the means can drop is limited.

Question 10: What is the smallest clinically meaningful difference?

When reporting and interpreting results, the effect size (say, the differences between means) is usually more important than the *P* value. The effect size can be interpreted clinically, whereas a *P* value cannot.[3]

The authors revealed that after the intervention, the difference between the means of the treatment and control groups was 3 points. However, a difference, to be a difference, must make a difference. The "critical effect size" (the minimum clinically important difference) for the Beck Inventory was not given. (It turns out to be 5 points.[6] More on this later.)

What are we to make of this 3-point difference? Does it matter whether the difference crosses one of the threshold scores that define a different degree of depression? Does it matter whether the difference occurs at the low end or high end of the scale? Pain measured on a 10-point scale may be nonlinear; that is, a reduction from 9 to 8 may be greater than a reduction from 4 to 3.[7] We don't have to know whether the scale reflects a linear relationship among scores, however, we just have to ask authors if it is (Figure 3). Don't be surprised if they don't know.

Question 11: What were the actual mean values of both groups?

Now the illusion became apparent. A table showed that the mean score was 5 in the treatment group and 8 in the control group. These values are consistent with the 3-point difference between means and with the 38% lower score of the treatment group ([8 – 5]/8 = 0.38 Î 100 = 38%). However, both means are in the normal range (scores less than 9; Figure 4), so describing the result as "one group is less depressed than the other" is incorrect and misleading. The scores also differ by less than 5 points, so the difference is not clinically important.[6] The authors seemed to have based their interpretation solely on a significant *P* value, without considering the clinical implications of the results. "Why" could be a most interesting question here.

Question 12: What was the proportion of patients in each group who were still depressed after treatment?

The example compared the means of two groups. However, a common error in clinical research is to report changes or differences in means rather than indicating how many

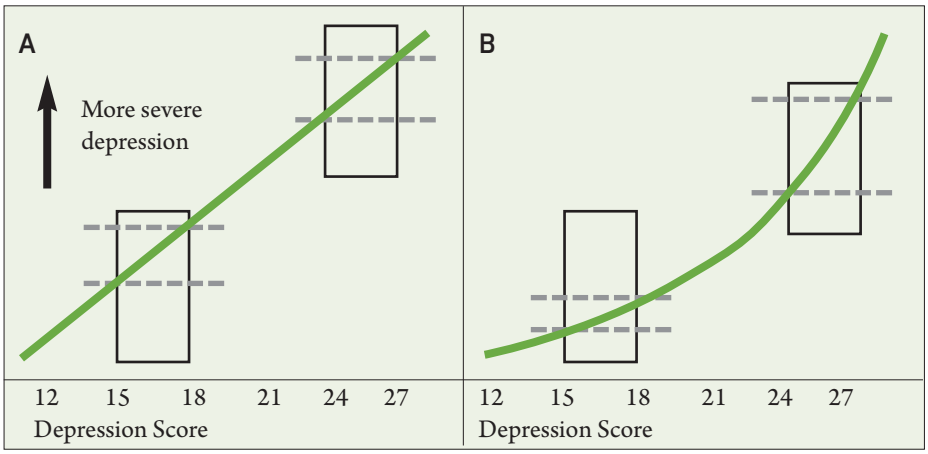


Figure 4. Measurement scales may or may not be linear. A. If the scale is linear, a 3-point change at the high end means the same thing as a 3-point change on the low end: the distance between the dotted lines is the same in both rectangles. B. If the scale is not linear, where the distance between the dotted lines in the two rectangles is different, the importance of a 3-point change depends on where that change occurs on the scale.

patients got better or worse (Figure 2). It would have been helpful to know how many patients were no longer depressed by the end of the study.

The issue here is the "unit of observation." I once edited a manuscript describing a study of 25 eyes, but it never said how many patients were involved. The unit of observation was eyes, not patient. The primary outcome of interest – the unit of observation – is in the protocol, but, as in the example, how patients responded is often and surprisingly not given.

Question 13: Is the drug likely to be generally effective?

Determining the effectiveness of the drug was the purpose of the study. The authors' claim that "*one group of patients was significantly less depressed than the other*" was supposed to mean that the drug was effective. They should have written something like: "After treatment, 72% (38/53) of the treated patients and 49% (27/55) of the control patients scored 9 or below on the Beck Inventory (95% CI for the 23% difference, 2% to 41%)," but they didn't. Instead, given the small effect size (3 points on the Beck scale in which 5 points is the smallest important difference), the fact that both means were in the healthy range, the lack of a confidence interval, and not knowing how many patients were no longer depressed at the end of the study, it does not seem reasonable to agree with the authors that the drug was effective.

However, we also can't conclude that the drug was ineffective. The difference was statistically significant, if clinically irrelevant. The drug did reduce the mean of the treatment group from well above 19 to 3, which supports the claim of efficacy, but the mean in the control group may have been reduced to a similar degree. All we can say is that they study was not well conducted, not well reported, or both.

Conclusions

Not all sentences are this involved, but many are and require analysis as detailed as the example presented here. Analytical editing can take time – and skill, training, and experience. What makes good writing and editing valuable is that they reduce readers' time, effort, and uncertainty about the meaning of a text, and they don't create the illusion of clarity. The problem is that many scientific articles are poorly written and poorly edited. Worldwide, authors are generally not skilled in communicating technical information in writing and do not receive adequate editorial support, and most journals provide only superficial copy-editing. This situation pretty much assures that readers of the scientific literature will regularly encounter the "illusion of certainty" and therefore must be prepared to accept the "certainty of illusion."

Acknowledgement

My thanks to Bart Harvey, MD, PhD, MEd, FRCPC, FACPM, for his review of the manuscript and help with the calculations.

References

1. Cited in: Caroselli M. Leadership skills for managers. New York: McGraw Hill Professional; 2000.
2. Beck Depression Inventory. 2013 [cited June 6, 2013]. Available from: http://en.wikipedia.org/wiki/Beck_Depression_Inventory.
3. Lang T. How to write, publish, and present in the health sciences: a guide for clinicians and laboratory researchers. Philadelphia: American College of Physicians; 2010.
4. Lang T, Secic M. How to report statistics in medicine: annotated guidelines for authors, editors, and reviewers, 2nd ed. Philadelphia: American College of Physicians; 2006.
5. Style Manual Committee, Council of Science Editors. The CSE manual for authors, editors, and publishers. 7th Ed. Reston, VA: Council of Science Editors in cooperation with the Rockefeller University Press; 2006.
6. Hiroe T, Kojima M, Yamamoto I, Nojima S, Kinoshita Y, Hashimoto N, *et al*. Gradations of clinical severity and sensitivity to change assessed with the Beck Depression Inventory-II in Japanese patients with depression. Psychiatry Res. 2005;135(3):229-35.
7. Aubrun F, Langeron O, Quesnel C, Coriat P, Riou B. Relationships between measurement of pain using visual analog score and morphine requirements during postoperative intravenous morphine titration. Anesthesiology. 2003;98(6):1415-21.

Author information

Tom is an international consultant and educator in medical writing/editing, critical appraisal of clinical research, and scientific publications. He is past President of the Council of Science Editors and current Treasurer of the World Association of Medical Editors. He also served on the CONSORT and PRISMA groups and wrote *How to Write, Publish, and Present in the Health Sciences*.

Never *P* alone: The value of estimates and confidence intervals

Tom Lang
Tom Lang Communications and Training International, Kirkland, WA, USA

Correspondence to:
Tom Lang
10003 NE 115th Lane
Kirkland, WA 98933
USA
tomlangcom@aol.com
+1 425 242 1370

Abstract
How results are reported influences how they are interpreted. Although *P* values have been granted great importance, they have no clinical interpretation. Rather, they are a measure of chance as an explanation for the results. Their either-or interpretation takes attention away from the results themselves–the difference between groups or the effect size–which are more important. Effect sizes are also estimates. Estimates are only useful if they are accompanied by a measure of precision. In medicine, this measure is usually the 95% confidence interval (CI). This article explains the concepts underlying CIs and illustrates how they are more useful than *P* values in reporting research. As such, journals are increasingly asking for CIs, instead of, or at least in addition to, *P* values.

Introduction
Statistics can be divided into two broad areas: **descriptive statistics**, in which data are summarised in a few numbers to make them more manageable, such as percentages and medians, and **inferential statistics**, in which measurements of a sample are generalised to the population from which the sample was drawn. This article is concerned with inferential statistics; in particular, the reporting of estimates and confidence intervals.

Most medical research is done on samples, but the findings are actually estimates of what we would expect if the treatment were to be given to the population from which the sample was drawn. For example, we can't study all patients with, say, epilepsy, we can only study a sample of such patients. When we're done, we hope that what we have learned from the sample will also be true for all patients who have epilepsy.

However, the sample is almost always only a tiny fraction of the population, so we need to know how good our estimate is. In medicine, this measure of precision is most often expressed as a confidence interval (CI), usually a 95% CI, although the "confidence coefficient" (the 95%) may be 90% for smaller samples and theoretically can be any number. Thus, understanding estimates and confidence intervals is important to understanding the medical literature.

In this article, I illustrate the concepts underlying estimates and confidence intervals with a hypothetical example. Hypothetical, because the concepts involved differ from the actual research methods and mathematics used to compute the confidence intervals. After giving the example, I'll explain how the confidence interval is actually determined. Interested readers are

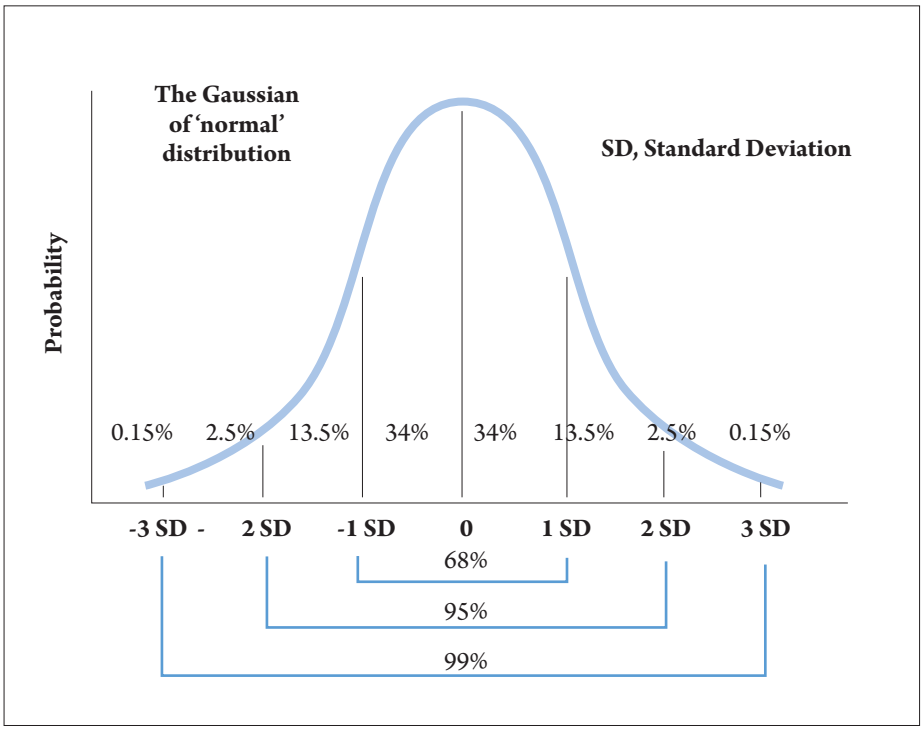


Figure 1. The relationship between the standard deviation and the area under the normal curve holds for all normal distributions, no matter how flat or peaked.

In a distribution of data, the SD is the preferred "measure of dispersion," or spread of the data. Other normal distributions have an SD, but the name changes to connect it with the distribution. In a distribution of all possible sample means, as described below, the SD is called the standard error of the mean (SE). It has the same mathematical properties as the SD, it's just associated with a different distribution.

invited to read *Statistics without Tears*, by Rowntree[1] for a fuller description of the approach taken here and *How to Report Statistics in Medicine*, by Lang[2] for more information about reporting estimates and confidence intervals.

Background information

Before we can talk about estimates and confidence intervals, we have to review some basic concepts of probability. In particular, we need to review the properties of the "normal distribution" or the gaussian or bell-shaped curve.

In any normal distribution, the mean value equals the median value equals the modal value, and the curve is symmetrical about the mean. It also has two "inflection points" where the curve changes direction to give it its bell shape. Most importantly, the area under the curve can be described in units of standard deviation (SD), and this relationship holds for any normal distribution *(Figure 1)*. Importantly, this relationship allows us to compare values on any normal distribution with those of any other, no matter how peaked or flattened the curves.

Suppose we wanted to compare patient survival in two groups of different sizes. It wouldn't be fair to compare the raw numbers of survivors between groups because one group is larger than the other. Instead, we convert the raw numbers into a common unit – percentages – to accommodate the difference in group size and then compare the percentages.

Now, suppose we want to compare two different normal distributions. Linda took the final exam in her law class, and Bill took his in economics. We want to compare their scores to determine who is the better student *(Figure 2)*. We can't compare Bill's score of 90 to Linda's score of 80 because each test has a different distribution of values; one test had more questions than the other, which changes the range of possible scores, or maybe one class had more variability than the other because more people did well and more people did poorly on the test.

As we did with percentages, however, we can compare scores from different distributions if we can express the values in a common measure. We do this by converting raw scores into units of SD (a "standard score," or z-score), which we can then compare on a common distribution. A score equal to the mean value of the common distribution (or "standard normal distribution") has an SD of zero; half the values are less than the score and half are greater. A score 1 SD above the mean is greater than about 84% of the values (50% to the left of the median or center value plus 34%) and less than about 16%, whereas a score of -1 SD below the mean is greater than about 16% of the values and less than about 84% (*Figure 1*).

Getting back to Bill and Linda, if we now express the two scores in terms of SDs, we see that Bill's score of 90 was 2 SD above the mean in his class, and Linda's was 3 SD above the mean in her's. So, Bill did better than about 97.5% of his classmates, but Linda did better than about 99.9% of hers. Linda did relatively better, even though her raw score was less than Bill's.

It is important to remember that the SD indicates these proportions only for normal distributions. So, normal distributions can be appropriately summarised with means and SDs, but distributions of other shapes should be summarized with different descriptive statistics.

Estimating a population value

An **estimate** is a *probable* value for a population that is inferred from a *measured* value of a sample. In medicine, we sometimes want to estimate the value of a physical trait in a population, such as average birth weight. We might also want to estimate the response to an intervention, such as differences between groups ("between-group comparisons") or in the same group before and after treatment ("with-in group comparisons").

Here's the hypothetical example. Imagine a gnome, a mythical being that guards the earth's underground treasures. Gnomes have

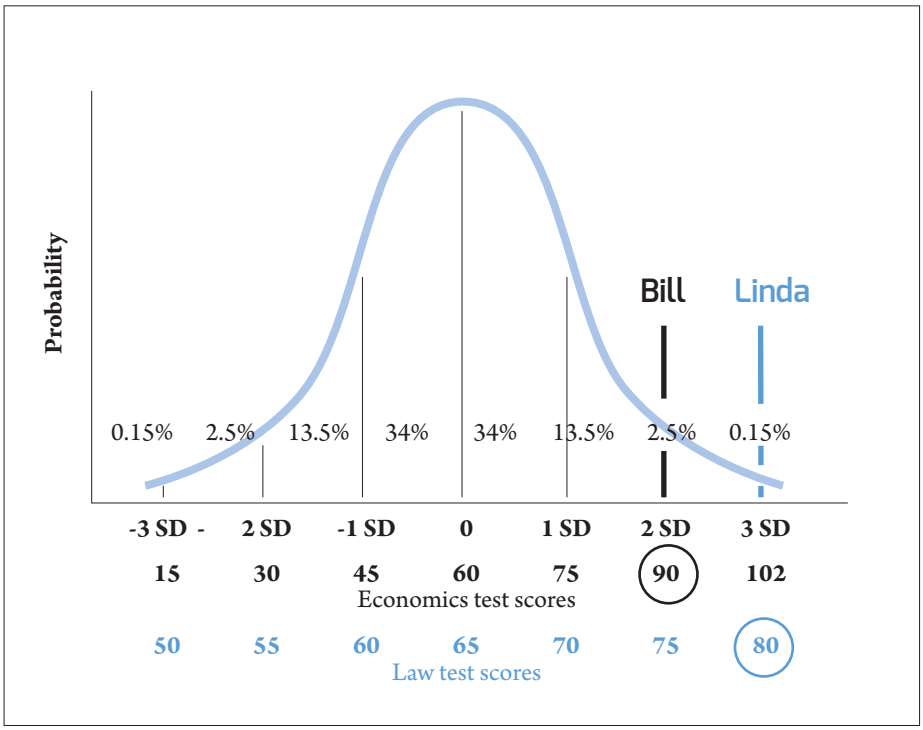


Figure 2. Comparing two distributions of different proportions with the standard deviation. The distributions of scores on the law and economics tests are shown below the standard normal distribution. Linda's raw score of 80 is 3 SD above the mean in her class, and Bill's is 2 SD above the mean in his. Clearly, Linda did relatively better than Bill on her test.

only been seen in small groups, however, so no one knows how tall the average gnome is. Thus, our research question is "How do we estimate the average height of all gnomes if we can only measure a few of them?"

Suppose that a gnome magically appears on your desk. You measure him and find that he is exactly 10 cm tall *(Figure 3)*. What's our best guess about the average height of all gnomes? The answer is 10 cm, because it's all the information we have in a sample size of 1.

Now suppose a second gnome appears beside the first one. This gnome could be 10 cm, but probably he will be a little bigger or a little smaller. Supposed he is 11 cm tall. Now, what is our best guess about the average height of all gnomes? The answer is: 10.5 cm, because it's all the information we have. That is, we average the heights of our sample of two, which is 10.5 cm. In fact, **the sample mean is the best estimate of the population mean** because it uses all the available data.

We could repeat this process if, say, 10 gnomes were to appear: measure the height of each gnome and then calculate their mean height. This sample mean will, again, be the best estimate of the mean height of the gnome population. The same is true for other characteristics of the sample as well: for medians, ranges, and standard deviations, for example.

Figure 3. A gnome 10 cm tall. If this gnome is the only one we've measured, our best estimate of the average height of all gnomes is 10 cm because that is all the information we have. The mean of the sample is best estimate of the mean of the population.

Notice that our sample was small: 10 gnomes out of a population of several thousand gnomes (or so I've been told). With so many gnomes, how likely is it that our estimate, based on 10 gnomes, is accurate? If we happened to get a single sample containing the smallest gnomes, we would underestimate the average height in the population, and if we happened to get a sample containing the largest gnomes, we would overestimate it. What we need is a way to determine how precise our estimate might be. This measure is the confidence interval.

The hypothetical example illustrating confidence intervals

Suppose we have unlimited resources and unlimited cooperation of all the gnomes, such that we can take all possible random samples of, say, 10 gnomes. In other words, we draw a sample of 10 gnomes, measure the height of each, calculate the sample mean, graph the mean, and then return the gnomes to the population. We then draw another sample of 10 gnomes and repeat the process: measure each one, calculate the sample mean, graph the mean, and return the gnomes to the population. We repeat this process until we have taken samples of every possible combination of 10 gnomes *(Table overleaf)*. (You can see why the example is fictitious: agencies funding research into gnomes won't pay for this kind of sampling.)

We repeat this process until we have taken samples of every possible combination of 10 gnomes. (You can see why the example is fictitious: agencies funding research into gnomes won't pay for this kind of sampling.)

When we graph the means of all our samples *(Figure 4)*, we find that they are normally distributed. (This

Sample No.	Height of each of 10 gnomes in the sample, cm										Sample means
	1	2	3	4	5	6	7	8	9	10	
1	16	11	5	14	7	13	12	13	15	20	12.6
2	16	12	12	2	4	5	14	7	11	8	9.1
3	1	9	2	6	8	10	4	7	2	10	5.9
4	2	8	3	19	13	9	6	6	14	5	8.5
5	14	4	18	13	12	5	19	11	8	8	11.2
6	14	11	2	2	9	17	11	10	8	16	10
7	5	3	13	11	1	14	13	3	8	7	7.8
8	6	15	13	11	9	13	6	7	15	2	9.7
9	18	14	3	8	14	9	12	7	2	17	10.4
10	3	5	5	2	20	7	14	4	7	7	7.4

Table. The Heights of 100 Gnomes as Collected in 10 Samples of 10 Gnomes.
The overall mean (SD) of the 10 sample means is 9.3 (2.0) cm, which is the best estimate of the mean height (and SD) of the gnome population. The SE equals the standard deviation of the sample (2.0) divided by the square root of the sample size of 10 (3.2), or 0.63. Twice the SE is 1.3, so the mean -2 SE = 8 and the mean +2 SE = 10.6 cm, giving us an estimated mean of 9.3 cm (95% CI, 8 to 10.6 cm). See text for details.

result is explained by what is called "the central limit theorem," which I won't address here.) Remember that the "area under the curve" can be expressed in units of standard deviation. More importantly, the mean of this graph of sample means is, again, our best estimate of the population mean. Now, however, instead of a single sample mean, we have a *distribution* of sample means. When we had a sample of data, we called the measure of dispersion the standard deviation (SD). Now we have a distribution of sample means, so we are going to call the standard deviation the "standard error of the mean (SE)."

The SD and SE represent the same concept and have the same mathematical properties: both can be used to indicate the area under a normal curve. The only difference is that the standard deviation is a descriptive statistic that indicates the variability of a distribution of data, whereas the standard error of the mean is an inferential statistic that indicates the variability of an estimate; that is, the variability of the distribution of the means of all possible samples of the same size.

Remember that about 68% of the data will be included in the range defined by -1 SD below the mean to +1 SD above the mean, and that about 95% will be included between -2 SD and +2 SD. These relationships are the same for the SE: about 68% of the sample means will be included in the range defined by -1 SE below the mean of the sample means to +1 SE above the mean, and about 95% will be included between -2 SE and +2 SE (*Figure 5*).

The mean of this distribution of sample means is the best estimate of the population mean, and the range given by plus or minus 2 SEs is a 95% CI. In other words, we measured only samples of gnomes, and the estimate varied from sample to sample. However, the mean in 95 of 100 samples of 10 will probably fall within the range defined by 2 SEs above and below the mean of our distribution of sample means.

Calculating confidence intervals

In reality, we generally measure only a single sample. The (measured) sample mean is the best estimate of the population mean, and the 95% CI is calculated from the SE with the simple formula:

$$\text{SE} = \frac{\textbf{Standard deviation of the sample}}{\textbf{Square root of the sample size}}$$

One SE on either side of our mean of sample means is about a 68% CI. To get the

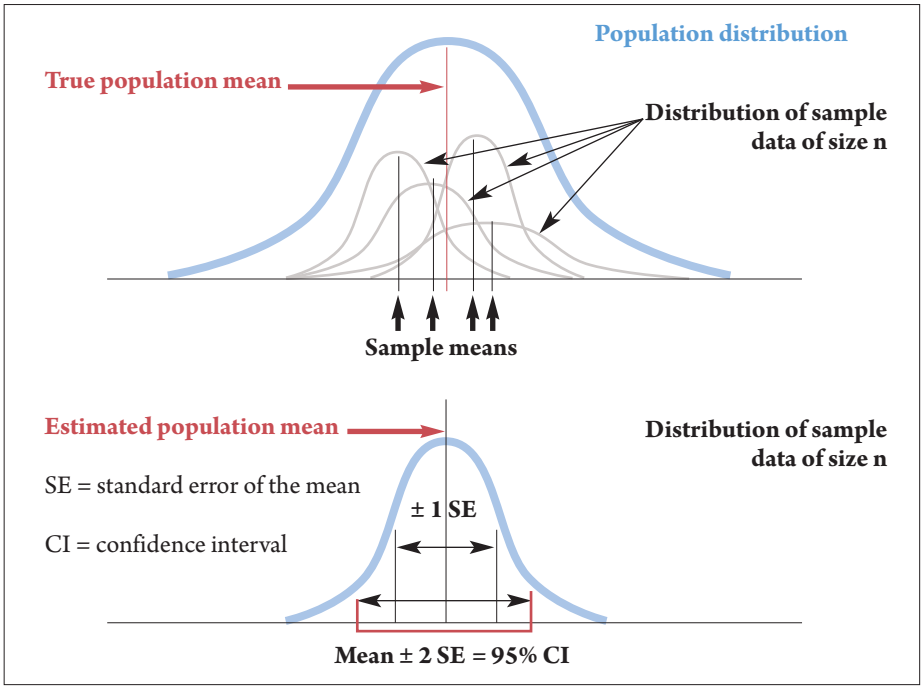


Figure 4. The conceptual process of creating a 95% confidence interval around the estimated mean height.
Upper panel: we take all possible samples of the same size from the population of interest, compute the mean height of each sample, and graph the means. Lower panel: the new distribution of means will be normally distributed, so 95% of the samples we drew had means that ranged between two SEs above and below the overall mean of the new distribution. The overall mean is the estimated height, and the range between the mean plus and minus 2 SEs is the 95% interval for the estimate.

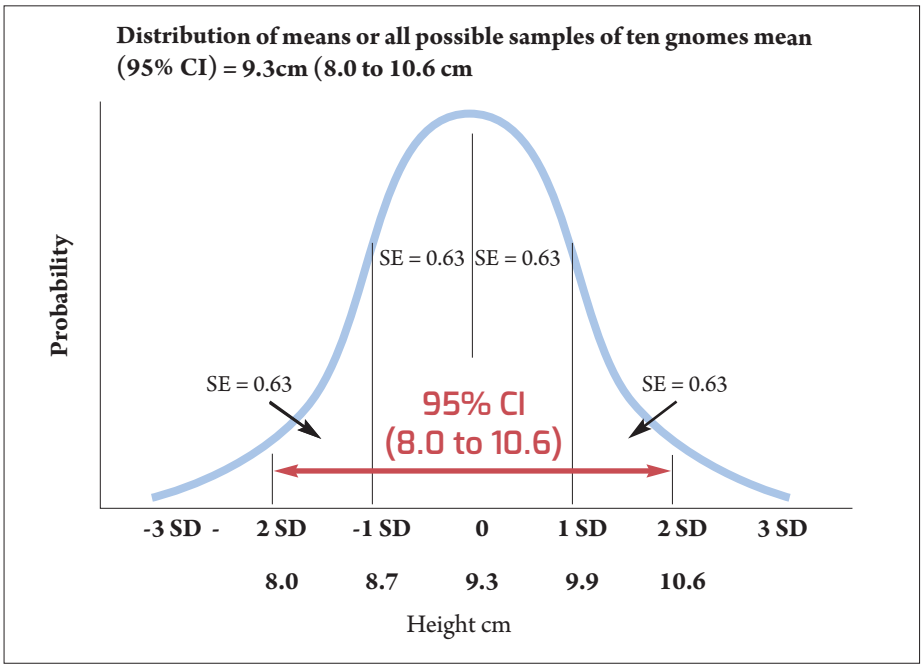


Figure 5. The distribution of sample means in the example (summarised in the Table) of estimating the average height of gnomes.

95% CI, we essentially double the SE, which gives the range of values in which we expect the mean height to fall in 95 of 100 similar samples.

Using data from the example in the table, the mean of the distribution of all possible samples of the same size (although only 10 are shown here) is 9.26 cm. The SE is 1.96, and 2 SEs equal about 3.8. Adding and subtracting the 3.8 to the mean of 9.26 gives us an estimated height of 9.26 cm with a 95% CI of 6.2 to 13.8 cm.

The value of confidence intervals

Confidence intervals have enormous value in reporting the results of medical research. The results of most biomedical studies (that is, the "effect size") are actually estimates and so should be accompanied by CIs. In addition, CIs are increasingly preferred to *P* values when reporting results. The *P* value is a mathematical measure of chance as an explanation and has no biological interpretation. On the other hand, CIs keep the interpretation focused on the biological implications of the effect size.

Here's an example of the value of confident intervals. Consider this sentence:

"The drug reduced diastolic blood pressure (DBP) by a mean of 15 mm Hg (95% CI = 3.5 to 26.5 mm Hg; *P* = 0.01)."

In this particular study, the effect size was a reduction in DPB of 15 mm Hg, and the reduction was statistically significant. That is, if the drug did nothing (the assumption of the null hypothesis), we would expect to get a reduction in DBP of 15 mm Hg or higher *by chance* in only 1 of 100 similar studies. Given that low probability, we decide that the drug was probably responsible for the reduction (we "reject the null hypothesis").

Let's assume that the 15-mm Hg reduction in DBP is clinically important. Although this result is statistically significant and clinically important in this particular study, the 95% CI tells us that the reduction in DBP would probably range from 3.5 to 26.5 mm Hg in 95 of 100 similar studies. A drop of 26.5 mm Hg is clinically important, but a drop of only 3.5 mm Hg probably is not. That is, the confidence is "heterogeneous": it contains both clinically important and clinically unimportant values. So, we can't really say for sure that the drug is effective in 95 of 100 trials; our 15-mm Hg estimate is not precise enough. We need to do the study again, probably with a larger sample, to improve (narrow) the precision of the estimate. When all the values in the CI are clinically important (or when all are not clinically important)–that is, when the CI is "homogenous"–we have a more definitive answer to our question about the efficacy of the drug.

The misuse of the standard error of the mean

The SE is often used incorrectly as a descriptive statistic. Especially in the basic life sciences, measurements are routinely reported as means and SEs. This practice is established and poses no problem to those who are used to seeing measurements presented this way. However, because the SE is always smaller than the SD, it makes measurements look more precise than they would look if they were reported with SDs, so this distortion needs to be kept in mind when interpreting the SE. My research, which mostly concerns statistical reporting in clinical medicine, indicates that the SE is appropriately reported in only a few circumstances, such as in tables reporting regression analysis. The SD is preferred to describe a distribution of data, and the 95% CI is the preferred measure of precision for an estimate.

References

1. Rowntree D. Statistics without tears: an introduction for non-mathematicians. London: Penguin Books; 2000.
2. Lang TA, Secic M. How to report statistics in medicine: annotated guidelines for authors, editors, and reviewers. 2nd ed. Philadelphia: American College of Physicians; 2006.

Author information

Tom is an international consultant and educator in medical writing/editing, critical appraisal of clinical research, and scientific publications. He is past President of the Council of Science Editors and current Treasurer of the World Association of Medical Editors. He also served on the CONSORT and PRISMA groups and wrote *How to Write, Publish, and Present in the Health Sciences.*

A medical writer's guide to meta-analysis

Adam Jacobs
Premier Research, Wokingham, UK

Correspondence to:
Adam Jacobs
Senior Principal Statistician
Premier Research
Fishponds Road
Wokingham
RG41 2GY
UK
adam.jacobs@premier-research.com

Abstract
Meta-analysis is a statistical technique for summarising the results of multiple studies in a quantitative manner. It should not be confused with a systematic review, though in practice the two are often found together. The main pitfalls with meta-analyses are being sure that the studies being combined are similar enough that it makes sense to combine them, and being sure that all relevant studies have been included.

Meta-analysis is a statistical technique for combining the results of more than one study. It should be immediately obvious how useful this is: it is very rare that a single study gives us a definitive answer in medicine. To get a good idea of whether an intervention works to treat or prevent disease, or whether a particular environmental factor is associated with an increased risk of disease, for example, it is frequently necessary to take account of many studies to get a better overall picture.

By combining studies in this way, not only can we reduce the risk of being fooled by a study with unusual results as a result of a statistical fluke or bad study design, we can also get more precise estimates of the magnitude of effects. It is entirely possible, for example, that several individual studies have looked at a particular intervention but been underpowered to detect its effects, and each of them alone failed to find a significant effect, but if you combine all the studies in a meta-analysis you could find that the overall result is that a statistically significant effect can be confirmed.

Meta-analysis should not be confused with systematic review, although the two often go together. A systematic review is an attempt to find and review the entirety of literature on a particular topic using a thorough literature search, often looking for unpublished as well as published studies. This guards against any cherry-picking (at least in theory) and ensures that decisions are made on the totality of evidence.

Often, a systematic review will include a meta-analysis. Once all the relevant studies have been identified, their results can be combined using a meta-analysis to give a numerical summary. However, it is possible to do a systematic review without a meta-analysis: typically, results will be presented in narrative form with no attempt made to produce a precise numerical summary of the results. This might be done, for example, if all the studies identified had such different methods, interventions, or study populations that trying to combine them into a single estimate does not make sense.

Equally, it is possible to do a meta-

analysis without a systematic review. Sometimes studies may be chosen in a non-systematic way and yet still combined in a meta-analysis. Obviously when interpreting the results of such an analysis it is important to ask questions about what other studies might exist and why they were not included, but there may sometimes be legitimate reasons for meta-analysis of data that have not been chosen through the methods of systematic review.

One of the most important decisions for the meta-analyst is when it makes sense to combine data and when it doesn't.

I say that "in theory" a systematic review guards against cherry-picking, but in practice a systematic review is not an absolute guarantee. An important process in a systematic review is setting the inclusion criteria for the studies that will be included. There are no hard and fast rules about what inclusion criteria should be, and some judgement is always required. For example, do you require a minimum sample size for each study, and if so, what size? Will you include just trials against placebo or also trials against active comparators? Will you only include randomised trials or will you also include observational research? Should there be a minimum study duration? Will you include studies on all patients with cancer, all patients with advanced cancer, or only on those with confirmed metastatic disease? The possibilities are endless, and there are no right answers: the best choice will depend very much on individual circumstances.

And here is the problem. If you know the literature in a particular area well – as many systematic reviewers do – you will know what the important studies are. You will therefore know, when you decide on your inclusion criteria, that a particular choice of inclusion criteria will exclude specific studies that you already know about. If you have an agenda, then you can still cherry pick your data subtly by choosing inclusion criteria to exclude the studies that you don't like. So just because a systematic review has been conducted thoroughly and scrupulously in accordance with its inclusion criteria, there is still no guarantee that all relevant trials have been included. It's always worth reading the inclusion criteria carefully and making your own mind up about how reasonable they are.

One of the most important decisions for the meta-analyst is when it makes sense to combine data and when it doesn't. By combining a wide range of studies you can get apparently more statistically precise estimates, as you have more data. However, that statistical precision may be illusory. If you are investigating the efficacy of a particular treatment in different study populations, for example, an overall estimate may conceal the fact that the treatment works really well in some patients and is harmful in others. So when looking at a meta-analysis it is always worth looking at the detail of the individual studies and asking if they are investigating the same thing. If they are not, then an overall estimate may be meaningless.

Happily, this question of how comparable different studies are can be investigated statistically. A good meta-analyst will look for a measure of heterogeneity among the studies. It is expected that not all studies will give exactly the same result just because of normal random variation, but do the studies vary more than would be expected by chance? That's a simple question to ask, though not so simple to answer. Although it is possible to calculate a simple statistical test and calculate a P value, where a significant P value shows significant heterogeneity, the results of such a test are not straightforward to interpret, as there is a high risk of both false positive and false negative conclusions.

Higgins *et al.*[1] have proposed an alternative approach to quantifying heterogeneity, by calculating a measure known as the I^2 statistic, where 0 means that the studies are all identical and higher values (with a maximum of 100%) show increasing heterogeneity.

If you observe substantial heterogeneity, then it is reasonable to question the relevance of an overall estimate.

If you are looking at meta-analysis results you will come across things called "fixed effects estimates" and "random effects estimates". These are alternative statistical approaches for combining multiple studies, and are based on different assumptions.

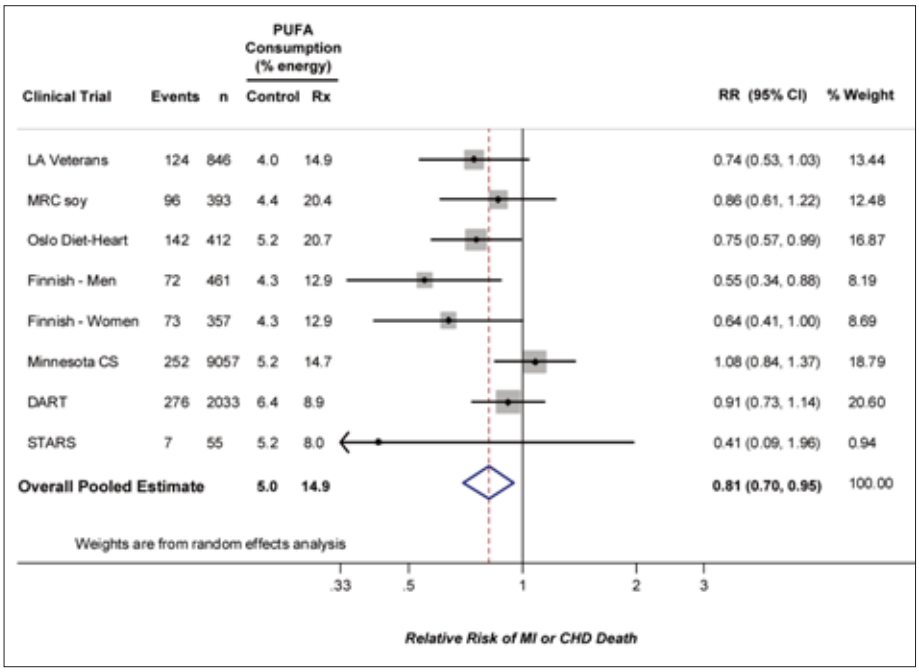
The fixed effects method makes the assumption that there is no important heterogeneity, and that all studies are essentially measuring the same thing. In other words, it assumes that any differences in estimates of treatment effects from one study to the next are due purely to statistical random variability. If in fact you observe that heterogeneity is low, then the fixed effect measure gives you a good summary of the results.

The random effects method assumes that heterogeneity is present, and the differences among studies are due partly to statistical random variability, but also due to differences in the "true" treatment effect that each study is measuring, as it is not assumed that all studies are measuring the same thing. In practical terms, the main difference between the two methods is that random effects estimation gives more weight to small studies that give different results to the average effect.

Interpreting the results of random effects meta-analyses is, as mentioned above, difficult. Although it gives you an estimate of the average effect, that treatment effect may depend on specific characteristics of the studies. If you want to apply the results to a real life situation, there is no guarantee that you will be applying it in an average situation. Your situation may match some studies far better than others.

If you observe substantial heterogeneity, then it is reasonable to question the relevance of an overall estimate.

For example, some studies may have used different doses. You may find that the high dose studies give greater treatment effects



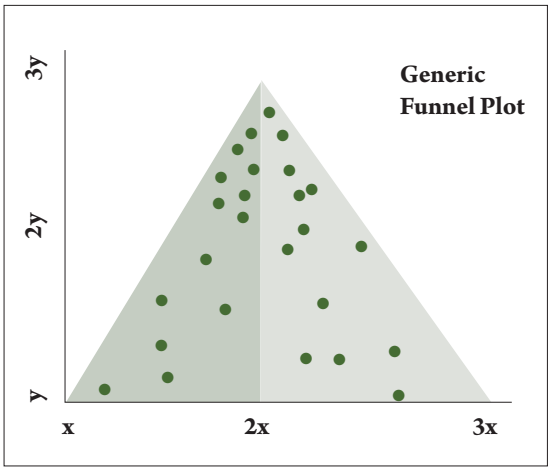
Source: Mozaffarian et al.[4]

Figure 1: Forest plot of the effects of replacing saturated fat with polyunsaturated fat on coronary heart disease. *Abbreviations: CHD: coronary heart disease; MI: myocardial infarction; PUFA: polyunsaturated fatty acid; RR: relative risk*

than the low dose studies. The relevant estimate is therefore not an average, but the treatment effect for the dose level that you are interested in. That's a fairly obvious example, but there can be many other more subtle factors that can affect treatment effects, such as the inclusion criteria for the study, treatment duration, concomitant medications, healthcare setting, etc.

One way to deal with the problem of heterogeneity is to determine the major cause of heterogeneity and to present separate estimates for different groups. For example, Annane *et al.*[2] did a systematic review and meta-analysis to investigate the effects of corticosteroids on overall mortality at 28 days in patients with severe sepsis and septic shock. Their overall meta-analyses did not find a significant effect on mortality (relative risk 0.92, 95% confidence interval 0.75 to 1.14, P = 0.46), but it also found significant heterogeneity (I^2 = 58%, P = 0.003). When they divided their studies into those that had used long courses of low dose corticosteroids or short courses of high dose corticosteroids, they found that there was indeed a significant reduction in mortality in the studies that had used long courses of low doses (relative risk: 0.80, 0.67 to 0.95, P = 0.01), but not in the studies with short courses of high doses. Ignoring the heterogeneity would have meant missing the important difference between the difference dosing regimens.

That said, use of corticosteroids in sepsis is complex and controversial, and Annane *et al*'s analysis is unlikely to be the last word. Although a meta-analysis can give more reliable results than a single study, even a meta-analysis is often not sufficient to settle a medical question once and for all. There is probably considerably more heterogeneity that needs to be unpicked in this case, including genetic features of the patient and the nature of the infecting organism.[3]



Source: Wikipedia

Figure 2: Hypothetical symmetric funnel plot

One very common way in which the results of results of meta-analyses are presented is with a graph known as a forest plot. The example in Figure 1 is typical.

This shows the results of a meta-analysis on the effects on coronary heart disease (CHD) of increasing polyunsaturated fat in place of saturated fat.[4] There is a lot of information in that one graph. We can see details of each study, including the name of the study, the number of patients, and the number of CHD events. We also see how extensive the dietary changes were in each study as figures for % polyunsaturated fatty acid consumption in the control and intervention groups. We then see the results presented both graphically and in text. The central blob of each line shows the estimated relative risk from each study, and the extent of the horizontal line shows the 95% confidence interval. The size of the central blob shows how much weight the study provides (mainly a function of the number of patients in each study), the bigger the blob, the more that study contributes to the overall analysis. We then get the same information in text form to the right of the graph.

At the bottom, we see the overall estimate. Again, we see the relative risk and its confidence interval, presented both graphically and in text form. That's the important number to take away from meta-analyses, though as stated previously, it may be hard to interpret in the presence of significant heterogeneity among studies. The forest plot gives us another means of assessing heterogeneity by simply eyeballing the spread of the estimates from the individual studies.

Lastly, no discussion of meta-analyses would be complete without a few words about publication bias. Meta-analyses will never give a true summary of all the research that has been done if some studies are excluded. We know that not all studies are published. The claim by the All Trials campaign that only 50% of studies are published is of course nonsense and the real

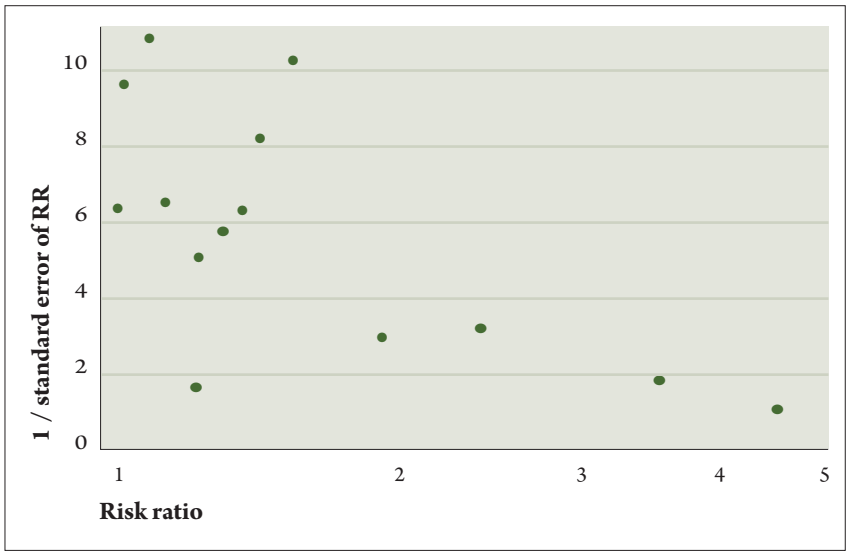


Figure 3: Funnel plot of studies investigating link between industry sponsorship and results favourable to the sponsor's product. *Abbreviations: RR: relative risk*

figure is probably much higher,[5] but nonetheless, the proportion of trials that are published is certainly less than 100%, and we know that studies reporting negative results are less likely to be published than positive studies.[6] If a meta-analysis includes only positive trials and ignores negative ones, then it will give an over-optimistic estimate of the true treatment effect.

A careful meta-analyst will therefore try to tell whether there is any evidence that publication bias has occurred. One way to do this is with a funnel plot, in which the treatment effect of individual studies is plotted on the x axis against the size of the study on the y axis. If all studies are published, the results would look roughly like an inverted funnel, with a greater spread of studies towards the bottom of the plot, where small sample sizes means that considerable variation in results is likely, and a smaller spread towards the top, where large sample sizes would keep results close to the "true" result (Figure 2).

If there is publication bias, it is likely that small negative studies will be unpublished, whereas small positive studies will be published. Large studies are more likely to be published whatever they show, as once you've gone to all the trouble of doing a large study you are more likely to be motivated to write it up. This can give rise to asymmetry in the funnel plot. Figure 3 shows one example of an asymmetric funnel plot.

I created this funnel plot from data provided in a Cochrane review of the effect of pharmaceutical industry sponsorship on publications.[7] The reviewers claimed that trials sponsored by pharmaceutical companies were more likely to be favourable to the sponsor's product than independent studies. Certainly the results of their meta-analysis showed that very strongly, but how much can we trust that result with such strong evidence of publication bias?

Meta-analysis is undoubtedly a useful technique that can provide important insights when summarising the medical literature. However, it is not a magic bullet, and must be interpreted with the same caution you would apply to any other results. Obviously if a meta-analysis is based on poor quality studies, the result will also be questionable. But in addition, it is also important to be aware of whether the studies are sufficiently similar that a meta-analysis makes sense, and crucially, whether all relevant studies have been included.

References

1. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327:557–60.
2. Annane D, Bellissant E, Bollaert E, Briegel J, Keh D, Kupfer Y. Corticosteroids for severe sepsis and septic shock: a systematic review and meta-analysis BMJ. 2004;329(7464):480.
3. Morgan, Paul (@drpaulmorgan). "@statsguyuk Very! The host-response mechanism is a mix of genetics (cytokine response) and the individual bug. Lots of interesting research!". [Tweet,10 June 2016, 8:12 AM]. Available from: https://twitter.com/drpaulmorgan/status/741166150147538944.
4. Mozaffarian D, Micha R, Wallace S. Effects on Coronary Heart Disease of Increasing Polyunsaturated Fat in Place of Saturated Fat: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. PLoS Med. 2010;7(3):e1000252.
5. Jacobs A. Zombie statistics on half of all clinical trials unpublished. [Blogpost; 29 August 2015]. Available from: http://www.statsguy. co.uk/zombie-statistics-on-half-of-all-clinical-trials-unpublished.
6. Song F, Parekh S, Hooper L, Loke YK, Ryder J, Sutton AJ, *et al.* Dissemination and publication of research findings: an updated review of related biases. Health Technol Assess. 2010;14(8).
7. Jacobs A. Cochrane review on industry sponsorship. [Blogpost; 20 December 2012]. Available from: http://dianthus.co.uk/cochrane-review-on-industry-sponsorship.

Author information

Adam Jacobs was previously a medical writer, and was president of EMWA in 2004 to 2005. He now works as a medical statistician at Premier Research. He still teaches regular workshops for EMWA on statistical topics. You can find him on Twitter at @statsguyuk.

Study design made easy

Diogo Bruno
Arteriae, Lisbon, Portugal

Correspondence to:
Diogo Bruno, Arteriae Unip. Lda.
R. Cid. Cádiz 14 4.º Dto-A, 1500-157
Lisboa, Portugal
diogobruno@arteriae.net
+351 966 625 510

Abstract
Analysis of statistical data is an important part of any medical writer's skill set, especially those professionals working in publication and regulatory areas. Understanding the various study designs is key to a thorough understanding of study methodology. Nevertheless, many medical writers come from a non-clinical background and have a knowledge gap when it comes to study design options. This article describes the main types of study design. Case report, cross-sectional, case control, cohort, quasi-experimental, randomised controlled trials, and systematic reviews and meta-analyses studies are explained and their uses, advantages, and limitations discussed.

The naked truth is that mankind still lacks time machines. If we had them, there would be no need for epidemiology as one could, for example, easily observe a group of individuals exposed to smoking over the course of their lives and, then, travel back in time and reobserve them after persuading them to stop smoking. Epidemiologists try to determine whether an exposure (i.e. risk factor) is associated with an outcome (i.e. disease), such as smoking and lung cancer in this example.[1]

The first step in a study is to define the hypothesis to be tested. After this, one must determine which study design is the most appropriate and/or feasible to test this hypothesis.

Overview of study designs
Broadly-speaking there are two approaches to study the association between an

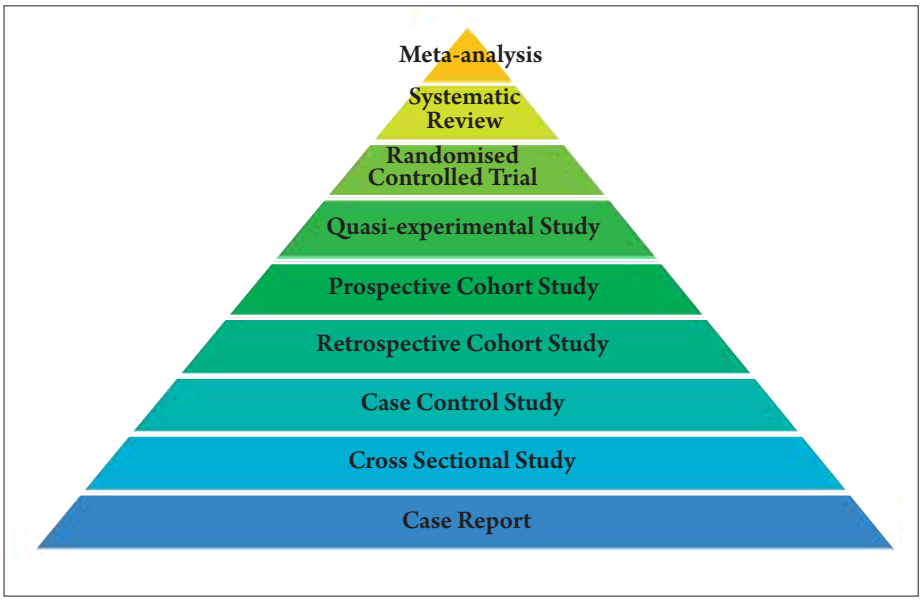


Figure 1: Main types of study design.
Study designs organised in order of statistical validity from the highest validity on the top of the pyramid to the lowest validity in the bottom of the pyramid.

exposure and an outcome: 1. **interventional or experimental** and 2. **observational or non-experimental** studies.[2] When analysing their scientific validity, experimental studies are of higher quality when compared to observational studies. They usually involve the study of a factor that can be controlled by the investigator and enrolled individuals are randomly assigned to being exposed or not to that factor. Observational studies, on the other hand, lack randomisation and, as such, various other factors might be unevenly distributed between the studied groups; as a consequence of these **confounding factors**, a true association is more difficult to ascertain.[2]

In addition, studies can be characterised as **retrospective** or **prospective** based on when the subjects are enrolled into the study.[3] These differences will be further explained when we explore cohort studies. As a consequence of these differences, study designs are often organised as a pyramid in order of validity (Figure 1).[4] Unfortunately, the most valid studies are often more expensive, more time-consuming, and more difficult to manage.

In the next sections, I will describe each study design further with a special emphasis on the most important study design for medical writers, the randomised controlled trial (RCT).

Unfortunately, the most valid studies are often more expensive, more time-consuming, and more difficult to manage.

Case report

Case report articles are considered the lowest level of evidence and findings usually require formal verification through robust epidemiological studies. However, they can represent the emergence of new issues and key ideas. Namely, they can provide important information for patient care that is not detected in clinical trials or other studies seen as more robust in design. They usually describe in detail an individual clinical case that shows: 1. a rare variation of a condition, 2. an unexpected drug adverse event, 3. clues on the pathogenesis of a disease, 4. an unexpected association between factors, 5. a unique therapy, or 6. a unique anatomical variation.

Cross-sectional study

Cross-sectional studies analyse data taken from a sample at a specific point in time. They are usually applied for public health purposes as they give a snapshot of the rate of an outcome of interest (i.e. prevalence of a condition) in a population. Moreover, researchers also describe patient characterstics and important risk factors thought to be associated with the outcome. Another use for this design is in the case of descriptive survey studies when the main aim is to describe a population in a given time period.[5]

Cross-sectional design lends itself well to **descriptive statistics**, where no association between exposure and outcome or causal relationship is sought, and the intention is solely to describe the properties of the observed data. On the other hand, **inferential statistics** aims to drive an association between exposures and outcomes through hypotheses and estimates. The designs described in the next sections are better approaches to describe these associations as cross-sectional studies give no indication of the sequence of events and are prone to prevalence-incidence bias (e.g. high mortality conditions will be under-represented as they will have low prevalence even in the case of high incidence).

Case control study

Contrary to the cross-sectional studies, the case control design aims to establish an association between risk factors and disease (i.e. uses inferential statistics). A group of patients with the study disease (**cases**) is selected and compared to a group of healthy individuals similar to the group of cases in every other aspect (**controls**). Information about risk factors is then collected retrospectively and is used to compare both groups and to find measures of association.[6]

Consequently, this design is often used to study infrequent or rare diseases in which prospective studies would be difficult to perform. To study rare diseases in prospective designs, a great number of patients would have to be enrolled, rendering them unfeasible.[7] Additionally, case control studies may have more power than cohort studies as it is easier to have larger samples.

Finally, instead of measuring the risk of disease based on exposure, we measure the odds of exposure based on disease. Therefore, relative risk is not applicable as a measure of association. It is the disease that is selected at the study onset, so the odds ratio is used.

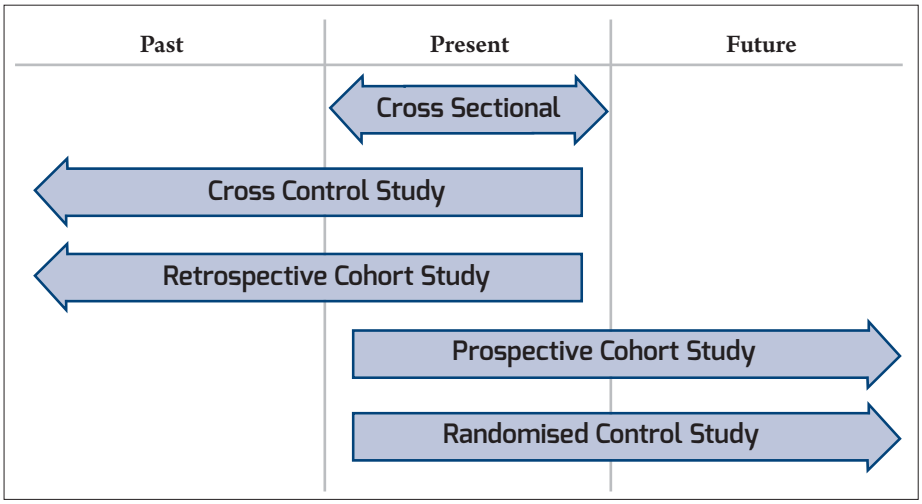


Figure 2: Retrospective and prospective study designs.
Main types of study design and their relation to the studied time-points.

Cohort study

By definition, in cohort studies a group of subjects comprising a sample deemed to represent the population of interested is followed over time whilst collecting data on risk factors and outcomes. It differs from cross-sectional studies in the sense that, although risk factors and outcomes are studied, subjects are studied over time. Moreover, unlike case control studies, individuals are disease-free at the outset of the study and the risk of development of the disease (outcome) is the measure of interest.

As discussed previously, cohort studies can either be retrospective or prospective when relating to time of subject enrolment (Figure 2). **Retrospective** studies are also called **historical cohort studies** and study events from the past up until the present time. The obvious advantage is that the information is readily available, however tracing subjects might prove difficult and investigators have to rely on the quality of the recorded information (e.g. electronic health records, patient recollection) which is often low.[3]

On the other hand, prospective cohort studies are those studying events from the present time until a time in the future. The study design allows investigators to incorporate any exposure or baseline characteristics to be studied so the study is more complete; however, the follow-up time can be long, especially for infrequent outcomes, and such studies can have a high loss to follow-up (dropout) rate.

Less often, **ambispective** or **ambidirectional cohort studies** are performed that, as the name implies, combine retrospective and prospective information including past, present, and future timepoints.[8]

Regarding the data analysis, unlike case control studies, the most usual measure is the risk ratio of the outcome of interest, calculated by the risk of the outcome in exposed subjects relative to those not exposed to the risk factor.

Quasi-experimental study

Quasi-experimental studies are sometimes also called **nonrandomised** or **pre-post intervention studies** and are used to evaluate the effects of specific interventions or policy changes. It is a design chosen when it is not logistically feasible or ethical to conduct a RCT. By definition, these studies lack randomisation and are conducted after a policy change comes into effect. As policy changes can inadvertently be non-beneficial, investigators compare specific outcomes before and after a policy change to determine if it was of benefit.[9]

Unlike previous designs, these studies include an intervention chosen by the researchers (policy change at a given time-point). However, they lack randomisation and, consequently, are sensitive to confounding effects and causal relationships, and are, therefore, less valid than RCTs.[9]

Randomised controlled trial (RCT)

RCTs are the only studies truly experimental in nature as the effect of an investigator-chosen intervention is studied in randomly assigned subjects from a study sample deemed to be representative of a population. Frequently, a study group is exposed to an intervention (e.g. a novel treatment) and its effects on one or more outcomes of interest are studied by comparing the exposed group to a control group not exposed to the intervention (e.g. placebo) or exposed to a standard intervention (e.g. already established therapy or standard of care).[10] Figure 3 gives a representation of this study design.

The RCT's inherent characteristics make it a robust study design. **Randomisation** of intervention allocation is used to decrease confounding effects and allocation bias. The characteristics that might affect the relationship between intervention and outcome measures will be roughly equal between study and control groups. **Blinding** or **masking** is also frequently used to decrease study bias. In single-blinded studies, subjects are unaware of their group assignment, decreasing the performance bias that could occur as this knowledge can affect the subjects' response to the intervention. In double-blinded studies, group allocation is not known to both study subjects and investigators. This further decreases bias by avoiding differences in treatment administration between treatment arms (performance bias) or the over- or under-estimation of the effects of an intervention (assessment bias).[11] Studies that have no blinding are characterised as **open-label**.

The RCT allows investigators to control the intervention and establish causality with a good degree of certainty providing the strongest evidence of an association (efficacy or safety data). However, to calculate the sample size, researchers must have prior knowledge about the expected effect size and sometimes ethical issues prevent the comparison of an intervention with a **placebo** (an inert treatment in blinded

RCTs frequently give rise to higher quality evidence, however researchers and medical writers have to be aware of the limitations of RCTs when analysing the results

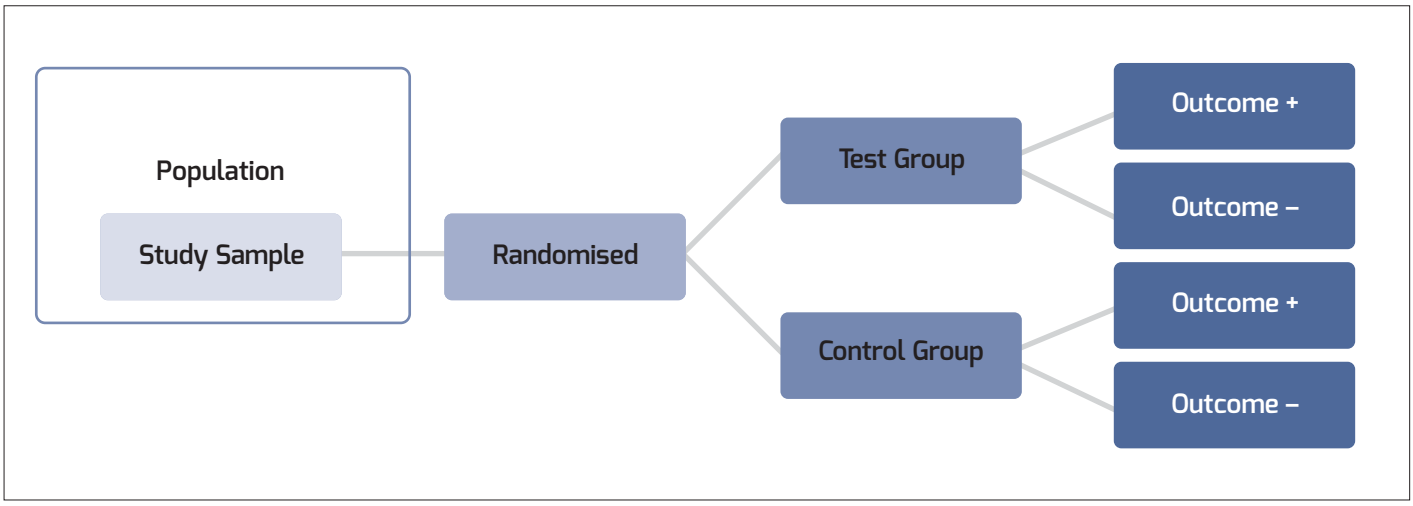


Figure 3: Schematic representation of a randomised controlled trial.
Subjects in a sample are randomly assigned to different interventions (test and control groups) and then followed to compare the risk of the outcome between groups.

studies) or no intervention (in open-label studies).[12]

RCTs of new drugs are often classified in **phases**. Phase I trials involve testing in healthy volunteers (except for novel oncology drugs) with dose escalation to assess safety (i.e. side effects and toxicity) and to determine if it is appropriate to check for efficacy. Phase II trials involve a small group of patients to assess safety and efficacy. Phase III trials involve a large group of patients to further assess and establish safety, efficacy, and effectiveness. Phase IV trials are those performed during post-marketing surveillance.[13]

Systematic review and meta-analysis

Systematic reviews are studies that try to collect all available evidence about a subject of interest and critically appraise it. Systematic reviews of RCTs are often used to guide guidelines and other aspects of evidence-based medicine. They usually involve a thorough search on a research question in multiple article databases and indexes, such as Web of Science, Embase, and PubMed.[14]

Meta-analyses not only try to collect all available evidence but also combine the results of similar papers to give an approximate pooled measure of association (e.g. odds ratio or risk ratio) using specific statistical methodology. In summary, systematic reviews give a qualitative evaluation, whilst meta-analyses aim at a quantitative appraisal.[14] While meta-analyses virtually always include systematic reviews of the literature, this is not always true for systematic reviews.

Discussion and conclusion

It is true that observational studies are more prone to bias and confounding effects than RCTs and that RCTs frequently give rise to higher quality evidence, however researchers and medical writers have to be aware of the limitations of RCTs when analysing the results. They have more **internal validity** when compared to prospective cohort studies; that is, the causal inference or relation is properly demonstrated in the study sample, as bias and confounding factors are often adequately controlled. However, the strategies aimed at increasing internal validity, such as controlling the intervention and the strict inclusion and exclusion criteria, may undermine the **external validity** or **generalisability** of the study findings. One pivotal example is the finding of increased rates of coronary heart disease (CHD) in postmenopausal women taking hormone replacement therapy in the Women's Health Initiative (WHI) RCT.[15] Contrary to the WHI, two previous prospective cohort studies based on the Nurses' Health Study Cohort suggested a reduced CHD risk.[16,17] At first glance, differences were attributed to a lack of randomisation in the observational studies. As a consequence of the WHI study, millions of women worldwide stopped taking hormone replacement therapy. More recent studies, however, attribute the differences to a lack of external validity of the WHI study that had an older study population (average age of 63 vs. 57-59 years) and a higher percentage of users who had gone through menopause more than 10 years previously.[18]

In the end, the pyramid shown in Figure 1 stands true for most study examples and studies higher in the pyramid have more valid and robust findings. However, medical writers have an increasingly active role as consultants and in literature review to properly counsel clients in strategies and evidence-based medicine. A sound knowledge of statistical methods is therefore essential for any contemporary medical writer, and understanding the key advantages and limitations of the several study designs presently at our disposal is another small step to achieve that goal.

References

1. Adami HO, Hunter D, Trichopoulos D, editors. Textbook of Cancer Epidemiology. 2nd ed. Oxford: Oxford University Press; 2009.
2. dos Santos Silva I. Cancer Epidemiology: Principles and Methods. 1st ed. Lyon: IARC Press; 1999.
3. Levin KA. Study design IV. Cohort studies. Evid Based Dent. 2003;7(2):51–2.

4. Haynes B. Of studies, syntheses, synopses, summaries, and systems: the "5S" evolution of information services for evidence-based healthcare decisions. Evid Based Nurs. 2007;10(1):6–7.
5. Levin KA. Study design III: Cross-sectional studies. Evid Based Dent. 2006;7(1):24–5.
6. Levin KA. Study design V. Case-control studies. Evid Based Dent. 2006;7(3):83–4.
7. Schulz KF, Grimes DA. Case-control studies: research in reverse. Lancet. 2002;359(9304):431–4.
8. Mortensen PB. Ambidirectional studies – an extension of longitudinal studies in psychiatry. Psychiatr Dev. 1988;6(2):173–81.
9. Harris AD, Bradham DD, Baumgarten M, Zuckerman IH, Fink JC, Perencevich EN. The use and interpretation of quasi-experimental studies in infectious diseases. Clin Infect Dis. 2004;38(11):1586–91.
10. Levin KA. Study design VII. Randomised controlled trials. Evid Based Dent. 2007;8(1):22–3.
11. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med. 2013;158(3):200–7.
12. Wechsler ME, Kelley JM, Boyd IO, Dutile S, Marigowda G, Kirsch I, et al. Active albuterol or placebo, sham acupuncture, or no intervention in asthma. N Engl J Med. 2011;365(2):119–26.
13. Friedman LM, Furberg CD, DeMets DL. Fundamentals of clinical trials. New York: Springer; 2010.
14. Higgins J, Churchill R, Cumpston M, Chandler J, editors. Cochrane Handbook for Systematic Reviews of Interventions. Ver. 5.0.2. New Jersey: Wiley-Blackwell; 2008.
15. Group TWHIS. Design of the Women's Health Initiative clinical trial and observational study. Control Clin Trials. 1998;19(1):61–109.
16. Grodstein F, Manson JE, Colditz GA, Willett WC, Speizer FE, Stampfer MJ. A prospective, observational study of postmenopausal hormone therapy and primary prevention of cardiovascular disease. Ann Intern Med. 2000;133(12):933–41.
17. Grodstein F, Stampfer MJ, Manson JE, Colditz GA, Willett WC, Rosner B, *et al.* Postmenopausal estrogen and progestin use and the risk of cardiovascular disease. N Engl J Med. 1996;335(7):453–61.
18. Hernán MA, Alonso A, Logan R, Grodstein F, Michels KB, Willett WC, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. Epidemiology. 2008; 19(6):766–79.

Acknowledgements

The author would like to express gratitude to Joana Carlos for the proofreading of this article.

Conflicts of Interest and Disclaimers

The author declares no conflicts of interest.

Author information

Diogo Bruno, MD has been a medical writer, translator and proofreader for three years. He is a senior OB/GYN Specialty Registrar at *Hospital Prof. Dr. Fernando Fonseca*, Portugal, has a Harvard Medical School certificate in clinical research, and has presented works in obstetrics and gynaecology at several meetings.

Statistical analyses and methods in the published literature: The SAMPL guidelines

Thomas A. Lang[1] and **Douglas G. Altman**[2]

1 Tom Lang Communications and Training International, Kirkland, WA, USA

2 Centre for Statistics in Medicine, Oxford University, Oxford, UK

Correspondence to:

Tom Lang
10003 NE 115th Lane
Kirkland, WA 98933
USA
tomlangcom@aol.com
+1 425 242 1370

Abstract

Despite calls for guidelines on reporting statistical aspects of studies, most journals have still not included in their instructions for authors more than a paragraph or two about reporting statistical methods and results. However, given that many statistical errors concern basic statistics, a comprehensive – and comprehensible – set of reporting guidelines might improve how statistical analyses are documented. The SAMPL guidelines are designed to be included in a journal's Instructions for Authors. These guidelines tell authors, journal editors, and reviewers how to report basic statistical methods and results. Although these guidelines are limited to the most common statistical analyses, they are nevertheless sufficient to prevent most of the reporting deficiencies routinely found in scientific articles.

This article is an update of an article originally published in the *Science Editors' Handbook* (Lang T, Altman D. Basic statistical reporting for articles published in clinical medical journals: the SAMPL Guidelines. In: Smart P, Maisonneuve H, Polderman A (eds). Science Editors' Handbook. European Association of Science Editors; 2013). References documenting the incidence of statistical errors have been updated in this revision.

Have they reflected that the sciences founded on observation can only be promoted by statistics? … If medicine had not neglected this instrument, this means of progress, it would possess a greater number of positive truths, and stand less liable to the accusation of being a science of unfixed principles, vague and conjectural.
Jean-Etienne Dominique Esquirol, an early French psychiatrist, quoted in *The Lancet*, 1838.[1]

Introduction

The first major study of the quality of statistical reporting in the biomedical literature was published in 1966.[2] Since then, dozens of similar studies have been published, every one of which has found that large proportions of articles contain errors in the application, analysis, interpretation, or reporting of statistics or in the design or conduct of research. (See, for example, references 3 through 19.) Further, large proportions of these errors are serious enough to call the authors' conclusions into

question.[5,18,19] The problem is made worse by the fact that most of these studies are of the world's leading peer-reviewed general medical and specialty journals.

Although errors have been reported for more complex statistical procedures,[19-22] paradoxically, many errors are in basic, not advanced, statistical methods.[23] Perhaps advanced methods are suggested by consulting statisticians, who perform the analyses competently, but it is also true that authors are far more likely to use only elementary statistical methods, if they use any at all.[23-26] Still, articles with even major errors continue to pass editorial and peer review and to be published in leading journals.

The truth is that the problem of poor statistical reporting is long-standing, widespread, potentially serious, concerns mostly basic statistics, and yet is largely unsuspected by most readers of the biomedical literature.[27]

More than 30 years ago, O'Fallon and colleagues recommended that "Standards governing the content and format of statistical aspects should be developed to guide authors in the preparation of manuscripts."[28] Despite the fact that this call has since been echoed by several others,[29-32] most journals have still not included in their Instructions for Authors more than a paragraph or two about reporting statistical methods and results.[33] However, given that many statistical errors concern basic statistics, a comprehensive – and comprehensible – set of reporting guidelines might improve how statistical analyses are documented.

The SAMPL guidelines are designed to be included in a journal's Instructions for Authors. These guidelines tell authors, journal editors, and reviewers how to report basic statistical methods and results. Although these guidelines are limited to the most common statistical analyses, they are nevertheless sufficient to prevent most of the reporting deficiencies routinely found in scientific articles.

Unlike most of the other guidelines in this book, the SAMPL guidelines were not developed by a formal consensus-building process, but they do draw considerably from published guidelines.[27,34-37] In addition, a comprehensive review of the literature on statistical reporting errors reveals near universal agreement on how to report the most common methods.[27]

Statistical analyses are closely related to the design and activities of the research itself. However, we do not address these issues here. Instead, we refer readers to the EQUATOR Network website (www.equator-network.org) where guidelines for reporting specific research designs can be found. (For example, see CONSORT,[38] TREND,[39] and STROBE[40]) These guidelines for reporting methodologies all include items on reporting statistics, but the guidelines presented here are more specific and complement, not duplicate, those in the methodology guidelines.

We welcome feedback and anticipate the need to update this guidance in due course.

The truth is that the problem of poor statistical reporting is long-standing, widespread, potentially serious, concerns mostly basic statistics, and yet is largely unsuspected by most readers of the biomedical literature.

Guiding principles for reporting statistical methods and results

Our first guiding principle for statistical reporting comes from The International Committee of Medical Journal Editors, whose Uniform Requirements for Manuscripts Submitted to Biomedical Journals include the following excellent statement about reporting statistical analyses:

"Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results. [Emphasis added.] When possible, quantify findings and present them with appropriate indicators of measurement error or uncertainty (such as confidence intervals). Avoid relying solely on statistical hypothesis testing, such as *P* values, which fail to convey important information about effect size. References for the design of the study and statistical methods should be to standard works when possible (with pages stated). Define statistical terms, abbreviations, and most symbols. Specify the computer software used."[33,41]

Our second guiding principle for statistical reporting is to **provide enough detail that the results can be incorporated into other analyses.** In general, this principle requires reporting the descriptive statistics from which other statistics are derived, such as the numerators and denominators of percentages, especially in risk, odds, and hazards ratios. Likewise, *P* values are not sufficient for re-analysis. Needed instead are descriptive statistics for the variables being compared, including sample size of the groups involved, the estimate (or "effect size") associated with the P value, and a measure of precision for the estimate, usually a 95% confidence interval.

General principles for reporting statistical methods

Preliminary analyses

- Identify any statistical procedures used to modify raw data before analysis.

Examples include mathematically transforming continuous measurements to make distributions closer to the normal distribution, creating ratios or other derived variables, and collapsing continuous data into categorical data or combining categories.

Primary analyses

- Describe the purpose of the analysis.
- Identify the variables used in the analysis and summarize each with descriptive statistics.
- When possible, identify the smallest difference considered to be clinically important.
- Describe fully the main methods for analysing the primary objectives of the study.
- Make clear which method was used for each analysis, rather than just listing in one place all the statistical methods used.
- Verify that that data conformed to the assumptions of the test used to analyse them. In particular, specify that 1. skewed data were analysed with non-parametric tests, 2. paired data were analysed with paired tests, and 3. the underlying relationship analysed with linear regression models was linear.
- Indicate whether and how any allowance or adjustments were made for multiple comparisons (performing multiple hypothesis tests on the same data).
- If relevant, report how any outlying data were treated in the analysis.
- Say whether tests were one- or two-tailed and justify the use of one-tailed tests.
- Report the alpha level (e.g., 0.05) that defines statistical significance.
- Name the statistical package or programme used in the analysis.

Supplementary analyses

- Describe methods used for any ancillary analyses, such as sensitivity analyses, imputation of missing values, or testing of assumptions underlying methods of analysis.
- Identify post-hoc analyses, including unplanned subgroup analyses, as exploratory.

General principles for reporting statistical results

Reporting numbers and descriptive statistics

- Report numbers – especially measurements – with an appropriate degree of precision. For ease of comprehension and simplicity, round to a reasonable extent. For example, mean age can often be rounded to the nearest year without compromising either the clinical or the statistical analysis. If the smallest meaningful difference on a scale is 5 points, scores can be reported as whole numbers; decimals are not necessary.
- Report total sample and group sizes for each analysis.
- Report numerators and denominators for all percentages.
- Summarise data that are approximately normally distributed with means and standard deviations (SD). Use the form: mean (SD), not mean ± SD.
- Summarise data that are not normally distributed with medians and inter-percentile ranges, ranges, or both. Report the upper and lower boundaries of inter-percentile ranges and the minimum and maximum values of ranges, not just the size of the range.
- Do NOT use the standard error of the mean (SE) to indicate the variability of a data set. Use standard deviations, inter-percentile ranges, or ranges instead. (The SE is an inferential statistic – it is about a 68% confidence interval – not a descriptive statistic.)
- Display data in tables or figures. Tables present exact values, and figures provide an overall assessment of the data.[42,43]

Reporting risk, rates, and ratios

- Identify the type of rate (e.g., incidence rates; survival rates), ratio (e.g., odds ratios; hazards ratios), or risk (e.g., absolute risks; relative risk differences), being reported.
- Identify the quantities represented in the numerator and denominator (e.g., the number of men with prostate cancer divided by the number of men in whom prostate cancer can occur).
- Identify the time period over with each rate applies.
- Identify any unit of population (that is, the unit multiplier: e.g., x 100; x 10,000) associated with the rate.
- Consider reporting a measure of precision (a confidence interval) for estimated risks, rates, and ratios.

Reporting hypothesis tests

- State the hypothesis being tested.
- Identify the variables in the analysis and summarize the data for each variable with the appropriate descriptive statistics.
- If possible, identify the minimum difference considered to be clinically important.
- For equivalence and non-inferiority studies, report the largest difference between groups that will still be accepted as indicating biological equivalence (the equivalence margin).
- Identify the name of the test used in the analysis. Report whether the test was one- or two-tailed (justify the use of one-tailed tests) and for paired or independent samples.
- Confirm that the assumptions of the test were met by the data.
- Report the alpha level (e.g., 0.05) that defines statistical significance.
- At least for primary outcomes, such as differences or agreement between groups, diagnostic sensitivity, and slopes of regression lines, report a measure of precision, such as the 95% confidence interval.
- Do NOT use the standard error of the mean (SE) to indicate the precision of an estimate. The SE is essentially a 68% confidence coefficient: use the 95% confidence coefficient instead.
- Although not preferred to confidence intervals, if desired, *P* values should be reported as equalities when possible and to one or two decimal places (e.g., $P = 0.03$ or 0.22 not as inequalities: e.g., $P < 0.05$). Do NOT report "NS"; give the actual P value. The smallest *P* value that need be reported is $P < 0.001$, save in studies of genetic associations.

- Report whether and how any adjustments were made for multiple statistical comparisons.
- Name the statistical software package used in the analysis.

Reporting association analyses

- Describe the association of interest.
- Identify the variables used and summarise each with descriptive statistics.
- Identify the test of association used.
- Indicate whether the test was one- or two-tailed. Justify the use of one-tailed tests.
- For *tests* of association (e.g., a *chi*-square test), report the *P* value of the test (because association is defined as a statistically significant result).
- For *measures* of association (i.e., the *phi* coefficient), report the value of the coefficient and a confidence interval. Do not describe the association as low, moderate, or high unless the ranges for these categories have been defined. Even then, consider the wisdom of using these categories given their biological implications or realities.
- For primary comparisons, consider including the full contingency table for the analysis.
- Name the statistical package or program used in the analysis.

Reporting correlation analyses

- Describe the purpose of the analysis.
- Summarise each variable with the appropriate descriptive statistics.
- Identify the correlation coefficient used in the analysis (e.g., Pearson, Spearman).
- Confirm that the assumptions of the analysis were met.
- Report the alpha level (e.g., 0.05) that indicates whether the correlation coefficient is statistically significant.
- Report the value of the correlation coefficient. Do not describe correlation as low, moderate, or high unless the ranges for these categories have been defined. Even then, consider the wisdom of using these categories given their biological implications or realities.
- For primary comparisons, report the (95%) confidence interval for the correlation coefficient, whether or not it is statistically significant.
- For primary comparisons, consider reporting the results as a scatter plot. The sample size, correlation coefficient (with its confidence interval), and *P* value can be included in the data field.
- Name the statistical package or programme used in the analysis.

Reporting regression analyses

- Describe the purpose of the analysis.
- Identify the variables used in the analysis and summarize each with descriptive statistics.
- Confirm that the assumptions of the analysis were met. For example, in linear regression indicate whether an analysis of residuals confirmed the assumptions of linearity.
- If relevant, report how any outlying values were treated in the analysis.
- Report how any missing data were treated in the analyses.
- For either simple or multiple (multivariable) regression analyses, report the regression equation.
- For multiple regression analyses: 1. report the alpha level used in the univariate analysis; 2. report whether the variables were assessed for a. co-linearity and b. interaction; and 3. describe the variable selection process by which the final model was developed (e.g., forward-stepwise; best subset).
- Report the regression coefficients (beta weights) of each explanatory variable and the associated confidence intervals and *P* values, preferably in a table.
- Provide a measure of the model's "goodness-of-fit" to the data (the coefficient of determination, r^2, for simple regression and the coefficient of multiple determination, R^2, for multiple regression).
- Specify whether and how the model was validated.
- For primary comparisons analysed with simple linear regression analysis, consider reporting the results graphically, in a scatter plot showing the regression line and its confidence bounds. Do not extend the regression line (or the interpretation of the analysis) beyond the minimum and maximum values of the data.
- Name the statistical package or programme used in the analysis.

Reporting analyses of variance (ANOVA) or of covariance (ANCOVA)

- Describe the purpose of the analysis.
- Identify the variables used in the analysis and summarize each with descriptive statistics.
- Confirm that the assumptions of the analysis were met. For example, indicate whether an analysis of residuals confirmed the assumptions of linearity.
- If relevant, report how any outlying data were treated in the analysis.
- Report how any missing data were treated in the analyses.
- Specify whether the explanatory variables were tested for interaction, and if so how these interactions were treated.
- If appropriate, in a table, report the *P*

value for each explanatory variable, the test statistics and, where applicable, the degrees of freedom for the analysis.
- Provide an assessment of the goodness-of-fit of the model to the data, such as R^2.
- Specify whether and how the model was validated.
- Name the statistical package or programme used in the analysis.

Reporting survival (time-to-event) analyses
- Describe the purpose of the analysis.
- Identify the dates or events that mark the beginning and the end of the time period analysed.
- Specify the circumstances under which data were censored.
- Specify the statistical methods used to estimate the survival rate.
- Confirm that the assumptions of survival analysis were met.
- For each group, give the estimated survival probability at appropriate follow-up times, with confidence intervals, and the number of participants at risk for death at each time. It is often more helpful to plot the cumulative probability of not surviving, especially when events are not common.
- Reporting median survival times, with confidence intervals, is often useful to allow the results to be compared with those of other studies.
- Consider presenting the full results in a graph (e.g., a Kaplan-Meier plot) or table.
- Specify the statistical methods used to compare two or more survival curves.
- When comparing two or more survival curves with hypothesis tests, report the P value of the comparison
- Report the regression model used to assess the associations between the explanatory variables and survival or time-to-event.
- Report a measure of risk (e.g., a hazard ratio) for each explanatory variable, with a confidence interval.

Reporting Bayesian analyses
- Specify the pre-trial probabilities ("priors").
- Explain how the priors were selected.
- Describe the statistical model used.
- Describe the techniques used in the analysis.
- Identify the statistical software program used in the analysis.
- Summarise the posterior distribution with a measure of central tendency and a credibility interval
- Assess the sensitivity of the analysis to different priors.

References

1. Esquirol JED. Cited in: Pearl, R. Introduction to Medical Biometry and Statistics. WB Saunders, Philadelphia; 1941.
2. Schor S, Karten I. Statistical evaluation of medical journal manuscripts. J Am Med Assn. 1966;195:1123-8.
3. Prescott RJ, Civil I. Lies, damn lies and statistics: errors and omission in papers submitted to Injury 2010-2012. Injury. 2013;44:6-11.
4. Fernandes-Taylor S, Hyun JK, Reeder RN, Harris AH. Common statistical and research design problems in manuscripts submitted to high-impact medical journals. BMC Res Notes. 2011;4:304.
5. Bosker T, Mudge JF, Munkittrick KR. Statistical reporting deficiencies in environmental toxicology. Environ Tox Chem. 2013;32(8):1737-9.
6. Vesterinen HM, Egan K, Deister A, Schlattmann P, Macleod MR, Dirnagl U. Systematic survey of the design, statistical analysis, and reporting of studies published in the 2008 volume of the Journal of Cerebral Blood Flow and Metabolism. J Cerebr Blood Flow Metab. 2011;31:1064-72.
7. Kim JS, Kim DK, Hong SJ. Assessment of errors and misused statistics in dental research. Int Dent J. 2011;61:163-7.
8. Lee HJ, Jung SK. The use of statistical methodology in articles in medical journals and suggestions for the quality improvement of the Pediatric Allergy and Respiratory Disease. Pediatr Allergy Respir Dis. 2011;21:144-55.
9. Yim KH, Nahm FS. Han KA, Park SY Analysis of statistical methods and errors in the articles published in the Korean Journal of Pain. Kor J Pain. 2010;23:35-41.
10. Al-Benna S, Al-Ajam Y, Way B, Steinstraesser L. Descriptive and inferential statistical methods used in burns research. Burns. 2010;36(3):343-6.
11. Robinson PM, Menakuru S, Reed MW, Balasubramanian SP. Description and reporting of surgical data – scope for improvement? Surgeon. 2009;7(1);6-9.
12. Afshar K, Jafari S, Seth A, Lee JK, MacNeily AE. Publications by the American Academy of Pediatrics Section on Urology: the quality of research design and statistical methodology. J Urol. 2009;182(4 Suppl):1906-10.
13. Barbosa FT, de Souza DA. Frequency of the adequate use of statistical tests of hypothesis in original articles published in the Revista Brasileira de Anestesiologia between January 2008 and December 2009. Rev Brasil Anestesiol. 2010;60:528-36.
14. Neville JA, Lang W, Fleischer AB. Errors in the Archives of Dermatology and the Journal of the American Academy of Dermatology from January through December 2003. Arch Dermatol. 2006;142:737-40.
15. Kurichi JE, Sonnad SS. Statistical methods in the surgical literature. J A Coll Surg. 2006;202:476-84.
16. Scales CD, Norris RD, Preminger GM, Vieweg J, Peterson BL, Dahm P. Evaluating the evidence: statistical methods in randomized controlled trials in the urological literature. J Urol. 208;180 1463-7.
17. Gaskin CJ, Happell B. Power, effects, confidence, and significance: an investigation of statistical practices in nursing research. Int J Nurs Stud. 2014;51(5):795-806.
18. Jaykaran YP. Quality of reporting statistics in two Indian pharmacology journals. J Pharmacol Pharmacother. 2011;2(2):85-9.

19. Mikolajczyk RT, DiSilvestro A, Zhang J. Evaluation of logistic regression reporting in current obstetrics and gynecology literature. Obst Gynecol. 2008;111(2 Pt 1):413-9.
20. Burton A, Altman DG. Missing covariate data within cancer prognostic studies: a review of current reporting and proposed guidelines. Br J Cancer. 2004;91:4-8.
21. Mackinnon, A. The use and reporting of multiple imputation in medical research – a review. J Intern Med. 2010;268:586-93.
22. Abraira V, Muriel A, Emparanza JI, Pijoan JI, Royuela A, Plana MN. Cano A, Urreta I, Zamora J. Reporting quality of survival analyses in medical journals still needs improvement. A minimal requirements proposal. J Clin Epidemiol. 2013;66:1340-6.
23. Kim M. Statistical methods in Arthritis & Rheumatism – current trends. Arthr Rheum. 2006;54:3741-9.
24. Reed JF, Salen P, Bagher P. Methodological and statistical techniques: what do residents really need to know about statistics? J Med Syst. 2003;27(3):233-8.
25. Aljoudi AS. Study designs and statistical methods in the Journal of Family and Community Medicine: 1994-2010. J Fam Commun Med. 2013;20(1):8-11
26. Lee CM, Soin HK, Einarson TR. Statistics in the pharmacy literature. Ann Pharmacother. 2004;38(9): 1412–8
27. Lang T, Secic M. How to Report Statistics in Medicine: Annotated Guidelines for Authors, Editors, and Reviewers, 2nd edn. American College of Physicians, Philadelphia; 2006.
28. O'Fallon, J.R., Duby, S.D., Salsburg, D.S. et al. Should there be statistical guidelines for medical research papers? Biometrics. 1978;34:687–95.
29. Shott S. Statistics in veterinary research. J Am Vet Med Association 1985; 187:138–41.
30. Hayden GF. Biostatistical trends in Pediatrics: implications for the future. Pediatrics. 1983;72:84–7.
31. Altman DG, Bland JM. Improving doctors' understanding of statistics. J Royal Stat Soc SerA. 1991;154, 223–67.
32. Altman DG, Gore SM, Gardner MJ, Pocock SJ. Statistical guidelines for contributors to medical journals. BMJ. 1983;286, 1489–93.
33. Bailar JC, Mosteller F. Guidelines for statistical reporting in articles for medical journals. Amplifications and explanations. Ann Intern Med. 1988;108 (2), 266–73.
34. Bond GR. Mintz J, McHugo GJ. Statistical guidelines for the Archives of PM&R. Arch Phys Med Rehab. 1995;76:784–7.
35. Wilkinson L, Task Force on Statistical Inference. Statistical methods in psychology journals. Guidelines and explanations. Am Psych. 1999;54: 594–604.
36. Curran-Everett D, Benos DJ. American Physiological Society. Guidelines for reporting statistics in journals published by the American Physiological Society. Am J Physiol Endocrinol Metab. 2004;287: E189–91.
37. Curran-Everett D, Benos DJ. Guidelines for reporting statistics in journals published by the American Physiological Society: the sequel. Adv Physiol Ed. 2007;31:295–8.
38. Moher D, Schulz K, Altman DG, for the CONSORT Group. CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. Ann Intern Med. 2001;134:657–62.
39. DesJarlais DC, Lyles C, Crepaz N, Trend Group. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Publ Health. 2004;94(3):361–6.
40. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. Ann Inter Med. 2007;147 (8), 573–7.
41. International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals: writing and editing for biomedical publication. 2011 [cited 12 Dec 2012]. Available from: http://www.icmje.org.
42. Schriger DL, Arora S, Altman, DG. The content of medical journal instructions for authors. Ann Emerg Med. 2006;48:743–9.
43. Lang T. How to Write, Publish, and Present in the Health Sciences: A Guide for Clinicians and Laboratory Researchers. Philadelphia: American College of Physicians; 2010.

How to interpret and report the results from multivariable analyses

Neus Valveny and **Stephen Gilliver**

TFS Develop, Medical Writing Unit, Barcelona, Spain / Lund, Sweden

Correspondence to:

Neus Valveny
TFS Develop
Consell de Cent, 334-336, 4th floor
ES-08009 Barcelona, Spain
+ 34 617 414 124
neus.valveny@tfscro.com

Abstract

Multivariable analyses are some of the central statistical methods of clinical trials, and yet some medical writers may be unsure as to what they are and how best to interpret and report the results. In this article we provide an overview of multivariable analyses, introducing some of the core models biostatisticians use to analyse trial data. We focus on odds ratios, hazard ratios, and β coefficients as key parameters and provide guidance on important considerations when reporting them.

What is a multivariable analysis?

Univariate analyses – analyses involving only a single variable – are descriptive by nature. They allow us to describe the distribution of a variable in a sample of *n* individuals or *n* tumour biopsies, for example. In univariate analyses we commonly use parameters such as the median, mean, and standard deviation to describe *quantitative* (or *continuous*) *variables* and frequencies and percentages to describe *categorical variables*. We can also estimate population parameters by calculating 95% *confidence intervals* (CIs) for the aforementioned summary statistics (median, mean, percentage). With univariate analyses we can only answer "descriptive questions" in a single arm or cohort, such as "What is the rate of responders to drug X?" or "What is the mean survival time in patients treated with drug Y?"

But what about situations where we wish to analyse more than one variable at a time? The purpose of *bivariate* and *multivariable analyses* is to probe the relationships between two (bivariate) or more than two (multivariable) variables. These types of analyses allow us to test a previously defined hypothesis (e.g. the primary efficacy analysis of a confirmatory study) or to explore the existing relationships between the collected variables (e.g. between-arm analyses, subgroup analyses, exploratory analyses). With bivariate and multivariable analyses we can answer "analytical questions" in one or more cohorts, such as "What is the overall survival with drug X compared with drug Y?", "What is the efficacy of drug Z, based on the reduction in cholesterol levels, compared with placebo?", or "What is the relationship between response rate to drug X and the level of biomarker Y?"

In both bivariate and multivariable analyses the participating variables can be classified into:

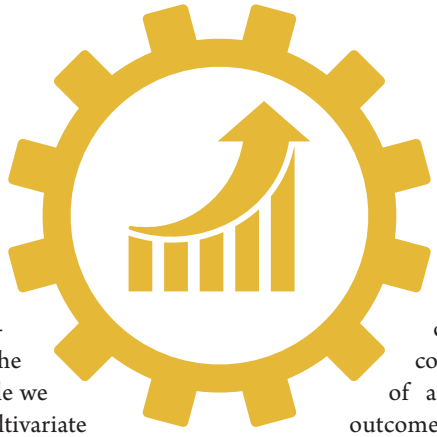
- *Dependent* (or *outcome* or *predicted*) *variables* and
- *Independent* (or *predictor* or *explanatory*) *variables*, which in some models can be further classified into factors and covariates (or confounding factors).

In a bivariate analysis (sometimes referred to as *univariate* – see Box 1 below) there is only one independent and one dependent variable.

In a multivariable analysis there are:

- **One** dependent variable and
- **Two or more** independent variables.

BOX 1: Bivariate analyses that analyse the relationship between one independent variable and one dependent variable are often referred to as "univariate" analyses to distinguish them from multivariable analyses, in which two or more independent variables are assessed in relation to a dependent outcome. In this context, the term "univariate" is correct and replaces the term "bivariate".



Multivariable analyses should not be confused with *multivariate analyses*, which are used to assess the relationships of several predictors with two or more dependent variables or outcomes at the same time. In this article we will not review multivariate analyses. However, medical writers should be aware that the terms multivariate and multivariable are often used interchangeably. Do not be surprised to see multivariable analyses described as multivariate.

To correctly interpret a multivariable analysis it is highly recommendable to first look at the bivariate analyses between the variables that were involved in the multivariable modelling. They show you: 1. the raw relationships between the dependent and independent variables (which allow the unadjusted associations to be quantified) and 2. correlations or associations between independent variables (which, if present, may require changes to the model).

Variables: Dependent vs independent / Quantitative vs categorical

It is very important to note that **both the dependent and independent variables can be either quantitative or categorical, and correct identification of these statistical properties is essential for the medical writer to correctly interpret and report the results**.

Common *quantitative outcomes* include cholesterol levels, blood pressure, and questionnaire scores and common *categorical outcomes* are survival (yes/no), response to treatment, and presence/absence of a specific event (e.g. cardiovascular event, relapse).

Common *quantitative predictors* include age, BMI, and baseline values for the outcomes. Common *categorical predictors* include treatment arm, gender, and baseline disease severity.

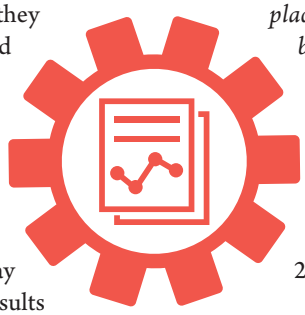
Note that categorical variables with only two categories are referred to as *dichotomous*.

The dependent variable is the one that is assessed with the study. Sometimes it is referred to as the *endpoint*. Usually this term is reserved for the combination of the outcome plus the timepoint(s) of assessment (e.g. if the outcome is "mortality", the endpoint could be "mortality rate at 6 months").

The independent variables define the subgroups of patients in which the outcome will be compared (e.g. treatment arms).

- If the independent variable is categorical (e.g. treatment arm, gender), the parameters of the multivariable models we will review in later sections – the odds ratio (OR), hazard ratio (HR), and beta coefficient (β) – always estimate the effect on the outcome of **one or more categories versus a *reference* category** (e.g. placebo or female gender), which must be defined *a priori*.
- If the independent variable is quantitative (e.g. age), no subgroups are compared and the OR, β, and HR estimate the effect on the outcome of **each 1-unit increase in the independent variable** (e.g. "for each 1 mg/dl increase in baseline cholesterol").

It is very common for continuous predictors to be transformed into categorical variables prior to the multivariable analysis using a previously defined cut-off point (from the literature). This is because the parameters of the models are much easier for physicians to interpret if they compare one category to another than if they inform about the risk associated with a 1-unit increase in the predictor. However, this leads to a loss of statistical power and to the risk of not finding significant results. If the model includes the original continuous predictor, the medical writer may facilitate interpretation of the results by reporting the risk associated with, for example, a 10-unit increase in the predictor.



In interpreting a multivariable analysis we must also consider that some independent variables may be entered in the model because they are *confounding variables* (sometimes also denoted as *covariates*). Confounding variables are factors related to both the dependent and independent variables. Unless we adjust our multivariable analysis for confounding variables, we may end up with an inaccurate or incorrect representation of the true relationships between the dependent and independent variables. For example, in many clinical trials the baseline value for a quantitative outcome (e.g. baseline blood pressure in a hypertension trial) is a potential confounding variable if it is not fully balanced between the two treatment arms, despite randomisation of the patients, because it is also related to the outcome. For this reason, the primary efficacy analysis should always include the baseline value for the quantitative outcome as a covariate.

When to apply a multivariable analysis

A multivariable analysis is needed in the following cases:

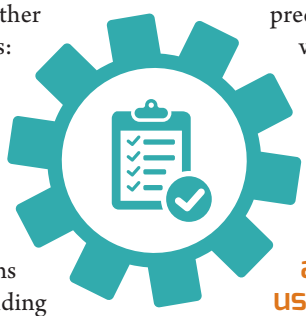
1. If there is **one** main independent variable of interest (the other independent variables being secondary factors):
   a. To evaluate the relationship between the variable of interest and the outcome **after adjusting (or controlling) for other independent variables that may also be related to the outcome (confounding factors or covariates)**.
   Examples:
   "Patients treated with drug A had significantly higher cholesterol levels at 6 months compared to patients treated with placebo, after adjustment for baseline cholesterol."
   "Higher biomarker X levels were significantly associated with a higher response rate, independently of/after adjusting for age and gender."
   (Box 2 opposite)
2. If there are **two or more** main independent variables of interest:
   a. To explore which of the independent variables are **independently** associated

with the outcome, i.e. they keep a significant p-value in the model despite the inclusion of other independent variables: ***exploratory models***.
These models are commonly used to look for "causal relationships", although the results must always be interpreted with caution because associations may be due to confounding factors that were not accounted for.
Example:
"In patients with disease Z, male gender and higher blood pressure were independently associated with higher mortality."

b. To predict an outcome with independent variables that are known to be associated with the outcome: ***predictive models***.
These models are commonly used in oncology to establish prognostic factors that may be useful to select candidate patients for more aggressive therapies. They can also be used to predict response, compliance, and quality of life.
Example:
"In patients with disease Z, the independent factors predicting response to drug X were tumour stage at diagnosis and baseline beta-2-microglobulin level. The model with these two variables correctly predicted the response in 65% of patients".

BOX 2: When describing associations between different variables, a common mistake is to not give the direction of the association, e.g. "Biomarker X levels were significantly associated with the response rate." From this sentence, the reader cannot ascertain whether a higher response rate is associated with high or low biomarker X levels. Although, if not otherwise indicated, such an association would usually be interpreted as positive, a good medical writer should clearly indicate the direction of the association.

Please note that the words "independently associated" and "independent factor/predictor" imply that a multivariable model has been used and that the described relationship has been adjusted for at least one additional factor.



Multivariable analyses commonly used in biomedical studies

There are several different types of multivariable analysis. Three of the most commonly used analyses are *multiple logistic regression*, *multiple Cox regression*, and *multiple linear regression/multiple analysis of variance (ANOVA)/analysis of covariance (ANCOVA)* (Table 1 overleaf). It is important to note that multiple regression and *multivariate regression* are not the same thing. In multiple regression there is only one dependent variable; multivariate regression involves two or more main dependent variables and is less commonly used.

With multiple logistic regression the aim is to determine how one **dichotomous dependent variable** varies according to two or more independent (quantitative or categorical) variables. Multiple logistic regression might, for example, be used to test the relationships of weekly alcohol consumption at age 30 and gender (independent variables) with probability of developing liver cancer during a 10 year period (dependent variable). Liver cancer is a categorical variable with two categories at the end of the follow-up period: "cancer" and "no cancer".

Multiple Cox regression is similar to multiple logistic regression but it explores the relationships between independent variables and a **time-to-event dependent variable** (dichotomous), e.g. time to death. If we wanted to determine whether a new treatment (independent variable) affects probability of disease progression (dependent variable) in patients with renal cell carcinomas of different clinical stages at baseline (second independent variable that may be considered a covariate), we could potentially use multiple Cox regression.

Finally, multiple linear regression, multiple ANOVA, and ANCOVA are multivariable models in which the **dependent variable is continuous**, i.e. it can theoretically take any value in its given range. Despite being slightly different from each other, these models can be considered equivalent from a medical writer's point of view. An example scenario would be to determine whether a new treatment (independent variable) reduces the score for disease index X (dependent variable) after adjusting for country and baseline disease index X score (independent variables considered as covariates).

These multivariable analyses will be discussed in further detail below. The aim is not to explain how to run the analyses, rather how to interpret and report the results they give. The focus will be on ORs, HRs, and β coefficients.

Multiple logistic regression: What is an odds ratio?

What is an OR? Let's define two groups of subjects: a test group we are interested in and a reference group we wish to compare the test group to. The OR is the **ratio of two sets of odds**: the *odds* of an event occurring *in the test group* divided by the *odds* of the same event *in the reference group*. Note that odds are *not* the same as probability: the odds are the probability of an event (e.g. death) occurring *divided by* the probability of it not occurring. While probability ranges from 0 to 1, the odds may range from 0 to positive infinity.

Going back to our example above, how do weekly alcohol consumption and gender affect the odds of developing liver cancer? Here we can define two reference groups: one for weekly alcohol consumption and one for gender. The reference group for weekly alcohol consumption might be "0 units" and let's say the one for gender is "female". (If you're wondering how a categorical independent variable such as gender may be entered into a mathematical model, this can be achieved by creating a *dummy variable* with a value of 0 or 1. In the present example, females may be given a value of 0 and males 1.)

	Multiple logistic regression	**Multiple Cox regression**	**Multiple linear regression / Multiple ANOVA / ANCOVA**
Dependent variable	**Dichotomous** (no information about timepoint) *Example: Treatment response (yes/no)*	**Time to event** (dichotomous with information about timepoint) *Example: Overall survival*	**Quantitative** *Example: Blood pressure*
Independent variables	2 or more quantitative or categorical variables	2 or more quantitative or categorical variables	2 or more quantitative or categorical variables[a]
Equation[b]	$\text{logit}(p) = a + b_1x_1 + b_2x_2 \ldots$	$\log(h_i(t)) = a + b_1x_1 + b_2x_2 \ldots$	$y = a + b_1x_1 + b_2x_2 \ldots$
Parameter	OR (= Exp(b))	HR (= Exp(b))	β (= b)
Interpretation	Odds for: • Category X vs reference category (if independent variable is categorical) • A 1-unit increase (if independent variable is quantitative)	Instantaneous risk/hazard (hazard per unit time) for: • Category X vs reference category (if independent variable is categorical) • A 1-unit increase (if independent variable is quantitative)	Size of the effect on the outcome (in outcome units) for: • Category X vs reference category (if independent variable is categorical) • A 1-unit increase (if independent variable is quantitative)
Example of reporting	*"… odds of treatment failure were 3 times higher in men than in women"*	*"… risk of death was 3 times higher in men versus women"*	*"… systolic blood pressure was 3 mmHg higher in men than in women"*

[a] For ANOVA and ANCOVA at least 1 categorical variable is needed

[b] $\text{logit}(p)$ is $\log(p/1\text{-}p)$, where p is the probability of the outcome; a denotes a constant, b_n denotes the coefficient for each independent variable, x_n denotes an independent variable, $h_i(t)$ is the hazard to individual i at time t, and y denotes a dependent variable

Table 1. Types of multivariable models commonly used in biomedical studies

Say we obtain an OR for liver cancer of 1.68 for people who consume 40+ units of alcohol per week versus those who consume 0 units per week. This means that the odds of liver cancer are 1.68 times as high (or 68% higher) for those consuming 40+ units of alcohol per week than for teetotallers. Similarly, an OR of 1.22 for males versus females would mean that males have 22% higher odds of developing liver cancer compared to females.

ORs are typically presented with CIs. In general terms, the CI is a range of values within which the true value of a parameter **in the population** (*not* in the study sample) is expected to lie. A narrow CI indicates good precision in our OR estimate; a wider CI would indicate more uncertainty. Narrower intervals are obtained with larger samples. For an OR, a CI that includes 1 (e.g. 0.9 to 2.5) prevents us from inferring a significant difference between groups.

If we adjust our multiple logistic regression model for confounder variables, then the ORs we obtain will be referred to as *adjusted ORs*. If in the present example we calculate a 95% CI of 1.25 to 2.13 for our OR of 1.68, we could describe the results of the multiple logistic regression thus:

> *Compared to teetotallers, those who consumed 40+ units of alcohol per week at age 30 had higher odds of developing liver cancer (adjusted OR=1.68, 95% CI=1.25 to 2.13). Males had higher odds of liver cancer than females (adjusted OR=1.22, 95% CI=1.03 to 1.44).*

Note that we are not claiming that alcohol consumption *causes* liver cancer (although there is ample evidence to suggest this is the case). Rather, we are merely saying that excessive alcohol consumption *is associated with* liver cancer; it may or may not cause liver cancer.

Risk has a particular meaning in statistics, with *relative risk* (RR) implying a comparison of probabilities, not odds. In the above example, the odds of liver cancer were 1.22 times higher in males compared to females; we should not write that the "risk" of liver cancer was 1.22 times higher in males, because this would be inaccurate. Phrases that indicate or imply probability, such as "X times as likely to" and "a 50% higher probability of", should also be

avoided when reporting ORs. Note that the OR gives a reasonable approximation for the RR when the event is rare, but not when the event is common.

Multiple Cox regression: What is a hazard ratio?

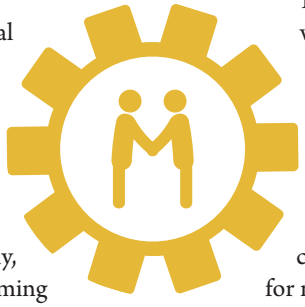
Multiple Cox regression is used to calculate HRs. An HR indicates the *instantaneous risk* or *hazard* (hazard per unit time, usually 1 day) of an event (e.g. death) in a test group relative to a reference group. Let's return to the example of the new treatment for renal cell carcinoma. The new treatment (test group) gives an HR for death of 0.5 versus the existing gold standard treatment (reference group). How do we interpret this?

In this example, the HR indicates the relative rates of death per day in the two treatment groups. The value of 0.5 indicates that the rate of death at any time during the follow-up period is twice as high with the gold standard treatment compared to the new treatment. A value of 1.0 would indicate no difference in rate of death between the two treatments.

Like ORs and β coefficients, HRs are typically presented with CIs. Assuming we adjust our multiple Cox regression model for several confounder variables, we could report the results in the present example as:

Compared to the gold standard treatment, the new treatment was associated with a significantly lower rate of death (adjusted HR=0.5, 95% CI=0.25 to 0.75).

In descriptions of survival analyses, RR cannot be used instead of HR, since the two terms are not synonymous. Though they can be interpreted in more or less the same way, HRs and RRs are calculated differently. Notably, RRs do not account for the timing of the events of interest. Don't write relative risk when you mean hazard ratio!



Multiple linear regression / Multiple ANOVA / ANCOVA: What is the β coefficient?

In multiple linear regression, multiple ANOVA, and ANCOVA, the dependent variable is continuous. One such variable is height at age 18. What is its relationship with birth length and age at puberty onset (independent variables)?

In addressing this question by multiple linear regression we obtain one *β coefficient* for each quantitative independent variable and for each non-reference category of each categorical independent variable.

For continuous independent variables such as birth length the β coefficient indicates how a *1-unit change in the value of the independent variable* would affect the value of the dependent variable if all other variables in the model were held constant, and the units for β are the units for the dependent variable divided by those for the independent variable. For categorical independent variables the units of the β coefficient are the same as those of the dependent variable. It is very important to understand this to correctly describe the results of the model.

If the β coefficient for birth length is positive (e.g. 1.2 cm/cm), then a higher birth length will be associated with a greater height at age 18. A negative β coefficient for age at puberty onset (e.g. -0.3 cm/year) indicates a negative association between age at puberty onset and height at age 18. The statistical significance of these results should be reported using the p-value associated with each β coefficient.

For categorical independent variables such as gender the β coefficient and the **corresponding** p-value will indicate whether the category is associated with greater height at age 18 compared to the reference category. A positive β coefficient for males relative to females with a p-value of <0.05 would indicate that males are likely to be taller than females at age 18.

β coefficients are often presented with corresponding CIs; sometimes the CI is replaced by the standard error (SE) and the p-value. **If a CI does not include 0**, the association between the independent variable and the dependent variable (after adjusting for covariates) is significant. Thus we could describe the results of the current analysis as:

Higher birth length was associated with greater height at age 18 (β=1.2 cm/cm, 95% CI=0.93 to 1.49). Age at puberty onset was inversely associated with height at age 18 (β=-0.3 cm/year, 95% CI=-0.19 to -0.45).

As a final remark regarding β coefficients, please be aware that they are sometimes also provided for multiple logistic regression and Cox regression models. In such cases, β is simply the natural logarithm (ln) of the OR (logistic regression) or HR (Cox regression).

How to report the results from multivariable models

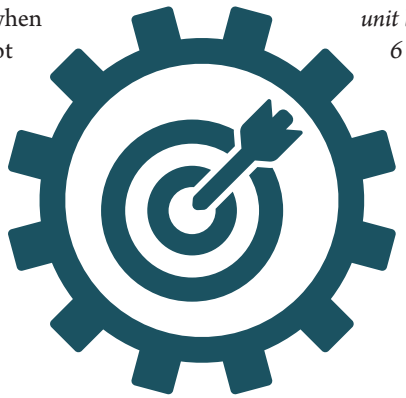
Whatever the model used, good medical writing practice is to list all the factors that were taken into account in the multivariable analysis, including those that were discarded during the modelling process. Also, remember always to include the parameter that indicates the strength **and direction** of the association (i.e. the OR, HR, or β coefficient), preferably with the 95% CI and/or the p-value **for the variable** (different from the overall p-value for the model). If you are at all unsure as to the direction of a particular association, ask a statistician for clarification.

When reporting the parameter, the writing differs depending on the direction of the association. We round off our introduction to multivariable analyses with some illustrative examples:

- For ORs (logistic regression) and HRs (Cox regression), results are significant when the 95% CI does not include 1:
  - A value **<1** implies that the factor is negatively associated with (i.e. protects against) the outcome. The percentage decrease in the odds (OR) or risk (HR) is **(1 - OR or HR) × 100**. Example: "*Category X protected against mortality (adjusted OR=0.8, 95% CI=0.6 to 0.9 versus reference category*

Y)" or "Compared to Y, X was associated with a 20% reduction in the odds of death."

- A value **>1** implies that the factor is positively associated with (i.e. increases the risk of) the outcome. The percentage increase in the odds (OR) or risk (HR) is **(OR or HR - 1) × 100**.
  Example: "*Category X was a risk factor for mortality (adjusted HR=1.5, 95% CI=1.1 to 1.9 versus reference category Y)" or "Compared to Y, X was associated with a 50% increase in the risk of death."*
  When the percentage is ≥100, the "number of times" construction is often used:
  Example: "*Patients with X had a risk of death approximately three times higher compared to those with Y (adjusted HR=3.2, 95% CI=2.1 to 4.9).*"
- For β coefficients (multiple regression, multiple ANOVA, ANCOVA), results are significant when the 95% CI does not include 0:
- A value **<0** implies that the factor is *negatively associated* with the outcome.
  Examples:
  Quantitative factor: "*A 1-unit increase in X was associated with a decrease of 3 mmHg in systolic blood pressure at 4 weeks (β=-3, 95% CI=-2.1 to -3.9)".*
  Categorical factor: "*Compared to placebo, treatment with drug Z was associated with a decrease of 3 mmHg in systolic blood pressure at 4 weeks (β=-3, 95% CI=-2.1 to 3.9).*"
- A value **>0** implies that the factor is *positively associated* with the outcome.
  Examples:
  Quantitative factor: *"A 1 mg/dl increase in X was associated with a 2-unit increase in quality of life score at 6 months (β=2, SE=0.3, p=0.025)."*
  Categorical factor: *"Compared to patients with mild disease at baseline, severe disease was associated with a 2-unit lower quality of life score at 6 months (β=–2, SE=0.3, p=0.025)."*



Acknowledgements

We would like to thank Elena Stolyarova, a biostatistician at TFS, for her insightful comments on an early draft of this article. We also thank Adam Jacobs, Senior Principal Statistician at Premier Research, for providing constructive feedback and suggesting improvements.

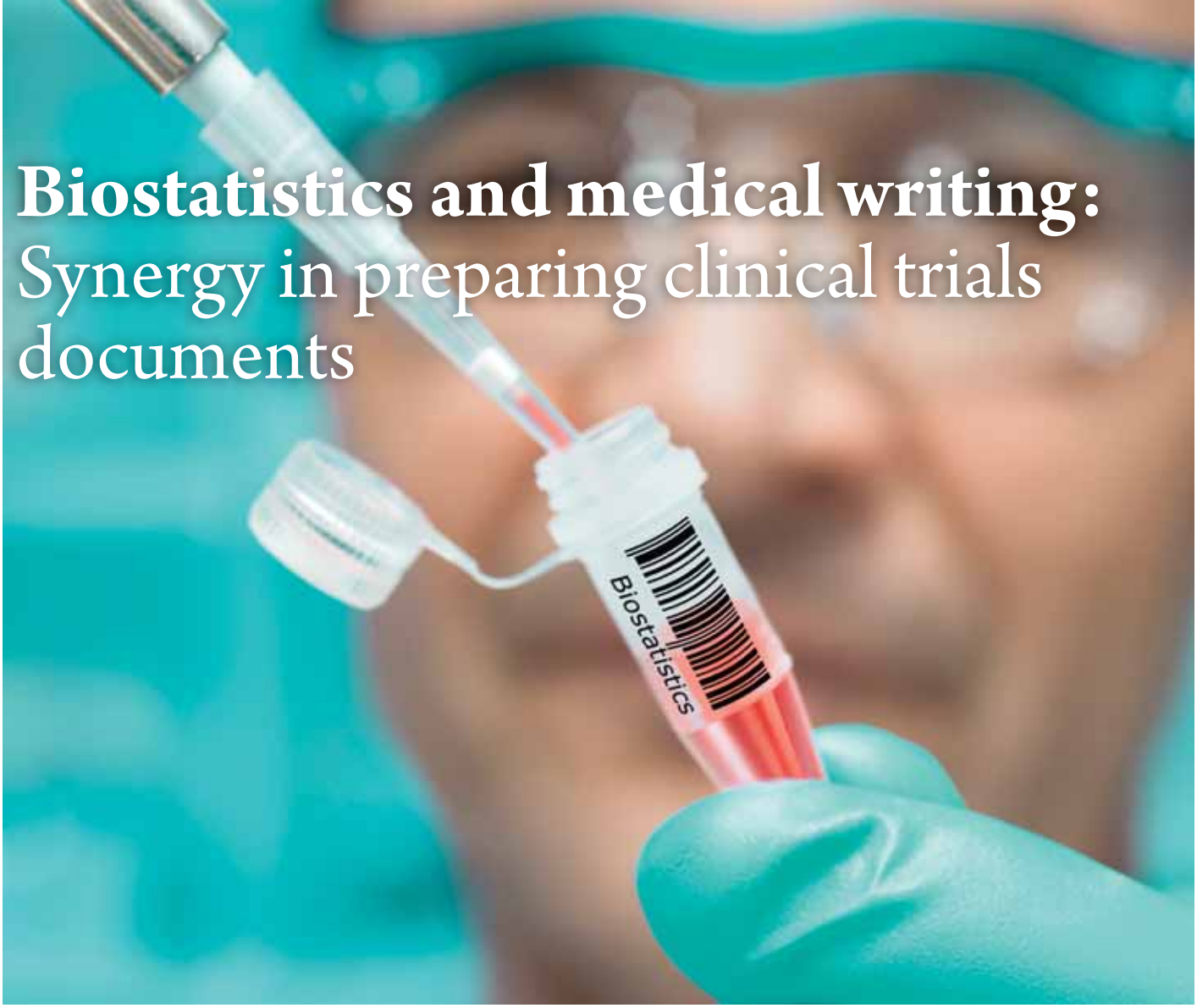
Conflicts of Interest and Disclaimers

The content of this article reflects the view of the authors, and not those of TFS. There are no conflicts of interest.

Author information

Neus Valveny has been a medical writer for the last 15 years, preparing more than 40 manuscripts, 50 abstracts and posters, 20 study protocols, and 15 study reports. She has also worked as a statistician. She is now Associate Director of Medical Writing at TFS, a medium-sized international CRO founded in Sweden.

Previously a full-time science editor and part-time copy editor, **Stephen Gilliver** is now a medical writer at TFS. He is also Co-Editor of *Medical Writing*.



Biostatistics and medical writing: Synergy in preparing clinical trials documents

Scott Miller[1] and **Raquel Billiones**[2]

1 Clinipace Worldwide, Morrisville, NC, USA
2 Clinipace Worldwide, Zurich, Switzerland

Correspondence to:
Raquel Billiones
Clinipace Worldwide
Chriesbaumstrasse 2
8604 Volketswil (Zurich)
Switzerland
rbilliones@clinipace.com

Abstract
Biostatisticians and medical writers are among the key people who develop important documents for clinical trials. These documents include clinical study protocols, statistical analysis plans, statistical outputs, and clinical study reports. This article demonstrates how biostatisticians and medical writers should work together to streamline the document preparation process and ensure the quality of these documents.

Introduction
Biostatisticians (BSTs) and medical writers (MWs) play key roles in clinical trials (CTs) without visiting a study site or seeing a patient. In spite of their 'back office' positions, their roles are nevertheless crucial to study design, study conduct, and data analysis as they deal with a wide range of interrelated CT documents that include clinical study protocols (CSPs), statistical analysis plans (SAPs), statistical outputs, and clinical study reports (CSRs).

Generally, the MW's core competencies lie in producing words and text, whereas the BST's expertise is in numbers and analysis of data. Though divergent at first glance, the MW and BST skills sets actually have a powerful synergy that can have a major impact on the execution of a CT.

In this article, we describe how the BST and the MW should work together on a CT project. The scenario we describe comes from full-service projects in a global contract research organisation (CRO) environment but the principles are applicable to many CT project configurations.

Communication
At the start of the study, the BST and the MW should get to know each other's names, exchange contact details and time zones/work schedules, and discuss timelines. In today's digital global office

environment, having a strong professional affinity and an open line of communication is especially important.

Collaboration

The CT documents that the BST and the MW produce are all interrelated, as shown in Figure 1. All these documents revolve around common themes: the study design, the study objectives, and the corresponding study endpoints – interconnected by the so-called 'golden thread'.[1] Working together, the BST and the MW need to ensure that the objectives and endpoints are well-defined, congruent, and remain consistent throughout the different documents produced as the study proceeds.

Too often, the transfer of content from one document to the next inadvertently results in errors and loss of information (e.g. during copy and paste). However, this can also happen during the document revision process. To avoid this, any changes to the documents and the rationale behind these changes should be discussed within the team and clearly documented. Both the BST and the MW should be involved in the review of each document.

CSP

The protocol is the main starting point of a CT. In drafting the CSP, the BST and the MW should work together to ensure that the study objectives and endpoints are aligned. The BST should complete the statistical sections of the protocol, including the sample size calculations; the MW should review them. Any ambiguities should be clarified, and any changes that need to be implemented as the study proceeds should be documented.

SAP

Developed early on in the trial, the SAP is the responsibility of the BST. The MW is one of the downstream end users, i.e. during CSR development. Hence, the MW should be able to review and provide feedback on the SAP and the shell ('mock') statistical outputs before their finalisation to ensure consistency between the CSP and the SAP.

Statistical outputs

The BST delivers the statistical outputs in the form of tables, figures, and listings (TFLs), to be used as the primary data source for the CSR. The MW's involvement in TFL review, which started during SAP and mock TFL development, continues with the real statistical outputs. The MW should thoroughly review the draft TFLs and request any necessary revisions or additional TFLs as early as possible so that the BST has sufficient time to deliver them without impacting the CSR delivery date.

The end users of the TFLs will include medical reviewers, investigators, and regulators. It is very important that the MW reviews the outputs from the end user perspective; the individual tables and listings should, as a rule of thumb, be standalone documents. The BST should work with the statistical programmer and the MW to ensure that the TFLs meet the specifications defined in the CSP and in the SAP.

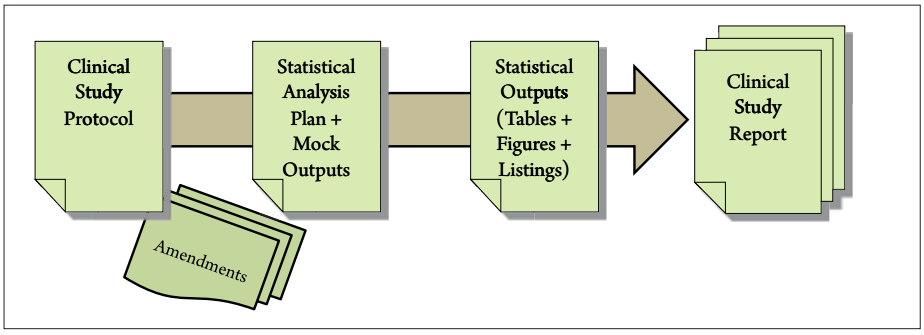


Figure 1. The different documents that the BST and the MW develop during a clinical trial.

CSR

The CSR is the responsibility of the MW. However, the BST should take an active role in providing input, not only on the statistical sections, but also on the results sections with respect to the endpoints and their interpretation. There was a time when a separate statistical analysis report was issued by the BST. The industry trend nowadays is to integrate the clinical and statistical text and analyses into a single document – the CSR as we know it today.[2,3] The BST accompanies the MW throughout the CSR review cycles, always ready to answer questions and clarify queries.

At the end of the study, the MW and the BST should produce an 'integrated' CSR that is actually a whole dossier containing all the CT documents they worked on during the study. And all throughout, the golden thread connecting the initial CSP to the final CSR and the other documents in between should remain unbroken.

Sharing information

The BST and the MW should keep each other in the loop. In full-service CRO CT projects, the BST is generally involved with the trial on an ongoing basis while the MW is often brought back in near database lock. As a result, it is possible that the BST will become aware of issues in trial conduct (e.g. delayed enrolment, early trial termination, protocol amendments, randomisation issues, and protocol deviations) which could impact the timeline or content of the CSR. It is vital that the BST passes this information to the MW as it becomes available to ensure that the MW becomes aware of these critical issues.

During review of the CSP and the CSR, the MW is responsible for addressing reviewers' comments. The MW needs to keep an eye on any changes that have an impact on the statistical methods and data analysis and should immediately flag these changes to the BST for re-validation.

Knowing each other's procedures and processes

The BST and the MW standard operating procedures (SOPs) should be aligned and not contradict one another. It is best for the

BST and MW teams to consult each other when developing and revising SOPs and processes. If this isn't done, the two teams should at least share with each other their relevant SOPs and process guidelines.

The BST and the MW should inform each other of their expectations in terms of templates and style guides, number of review cycles, expected review time, and level of review (e.g. text/content only or formatting/grammar, full document or only certain sections).

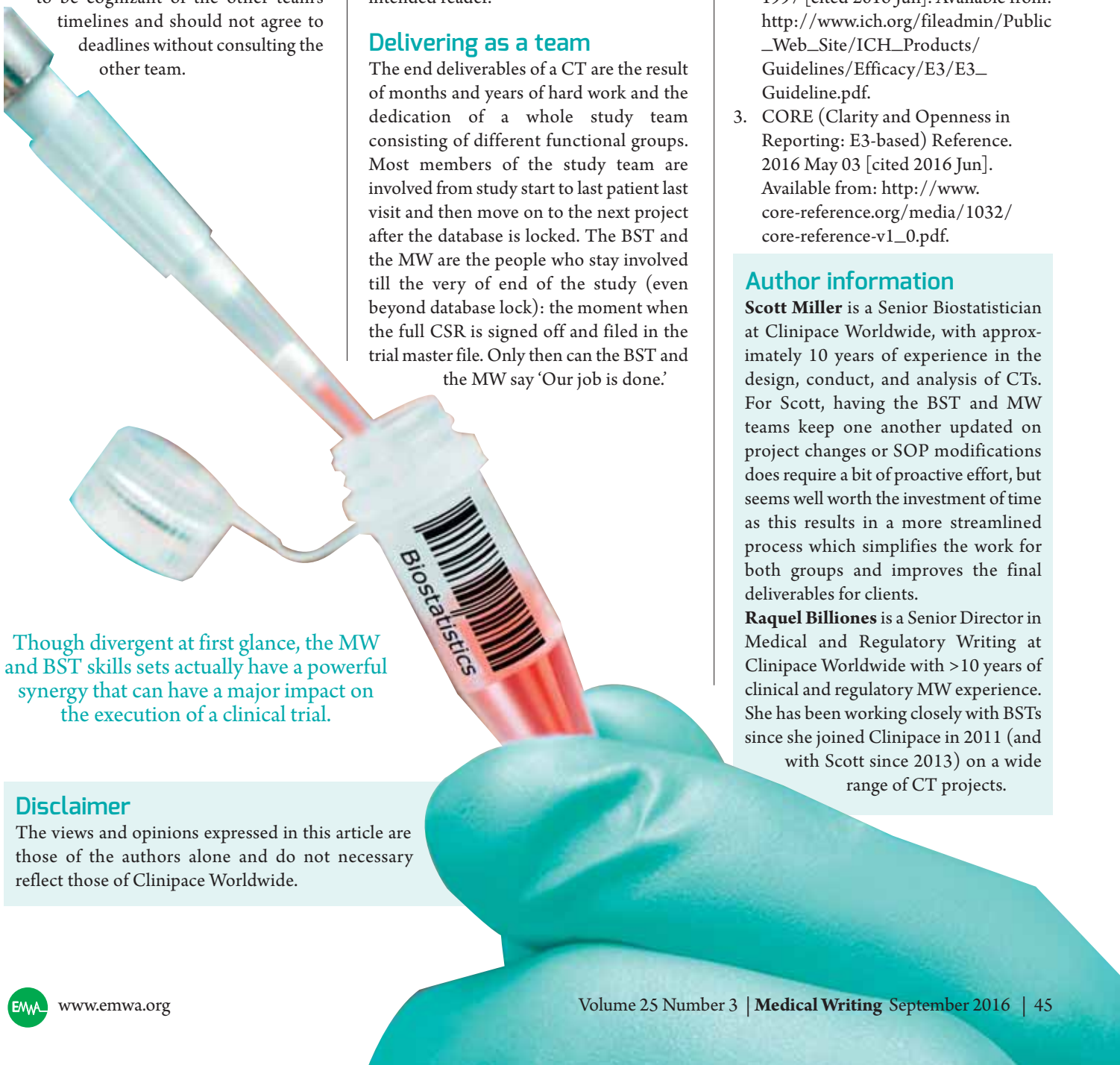
The deliverables of the BST and the MW are interdependent. Each team has to be cognizant of the other team's timelines and should not agree to deadlines without consulting the other team.

Leveraging each other's expertise

The MW should not hesitate to ask statistical questions, even if they seem basic or have been discussed before. For their part, the BST should consult the MW on textual, content, and formatting issues, as well as for guidance on regulatory requirements, if necessary. The MW should be cognizant of the needs of a document's target audience; the BST should take advantage of this expertise and collaborate with the MW to customise technical documents to the level of the intended reader.

Delivering as a team

The end deliverables of a CT are the result of months and years of hard work and the dedication of a whole study team consisting of different functional groups. Most members of the study team are involved from study start to last patient last visit and then move on to the next project after the database is locked. The BST and the MW are the people who stay involved till the very of end of the study (even beyond database lock): the moment when the full CSR is signed off and filed in the trial master file. Only then can the BST and the MW say 'Our job is done.'



Though divergent at first glance, the MW and BST skills sets actually have a powerful synergy that can have a major impact on the execution of a clinical trial.

Disclaimer

The views and opinions expressed in this article are those of the authors alone and do not necessary reflect those of Clinipace Worldwide.

References

1. ICH E9(R1) Final Concept Paper. Addendum to Statistical Principles for Clinical Trials on Choosing Appropriate Estimands and Defining Sensitivity Analyses in Clinical Trials. 2014 Oct 22 [cited 2016 Jun]. Available from: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/E9_R1_Final_Concept_Paper_October_23_2014.pdf.
2. ICH E3 Guideline. Structure and Content of Clinical Study Reports. 1997 [cited 2016 Jun]. Available from: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf.
3. CORE (Clarity and Openness in Reporting: E3-based) Reference. 2016 May 03 [cited 2016 Jun]. Available from: http://www.core-reference.org/media/1032/core-reference-v1_0.pdf.

Author information

Scott Miller is a Senior Biostatistician at Clinipace Worldwide, with approximately 10 years of experience in the design, conduct, and analysis of CTs. For Scott, having the BST and MW teams keep one another updated on project changes or SOP modifications does require a bit of proactive effort, but seems well worth the investment of time as this results in a more streamlined process which simplifies the work for both groups and improves the final deliverables for clients.

Raquel Billiones is a Senior Director in Medical and Regulatory Writing at Clinipace Worldwide with >10 years of clinical and regulatory MW experience. She has been working closely with BSTs since she joined Clinipace in 2011 (and with Scott since 2013) on a wide range of CT projects.

Best friends forever: A pattern of collaboration between medical writers and biostatisticians within the Russian CRO

Eugenia Radkova[1] and
Ivan Dobromyslov[2]

[1] Eugenia Radkova, OCT Rus (OCT Group), Saint-Petersburg, Russia
[2] Ivan Dobromyslov, Data MATRIX (OCT Group), Saint-Petersburg, Russia

Correspondence to:

Eugenia Radkova
OCT Rus
5 Kovensky Lane
6-7 floor
Saint-Petersburg, Russia 191014
+7 (812) 449 86 34

Abstract

The Russian clinical trial industry and Russia's local regulatory requirements are developing rapidly. Within Russian contract research organisations, medical writers must take on non-traditional roles and, in particular, must collaborate closely with biostatisticians within their organisations. This article describes the special relationship between medical writers and biostatisticians within Russian contract research organisations and the special expertise that medical writers in Russia need to develop.

Clinical trials industry in Russia

The world is undergoing a boom in clinical trials, with annual increases between 1% and 14%.[1] Considering that the average annual growth in recent years has been 4-5%, the number of approved clinical trials may reach 30,000 by 2020 (Figure 1).

The Russian Federation is an attractive and fast-growing market for pharmaceutical products. The number of approved clinical trials could increase by 20% by 2020 to reach 1,000 (Figure 2). The Russian government continues to invest in pharmaceutical manufacturing and drug development, with implementation of many economic measures and programmes in the last few years.

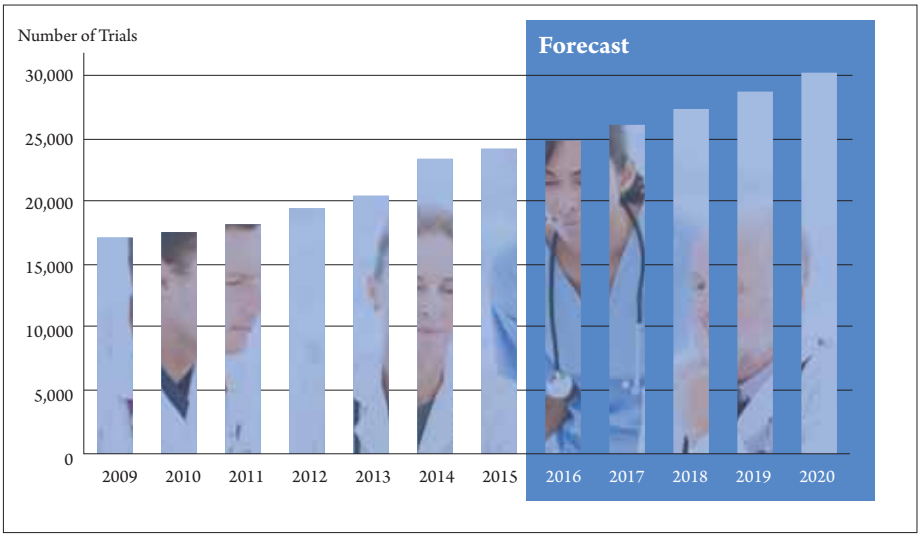


Figure 1. Number of clinical trials approved worldwide between 2009 and 2015 and forecast for 2016 to 2020. Data were from ClinicalTrials.gov.

A main reason for this rapid growth in clinical trials in Russia is the 2010 Federal Law on Drug Circulation.[2] This law imposed new regulatory requirements on the clinical trials industry and stipulated that well-controlled, evidence-based confirmative clinical trials must be performed as part of the drug registration process in Russia. This law applied to most drug categories and included products that had already been investigated and registered in other countries. The result has been a considerable increase of so-called "local registrational" phase III studies. These now represent approximately two-thirds of all clinical trials carried out in Russia (Figure 2).

These new regulations resulted in a rapid increase in demand for regulatory and trial-related documents and therefore the development of the medical writing industry in Russia.[3] At the same time, pharmaceutical companies have lacked experience in planning and conducting clinical trials, have not been able to liaise with and obtain scientific advice from regulatory authorities, have not had enough personnel knowledgeable about study design and methodology.

Medical writers in Russia therefore need to be able to provide not only writing but also scientific advice on drug development and regulatory affairs as well as different aspects of study design and methodology. This means that medical writers must have a good understanding of basic biostatistics and statistical methodologies applied to clinical trials. This demand for high-level knowledge of biostatistics means that medical writers need to work in close

collaboration with biostatisticians, especially within the contract research organisations (CROs).

In this article, we describe the collaborative work between medical writers and biostatisticians within Russian CROs and give a brief overview of local regulatory environment in which these specialists operate.

Key regulatory documents in Russia

Federal Law #61-FZ "On circulation of medicines", which was passed in 2010, is the main act controlling drug manufacturing, non-clinical and clinical studies, pharmacovigilance, and drug registration in Russia.[1] This regulation is still developing, and more than 10 amendments have been made to date. This law defines the essential documents required for clinical trial approval that must be submitted to the Ministry of Healthcare of the Russian Federation. These documents are assessed by the Ethics Council and the Federal State Institution Scientific Center for Expertise of Medical Products. Since 2012, the Scientific Center for Expertise of Medical Products has published a series of guidelines on different aspects of study planning and conduct, including the content and structure of study protocols and clinical study reports (CSRs), statistical principles for clinical studies, and study design and methodology for different therapeutic areas.[4,5]

In addition to Federal Law #61-FZ, the National Standard of the Russian Federation GOST 52379-2005 contains key guidance for the conduct of clinical trials in Russia.[6] It stipulates that all clinical trials in Russia must be conducted according to Good Clinical Practice and includes a translation of the International Conference on Harmonisation Guideline for Good Clinical Practice.[7] This act harmonises Russian clinical trials with the rest of the world.

Besides these documents, the Eurasian Economic Union (EAEU), which includes the Republic of Armenia, the Republic of Belarus, the Republic of Kazakhstan, the Kyrgyz Republic, and the Russian Federation, is developing legislation on pharmaceutical drug development. The EAEU member-states have agreed to establish a single pharmaceutical market and are in the process of developing unified regulatory principles and rules for drug development, approval, and marketing. The first EAEU regulations on clinical studies came into force in December 2015 and included detailed regulations for drug expertise and registration in the EAEU and Good Clinical Practice, along with requirements for the content and structure of CSRs.[8,9]

Generally, the Russian and EAEU regulations and guidance are in harmony with the international guidelines and standards implemented by International Conference on Harmonisation, the US Food and Drug Administration, and the European Medicines Agency. At the same time, there are some local differences that can lead to challenges and concerns at different stages of drug development. Moreover, the Russian legislation is still being developed, so to provide optimal services for clients, CROs should continue to track the changes and trends.

Medical writers and biostatisticians in the Russian CRO: collaboration towards study success

A successful clinical study is one that provides accurate, reliable, and valid data and that allows regulatory authorities to make accurate regulatory decisions about the safety and efficacy of the medicinal product. To achieve this, the multi-disciplinary team must collaborate effectively from development of the study concept to data analysis and reporting. In Russian CROs, medical writers and biostatisticians must take on non-traditional roles and collaborate closely to reach the objective of a successful clinical study.

Collaboration during study concept development

Developing the scientific concept is one of the most challenging aspects of a clinical study but is also the most important for determining its regulatory, scientific, and financial success. In the Russian CRO, medical writers and biostatisticians are responsible for ensuring that the study is in compliance with

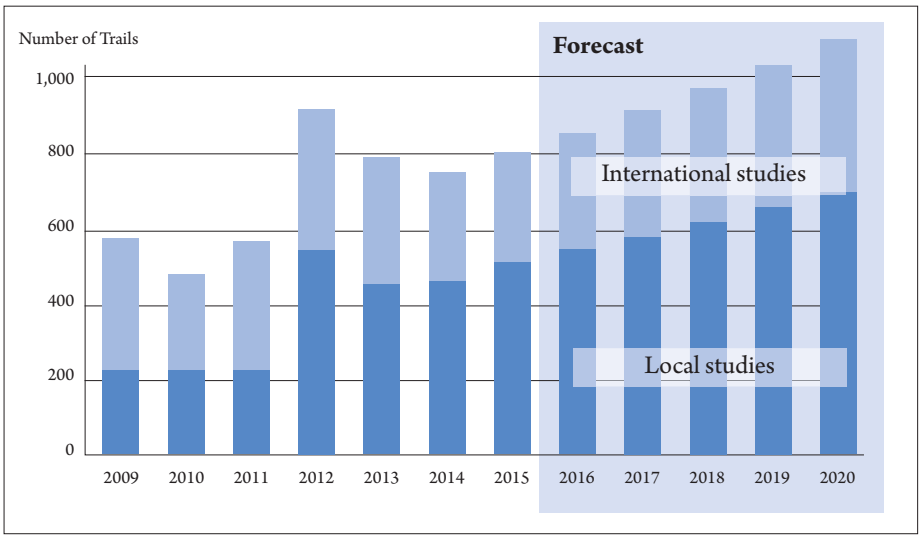


Figure 2. Number of clinical trials approved in Russia between 2009 and 2015 and forecast for 2016 to 2020. Data were from the Russian Registry of Approved Clinical Trials (http://grls.rosminzdrav.ru/CIPermitionReg.aspx) and the Association of Clinical Trials Organisations (http://acto-russia.org).

local and international regulatory requirements, applicable scientific guidelines, and trends and tendencies in the specific therapeutic area. In addition, medical writers and biostatisticians are responsible for ensuring not only that the study is feasible and time – and cost-effective but also that the results it produces are internally and externally valid. The medical writer and the biostatistician must work together to research and select the study population, the primary and secondary endpoints, and the types and time points for study assessments in the context of the requirements for the specific phase of clinical development.

As part of this, medical writers within Russian CROs need to be aware of the basic approaches and regulatory requirements for sample size calculations. To this end, Russian medical writers need to be aware that the Russian regulatory authority has been focusing much more attention on the details of study design and on the statistical aspects of the study, including assumptions used for sample size calculation. For instance, for phase III confirmatory studies, the most common concerns raised during regulatory review are the choice of primary endpoint, study hypothesis, justification of non-inferiority/equivalence margin, and clinical and statistical assumptions supporting a sample size calculation. This part of the study concept often becomes particularly challenging when developing "local registrational" phase III studies for products that have been registered in other countries for a long time and for which clinical trial data may be limited or absent. Such cases usually require extensive literature searches and a great deal of creativity and thought to develop arguments to support the study concept and design.

Collaboration during protocol development

Once the study outline is finalised, medical writers are responsible for developing the clinical study protocol, which defines all aspects of the study and, in large measure, influences the quality of future data resulting from the study.[10] The statistical part of the protocol is usually written by a biostatistician and reviewed by the medical writer to ensure that the terminology, text style, and formatting are consistent and follow the appropriate templates and style guidelines. Therefore, within Russian CROs, medical writers must be able to understand the main statistical aspects of the study to be able to provide comments and suggestions related to the statistical methods in the protocol.

After the clinical study is approved by the regulatory authority, the medical writer and the biostatistician must continue to work together to development a statistical analysis plan, and to review of the case report form, which is usually generated by the CRO's data management department. Medical writing review of the statistical analysis plan is essential for planning and outlining the CSR and for avoiding late changes to statistical outputs.[11]

Collaboration during study conduct

Medical writers and biostatisticians within the Russian CRO continue to be a part of the process after the protocol has been implemented at the clinical study sites. The medical writer and biostatistician may consult with the clinical trial team and sponsor on questions and difficulties in the practical application of the protocol during clinical research, such as a high rate of premature withdrawal loss to follow-up, difficulty in performing assessments, and problems related to data analysis or reporting. The participation of medical writers and biostatisticians in these discussions helps guarantee that decisions are made in accordance with the protocol and are aligned with needs of future statistical analysis and data reporting.

Collaboration during data review

After a clinical study is complete and database is cleaned, medical writers and biostatisticians within the Russian CRO participate in data review before database lock and before starting statistical analysis. This is an important step that should not be underestimated because it helps to ensure the data are clean and complete for the final analysis. A detailed preliminary review of raw data can save the medical writer and biostatistician time later by avoiding problems with final data analysis and interpretation. Although they collaborate, the biostatistician and medical writer have different roles during data review: the biostatistician performs statistical review checks to identify outliers and to find missing or inconsistent data, while the medical writer searches for errors and inconsistences in coding of adverse events, medical history, and concomitant medications and looks for underreporting of clinical descriptions, which could complicate interpretation of the collected data.

The medical writer and biostatistician within Russian CROs must also take part in data review meetings during which final decisions about data issues are made in conjunction with the various stakeholders.

For phase II and III studies, the medical writer is expected to work in close collaboration with the biostatistician on issues related to the distribution of patients in data analysis sets, especially in cases requiring clinical judgement and opinion.

Collaboration during development of the CSR

The success of the CSR depends on having carried out effective reviews of the statistical analysis plan, the final raw data, and the statistical output. To allow timelines to be respected, issues and concerns related to the database and analysed data must be resolved before statistical output can be included in the core text and appendices of the CSR. Substantial time can be saved by having the medical writer review the prepared statistical output before they start writing the CSR, even if an output is considered final after the quality check procedures. This allows discrepancies, errors, and confusing results to be identified, discussed, and corrected, reducing complications and confusion during writing. Once the first draft of the integrated CSR is prepared, the medical writer and the biostatistician carefully go through the core text to ensure that the results are interpreted correctly from both a statistical and clinical perspective.

Factors of successful collaboration

Working conditions have a great influence on team effectiveness. Having the majority of employees of the medical writing and biostatistics departments work at the same office facilitates communication and allows issues to be resolved quickly so that project timelines are respected. Regular face-to-face meetings beginning from study start-up and shared training on standard operation procedures enhances understanding about project plans, milestones, and specific project requirements.

When team members work closely, professional skills and knowledge are easier to attain. Within the CRO, a collaborative learning environment can be maintained by sharing and discussing useful literature publications, conference materials, and new regulatory information.

Conclusion

Thanks to new regulations and on-going changes, the Russian pharmaceutical industry is rapidly developing. Because of these changes, medical writers and biostatisticians within Russian CROs must play non-traditional roles: in addition to their usual functions, they act as a source of expert knowledge on the scientific and regulatory aspects of study design and methodology. Close collaboration between them improves the efficiency and quality of the drug development process. This can be fostered by creating an environment of discussion, support, and shared learning.

Conflicts of Interest and Disclaimers

The opinions expressed in the article are those of the authors and do not necessarily reflect the views of their employers.

References

1. ClinicalTrials.gov. Trends, Charts, and Maps [cited 2016 Jun 15]. Available from: https://clinicaltrials.gov/ct2/resources/trends.
2. The Russian Federation, Federal Law #61 on Circulation of Medicines [Федеральный закон от 12.04.2010 года №61-ФЗ 'Об обращении лекарственных средств']. Available from: State Law Publisher.
3. Davydova A. Medical writing in the Russian Federation: Promises and pitfalls. Medical Writing 2013;22(2):121-123.
4. Guidance for the conduct of clinical trials. Federal State Budgetary Institution "Scientific Center for Expertise of Medical Application Products";2012.
5. Guidance for expertise of medications. Federal State Budgetary Institution "Scientific Center for Expertise of Medical Application Products"; 2013, 2014.
6. National standard of the Russian Federation GOST-P 52379-2005 "Good Clinical Practice" [Национальный стандарт Российской Федерации ГОСТ Р 52379-2005]. Available from: State Law Publisher.
7. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R1_Guideline.pdf.
8. https://docs.eaeunion.org/docs/ru-ru/0149249/clco_30122015_170
9. EAEU Good Clinical Practice [Правила надлежащей клинической практики Евразийского экономического союза], 2015 [cited 2016 Jun 15]. Available from: https://docs.eaeunion.org/docs/enus/0149256/clco_30122015_172_att.
10. No authors listed. Clinical study protocol is the key to data quality [Протокол клинического исследования – ключ к качеству данных]. Pharmaceutical Vestnik, 15 March 2016 [cited 2016 Jun 15]. Available from: www.pharmvestnik.ru.
11. McPhail D, Goodwin I, Gordon K. Reviewing Statistical Analysis Plans A Guide for Medical Writers. Drug Information Journal 2006;40:197-202.

Author information

Eugenia Radkova, MD, PhD, is a medical writer at a full-service contract research organisation OCT Rus (OCT Group) operating in Russia, Ukraine, Belarus, Latvia, Lithuania, Estonia. Her responsibilities include both regulatory writing and medical advising. Before joining OCT Rus she had a strong educational and research experience in pharmacology and preventive medicine. **Ivan Dobromyslov**, MD, MBA has been in clinical trials since 2006 and is the CEO of Data MATRIX Ltd (OCT Group) which provides biostatistics, data management services, and software, including EDC, IWRS, Pharmacovigilance and other CTMS instruments for global users.

Where have all the UK entry level pharmaceutical regulatory medical writing jobs gone?

Elsa Lewis and **Jonathan Oliver**

Lioness Writing Ltd, Canterbury, UK

Correspondence to:

Elsa Lewis
Lioness Writing Ltd
16 Ersham Road
Canterbury, Kent, UK
lionesswritingltd@gmail.com
+44 1227 453682

Abstract

The ways onto the regulatory medical writing ladder appear to be disappearing. Is there a reason for this recent scarcity of entry level jobs in the UK? Is this indicative of a larger problem – that of an impending danger of a skills drain from the UK in this field? The authors examined advertised job vacancies, conducted interviews and canvassed opinion on social media to explore a possible skills drain in the context of the outsourcing of regulatory medical writing. Partnerships between experienced individual independent medical writers and industry brokered by professional organisations might be a solution.

Where have all the UK entry level pharmaceutical regulatory medical writing jobs gone?

Introduction

At careers fairs for life sciences graduates and postgraduates, there has been a noticeable absence of entry level jobs for graduating students, working clinical technicians, or even for experienced regulatory affairs personnel looking for a sideways career move.

In the UK, an active seeker of medical writing posts through online job searches and recruitment agencies will discover 97% currently advertise for posts openly requiring at least two years' medical writing experience. The remaining 3% give veiled impressions that both a PhD and prior experience is still likely to be required. The only posts not requiring previous experience seem to be in the US, or elsewhere in Europe and India.

The ways onto the regulatory medical writing ladder appear to be disappearing. Is there a reason for this recent scarcity of entry level jobs in the UK? Is this indicative of a larger problem - that of an impending danger of a skills drain from the UK in this field? The authors examined advertised job vacancies, conducted interviews and

a possible skills drain in the context of the outsourcing of regulatory medical writing.

Is regulatory medical writing for the pharmaceutical industry a proper career?

For many careers within a highly qualified industry, most graduates or postgraduates starting a post might expect a probation period as they become accustomed to the workplace and the role they are expected to fulfil. Law students have articles, medical doctors have the hierarchy of a hospital to guide them and architects would have work experience before practising unsupervised. For scientific or technology teams, there would be a team for new starters to join, mentors on hand to help and training to undertake. Summer jobs and internships also help new graduates taste and see whether this might be the job that would suit them, while equipping them with experience and basic training for the post. These new starter training programmes/constructs appear to be on the decline for the medical writer, despite an increase in demand for the role.

Various medical writer journal articles give the impression that a medical writing

post is something that is a temporary position or a post occupied while one is between jobs – perhaps something to try once all other avenues are exhausted.[1] This is exacerbated when higher education and academic organisations struggle to identify what the role actually entails.

Nevertheless, the introduction of the International Council for Harmonisation (ICH) clinical study report guideline (E3), adopted in 1995, formalised the content requirements of a study report, together with specifying associated appendices and the requirement for a documented quality control and quality assurance process. This resulted in a rising demand for the medical writer role within a drug development team and the subsequent regulatory requirements to write up reports for every study enrolling patients further increased this demand.

Prior to 1995, drug development teams might only have included a medical writer as an afterthought once a project neared a study's end. Today's teams include at least one writer throughout the clinical phases. Documents now written by regulatory medical writers include protocols, which describe how each study will be run, through to submissions to regulatory authorities, which describe the benefits and risks associated with a new drug and present a case for why it should be approved for use. As identified on the ABPI website,[2] in recent years, the value of medical writing has been increasingly recognised, the status of medical writers has risen, and the range of opportunities for experienced writers has grown.

The competitive salary and volume of work for experienced medical writers in the field reflect this picture, suggesting more that it is a recognised and significant role throughout the clinical phases of drug development.[2] Opportunities for medical writing are clearly shown from the increase in advertised posts and the commitment of recruitment agencies dedicated to finding suitably qualified individuals.

The profession entails a unique mixture of technical writing and project management in a team environment and is well suited for life sciences graduates who prefer the communication aspects of the field. An effective medical writer needs experience to understand the 'shepherding' role that is required, the regulatory environment with its guidelines, and a team's working practices.

What was the entry route for graduates previously? What has changed?

As departments for medical writing emerged in pharmaceutical companies at the end of the last century, entry level jobs, internships and secondments were available for life sciences graduates and also for individuals moving from other departments within the pharmaceutical industry. Employee development structures appeared for regulatory and writing training, mentoring, development of good practice and quality control. Hierarchies emerged and managers organised lead, principal, senior and junior posts: the medical writer role became established.

However, as is also the case for the statistics role, there is a natural rise and fall in demand during the life cycle of a drug. There is a heavier workload for the medical writer as studies reach their conclusion and the final dossiers are compiled. In the increasingly difficult economic environment around the 2008 financial year, coupled with scientific challenges to the drug pipeline, the outsourcing of Research and Development roles became a way to make short-term savings.

What is the long-term future of the profession?

The release of ICH E3 in its modern form (1995) led to an increase in entry level jobs as medical writing departments grew to write the study reports and dossiers required for drug applications. The demands on the role within the development team has continued and expanded as regulatory requirements become more stringent, including, for instance, the mandatory requirement for a report for all clinical studies within 6 or 12 months of last patient visit.[3] The ICH agreement has English as the international language, which in theory would give native English speakers an advantage, especially when the UK has a good track record in providing well-qualified life science graduates and postgraduates. One would therefore expect that the UK would have little difficulty in filling a rising demand for entry level roles.

Medical writing appears to be a thriving profession with a healthy long-term future – so where are all the entry opportunities?

Outsourcing in the pharmaceutical industry

Although the need for medical writer output has increased over recent years, as with many pharmaceutical activities downstream of research, outsourcing has become the

model of choice due to perceived cost savings, risk reduction and the possibility of a more flexible response to changing demand depending on drug development success.[4] This is based on the premise that a medical writing contract is easier and cheaper to cancel than keeping a department during the times that a drug pipeline plateau is on a downward trajectory.

What is clear, however, is that solutions to a possible medical writer skills drain needs to be considered now before we lose these skills from the UK in the long-term.

Most of the medical writing departments in the pharmaceutical industry have shrunk considerably and in most pharmaceutical companies, a small group of outsourcing managers have replaced the in-house writers. The pharmaceutical company looking to outsource medical writing has three overall choices of who to outsource to:

- individual independent medical writers,
- small to intermediate medical writing specialist providers, or
- larger full service clinical research organisations (CROs)

So what are the current entry level job opportunities within these organisations?

Individual medical writers

There are currently no opportunities in development teams for an inexperienced independent medical writer. Gaining and undertaking training under the auspices of a professional organisation is currently expensive and still does not provide the on-the-job training required. Moreover, the two-year experience criterion would still be lacking. Internships and summer posts are in the decline as the medical writing departments within pharmaceutical companies disappear. At present, the only theoretical entry opportunity for a new starter in this area would be to find an experienced medical writer who is prepared to not only train and provide the experience, but also negotiate contracts that allow for a mentoring arrangement. In terms of business practice, however, this potentially looks as though individual writers would be investing a great deal in order to create their own competition.

Medium-sized medical writing specialist groups

With the transition into outsourcing models over the past 10 years or so, some forward-thinking independent writers (in some cases alongside statisticians and other departments under threat) set up small specialist organisations. Whole pharmaceutical industry medical writing departments were relocated off-site and re-employed by new independent service provider groups. Some of these specialist groups have grown to be nearly full service CROs and others have remained specialised (for example, purely medical writer or as an adjunct to statistics) and grown more established over the years. Based on observation of the lack of posts advertised, entry level jobs have existed in the past within these organisations, but in recent years in the UK it would appear that recruitment for many of these groups has been in decline or on hold. Entry opportunities in this area may well be diminishing as a result of increasing competition with the economies of scale possible in larger CROs. Experienced specialist expertise is the one staple competitive edge that such groups would offer. Indeed, these were the only organisations that were identified at the time of our research that currently offer any training posts for new graduates.

Clinical research organisations

These can be full service organisations providing for example a full team of clinical, statistical and regulatory writing support to the pharmaceutical industry. The largest 10 CROs are currently Quintiles, Parexel, Pharmaceutical Product Development (PPD), INC Research, Covance, Medpace, PRA Health Sciences, inVentiv Health, Meditrial Europe and Chiltern. A casual job search conducted in April 2016 revealed only one potential medical writer entry level position within the UK. A small number of jobs were available in USA, but not many at all in Europe. Vacancies appear to be in locations elsewhere than the UK, for a number of possible reasons: because head offices of these CROs are not in the UK, or the 2008 financial downturn is casting a long shadow, or in the US they have had more experience with outsourcing. Given the advantages that UK medical writing has to offer, the current situation seems inadequate and unsustainable.

What does this mean for the future?

The outsourcing model remains in a state of flux for drug research and development with differing levels of in house or outsourced roles for each company. If it is anything like the general global trend in outsourcing, it is likely to come under more frequent review. What appears to be the tendency is for roles further downstream, such as medical writing, to remain outsourced, which in the long-term will lead to a move away from using new medical writers based in the UK unless proactive, more long-term solutions are considered.

Universities have been challenged, for instance by the government's partnership approach for equipping the next generation with relevant suitable skills.[5,6] They have attempted to rise to the challenge, for instance, Worcester University had a Medical Writing Masters course, Canterbury Christ Church University was considering a regulatory module to its course, and University College London has a clinical trials module. However, none of these initiatives will lead to careers unless there are vacancies for new starters. Indeed the Worcester course has closed down and the regulatory module has now been refocused on research.

Possible solutions

The globalisation of recruitment and the advance of communication technologies is one trend that will continue to aid future experienced individuals. The role of a medical writer can be performed adequately remotely by an experienced writer – indeed it is quite common for development teams to be based on two or even three continents with large teleconferences and screen sharing technology bringing them together on a regular basis. But this still does not

address the problem of individuals within the UK, even those looking for a sideways career move, getting a foot on the ladder without some forward planning.

Partnerships between experienced individual independent medical writers and industry brokered by professional organisations might be a solution. Indeed, it looks as though moves are being made in this direction by larger organizations. Mentoring of new graduates over a two-year period is very costly to a small company, since extensive work shadowing, training and oversight is required. Only a handful of small specialist medical writing companies currently seem to have this environment open to new starters. Yet, if this effort could be shared within a collaborative structure, coupled with appropriate contractual agreements, there is more chance of it happening. If a highly organised, industry wide collaboration took place it would work to enhance a higher level of quality and add a more professional and competitive edge. Without such pressure, there is little incentive for independent medical writers to undertake this, as after the apprentice period is served, any enterprising experienced medical writer would currently leave to find their own contracts.

Perhaps one of the most forward-thinking solutions can be found through the recruitment agencies themselves. It is possible for a more specialist recruitment organisation to set about creating both a training academy and funding pool for graduates to gain the necessary experience and skills to join the industry.[7] Such programmes do exist and are in their infancy. Ensuring the necessary two years experience to take on such a role, however, remains something that is still being developed and would require considerable financial backing and resources.

Currently efforts are being made to bring industry and new graduates together at careers fairs and indeed regulatory professionals looking to move sideways into different roles. These efforts aid the process of communicating different vacancies that exist, and maybe could be expanded to seed partnering within the industry.

What is clear, however, is that solutions to a possible medical writer skills drain needs to be considered now before we lose these skills from the UK in the long-term.

References

1. University of Kent Careers and Employability Service. I Want to Work in … Science Writing and Medical Communications. 2012 [cited 2016 May 11]. Available from: https://www.kent.ac.uk/careers/workin/sciencewriting.htm.
2. Association of the British Pharmaceutical Industry. Medical Writing. 2016 [cited 2016 May 09]. Available from: http://careers.abpi.org.uk/working-in-the-industry/research/Pages/medical-writing.aspx.
3 European Commission. Detailed guidance for the request for authorisation of a clinical trial on a medicinal product for human use to the competent authorities, notification of substantial amendments and declaration of the end of the trial. 2005 [cited 2016 May 9]. Available from: http://ec.europa.eu/health/files/pharmacos/docs/doc2005/10_05/ca_14-2005_en.pdf.
4. KPMG. Outsourcing in the Pharmaceutical Industry: 2011 and beyond. 2012 [cited 2016 May 11]. Available from: https://www.kpmg.com/Ca/en/IssuesAndInsights/ArticlesPublications/Documents/Outsourcing-pharmaceutical-industry.pdf.
5. GOV.UK. Horizon Scanning Programme Team. 2016 [cited 2016 May 9]. Available from: https:// www.gov.uk/ government/ groups/horizon-scanning-programme-team.
6. Universities UK. The Funding Environment for Universities 2015. The Economic Role of Universities. 2015 [cited 19 Jul 2016]. Available from: http://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2015/the-economic-role-of-uk-universities.pdf.
7. Clinical Professionals. Clinical Professionals Training Academy. 2016 [cited 2016 May 11]. Available from: http://www.clinicalprofessionals.co.uk/graduateacademy.php.

Acknowledgements

Thank you to the following for contributing their views and/or help with this opinion piece. Claire Huntley, Tracey Wainwright, Justina Orleans-Lindsay, Gail Head, Chris Shilling, David Jolley, Tim Moody, Elizabeth Jenkins, Beth Gawthorpe, Tim Griffiths, Kerry Walker, Gauri Rao, Hazel Olway, Debbie Rogers, Yvette Cleland, Andrew Cooper, Erik Smit.

Author information

Elsa Lewis, Director and Regulatory Medical Writer with Lioness Writing Ltd with 18 years writing experience in the Pharmaceutical Industry.

Jonathan Oliver, an independent consultant for Lioness Writing Ltd, with experience in publishing and outsourcing.

Brussels 2016

The 43rd EMWA Conference will be held on 3 - 5 November 2016 at The Sheraton Brussels Hotel, Brussels, Belgium.

This conference will offer 28 foundation- and advanced-level workshops covering a wide range of topics. A number of special events outside the formal education programme are also included as part of the conference.

Thursday, November 3

- Welcome Event and Networking Reception

Friday, November 4

- Easy morning yoga session
- Freelancer session on sharing best practices in IT
- Freelance Business Forum
- Social programme

Saturday, November 5

- Easy morning yoga session
- Seminar on Introduction to Medical Writing
  (This event is open to the public and is free of charge)

The EMWA spring and autumn conferences provide forums for networking, active discussions, and extensive, cost-effective, professional training for EMWA members. The venues and career-enhancing programmes are chosen to offer the best possible learning environment. In addition, EMWA conferences offer an excellent opportunity to benefit from the experiences of other medical writers.

The conferences have a relaxed, friendly atmosphere that is ideal for networking opportunities and that encourages attendees to meet medical writers and communicators at all stages in their careers. Registration for 2016 Autumn Conference at Brussels is now open. To view the programme and sign up for workshops, please visit the EMWA website at www.emwa.org.

News from the EMA

More information can be found on the Agency's website: www.ema.europa.eu.

The articles included in this section are a selection from the European Medicines Agency's news and press release archive for April 2016 to July 2016.

Listening to the public's views on the safety of medicines

PRAC adopts rules of procedure on public hearings on selected safety reviews

April 15, 2016 – The European Medicines Agency's Pharmacovigilance Risk Assessment Committee (PRAC) has adopted the final rules of procedure for public hearings to be held by the Committee. The rules of procedure describe the process and practical arrangements for the preparation, conduct, and follow-up of public hearings.

As part of the implementation of these rules, the European Medicines Agency (EMA) will now organise an internal dry run exercise in order to test the process and procedures of public hearings. The dry run is scheduled to take place at the PRAC meeting in July 2016. Public hearings could take place as early as the fourth quarter of 2016, as soon as a relevant topic is identified.

Public hearings are a new tool for EMA to engage European Union (EU) citizens in the supervision of medicines and to listen to their views and experiences. The pharmacovigilance legislation has given the PRAC the possibility to hold public hearings as part of certain safety reviews of medicines, particularly in relation to their therapeutic effects and available therapeutic alternatives, as well as the feasibility and acceptance of proposed risk management and minimisation activities.

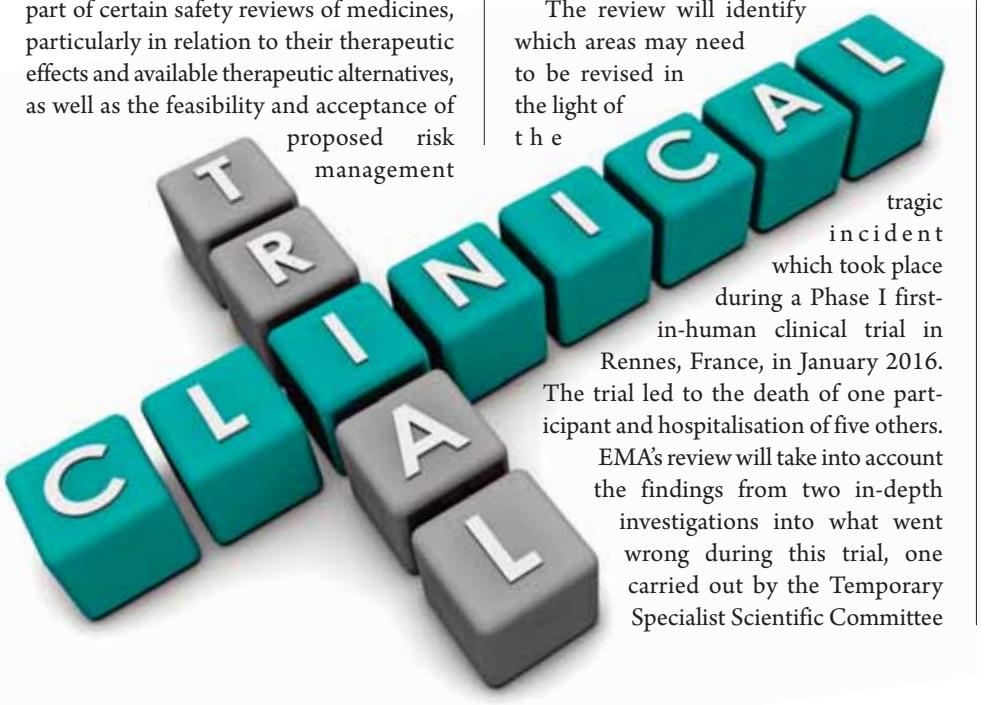
Contributions made by the public during a public hearing will be considered by the PRAC and inform the Committee's decision-making. Public hearings will be held on a case-by-case basis, where the Committee determines that collecting the views of the public would bring added value to its review. More details are outlined in the rules of procedure document.

Draft rules of procedure were published by the Agency for comments in July 2014 and drew 200 comments from 22 stakeholder contributions representing 25 organisations. The rules were updated and revised in light of the comments received.

Improving safety of first-in-human clinical trials

EMA starts EU-wide reflection on necessary changes to best practices

May 27, 2016 – The EMA has started a review of the guidelines that describe first-in-human clinical trials and the data needed to enable their appropriate design and allow initiation. This is being done in cooperation with the European Commission and the Member States of the EU.



The review will identify which areas may need to be revised in the light of the tragic incident which took place during a Phase I first-in-human clinical trial in Rennes, France, in January 2016. The trial led to the death of one participant and hospitalisation of five others. EMA's review will take into account the findings from two in-depth investigations into what went wrong during this trial, one carried out by the Temporary Specialist Scientific Committee (TSSC) set up by the French medicines agency ANSM, and the other by the Inspection Générale Des Affaires Sociales (IGAS), the inspectorate for socia,l affairs in France.

Both reports include a series of recommendations regarding the requirements for authorisation and conduct of first-in-human clinical trials for further examination by the international regulatory and public health community.

EMA's work will focus on best practices and guidance. The aim is to agree a concept paper by July identifying areas for change and proposals to further minimise the risk of similar accidents. The concept paper will form the basis for an EU-wide review of the guidelines. This process will include targeted discussions with stakeholders and a public consultation on proposed changes later in 2016.

The EMA review has started with two groups of experts who are carrying out preparatory work. One group is looking at pre-clinical aspects and the data needed from laboratory tests or animal studies to safely initiate first tests in humans. The other group is looking at clinical aspects of the design of first-in-human trials and how these could be improved to better ensure the safety of human volunteers taking part in these trials. This will lead into one EU-wide expert group discussion

SECTION EDITORS

Section Editor:
Anuradha Alahari
Anuradha.Alahari@parexel.com

EMA contact:
Monika Benstetter
press@ema.europa.eu

on revision of guidelines.

Clinical trials are essential for the development of medicines and without them patients cannot gain access to new potentially life-saving medicines. In the EU, the approval and conduct of clinical trials is within the remit of the relevant authorities of the European Member States.

EU guidelines are in place to ensure that these clinical trials are conducted as safely as possible. These guidelines include the requirement for extensive studies, including in animals, to gather information about a medicine before it is given to humans.

Severe adverse reactions in healthy volunteers such as those observed in the trial in Rennes are extremely rare during clinical trials. Since 2005, approximately 14,700 phase I clinical trials (with participation of 305,000 subjects) have been conducted in the EU, including 3,100 first-in-human studies. Only one other severe incident has been previously reported in that time in the EU.

Single, central platform now mandatory for all periodic safety update reports

PSUR repository facilitates information exchange on the safety of human medicines authorised in the EU

June 10, 2016 – As of June 13, 2016, all periodic safety update reports (PSURs) for human medicines authorised in the EU must be submitted to the PSUR repository, which has been developed by the EMA in close collaboration with EU Member States and the industry.

The PSUR repository is a single, central platform for PSURs and related documents to be used by all regulatory authorities and pharmaceutical companies in the EU. It was introduced by the EU pharmacovigilance legislation to facilitate the exchange of information on the safety of authorised medicines between regulators and pharmaceutical companies.

Marketing authorisation holders must now use the repository as a single point for all submissions and should no longer submit their PSURs to national competent authorities. The eSubmission Gateway is available on the eSubmission website.

The PSUR repository provides an important simplification for marketing authorisation holders allowing them to send all PSURs to a single recipient. It also facilitates the assessment of the reports by ensuring that national competent authorities, EMA and its scientific committees have timely and secure access to all relevant documents.

In June 2015, EMA's Management Board gave the green light for the use of the repository following an independent audit that confirmed that the tool meets the agreed functional specifications. Since the initial release of the PSUR repository in January 2015, EMA has been supporting companies and national competent authorities to ensure they are ready to use this new tool. The system has been implemented

Regulation of advanced therapy medicines

Report details concrete proposals to encourage development and authorisation of advanced therapy medicinal products (ATMPs) in the EU

June 3, 2016 – The EMA today published a report from a multi-stakeholder expert meeting held on May 27, 2016 to explore possible ways to foster the development of ATMPs in Europe and expand patients' access to these new treatments.

ATMPs comprise gene therapies, tissue engineered products and somatic cell therapies. These medicines have the potential to reshape the treatment of a wide range of conditions, particularly in disease areas where conventional approaches are inadequate. However, eight years since EU legislation on ATMPs entered into force in 2008, only five ATMPs are currently authorised. At the same time clinical trials investigating ATMPs appear to represent a fast-growing field of interest, underlining the need to better support innovation through a coherent and appropriate regulatory environment.

The meeting brought together leading academics and researchers, representatives from patients' and healthcare professionals' organisations, small and large pharmaceutical companies, the investment community, incubators and consortium organisations, health technology assessment (HTA) bodies, national competent authorities and the European Commission. In their discussions they focused on four key areas:

- Facilitating research and development
- Optimising regulatory processes for ATMPs
- Moving from hospital exemption to marketing authorisation
- Improving funding, investment and patient access

Ideas and solutions proposed by the different stakeholders are summarised in the meeting report published today. Some of the recurring themes include the need for early interaction and guidance from regulators, more transparency and information sharing, greater harmonisation between Member States on various aspects of the ATMP legislative framework and measures to tackle inequalities in patient access to ATMP treatments.

EMA and its scientific committees, together with the European Commission and the national competent authorities, have started discussing the proposals made during the meeting. Concrete actions will be determined over the next few months and shared with stakeholders.

Notes

- Although a total of seven ATMPs have received a marketing authorisation since 2009, only five ATMPs are currently authorised. One marketing authorisation for an ATMP was withdrawn by the marketing authorisation holder and the authorisation for another ATMP is currently suspended.
- For a recent analysis on clinical trials with ATMPs see Hanna E, Remuzat C, Auquier P, Toumi M. Advanced therapy medicinal products: current and future perspectives. J Mark Access Health Policy 2016;4.

The EU Medicins Agency

using a phased approach and feedback from users has been taken into account to improve the system. Guidance, interactive training sessions and links to all relevant documents have been made available on EMA's eSubmission website.

PSURs are reports providing an evaluation of the benefit-risk balance of a medicine. Marketing authorisation holders must submit PSURs at defined time points following a medicine's authorisation. PSURs include the results of all studies carried out with this medicine, both in its authorised and unauthorised uses.

EMA uses the information in PSURs to determine if there are new risks identified for a medicine or whether the balance of benefits and risks of a medicine has changed. It can then decide if further investigations need to be carried out or can take action to protect the public from the risks identified, for example by updating the information provided for healthcare professionals and patients.

EMA goes electronic for PDCO opinions and subsequent EMA decisions

PDCO opinions and subsequent EMA decisions will be transmitted to applicants electronically only

July 6, 2016 – From August 1, 2016 the EMA will transmit the opinions of the Paediatric Committee (PDCO) and subsequent EMA decisions to applicants in electronic format only. Applicants will no longer receive paper versions.

PDCO opinions and subsequent EMA decisions will be sent to applicants as a PDF via EudraLink – the European medicines regulatory network's secure file-transfer system. EMA decisions, as well as PDCO opinions will no longer contain a signature.

The date when the EudraLink message is opened by applicants for the first time will be considered as the day of the receipt of the document attached to the EudraLink message, for the purpose of calculating procedural timelines in accordance with Regulation (EC) No. 1901/2006. EudraLink automatically records this date as "access by".

Applicants should download and archive the attached documents upon receipt, as Eudralink preserves file attachments only for up to 90 days. EMA will retain a read-only version of the electronic documents in its electronic archives. Further information can be found on Paediatric investigation plans: questions and answers, under the section "Applying for a Paediatric Investigational Plan waiver or deferral". Applicants will be offered the possibility to opt-out and receive documents as hard copy instead.

The move from printouts to electronic documents responds to stakeholders' feedback collected over the years. Among the benefits of this change are: accelerated delivery of documents, more convenient receipt of documents as well as a shift towards greener solutions in line with EMA's environmental policy.

EMA statement on the outcome of the UK referendum

EMA's procedures and work streams continue as usual

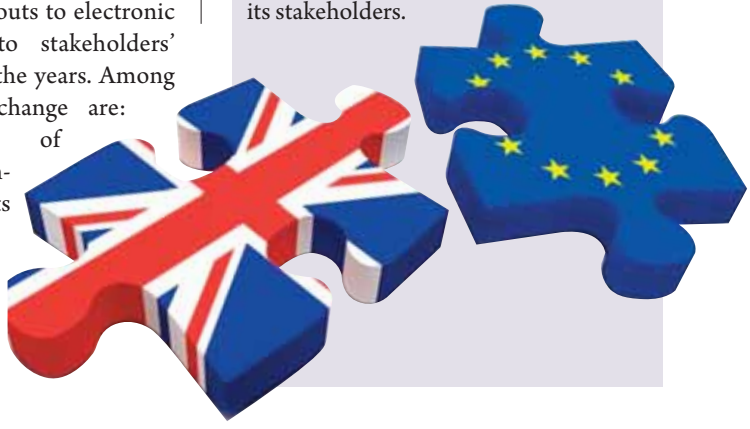
July 6, 2016 – The EMA acknowledges the outcome of the referendum of June 23, 2016. A majority voted against United Kingdom's (UK) continued membership of the EU and it is now up to the UK government to decide how to act upon the outcome of the referendum.

EMA would like to underline that its procedures and work streams are not affected by the outcome of the referendum. The Agency will continue its operations as usual, in accordance with the timelines set by its rules and regulations.

No Member State has ever decided to leave the EU, so there is no precedent for this situation. The implications for the seat and operations of EMA depend on the future relationship between the UK and the EU. This is unknown at present and therefore we will not engage in any speculations.

EMA welcomes the interest expressed by some Member States to host the Agency in future. The decision on the seat of the Agency will however not be taken by EMA, but will be decided by common agreement among the representatives of the Member States.

The European Regulatory Network as a whole is a very strong and flexible system that is able to adapt to changes without jeopardising the quality and effectiveness of its work. The Agency is in close contact with the EU institutions. As soon as concrete information will become available, EMA will share it with its stakeholders.



Journal Watch

Journal Watch is based on the French-language blog *Rédaction Médicale et Scientifique*, available at http://www.redactionmedicale.fr.

SECTION EDITOR

Hervé Maisonneuve
herve@h2mw.eu

Rather than reporting isolated P values, articles should include effect sizes and uncertainty metrics

This 8-page paper published in *JAMA* assessed the reporting of *P* values in the biomedical literature from 1990 to 2015. This huge piece of work used text mining to identify 4,572,043 *P* values in 1,608,736 MEDLINE abstracts and 3,438,299 *P* values in 385 393 PMC full-text articles. The reporting of *P* values in abstracts increased from 7.3% in 1990 to 15.6% in 2014. In 2014, *P* values were reported in 33.0% of abstracts (n = 29,725 abstracts), 35.7% of meta-analyses (n = 5,620), 38.9% of clinical trials (n = 4,624), 54.8% of randomised controlled trials (n = 13,544), and 2.4% of reviews (n = 71,529).

The distribution of reported *P* values in abstracts and in full-text articles showed strong clustering at *P* values of 0.05 and of 0.001 or smaller. *P* values reported in abstracts were in general lower (showing greater statistical significance) than *P* values reported in the full-text articles. Besides the substantial proportion of abstracts that report *P* values, a larger proportion of abstracts included qualitative statements about significance, mostly without any other quantitative information. Few articles included confidence intervals, Bayes factors, or effect sizes. The authors suggested that rather than reporting isolated P values, articles should include effect sizes and uncertainty metrics.

Reference: Chavalarias D, Wallach JD, Ting Li AH, Ioannidis JPA. Evolution of reporting P values in the biomedical literature, 1990-2015. JAMA 2016;315(11):1141-1148.

Further research should elucidate if and to what degree quotation errors are detrimental to scientific progress

The case is simple: citations are an essential element of manuscripts, but 25% do not serve their purpose! In a systematic review on quotation accuracy, 559 studies were screened, of which 28 were included in the main analysis, and the estimated major, minor and total quotation error rates were 11,9% (95% CI [8.4, 16.6]) 11.5% (95% CI [8.3, 15.7]), and 25.4% (95% CI [19.5, 32.4]), respectively. While heterogeneity was substantial, even the lowest estimate of total quotation errors was considerable (6.7%). Indirect references accounted for about one sixth of all quotation errors.

The strategies suggested for reducing quotation errors were: spot checks by editors and reviewers, correct placement of citations in the text, declarations by authors that they have checked cited material.

Reference: Jergas H, Baethge C. Quotation accuracy in medical journal articles – a systematic review and meta-analysis. Peer J. 2015;3:e1364.

There is poor performance and noticeable variation in the dissemination of clinical trial results across leading academic medical centers.

The objective of this study was to determine rates of publication and reporting of results within 2 years of completion for all clinical trials registered in ClinicalTrials.gov by leading academic medical centres in the United States. A total of 4,347 interventional clinical trials were identified across 51 US academic medical centers between October 2007 and September 2010. Overall, results were disseminated for 2,892 (66%) trials, with 1,560 (35.9%) within 24 months of study completion.

Additional tools and mechanisms are needed to rectify this lack of timely reporting and publication, as they impair the research enterprise and threaten to undermine evidence-based clinical decision making.

Reference: Chen R, Desai NR, Ross JS, Zhang W, Chau KH, Wayda B, et al. Publication and reporting of clinical trials results: cross sectional analysis across academic medical centers. BMJ 2016;352:i637.

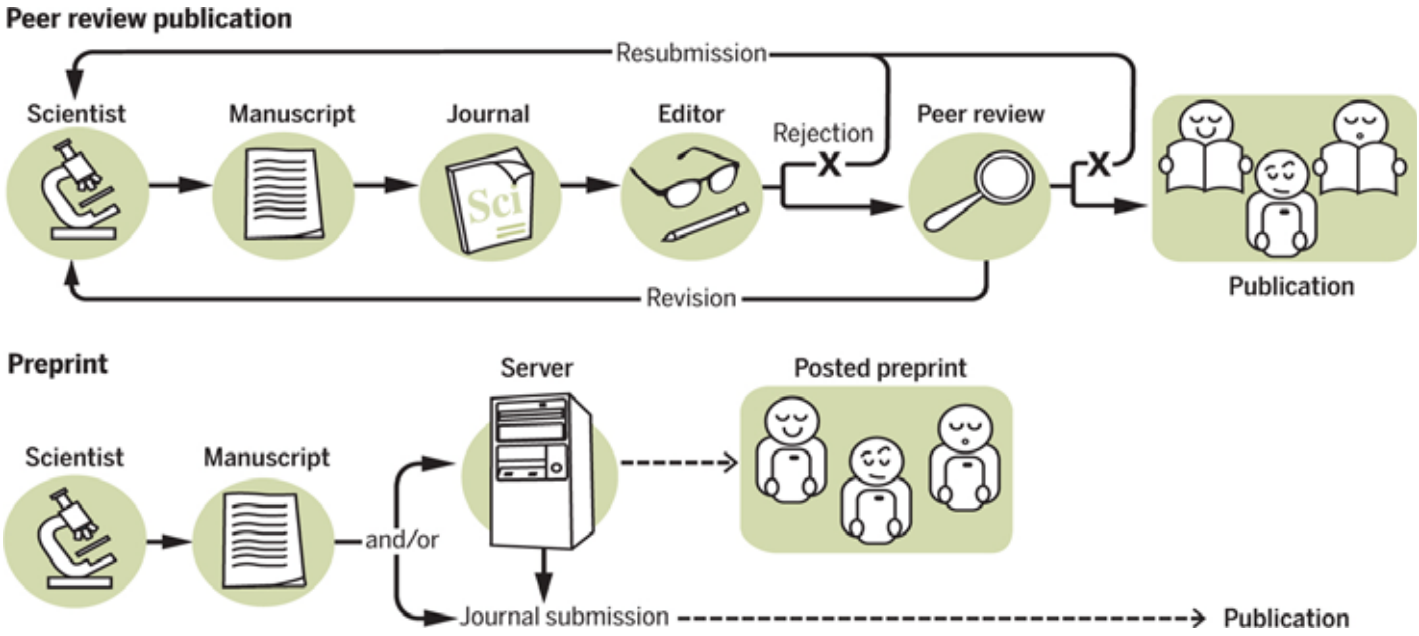


Figure 1: Peer review and preprints in the life science, as proposed by Accelerating Science and Publication in biology (http://asapbio.org/)

The time is right for biologists to post their research findings onto preprint servers: Accelerating Science And Publication in biology (ASAPbio)

The ASAPbio meeting (Feb 2016) was held to explore the wider use of preprints for disseminating ideas and results in the life sciences. *"A preprint is a complete scientific manuscript (often one also being submitted to a peer-reviewed journal) that is uploaded by the authors to a public server without formal review. After a brief inspection to ensure that the work is scientific in nature, the posted scientific manuscript can be viewed without charge on the web."*

The preprint server arXiv.org has been essential in the fields of physics, mathematics, and computer sciences for over two decades. Will such servers be implemented in other scientific fields?

This paper has 3 parts presenting the perspectives of Academics, Funders and Publishers. Stakeholders have different views and all suggest to rapidly change the publication system, moving to preprints. Servers are ready to serve such an objective, and biologists will see opportunities, as well as clinicians.

Reference: Berg JM, Bhalla N, Bourne PE, Chalfie M, Drubin DG, Fraser JS, *et al.* Preprints for the life sciences. Science 2016;352:899-901.

Authors of systematic reviews are on the front line to detect research misconduct

An analysis of 118 systematic reviews published in 4 journals (Ann Int Med, BMJ, JAMA, Lancet), and the Cochrane Library was carried out in 2013 to analyse application of procedures to counter-balance 6 forms of malpractices: 1. publication bias (through searching of unpublished trials), 2. selective outcome reporting (by contacting the authors of the original studies), 3. duplicate publications, 4. sponsors' and 5. authors' conflicts of interest on the conclusions of the review, and 6. ethical approval of the studies.

Overall, 59 (50%) reviews applied 3 or more procedures; 11 (9%) applied none. The extracted data were confirmed by 68% of the authors of the systematic reviews. Seven reviews suspected misconduct, of which 5 did not report it, and 2 reported it explicitly. The suspected cases were data falsification (3 reviews), data manipulation

(1 review), difference in data between the published trial and the re-analysed data posted on the FDA website (1 review), and selective reporting of outcomes (2 reviews). The risk related to double counting of participants due to duplicate publications and the risk of selective reporting of outcomes were recognised by most authors (69 %). In general, conflict of interest was underestimated.

Reference: Elia N, Elm E von, Chatagner A, Pöpping DM, Tramèr MR. How do authors of systematic reviews deal with research malpractice and misconduct in original studies? A cross-sectional analysis of systematic reviews and survey of their authors. BMJ Open 2016;6:e010442

The endorsement of CONSORT by high impact journals has increased over time

First published 20 year ago, the CONSORT reporting guidelines have received widespread attention. The 1996, 2001 and 2010 publication of the guidelines, the CONSORT statement and elaboration documents have been cited more than 12,000 times (Scopus, May 2015). Published in June 2016 in *Trials*, this is the third study evaluating the endorsement of CONSORT by journals. The mention of CONSORT in the online "Instructions to Authors" given by 168 high impact journals that were included in this study was examined (Table 1). CONSORT was mentioned in the "Instructions to Authors" by 63% of the journals, and was defined as mandatory by 42% for reporting of trails. The endorsement of CONSORT by high impact journals has increased over time, although the implementation is far from standardised (Table 1). There is still room for improvement to encourage compliance with CONSORT.

Table 1: Mention of CONSORT, ICMJE, and trial registration in the "Instructions to Authors" from the top impact factors journals in 2001, 2006 and 2012

	2003[a] N = 166 n (%)	**2007**[b] N = 165 n (%)	**2014**[c] N = 168 n (%)
CONSORT statement	36 (22 %)	62 (38 %)	106 (63 %)
ICMJE	72 (43 %)	69 (42 %)	130 (77 %)
Trial registration	Not collected	61 (37 %)	106 (63 %)

Abbreviations: IF: Impact Factor; ICMJE: International Committee of Medical Journal Editors; N =number of articles screened; n = number of articles that mentioned CONSORT, ICMJE, or trial registration in the "Instructions to Authors" .
[a]2001 IF; [b]2006 IF; [c]2012 IF
89 journals were included in each of the above 3 groups.

Reference: Shamseer L, Hopewell S, Altman DG, Moher D, Schulz KF. Update on the endorsement of CONSORT by high impact factor journals: a survey of journal "Instructions to authors" in 2014. Trials. 2016:17:301

Average number of authors per MEDLINE citation is still on the rise

The US National Library of Medicine has published extracts from the 2016 Statistical Reports on MEDLINE®/PubMed® Baseline Data (https://www.nlm.nih.gov/bsd/authors1.html). In 2015, they were on average, 5.48 authors on a paper, compared to 1.50 in the 1950s (Figure 2, orange line). The collective author names (also known as group names or corporate names) did not increase over time (Figure 2, blue line). For the top 25 publishing countries, the top 5 pairs of collaborating countries, based on author affiliations, were: 1. US and China (14,853 papers), 2. US and the United Kingdom (11,384), 3. US and Germany (8,421), 4. US and Canada (8,044), and 5. Germany and the United Kingdom (7,955).

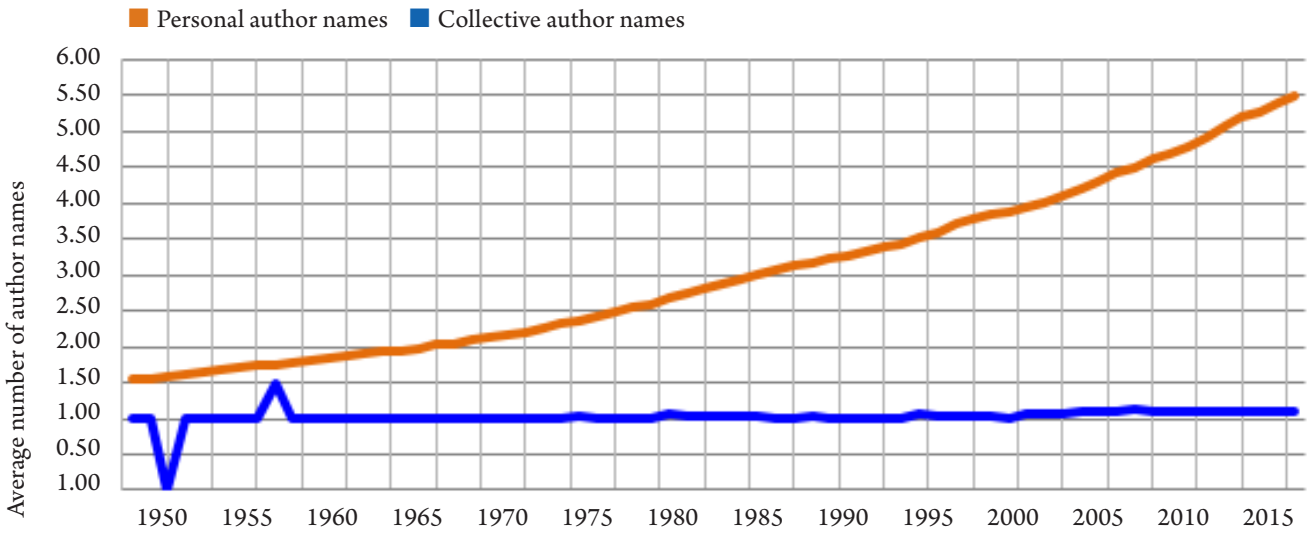


Figure 2: ***Average number of personal names or collective author names per MEDLINE/Pubmed citation per year from 1950 to 2015***

In the Bookstores

Essential Statistics for the Pharmaceutical Sciences (Second Edition)

By Philip Rowe;
John Wiley & Sons, 2016.
ISBN: 978-1-118-91339-0 (paperback).
£42.50. 308 pages.

SECTION EDITORS

Stephen Gilliver
stephen.gilliver@gmail.com
Alison McIntosh
alison.mcintosh@iconplc.com

This book aims to explain statistics to "those who have to use statistics, but have no ambition to become statisticians *per se*". The author, Philip Rowe, teaches at a university department of pharmacy and clearly understands that many students and researchers who need to understand statistics find it daunting. Although the book is aimed primarily at pharmacy students, much medical research involves both drugs and statistics, so medical writers should also find it a helpful introduction or refresher.

The book is divided into five sections, starting with presenting data, then covering the statistical tests that should be used with different types of data (continuous, nominal, and ordinal) and concluding with a section on other topics such as survival analyses and questionnaires. Most of the 25 chapters cover a specific statistical test such as t-tests, or an aspect of statistics such as confidence intervals.

Two aspects of the book I particularly liked were the way it uses diagrams to replace equations and the fact that complex (and potentially scary) mathematical calculations are firmly relegated to the appendices. These features make the text accessible to non-mathematically minded readers and should reduce the panic that such readers often feel when faced with pages of equations. This is important because, if that panic sets in, many readers give up and develop an allergy to statistics. As Rowe notes in the preface, many other books "place far too much emphasis on the mechanical number crunching of statistical procedures". In contrast, his book seeks to explain the important principles underlying the statistical tests while avoiding giving too much detail. The cleverly designed graphics will also help visual learners who, in my experience, often struggle with books on statistics that present either uninterrupted text or off-putting mathematical equations. Such presentation is not only daunting, but may be ineffective if it fails to convey the underlying principles. As the author comments, most people have access to statistical software packages, so most problems arise from failing to use and understand statistics correctly rather than "the number-crunching". Nevertheless, for anybody who needs to use statistical tests (rather than just write or edit material describing them), the book contains helpful guidance on how to enter data into common statistical software packages and provides a link to a free website developed by the author that offers help with Minitab and SPSS.

Two aspects of the book I particularly liked were the way it uses diagrams to replace equations and the fact that complex (and potentially scary) mathematical calculations are firmly relegated to the appendices.

Each chapter is highly structured, starting with a concise summary of what will be covered, and with text helpfully broken up by key points in boxes and plenty of headings, making it easy to navigate. Another nice feature is the use of "pirate boxes" – black-rimmed paragraphs accompanied by grinning skull and crossbone icons which alert readers to statistical dangers and trickery. Throughout the book the text has a chatty, informal style, and the pirate boxes, while highlighting real dangers, are often presented with humour (such as the one headed "Beware of drug companies bearing Odds Ratios").

This book doesn't set out to provide detailed guidance on the use of statistics or presentation of data in medical publications (for which my favourite resource remains Lang and Secic's excellent book "*How to Report Statistics in Medicine*").[1] Perhaps future editions could include more detail on this, such as guidance about the types of graphs that are suitable for journal articles or posters and the best ways to produce these. However, Rowe includes some insights into statistical language which should be useful for writers, such as a nice section on the meaning of the term "risk". The final chapters on multiple testing, survival analysis, and questionnaires are likely to be particularly relevant to writers. They also include some sharp criticism of the sloppy practices of many journals and the need for more rigorous statistical reporting.

Overall, although this book isn't written for medical writers, I think many would find it useful, and the refreshing approach will be especially appreciated by those, like me, who always feel they ought to know more about statistics, but find other texts impenetrable.

Reference

1 Lang TA, Secic M. How to Report Statistics in Medicine: Annotated Guidelines for Authors, Editors and Reviewers. 2nd ed. Philadelphia: American College of Physicians; 2006.

Reviewed by
Elizabeth Wager
Publications Consultant, Sideview
liz@sideview.demon.co.uk

The Webscout

The history of statistics

When did mankind start using statistics and for what purposes? The history of statistics includes names like Bernoulli, Laplace, Gauss, Bayes, and Pearson. I guess you will have heard of some or all of these famous people, whose theories still play an important role in our daily business. Think of Pearson's coefficient or the Gaussian distribution. Wikipedia gives you a comprehensive overview of the development of modern statistics and its main contributors: http://en.wikipedia.org/wiki/History_of_statistics.

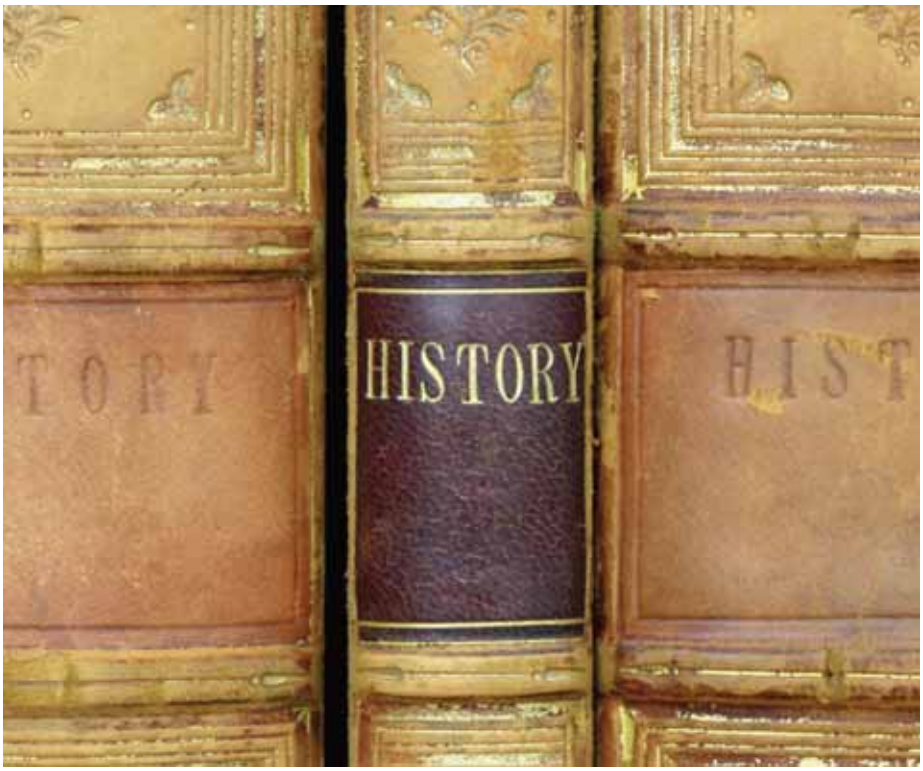
If you want to find more details, including important milestones by era, you can access http://www.economics.soton.ac.uk/staff/aldrich/figures.htm. According to this website, the origins of statistics lie in the period 1650 to 1700. For each subsequent period the website not only gives you the key events, but also a paragraph on the contributors, with details on their professional background, work, and achievements. For example, you can learn that the man behind Student's t-test was actually named William Sealy Gosset, but called himself "Student". He was a chemist who worked for the Dublin brewery Guinness. The Website further provides you with a huge number of links to other resources.

SECTION EDITOR

Karin Eichele
info@mediwiz.de

An appealing timeline of statistics is given by the American Statistical Association at http://www.statslife.org.uk/history-of-stats-science/1190-the-timeline-of-statistics. It explains that Gosset developed his t-test to ensure that every brew tastes equally good. The timeline goes far beyond the modern history of statistics. Indeed, statistics in some form was already being used in ancient times. The first event in the above-mentioned timeline is dated 450 BC, when Hippias of Elis used average values to estimate the date of the first Olympic Games.

Another way to look at the history of statistics is to review the history of a specific theory. Sharon Bertsch McGrayne, who has a professional background as a newspaper reporter and freelance scientific writer, does this in her popular book "The Theory That Would Not Die", which summarises the history of Bayes' theorem. The book's subtitle is "How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy". You can find great stories inside, like the one about a lost submarine that even inspired a famous Hollywood movie, "The Hunt for Red October". Bayes' theory, although an established standard approach nowadays, was once controversial. McGrayne's book establishes a link between statistical theories and the influence they can have on world history, society, and medicine. A summary of the book is given at http://lesswrong.com/lw/774/a_history_of_bayes_theorem/. Alternatively, you can listen to the author herself speaking about the book at Talks at Google: https://www.youtube.com/watch?v= 8oD6eBkjF9o.



A further example of the influence of statistics is the story of Florence Nightingale. She is often referred to as the founder of modern nursing. She was also a pioneer in statistical illustrations and statistics in health policy. She developed a polar area diagram, the so-called Coxcomb, to illustrate her statistical results on sanitary conditions in military hospitals for Queen Victoria. Her work led to health reforms in the United Kingdom. Nightingale became the first female member of the Royal Statistical Society and later of the American Statistical Association. Her life and work are summarised at http://blogs.sas.com/content/jmp/2013/02/04/celebrating-statisticians-florence-nightingale/ and http://en.wikipedia.org/wiki/Florence_Nightingale.

To close this Webscout, this link gives a nice summary of the history of statistics: https://www.youtube.com/watch?v=DeX5CQJ_S40.

Did you like this Webscout article? Do you have any questions or suggestions? Please feel free to get in touch and share your thoughts.

Good Writing Practice

Syntactic dissonance and impeded immediate comprehension
Coordination nonparallelism

SECTION EDITORS

Wendy Kingdom
info@wendykingdom.com
Amy Whereat
amy.whereat@speakthespeech.fr

Introduction

Coordination nonparallelism is the lack of structural symmetry between coordinated sentence constituents that are intended to be equivalent in importance. A classic example of such nonparallelism is "I love fishing, swimming, and to run."

In this article, examples of nonparallel coordination are adverb and adverbial (Part 1); noun and nominal, noun and noun phrase (Part 2); verb type (linking and intransitive or transitive; Part 3); and verb voice (active and passive; Part 4).

Part 1 – Adverb and adverbial

Example: 'reversibly' and 'specifically'

This example is from an introduction section: experimental approach.

> *This technique depends on biological molecules binding to other molecules reversibly and* with high specificity.

A nonparallelism exists between coordination of the adverb *reversibly* with an adverbial prepositional phrase *with high specificity*. In the revision, the adverbial prepositional phrase *with high specificity* is transformed into the adverb *specifically* to be parallel with *reversibly*.

> *This technique depends on biological molecules binding to other molecules reversibly and* ***specifically****.*

Notes

a. Nonparallelism between an adverb and an adverbial (a syntactic unit that functions adverbially) is a minor distraction of dissonance. However, when comparing a nonparallel example to a parallel revision in a principle-testing option exercise, most students select the revision.

b. Identification and revision of nonparallelism is often at the superficial level, simply identifying and revising the nonparallel structure of the coordinated unit. However, deeper insight into information meaning may elicit another pattern (not coordination) to more effectively match structure and rhetorical intent of the information. An evolution from coordination to another syntactic pattern is not unexpected because coordination is one of the first stages of syntactic fluency development.

c. The abundant amount of detail in research writing is amenable to coordination during which nonparallelism occurs between core sentence constituents (e.g., nouns, verbs) and between non-core constituents (e.g., modifiers). For most instances of non-parallelism, the rhetorical consequence is dissonance. In *I love fishing, swimming, and to run*, there is no impeded comprehension because all are sporting activities.

Part 2 – Noun and nominal

Example: 'epidemics and the eating of dinosaur eggs by early mammals'

This example is from an Introduction section: hypothesis.

> *Epidemics and the* eating of dinosaur eggs by early mammals were *two possible causes for dinosaur extinction.*

Nonparallelism between a noun and nominal (a syntactic unit that functions as a noun) may be a more serious distraction than the nonparallelism between an adverb and adverbial, possibly because nouns are core constituents of sentences, functioning as subjects, direct objects, and subject complements. In the example, *epidemics* the noun is nonparallel to *eating of dinosaur eggs by early mammals*. There are two suggested revisions; in the first, the second coordinated constituent is replaced by *mammalian-oophagy* a structure parallel to that of epidemics.

> *Epidemics and* ***mammalian-oophagy*** *were two possible causes for dinosaur extinction.*

In the second suggested revision, the coordination is eliminated by transforming (e.g., de-coordinating) the noun *epidemics* into the object of the phrasal preposition *in addition to*. By such de-coordination, the necessity for parallel structure is eliminated, and the resulting pattern may better express the interrelation of information. However,

by such de-coordination, *epidemics* is deemphasised as if it is the more understood of the causes, so that *the eating of dinosaur eggs by early mammals* is the focus of the hypothesis.

> ***In addition to epidemics**, a second possible cause for dinosaur extinction was the eating of dinosaur eggs by early mammals.*

Part 3 – Linking verb and instransitive verb and transitive verb

Example 1: 'is' and 'decreases'

This example is from an Introduction section: research problem context.

> *Lysine* **is** *an abundant muscle constituent, and* **decreases** *in amount during starvation.*

The most common nonparallelism in research writing occurs between coordinated verbs that differ in type (linking, transitive, intransitive). Revision often involves transformation into another syntactic pattern. In the Example, by the apparent equivalency of structure, the descriptive clause is overemphasised and as a result, the assertive clause is deemphasised. In the Revision, embedding the descriptive information as an appositive to the subject deemphasises the descriptive information and, concomitantly, emphasises the assertion. The syntactic reduction of the independent clause into the appositive *an abundant muscle constituent* enables the de-coordination.

> *Lysine,* ***an abundant muscle constituent,*** *decreases in amount during starvation.*

Note

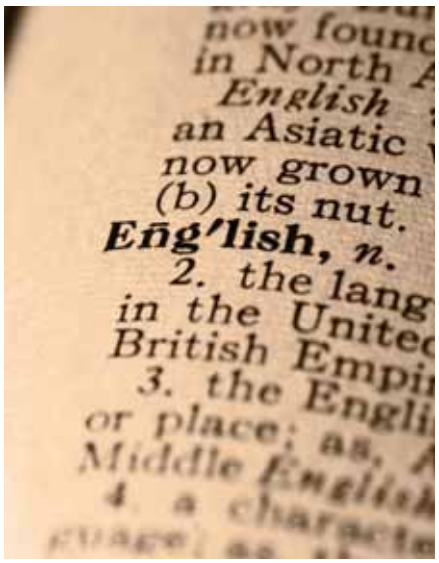
a. The coordination of nonparallel verbs often is mistakenly marked with a comma, but this ostensible apology for the nonparallelism and non-equivalency is itself distracting, because the comma disrupts the coordination.

Example 2: 'was' and 'had'

This example is from a Results section: data-based observation.

> *Spot no. 1* was *the most acidic and* had *the highest apparent molecular weight.*

Neither verb is emphasised in the Example because coordination renders them equivalent. However, the nonparallel verb type is a cue that the relation is non-equivalent. The linking verb *was* usually marks the less important descriptive information compared to the assertive information marked by



the transitive verb *had*. In the first suggested revision, the information marked by the linking verb is transformationally de-emphasised into a reduced adjectival phrase in apposition.

> *Spot no. 1,* ***the most acidic****, had the highest apparent molecular weight.*

In an alternative revision, the properties are listed because they are seemingly independent of each other and of equivalent importance.

> *The properties of spot no. 1 were as follows: most acidic; highest apparent molecular weight.*

Part 4 – Active and passive voice verbs

Example: 'are performed' and 'depend'

This example is from an Introduction section: research problem justification.

> *Traditional dose studies* are performed *at a pollutant concentration much higher than that observed in situ and* depend *on mortality as a final criterion.*

The nonparallelism resulting from coordination of the passive voice 'are performed' and the active voice verb 'depend' is a cue that an alternative syntactic structure (involving de-coordination) would be better matched to rhetorical intent. In the first revision, the verb phrase *are performed* is attenuated into the past participle *performed*.

> *Traditional dose studies,* ***performed at a pollutant concentration much higher than that observed in situ****, depend on mortality as a final criterion.*

In a second revision, *depend on mortality as a final criterion* is deemphasised as a present participial phrase *depending on mortality as a final criterion.*

> *Traditional dose studies,* ***depending on mortality as a final criterion,*** *are performed at a pollutant concentration much higher than that observed in situ.*

Selection of the first or second revision may depend on emphasis; that is, which information is intended to be emphasised. In both revisions, the phrase with the finite verb 'depend' (first revision) and 'are performed' (second revision) receives the emphasis, whereas the phrase functioning as a modifier is deemphasised.

Another determinant for selecting the first or second revision is the length of the modifying phrase initiated by 'performed' or 'depending'. Both intervene between the subject and verb of the sentence, the first by 12 words and the second by 7 words. Thus, based on the length of the disruptive phrase, the second revision may be the selection.

In a third revision, *depending on mortality as a final criterion* is placed before the subject *traditional dose studies* to avoid disruption between subject and verb.

> ***Depending on mortality as a final criterion****, traditional dose studies are performed at a pollutant concentration much higher than that observed in situ.*

Summary

Nonparallelism consists of two general types: 1. between equivalent or 2. between nonequivalent coordinated syntactic units. The nonparallelism between equivalent units is a dissonance, which can be revised by transforming the nonparallel into a parallel unit (See Part 1, Adverb and Adverbial; Part 2, Noun and Nominal).

The nonparallelism between nonequivalent units is impeded immediate comprehension, which can be revised syntactically by de-coordination. De-coordination involves transforming the less important unit into a lesser syntactic structure so that the other coordinated unit is emphasised. Thus, the nonparallelism is a cue that the units are not intended to be viewed as of equivalent importance (See Part 3: Nonparallel verb type and Part 4: Nonparallel verb voice).

Michael Lewis Schneir
Professor, Biomedical Sciences, Ostrow School of Dentistry, University of Southern California, Los Angeles, CA, USA
schneir@usc.edu

Medical Communications

SECTION EDITOR

Lisa Chamberlain James
lisa@trilogywriting.com

Editorial

Anyone working in pharmacovigilance (PV) will already have spent many months working their way through the ever-changing updates and reforms to the Risk Management Plan (RMP), and the newly legislated RMP summary. Those not working in PV will probably also have heard all about it (if only through the tortured wails of their PV colleagues!).

To everyone's delight and amazement, we survived the pilot phase; consultation comments have been received and a new and improved version is imminent. We all eagerly await the revision of the RMP summary in particular: will the original concerns be addressed? Will we still be asked to produce a single document that can satisfy both professional healthcare providers and the general public in one fell swoop? Will the RMP summary achieve its aim of increasing transparency for the lay audience??

We will find out in time, I'm sure. But in the meantime, I'm delighted to present to you a really excellent article from Tiziana von Bruchhausen and Stefanie Rechtsteiner. Tiziana and Stefanie chart beautifully the evolution of, and challenges posed, conquered, and still to be undertaken, by the RMP summary guidance.

This article really should be called 'Everything you wanted to know about the RMP summary but were afraid to ask' and I will be printing it out and pinning it to my wall!

Enjoy.

Bestest.

Lisa

RMP public summary reloaded: Revision 2 of GVP Module V

Transparency in PV: why the RMP public summary was introduced

The European Medicines Agency (EMA) describes on its website how the agency has been aiming at and working towards increasing the transparency of its processes and decisions ever since its formation in the 1990s. The European Public Assessment Report (EPAR) was one of the first tools used to provide information on a medicine and its use. Striving for transparency and openness, the EMA decided to go beyond what is legally required, in order to provide as much information as possible to all interested parties. However, marketing authorisation holders (MAHs), investigators, and other stakeholders need to have the assurance that their intellectual property, as well as their personal and commercially relevant information and data, are protected. Therefore, the EMA needs to carefully balance data protection against transparency.[1]

Besides providing as much information as possible to other health authorities, MAHs, investigators, and healthcare professionals (i.e. medical experts), amongst others, the EMA also strives to better inform the general public, and thus a lay audience. This initiative translated into the RMP public summary (Part VI.2), which was introduced with the new GVP legislation in 2012. The agency's goal is to involve patients more and to provide them with all relevant information available for a specific medicine, and this, in the case of the RMP public summary, in a language tailored to patients' needs.

A long journey: how the RMP public summary has evolved over time

The past: first introduction

In 2012, the EMA launched its 'EU Pharma Package' (Regulation (EU) No 1235/2010 and Directive 2010/84/EU) and the accompanying transparency initiative, with

the goal of enhancing public information on processes around a medicine's authorisation, its efficacy, and safety. With this, the RMP as a whole underwent a major overhaul, and, additionally, the new concept of the RMP public summary (RMP Part VI.2) came into existence. A 1-year pilot phase for the publishing of the RMP summary started in March 2014 for medicines authorised under the centralised procedure. For many medicines the RMP summary has since been made publicly available on the EMA website, and is intended for regulators, industry, and healthcare professionals, as well as for patients.

With the new GVP format, the RMP was now a comprehensive document with a broad spectrum of information provided, including epidemiology, non-clinical and clinical data, as well as post-authorisation data, based on which safety concerns, pharmacovigilance activities, and risk minimisation measures could be identified.[2, 3] The RMP Part VI with the public summary offered the most important information on a medicine's safety profile in a short and summarised form. This new approach, incorporating the publicly available RMP summary (with its inherent difficulty of ensuring transparency and data protection at the same time), immediately became a topic that was widely discussed amongst all stakeholders, and still remains the focus of interest.

The present: Revision 1

The RMP template was updated in July 2013.[4] A first revision of the GVP Module V was released in April 2014, addressing feedback that had been received from various stakeholders and providing more clarity on various aspects, such as definitions and terminology for safety concerns and triggers for RMP updates.[5] However, Revision 1 of both documents, which is currently valid and the basis of all RMP writing, was a minor one. It did not include results from the pilot phase on RMP summaries, which had just started at that time.

In general, Part VI of the RMP supports the overall goal of transparent, concise, and high-level communication of all relevant data and information. Part VI consists of two main parts:

- Part VI Section VI.1 'Elements for summary tables in the EPAR' provides tabular overviews of the medicine's safety concerns and of the related pharmacovigilance and risk minimisation measures. These tables are copied from the main body of the RMP and incorporated in the CHMP assessment report as well as in the EPAR public assessment report at the time of authorisation;
- Part VI Section VI.2 'Elements for a Public Summary' provides lay language summaries that are also partly incorporated in the EPAR summary for the public (summary on treatment benefits). Additionally, Section VI.2 is published as a stand-alone document (referred to in this article as RMP public summary). The summaries in Section VI.2 provide information on the disease epidemiology, the treatment benefits, the unknowns relating to treatment benefits, and the safety concerns. For medicines with additional risk minimisation measures proposed or in place, a further summary in lay language informs the public about these measures.

The format of the RMP public summary aims at providing condensed, clear, and understandable information on elements of the RMP. However, this task is very challenging for medical writers, as the RMP is a long, complex, and quite technical document. As previously described,[3,6,7] the major challenge posed is to tailor the complex information on the most relevant aspects of the RMP to a heterogeneous audience, encompassing healthcare professionals, industry stakeholders, and patients/patient organisations, while ensuring correctness, accuracy, and clarity. This task is even more challenging in view of the word count constraints imposed by the guidance for most of the lay language overviews in the RMP public summary.

From a regulator's perspective, communicating the important risks of a medicine and the associated risk minimisation measures to the public represents, in itself, a form of risk minimisation and may additionally be a valuable tool for healthcare professionals and patients to support decisions for or against use of a medicine. For this reason, it is crucial that the target audience of the RMP public summaries is able to understand the complex benefit-risk information presented, which means taking into consideration the health literacy of the readership. For the RMP public summaries in their current format, medical writers normally aim at a literacy level of 11-12 years old or below.[3]

From guidance to real life The package leaflet (PL) and the EPAR summary present key information in lay language on the benefits and the risks of a medicine. The RMP public summaries intend to provide a context for the risk-benefit evaluation of a medicine and to complement the EPAR and PL by providing information on the safety concerns of a medicine and the related post-authorisation studies. The introduction of the RMP public summaries was generally perceived as a positive measure to improve transparent communication and to contribute to a more patient-centred drug development process. However, there are inherent limitations due to format, requirements, and lay language, which, in combination with the complex contents, lead to the following two questions:

- how can the requirements and the format be adjusted to fulfil the needs of the targeted readership?
- is the lay public really the appropriate audience?

To explore the above questions, the EMA collected feedback from patients, healthcare professionals, and industry associations during the 1-year pilot phase on the RMP public summaries.

Industry feedback The industry welcomed the transparency initiative. However, the general perception of the industry was that, if the RMP public summary is mainly intended for patients, it should be improved and further adapted to meet the needs of this target audience. In particular, the suitability of the RMP public summary in its current format was critically questioned:[8]

- definitions for identified risks, potential risks, and missing information are not provided;
- there is no explanation on how the RMP public summary complements the SmPC, PL, and EPAR and what the differences are (e.g., important risks vs. side effects) between the concepts addressed in these documents;
- there are no explanations of pharmacovigilance and risk minimisation processes (post-authorisation plans, risk minimisation measures), with which the audience is not familiar.

In this context, RMP summaries containing numerous important risks and gaps in knowledge may lead to unjustified concerns and to the misleading perception that the product is more hazardous than it actually

is, and that the risks outweigh the benefits.

Patients' feedback The patients' feedback further challenged the RMP public summary format and language: in general, lack of clarity, context, and definitions were criticised, and the RMP public summary was considered to be hard for patients to understand due to its complex and technical contents. Moreover, among the public, knowledge and understanding of drug development, medicine safety monitoring, and health authority activities are generally limited. Therefore, it appears that the RMP public summaries are not perceived as a useful communication tool and do not effectively reach their target audience.[9]

In conclusion, despite the intention behind the lay language requirements, it is doubtful that the RMP public summary is indeed widely used. In addition, it does not provide the basic definitions to ensure understanding of the contents and of its relationship to the other publicly available documents. Therefore, it is questionable whether the RMP public summary fulfils criteria for effective, transparent communication.

The future: Revision 2

The objectives of the pilot phase on RMP public summaries were to confirm interest and usefulness for stakeholders, to confirm the target audience, to improve format and content based on the needs and expectations of the readership, and to streamline the process for preparation and update of the RMP public summaries.[8]

The pilot, which covered over 80 RMP public summaries, confirmed a wide interest from different audience groups and the need to improve format and contents to meet their demand and expectations. The main targets for the revision of the RMP public summary with regard to transparent communication are as follows:[8, 10]

- format, contents, and structure should be simplified with focus on the summary of safety concerns, risk minimisation measures, and planned post-authorisation development plan;
- while the PL and EPAR summary are the main primary source of information on benefits and risks of a medicine for patients, the RMP public summary should address an audience interested in additional background safety information provided in the PL;
- a plain language approach should be used; however, technical terms should not be avoided.

With Revision 2 of GVP module V, the RMP summary is now moving towards a rather professional audience and people seeking additional information, possibly with a slightly higher health literacy level. However, the RMP summary should still follow plain-language principles to facilitate readability by the general public:[11]

'The audience of RMP summaries is very broad. To ensure that the summary can satisfy the different needs, it should be written and presented clearly, using a plain-language approach. However, this does not mean that technical terms should be avoided. The document should clearly explain its purpose and how it relates to other information, in particular the product information (i.e. the SmPC, the PL and the labelling). It should contain the following information:

- the medicine and what it is used for;
- summary of safety concerns and missing information;
- routine and additional risk minimisation measures;
- additional pharmacovigilance activities.'[12]

The Revision 2 of GVP module V[12] and the RMP template[13] is a major one. The public consultation phase of this revision ended in May 2016; the publication of the final revision is expected in the third quarter of 2016. Although the revised RMP public summary considered many of the stakeholders' comments, it still does not seem to fully meet the needs of the diverse target audience. The contents of the revised RMP public summary are now very concise and limited to safety concerns, pharmacovigilance activities, and risk minimisation measures. The EPAR tables have been removed, as have the overviews on disease epidemiology and treatment benefits. Standard text has been proposed to define identified and potential risks, but not missing information. In addition, the definition of the 'importance' of a risk is still missing. Context is given with regard to the EPAR, the SmPC, and the PL; however, there is still no explanation about the difference between side effects/adverse events (terminology used in these documents) and important risks (terminology used in the RMP public summary). A definition has been provided for routine and additional risk minimisation measures as well as for routine pharmacovigilance activities, yet there is no explanation for additional pharmacovigilance activities

(post-authorisation studies).

In line with the objectives of template and process simplification, the proposed format of the RMP public summary maps the contents to the full RMP. This, however, gives the impression that its content is mainly taken verbatim from the body of the RMP. In this sense, the question then arises as to who should prepare the RMP public summaries in the new format: should it be the pharmaceutical company, or could it be the health authority when preparing the reader-friendly summaries?

The journey continues: open questions

Although Revision 2 of GVP Module V addresses many questions and concerns that were raised over the last two years, the following questions remain:

- the public summary is only available in English, which not everyone in the EU is able to understand. In addition, most people are likely not aware that an RMP summary, an RMP, or the EMA website exist, and therefore they do not have access to this information. Can the lay audience thus be reached at all with the RMP public summary?
- even if plain language is used, assuming the patients speak English, will they be able to understand the information provided and to consequently make appropriate decisions?
- does transparency require showing all details of the risk management process to an audience with low health literacy and no understanding of such processes?
- should the focus be shifted even more to patients' needs and readability, i.e. would user testing help to better meet patients' needs and to create a more reader-friendly document? Or should separate summaries be created for lay readers and expert readers?

Conclusions

In line with the transparency initiative and the efforts of the EMA to improve communication of clinical and safety information, the RMP public summary, four years after its first introduction, is currently undergoing a major revision based on feedback from all stakeholders. As a document that must address different needs and interests, and cover complex medical information, the RMP public summary has a major impact on how a medicine is perceived. Further interaction and exchange between all parties involved will likely be needed to reach the overall common goals: effective communication, transparency, and patients' safety.

Acknowledgements

We would like to thank Carol Sable, Sven Schirp, and John 'Chip' Sherrill for discussions and constructive feedback.

References

1. Transparency. European Medicines Agency [cited 2016 July 04]. Available from: http://www.ema.europa.eu/ema/ index.jsp?curl=pages/special_topics/ general/general_content_000526.jsp&mid=WC0b01ac0580789730
2. Von Bruchhausen T and Prechtel K. Pharmacovigilance medical writing: an evolving profession. Medical Writing 2015;24 (2):66-71.
3. Chamberlain-James L. Risk management plans – New challenges for a new era. Regulatory Rapporteur 2014;11(6):4-7.
4. Guidance on format of the risk management plan (RMP) in the EU – in integrated format. EMA/PRAC/613102/2015 Rev.2 accompanying GVP Module V Rev.1. European Medicines Agency; 2013 Jul 25 [cited 2016 Jul 04]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2012/11/WC500134650.pdf.
5. Guideline on good pharmacovigilance practices (GVP). Module V – Risk management systems (Rev 1). European Medicines Agency and Heads of Medicines Agencies; 2014 Apr 15 [cited 2016 Jul 04]. Available from: http://www.ema.europa.eu/docs/ en_GB/document_library/Scientific_guideline/2012/06/WC500129134. pdf.
6. Chamberlain-James L. Legislation and the lay audience: Challenges of communicating benefit and risk in the light of new regulations. Medical Writing 2015;24(4):195-9.
7. Prechtel K and Rechtsteiner S. Medical writing for two audiences – The RMP public summary. Medical Writing 2015;24(4):200-4.
8. Juan Garcia Burgos, EMA. Publication of Risk Management Plan (RMP) summaries: Analysis of the experience of the 1-year pilot phase. Lecture presented at PCWP meeting, 2015.
9. Raynor T. Does transparency in medicines information deliver benefits to patients? Lecture presented at DIA Euromeeting, Hamburg, 2016.
10. Caroline Voltz-Girolt, EMA. RMP template. 4th stakeholders platform – operation of EU pharmacovigilance legislation; RMP activities update session. Lecture presented at the 4th stakeholders platform, 2015.
11. Pharmacovigilance Risk Assessment Committee (PRAC). Minutes of the PRAC meeting on 11-14 January 2016. EMA/PRAC/92676/2016 Procedure Management and Committees Support Division [cited 2016 Jul 11] Available from: http://www.ema.europa.eu/docs/ en_ GB/document_library/Minutes/2016/03/WC500202623.pdf.
12. Guideline on good pharmacovigilance practices (GVP). Module V – Risk management systems (Rev 2) – draft for public consultation. European Medicines Agency and Heads of Medicines Agencies; 2016 Feb 24 [cited 2016 Jul 04]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2016/02/WC500202424.pdf.
13. Draft guidance on format of the risk management plan (RMP) in the EU – in integrated format. EMA/PRAC/613102/2015 Rev.2 accompanying GVP Module V Rev.2. European Medicines Agency; 2016 Feb 29 [cited 2016 Jul 04]. Available from: www.ema.europa.eu/.../en_GB/document_library/Regulatory_and_procedural_ guideline/2016/02/WC500202373.doc.

Tiziana von Bruchhausen and Stefanie Rechtsteiner
Boehringer Ingelheim Pharma Global Pharmacovigilance, Ingelheim, Germany
stefanie.rechtsteiner@boehringer-ingelheim.com

Getting Your Foot in the Door

SECTION EDITOR

Raquel Billiones
rbilliones@clinipace.com

Editorial

Welcome to Getting Your Foot in the Door or GYFD for short, the latest addition to MEW's regular sections.

It all started at the EMWA 2015 autumn meeting in The Hague. Derek Ho met up with the EMWA Executive Committee (EC) to talk about his idea of organising an activity aimed to connect potential interns with companies offering medical writing internships. At the same time, Danae Rokanas also contacted Phil Leventhal about her interest in a similar endeavour. The EC pledged support to this initiative and Derek and Danae, together with EMWA PR Officer Beatrix Doerr and Harald Meier got the ball rolling. Six months later, the first Live Internship Forum (IF) was held at the EMWA 2016 Spring Conference in Munich. Seven companies participated and a total of 50 intern applicants were present at the IF live event. Hats off to the IF team for pulling off this amazing feat in such a short period of time.

The event was a resounding success with very positive feedback from all parties involved. The IF was complemented by Helen Baldwin's regular not-for-credit short seminar '**Introduction to Medical Writing**' and another 'first' – Phil Leventhal's short seminar entitled **'Getting Your Foot in the Door: How to Build Experience to Get a First Medical Writing Job'.** The latter provided tips on how to get into the field without previous industry experience.

Job opportunities for new writers are scarce as Elsa Lewis laments in her article **'Where have all the UK entry level pharmaceutical regulatory medical writing jobs gone?'** on p. 50. The IF and the other EMWA offerings are witness to our organisation's commitment to help grow the medical writing field by engaging new graduates and bringing together entry level candidates and the industry. And EMWA will continue to do so through different platforms.

One such platform is this section in the MEW dedicated to this endeavour starting in September 2016, aptly adopting Phil's seminar title to be this section's name. Phil will continue to present GYFD in upcoming EMWA conferences and to complement this MedComms-focused presentation, another short seminar for newbies is planned, entitled from '**From academia to regulatory medical writing**'. However, most important of all, the IF will again be held at the 2017 Spring meeting in Birmingham, this time with a half day time slot dedicated to the forum.

In this maiden edition of GYFD, we would like to thank the Costello Medical team, Debbie Nixon and Sophie Pearson, for their insightful piece on the practical aspects of running a medical writing internship programme. In the upcoming issues, we will be covering internship from different perspectives.

Raquel

Successful internships: a company perspective

Medical communications is a rapidly growing industry and competition for experienced recruits is often fierce. Paradoxically, however, 'getting your foot in the door' can be very difficult for those without formal medical writing experience. Thus, for applicants wishing to embark on a career in medical communications, prior experience can strengthen a candidate's CV and provide them with an invaluable insight into the day-to-day life of a medical writer.

In lieu of previous employment in the medical writing field, a well-structured internship should also enable the intern to sharpen their writing and editorial skills, whilst learning more about the healthcare industry in general, ultimately increasing their chances of finding a suitable job in medical communications. There are also notable benefits for the host company, since the additional resource provided by interns can make important contributions to project work, easing the workloads of permanent staff, and provide opportunities to staff members looking for line management experience. Moreover, a successful period of on-the-job training and development of a core set of medical writing skills may result in the offer of a permanent position with the host company or elsewhere.

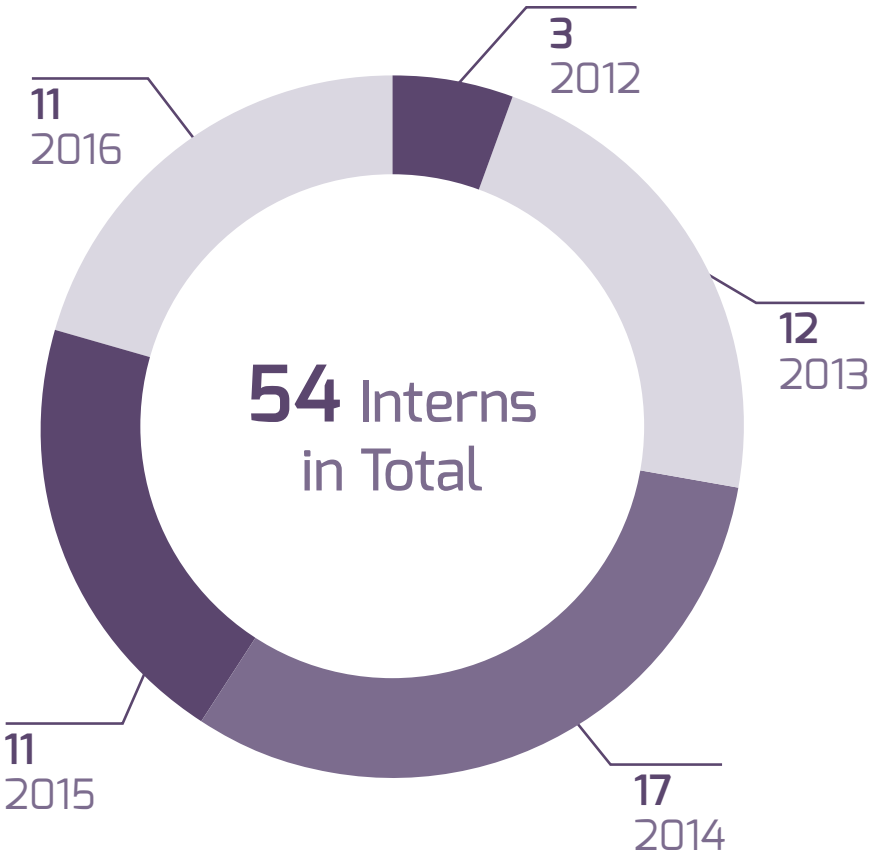
Establishing a successful internship programme

Since its launch in 2012, the **Scientific Internship Programme** at Costello Medical has offered internship positions for graduates and postgraduates from a range of disciplines to gain experience in the medical communications industry. Our primary goal had been to enable those interested in healthcare and with a passion for writing to make a more informed decision about whether the industry could provide a fulfilling and rewarding career for them. Now in its fifth consecutive year, with a total of 54 interns hosted to date, our internship programme has undoubtedly fulfilled its original ambition of giving individuals a hands-on experience of life in a busy medical communications agency. The programme has also supported the company's long-term growth strategy, with many of the interns transitioning into permanent positions at the end of their internship (see Figure 1 overleaf).

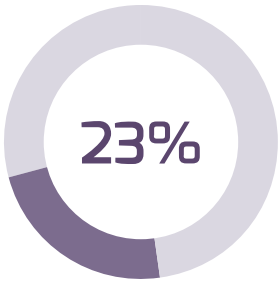
The duration of internships may vary greatly from company to company. At Costello Medical, we usually offer a three-month internship period, which allows sufficient time for an intensive two-week induction and training period, followed by approximately 10 weeks of project work where interns can contribute to a range of ongoing client deliverables. Subject to an internal review, interns may choose to extend their contracts for up to six months.

Costello Medical's Internship Programme

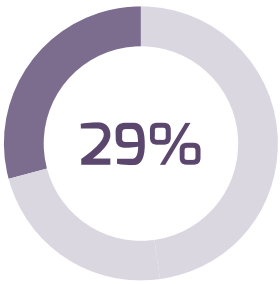
Number of Interns 2012-16



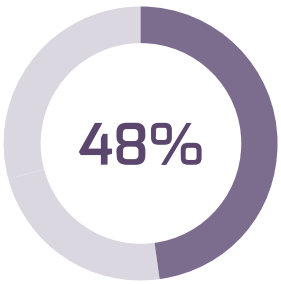
Alumni Destinations



Permanent Employee at Costello Medical



Further Study



Other Employment

Company Considerations for an Internship Programme

Establish clear and individualised objectives to maximise the internship experience

Carefully consider the content and distribution of your advertisement to target the right candidate

Assign a manager to provide guidance and support throughout the internship

Establish a thorough induction process and provide further project-specific training

Plan a variety of projects for the intern and flexibility for additional project support

Listen to intern feedback in order to maximise benefits of the programme for future interns

This provides a greater opportunity for the intern to take on more project responsibility and potentially experience work across other divisions of Costello Medical. Participants are typically based in one division of the company with their assigned line manager but are able to gain experience across a variety of project types, including: medical writing, evidence development, medical affairs and market access.

The internship programme at Costello Medical is constantly evolving as we implement feedback from every new round of interns. We have also defined a set of core programme objectives to make the experience as rewarding as possible for the intern (see box below).

Key objectives for a successful internship

- To experience a variety of projects reflective of a permanent position
- To see at least one project through from start to finish
- To gain an insight into project management
- To gain experience of communication with clients

Beyond these key objectives, individual intern programmes are designed to be managed flexibly and relatively informally; this enables the intern and their manager to match project work with the intern's own interests as new opportunities arise, and to facilitate the rapid delivery of client work when it is needed.

Practicalities for intern recruitment

Developing a compelling advertisement that communicates the benefits of the internship programme is a crucial first step in the recruitment process and is worth investing time in. The advertisement should include a clear outline of the tasks and responsibilities of the internship position, and list the type of hands-on experience that successful applicants will have the opportunity to gain. If there is scope for the internship to lead to a permanent role with the host company, this should also be mentioned. Ideally, candidates should be able to review the key criteria for a successful internship and quickly determine their suitability. Practical details relating to the internship should also be included, such as internship duration, possible start dates, hours of work, and salary.

Of course, a well-crafted internship advertisement will only yield results if it reaches the desired audience, so deciding where to post the role requires careful consideration. Most universities have a dedicated careers service, and many will advertise vacancies free of charge. For a broader reach, and to advertise to those already working in industry/academia or the medical profession, advertising on high-visibility recruitment sites can be a fruitful avenue to consider. Participation at organised recruitment events such as the EMWA Internship Forum, MedComms Networking events, university careers events, or in-house company open days can also provide good exposure for internship opportunities within your company, alongside opportunities for other permanent positions.

There is likely to be some overlap between established recruitment processes at the host company and those taken to recruit interns. However, given the short-term nature of the internship, these processes are unlikely to necessitate multiple assessment rounds or interview stages. Consider the most efficient and effective way of assessing how far the candidate meets the requirements for the internship – for example, completing a short writing test or attending an assessment day, as well as a formal interview. When interviewing intern candidates, take account of the fact that they are likely to have less industry experience than those applying for full-time roles, so their responses to questions about client management and commercial awareness in general may be less developed.

As we see greater globalisation of the pharmaceutical industry and the world of work overall, an increasing number of intern applications are likely to be received from overseas candidates. Therefore, on a practical level, it is important to verify the candidate's right to work in a particular country prior to making an offer, and also to be transparent during the recruitment process about the assistance available for obtaining visas or work permits.

Management and planning for internships

Establishing an effective system for intern management and resource planning can support long-term planning on a company-wide level, and also helps to ensure that the internship is mutually beneficial to both the intern and the host company. Below we summarise the approach taken at Costello Medical, although the most appropriate methods for management and planning may vary depending upon company size and structure.

Management and mentorship – Each intern is assigned a senior staff member (typically a Senior Medical Writer or Senior Analyst) to act as both line manager and mentor. The manager meets with the intern on their first day in the office and schedules weekly meetings to discuss ongoing projects, and address any questions or concerns the intern may have. Periodically, the intern and their manager also meet with an internship programme co-director to discuss how their individual programme is going and to identify any new objectives that may help the intern reach their full potential.

Intern training – Upon arrival at the company, each intern completes a two-week induction period. This training period provides an opportunity for the intern to gain an understanding of the company ethos and approach to work, as well as the breadth of projects that the company delivers. This also facilitates interaction with other company employees, across all divisions, and enables the intern to spend time with other new recruits. The intern will also receive more detailed guidance and training for any specific projects they are assigned to.

Planning intern resource – Prior to arrival, the intern will be assigned a variety of projects, with the aim to develop a varied and enjoyable workload throughout the internship. However, individual intern programmes are designed to be fairly flexible and planning typically allows for interns to provide support for additional projects. The ultimate goal is to design a tailored programme for the intern that takes into account the intern's own interests and skills, whilst factoring in the demands of current project work.

Project management – The intern is likely to be working within multiple project teams, led by different project managers. To ensure that the intern is not overwhelmed, a suitable amount of time is allocated on a weekly basis to work on each project. The intern manager plays a vital role in ensuring that all project managers are updated if there is a change in priority for any specific project.

The importance of intern feedback

Communication plays a vital role in generating a successful internship and project managers are encouraged to provide regular feedback on intern performance and attitude to work. Not only does this provide an opportunity to recognise high quality work and important contributions to a team project, it also helps to establish personal goals for the intern to work towards in order to maximise the success of their internship and to make the experience as rewarding as possible.

A formal review meeting around four to six weeks into the internship programme provides an opportunity for the intern manager to provide feedback on what the intern is doing well and where there is room for improvement. Another review meeting held at the end of the internship is incredibly valuable as this enables the intern, their manager and the programme co-director to discuss what went well and any aspects of the internship programme that could be improved. Suggestions for improvements to the programme are actively encouraged as part of the wrap-up process and are implemented where possible, playing a vital part in maximising the value of the programme for future interns. The intern is asked to prepare a short report in advance of the wrap-up meeting, summarising their internship experience. This report forms a great starting point for discussions and can also prove to be a useful document for the intern to refer to in any future job applications.

Post-internship support

As the internship comes to completion, the host company may offer the intern advice on alternative opportunities within medical communications or the wider pharmaceutical industry. The HR team and intern managers can also provide employment references and application advice.

For those who have enjoyed their internship experience, applying for a permanent role within the host company may be an attractive option. Promising interns who have demonstrated skills and commitment throughout their internship are offered support in applying for a permanent position with the company. If a successful intern does decide to apply for a permanent role, the host company benefits from a new staff member who has already undergone the induction and training process and who is familiar with the company's culture and ways of working, allowing them to make an immediate and positive impact upon their transition to permanent role.

Acknowledgements

We would like to thank Danielle Hart, Graphic Designer (and former Graphic Design intern), Costello Medical Consulting Ltd., for assistance with figures, and Simon Page, Head of Publications and EMWA Internship Programme Advisor, Costello Medical Consulting Ltd., for critical review and support.

Debbie Nixon
Publication Manager and Scientific Internship Programme Co-director, Costello Medical Consulting Ltd., Cambridge, UK
debbie.nixon@costellomedical.com

Sophie Pearson
Senior HR Coordinator, Costello Medical Consulting Ltd., Cambridge, UK
sophie.pearson@costellomedical.com

Lingua Franca and Beyond

Some thoughts on statistics and a bit more…

SECTION EDITOR

Maria Koltowska-Häggström
maria.koltowska-haggstrom@propermedicalwriting.com

In this issue of *Medical Writing*, I would like to recommend two articles. Szymon Musiol shares with us his thoughts on a statistician's role in research overall and also in the process of preparing publications. He takes a medical writer's perspective and convinces us that the risks for creating a mutual relationship between the author, the medical writer, and the statistician can be minimalised. Personally, I fully agree with Szymon, but on the other hand I understand that often communication between authors and statisticians can be very challenging. The former speak their clinical language and the latter – statistical *Hocus Pocus;* no way do they get into mutual communication. Here comes in or at least should come in the medical writer, just to translate so that the statistician understands the clinical question and the author comprehends the statistical answer. Not an easy task… The second article is from dear *Hotspur* – welcome back with your funny stories and observations.

In the March 2016 issue, we read about the unpredictable aspects of choosing a collaborator. Now, we can look further into refereeing mechanisms… Please, bear in mind the final conclusion! Very true!

Have a nice reading ☺

Maria

Statistics – are we doing it right?

Statistics has probably been a bone of contention of medical academia since the first time someone decided to calculate a p-value in support of their findings. Beloved by few, and dreaded by many, it remains a necessary evil for those wishing to engage in scientific endeavours of high quality. The mutual relationship between the author, the medical writer, and the statistician (not to mention the sponsors) is often underpinned by a power struggle, with each of the parties striving to vindicate their own agenda. I think resentment towards statistics by the non-statisticians has an important part to play here.

In my personal experience, the statistician is frequently involved too late in the process of research. The first person to mention it to me was a professor of statistics who tried to hammer this point home at every opportunity. At the time I thought he was merely bitter at missing out on dinners, but now I understand his advice was invaluable. Time and again, the statistician is approached with an often meticulously collected set of raw data and is curtly asked to 'run some stats on it'. Often the feedback states something along the lines of the study being underpowered by an order of magnitude to show the expected potential effect. Disappointment is followed by *plan B*, as the authors now ask the statistician to employ some mathematically dubious sorcery to shrink that p-value. The medical writers find themselves in an equally uncomfortable position having to word far-fetched conclusions from exotic maths they barely understand. One objectionable trick is to generate hypotheses *a posteriori*, based on the sample available.

By chance alone, if enough putative correlations are tested, some of them will be statistically significant. That's all we want, isn't it? A positive result with a nice p-value. Except, such data are not reproducible. If someone applies our devised model to their sample of the population, most likely there will be no correlation. And now no one wants to cite our paper.

That being said, *a posteriori* generation of hypotheses isn't always bad. I recently came across a report on Google being granted access to approximately 1.6 million patient records in the UK for the purpose of identifying predictors of *acute kidney injury*. Putting aside the Orwellian connotations this might prompt in some, it could lead to significant scientific advances. Why should it work then? Simply because of the enormous sample size. Any correlation identified, even based on a post hoc hypothesis, is very likely to hold true for the entire population. Most of us sadly can't dream of this level of statistical power, and to keep publishing high quality material we must adjust our methods appropriately.

All this stems from pressure to publish positive results only, something anyone ever applying for funding will be acutely aware of. The tabloid-like obsession with headines has led to significant publication bias in the literature. Fortunately, awareness of the problems it poses is now beginning to trickle into the minds of academics, reviewers, editors, and sponsors alike. With many meta-analyses laying bare this skew, The International Committee of Medical Journal Editors now advises periodicals to require pre-registration of clinical trials as a necessary condition for publication. This means hypotheses have to be devised *a priori* and negative results are also likely to be published. This spares the statistician having to dredge the data, and the medical writer the job of describing it. With a few exceptions, we may all remain in the safe sandbox of undergraduate level statistics with good old t-tests, linear regression, and correlation coefficients dominating the scene. Perhaps every now and then Mann and Whitney will pay us a visit. But all in all

it shouldn't be too bad, provided we abide by those ground rules.

My final advice is to always design the hypotheses to be tested in advance. Invite the statistician for coffee well before collecting the first batch of data. Let them explain what sort of sample size is needed to have sufficient power for the result you desire, and what statistical methods suit the purpose. Register your research whether original or secondary. And finally, unless you're Google, don't mine your data.

Szymon K. Musiol
szymon.musiol@propermedicalwriting.com

Refereeing: a humour-free occupation

Sometimes you find yourself in a mess through no fault of your own; on other occasions, the problem is self-inflicted because of character traits that have troubled you in the past and no doubt will do so again in the future. One such character trait that I possess is a tendency to be flippant and a little silly when feeling completely relaxed.

Well, I was very relaxed on a lovely summer evening a year or two ago in North America, the day's work was over, and I was in happy hour mode; I sensed upregulation of my flippancy gene. The purpose of my visit was to attend the annual meeting of the American Endocrine Society (AES); as far as I could tell, my presentation had been well received, and I was now enjoying myself at a dinner held for a small constituent society of the AES that always arranges a social function to coincide with the annual meeting.

I was searching for my second pre-dinner cocktail when I spotted Mike, an American colleague (the name has been changed to protect the guilty); I had known him for years. He was lively, intelligent, a man of strongly-held beliefs, and equipped with a great sense of humour. We had sparred many times before in a jocular good-hearted fashion. Furthermore, I was certain that he had just refereed one of our articles for Clinical Endocrinology; the style, the manner of expression, and the content of the referee's argument all pointed to Mike. The article was a tricky one and I have to admit that I was amazed that my research fellow had ever been able to conjure a manuscript out of such raw data; nonetheless, he had done a very professional job and no other data exist in the literature on this topic. So we were in the position of the one-eyed man in the land of the blind, and I anticipated that the journal would accept the manuscript.

Referee one, whom I guessed was Mike, started down the first page of his report by not liking the article and by the end of the page he hated it; boy, did he hate it! Vitriol was pouring from his pen as he became more and more agitated by our efforts. The journal behaved impeccably and allowed us to rebut the criticisms and comments, eventually accepting and publishing a revised version of the manuscript.

I had to pass in Mike's direction to get to the bar so I paused to greet him;

'How are you?'

'Fine', he said

'Tell me why did you hate our work with such intensity?'

'What are you talking about?' said Mike (my pulse rate quickened)

'You know, our recent manuscript for Clinical Endocrinology on pituitary disease.'

'I don't know what you are talking about, I have never refereed any manuscript of yours'.… (beads of sweat appeared on my forehead)

… 'but you refereed an article of mine recently that I sent to the Journal of Clinical Endocrinology and Metabolism' said Mike abrasively (my legs felt very heavy)

'Did I?'

Scrambling thoughts together; is he right? Oh god, I do seem to remember an article, the article presumably (brain no longer driving my side of the conversation).

'What happened to that article?' I queried

'Oh, it was rejected on your recommendation' he replied (now I was in desperate need of that drink)

'How do you know it was me that refereed your article?' (my legs were no longer capable of movement in any direction).

'The journal sent me the referee's comments with your name on the fax.' he stated.

Well, for the remainder of the AES meeting, as luck would have it, I ran into Mike every day, and without fail he reminded me that I had rejected his article. Even if on escalators moving in opposite directions, and too far apart for dialogue, he would simply look across at me with a baleful eye and then, Roman-style, give me a thumbs-down sign.

From that whole experience, I advise all readers to avoid all attempts at humour where the refereeing of manuscripts is concerned.

I was upset that my ability to referee-spot was not as good as I thought. In the past, I used to recognise the typeset, e.g. dropped 's' on typewriter in endocrine department of a famous London teaching hospital, an unhelpful talent in the computer era, but I had always told myself that I could identify the referee by the style and language used in the review. Thus, my illusion was shattered and self-esteem reduced.

Some 2 to 3 months had elapsed when out of the blue I received a fax from Mike; 'I have been reviewing my refereeing records and find that I have refereed two articles of yours, including the one on pituitary disease.'

Referee-spotting self-esteem restored instantly; I felt happier than if I had had a manuscript of my own accepted. More seriously, I was deeply impressed by his actions; he must have lain awake for a few nights tormented that he had been economical with the truth and driven by his conscience responded in the manner that he had.

From that whole experience, I advise all readers to avoid all attempts at humour where the refereeing of manuscripts is concerned. When an article is *criticised*, pain of varying intensity and duration is felt by the author however senior or junior.

I still relax after a day's work and remain a little silly but I am fonder of Mike, and I now understand that refereeing is a humour-free occupation.

Stephen M. Shalet
stephen.m.shalet@manchester.ac.uk

Gained in Translation

SECTION EDITOR

Laura C Collada Ali
laura.collada@teksema.com

Editorial

Welcome to the Translation Section editorial!

The Italian word *Itangliano* means Italian that is very much influenced by the English language and most of all it refers to the great presence of English words that are not adapted into Italian.[1] Unfortunately, the word is not new and it was first used back in the '70s when Italian business language started to be heavily influenced by English words that were not translated anymore, but rather used in English in an Italian discourse.

Biostatistics is no exception and I have asked an expert in the field to propose a short glossary of terms that are frequently used within the clinical research field. Francesca Paoloni, a biostatistician designing and analysing clinical trials on a daily basis, brings her expertise in the field as a technical language end user.

We do hope this article serves as a starting point for many other glossaries that could be added with time, as these are extremely useful tools for medical translators. Any damage to the *receiving* language represents important food for thought for translators…

Enjoy the article!

Laura C Collada Ali

English to Italian translation: 17 biostatistical terms that we are using in Itanglish that could just be used in proper Italian!

Background

Italians have integrated English words into their common vocabulary to such an extent that the word *Itanglish* – *Itangliano* in Italian – is well known among the community of linguists. Indeed, Italian is actually experiencing serious difficulties; the absence of norms that force the standardisation of new terms and the belief that a language is self-regulated is granting media and advertising agencies to unduly influence the language's development.

As you may be imagining already, science is not an exception and statistics is a field where *Itanglish* is due to win the first prize in this competition for the language of Dante.

So, is there anything that can be done to remedy the situation? One could say that what's done is done. It would be largely pointless to push out the many foreign expressions that have already entered the vernacular. Nonetheless, what about the many foreign words in use at present?

In our opinion, an effort is needed to translate widely-used technical expressions from English into Italian. Not as an attempt to counter globalisation – as many may think – but as a means of preserving our language, which is the most important vehicle we have for our local culture, identity and history.

The responsibility of translators towards language evolution

The point is not to force some manner of pure Italian, which is entirely free of *foreign barbarisms*. That certain interchanges take place between languages is an entirely natural consequence of globalisation. Yet it also seems absurd to condone the *refined ignorance* that prompts many Italians to pad their conversations with words from a language they barely even speak. Without pretending to be able to change the past, it seems that it would be enough for people to realise that the globalisation of language does not necessarily have to mean complete and unavoidable anglicisation.[2]

That said, translators are obviously responsible for preserving correctness of terms used and proper translation of new concepts into their own language. Translation is one of the most powerful means of communication between different cultures, yet attention needs to be paid if we want to preserve those different identities and cultures. Experienced translators are fully aware of this menace, but new professionals need to be particularly careful!

In February 2015, The Accademia della Crusca – the most authoritative body representing the Italian language – signed a petition to save their language (the so called #dilloinitaliano intervention in change.org). In fact, according to a recent study, the usage

of English terms has increased by 773% in the last 8 years. Quite worrying indeed…[3,4]

A proposal for proper translation of biostatistical terms into Italian

We first analysed a list of almost 200 frequently used terms that may present problems when translated into Italian and, from those, we have extracted the most frequently recurring ones that also coincide in frequently being used in English to Italian texts. Thus, we present a short list of English terms (in bold), the related definition and an Italian translation proposal (in green). We also remind readers of the most familiar translation, which is the one we are trying to avoid here (crossed out). The objective of this list is two-fold:

a. to provide a proposal for translation
b. to raise some polemics and hear our readers' opinions!

1. **attrition bias** > errore di esclusione (~~attrition bias~~)
   A systematic error caused by attrition (loss of participants), such as deviations and losses to follow-up, which may lead to a result that differs from true values.

2. **bias** > errore sistematico (~~bias sistematico~~)
   A systematic deviation from the real value.

3. **censoring** > censura/censorizzazione (~~censorizzazione~~)
   A subject leaving the observation before a certain event takesing place.

4. **cross-over design** > disegno incrociato (~~disegno cross-over~~)
   A study in which subjects receive a sequence of different treatments (also called *exposures*) at different time points. The design aims to evaluate these exposures.

5. **dropout** > abbandono (~~dropout~~)
   A subject included in a study fails, for any reason, to continue in the trial until the end of observation and, thus, drops out.

6. **dummy variable** > variabile binaria (~~variabile dummy~~)
   A variable that assumes values equal to 0 or 1 for absence or presence of a given condition.

7. **hazard ratio** > rapporto tra rischi (~~hazard ratio~~)
   The ratio of hazard rates at a single time, for two different kinds of subjects.

8. **intention to treat analysis** > analisi secondo l'intenzione al trattamento (~~analisi intention to treat~~)
   An analysis based on the initial treatment assignment, regardless of the treatment that patients actually receive.

9. **interquartile range** > ampiezza inter-quartile (~~range interquartile~~)
   The range between the 25th and the 75th percentiles.

10. **odds ratio** > rapporto delle proporzioni (~~odds ratio~~)
    The ratio of the odds (the proportion between the number of times that an event occurs and the number that it does not occur) for group A and group B.

11. **outlier** > valore anomalo (~~outlier~~)
    A value that lies outside of the expected range of the other values in a dataset.

12. **per protocol analysis** > analisi secondo protocollo (~~analisi per protocol~~)
    Analysis restricted to the population treated with the assigned treatment.

13. **range** > ampiezza (~~range~~)
    The interval between the lowest and highest values.

14. **recall bias** > errore da rievocazione o errore di memoria (~~recall bias~~)
    A systematic error caused by the difference between the cases and controls recalling information on exposure to a past factor.

15. **risk ratio/relative risk** > rischio relativo (~~risk ratio~~)
    The ratio of the probability of an event occurring in two different groups (i.e. exposed vs. non exposed).

16. **scatter plot** > grafico a dispersione (~~catterplot~~)
    A graph used to plot the data points for two different variables in order to show their relationship.

17. **statistical analysis plan** > piano dell' analisi statistica (~~statistical analysis plan~~)
    A document containing a detailed description of the planned analyses for a given clinical trial.

Conclusion

Obviously, not everyone agrees about the impact of dominant languages on other languages and language loss. Some may argue that it is all part of linguistic evolution and, for sure, not a matter of international concern. They contend that languages evolve based on their usefulness and consolidate according to the ability of a language to be a tool to communicate globally, without the burden of translation.

Yet, we are alarmed by the current rapid loss of technical Italian words used on a daily basis. We believe that this correlates with a decrease in the richness of our human expression and, thus, endangers the future of our linguistic identity.

References

1. Enciclopedia dell'Italiano. 2010 [cited 7 Jun 2016]. Available from: http://www.treccani.it/enciclopedia/ itangliano_ (Enciclopedia_ dell'Italiano)/.
2. Dani D, Baciu S. Intanglish: new Italian dialect? Cafébabel. 2012 Jan 5. [cited 7 Jun 2016]. Available from: http://www.cafebabel.co.uk/culture/article/itanglish-new-italian-dialect.html.
3. Rotta G. Itanglish: così parlò il Partigiano. La Stampa. 2013 Feb 17 [cited 7 Jun 2016]. Available from: http://www.lastampa.it/2013/02/17/cultura/opinioni/editoriali/itanglish-cosi-parlo-il-partigiano-mwzfdDpYxmZyfKLn3G6TeN/pagina.html.
4. Josh. An appropriate term for the 'contamination' of a language. English language and usage. 2016 [cited 7 Jun 2016]. Available from: http://english.stackexchange.com/questions/233187/an-appropriate-term-for-the-contamination-of-a-language/233209.

Laura C Collada Ali
Freelance translator & writer
laura.collada@teksema.com

Francesca Paoloni
Biostatistician at GIMEMA Data Centre
f.paoloni@gimema.it

Teaching Medical Writing

10 tips for organising a successful writing course

SECTION EDITOR

Claire Gudex
ClaireGudex@rsyd.dk

I teach medical writing at Copenhagen University, where I have successfully run the week-long Intensive Medical Writing Course for the last decade. Each time the course is held, there are three presenters, four tutors and 36 participants. In the first part of the programme on Monday afternoon, all participants attend three lectures: "Errors of grammar and usage", "Optimal presentation of figures" and "Medical writing seen from an editor's standpoint". The participants are then split into two groups and attend either the Tuesday and Wednesday or Thursday and Friday full-day sessions. This second part of the programme deals with the IMRAD structure for scientific articles, how to write clearly and concisely, style and vocabulary, punctuation, presentation of numbers, the publication process and the cover letter.

In 2007, I wrote an article for EMWA in which I described my early experiences regarding the setting up of the Intensive Medical Writing Course, whose aim was to increase chances of publication for non-native English speakers. The current article reflects my practical experiences since then. The following are 10 tips for organising a successful course.

1. **Send information to the participants three times**: as soon as the course is approved; 10 days beforehand; and immediately before the course starts. In my experience, participants tend to 'lose' emails and sending out three emails is not overdoing it.
2. **Make sure everyone can find the rooms.** Send out a map with instructions (third email) and then put up plenty of signs. A roller banner is a good investment as it's visible from a distance, attracts the participants' attention, and provides good publicity as people walk past. Copenhagen University's medical school, where our course is held, is a veritable rabbit warren of corridors, stairways and underground passages.
3. **Choose a room of an appropriate size for the opening session**: 40 people rattling around in a massive lecture theatre is far from ideal. In the smaller teaching rooms it's worth spending time re-arranging the tables in horseshoe formation. Participants can then see each other and this automatically generates a friendlier atmosphere. Looking at the back of someone's head is not conducive to any form of friendliness or interaction.
4. **Welcome participants as they arrive.** Direct early arrivals to the other end of the room so there are spaces near the door for latecomers. (Also welcome latecomers!) A quick round of introductions, where everyone mentions their name, institute and project, serves as an icebreaker and arouses curiosity; it's amazing how quickly participants then start interacting and building relationships.
5. **Mix lectures, presentations, exercises and small-group discussions**. The exercises and small-group discussions complement the lectures and presentations. There are short exercises on grammar and usage, punctuation and 'removing the dead wood', and a longer exercise on how numbers should be presented. In the small groups, where there are six participants and a tutor, participants' own texts are discussed. A checklist for evaluating manuscripts is provided. Interestingly, these small-group discussions and exercises are often considered the most valuable parts of the course.
6. **Keep up a fast pace: the programme should run seamlessly with minimal interruptions.** We keep to a tight schedule. Brief questions are encouraged, but longer discussions are kept for the coffee breaks and lunchtime. Note: All tutors are present throughout; they are always ready to answer questions and elaborate on points brought up during the sessions.
7. **Provide handouts that are useful and legible.** Provide answers to everything. There is nothing more frustrating than returning from a course and not being

able to make sense of the PowerPoint handouts.

8. **Spend time developing an effective evaluation form.** If participants mark down any of the items, ask them to explain why. Also, importantly, ask them to suggest improvements. (Ignore impossible requests.)
9. **Be prepared for all eventualities**. Try to fill places when there are last-minute cancellations. There are waiting lists for our courses and we do our best not to waste any places. Guest speakers can drop out unexpectedly; untimely failures of audio-visual equipment can try everyone's patience; and, worst of all, the coffee and cake can fail to arrive! Hence you should have the mobile numbers of important contacts, including the IT department and the canteen, written in indelible ink on the back of your hand. (Many things can go wrong – if anyone would like a comprehensive list they are welcome to get in touch.)
10. **Finally, send out a follow-up email with useful links and answers to questions that have required extra research.** Request additional feedback; this can be used to make the course even better next time.

The Intensive Medical Writing Course currently runs in January and June. In addition, longer medical writing courses, consisting of eight sessions with 12 participants, run in the spring and autumn. As a new venture – at the request of former participants – a one-day follow-up course was successfully established last November and is now scheduled to run twice a year. It should be noted that the texts submitted for the November follow-up course were light years ahead of those submitted for the preceding full-length courses, which illustrates the positive effect the writing courses are having.

Acknowledgements

Thank-you roll call: Joan Waddell, Gevene Hertz, Monika Schoell, Philip Hollingbery and Stephen Gilliver.

Suggested Reading

1. Møller C. New medical writing course in Copenhagen: Increasing chances of publication for non-native speakers. The Write Stuff. 2007;16(1):8–9.
2. Møller C, Schoell M. Language revisers/translators/editors: is there anyone out there? The Write Stuff. 2008;17(2):85.
3. Schoell M. Working as a medical editor and translator in a university environment in Germany. Medical Writing. 2012;21(3):257–9.

Christine Møller
APMIS (Acta Pathologica, Microbiologica et Immunologica Scandinavica) and Medical Manuscripts, Copenhagen, Denmark
apmis@apmis.dk or medicalmanuscripts@gmail.com

Profile

SECTION EDITOR

Beatrix Doerr
beatrix.doerr@gmail.com

An interview with Professor Peter Jüni

on methodology and statistics in scientific manuscripts

This issue of Medical Writing (MEW) is about statistics, so what is more appropriate than interviewing a research methodologist who focuses on epidemiology and statistics in clinical research? I am happy that we were able to win Professor Peter Jüni for this interview. Peter Jüni is a physician by education, has been a Professor of Clinical Epidemiology and the Director of the Clinical Trials Unit and the Institute of Primary Health Care at the University of Bern. In 2016, he moved to Toronto where he is a Professor of Medicine at the University of Toronto, and the director of the Applied Health Research Centre (AHRC) at the Li Ka Shing Knowledge Institute. The AHRC is a leading not-for-profit academic research organization fully integrated with the Li Ka Shing Knowledge Institute of St. Michael's Hospital and affiliated with the University of Toronto.

Peter Jüni has authored more than 270 peer-reviewed publications. Amongst them were several landmark trials and meta-analyses, various international guidelines (such as the 2014 ESC/EACTS guidelines on myocardial revascularization), and several articles on statistical topics such as systematic reviews, meta-analysis, and propensity score techniques. He has been a reviewer for major journals such as *The Lancet*, and was listed as highly cited researcher by Thomson Reuters.

Medical Writing (MEW): You review many manuscripts. What are the most common mistakes you see?

Peter Jüni (PJ): The most common mistakes I see is that the perspective of the reader is ignored and the manuscript is not structured logically and coherently. Thus, this is much more about a basic lack of structure and logic than about fancy statistics. A caveat: my observations are mostly related to working with fellows, PhD or MD students – they might not apply, or only to a lesser extent, to medical writers.

The introduction should clearly lead to the main question. The main question should then be reflected in the methodology, including the statistical section. All

descriptions should be transparent, consistent, and easy to understand. Often, I find analyses in the results section, which have not been described in the methods section or vice versa. In other cases, I find that the content does not reflect the structure of the manuscript, descriptions of methods end up in the results section, results in the methods section and things get mixed up quite a bit. 'What was done' belongs into the methods section, 'what was found' should be reported in the results section and 'how this should be interpreted' can be stated in the discussion.

Frequently, protocol-specified outcomes are missing from methods or results, or new outcomes are reported that were inexistent in the protocol. Randomisation lacks an appropriate description, important elements are lacking, such as the generation of the random sequence, including stratification and blocking, and more importantly, the reader does not understand the mechanism of concealment of allocation. However, all the high level stats are completely futile if randomisation was messed up in the first place. The subsequent methodological steps following randomization (blinding, follow-up of patients, intention-to-treat analysis) are ultimately deemed to maintain the experimental momentum introduced by randomization and should be described meticulously.

The discussion section if often a wild, completely unstructured experience, when in fact it can be structured into separate paragraphs describing main findings, context, strengths, weaknesses, clinical and scientific implications of the work. Display items, i.e. tables and figures, should be completely self explanatory, with a legend that makes sure that the reader will not have to go back to the main body of the manuscript to understand what is being reported. Following the CONSORT 2010 and related guidelines (see http://www.equator-network.org/) will help a great deal to get this right. However, I would recommend using these guidelines like a cook book – don't follow it too slavishly, but make sure to include most ingredients.

MEW: What are the most common mistakes you see related to statistics?

PJ: Well, on a more conceptual level, many of the mistakes I see probably start with our trouble in accepting uncertainty. People ignore that the probability of hypotheses depends on much more than just the p-value, and even worse, divide the world into significant and non-significant. Used in such a naïve way, statistics will not help us to quantify uncertainty appropriately.

Results of a trial should be interpreted in the light of the sample size consideration. So, a comprehensible and complete description of the power calculation, which is not too technical is crucial – simply copying and pasting the statements received from the statistician is not good enough. Reporting of results should include absolute numbers, percentages, estimated differences between groups with corresponding 95% confidence intervals. P-values would actually not be necessary, but if they are reported, they should be reported exactly, and not, as already stated above, as merely significant or non-significant. A frequently encountered tautology is the reporting of p-values for baseline comparisons in randomised trials – not really helpful at best, misleading at worst, please avoid! Other frequent mistakes include taking correlations as evidence for causation, choice of wrong statistical models, over-interpretation of secondary outcomes, over-interpretation of subgroup analyses and mixing up statistical significance with clinical relevance.

MEW: How should an ideal cooperation between a statistician and a medical writer look like?

PJ: Both parties need to understand clinical and biological context and basic statistical principles to properly interpret results from a statistical analysis – mere number crunching is not enough. Continued cooperation and mutual exchange is key.

The sophistication of a manuscript lays in its clarity, transparency, consistency and simplicity, and in its focus on the readers' perspective, not in complex writing styles.

Conclusion: Professor Peter Jüni shared some of his experience with us. I hope this will be valuable not only for inexperienced writers, but also for experienced ones. The sophistication of a manuscript lays in its clarity, transparency, consistency and simplicity, and in its focus on the readers' perspective, not in complex writing styles. And let's not forget our clinical judgement when we interpret statistical analyses!

Out on Our Own

SECTION EDITORS

Julie Charlesworth
julie@atreeoflifesciences.com
Satyen Shenoy
sshenoy@describescientific.de

Editorial

The theme for the FBF at the Munich conference was *Out On Our Own but Not Alone*. With an impressive attendance by more than 60 freelancers (FLs) from Europe and beyond, and also a special guest – Dr Andreas Lutz of the VGSD, a German association fostering freelancing – we rolled out new initiatives in building relationships with other associations that could help us in running our businesses (for those interested, minutes are available in the Freelance Resource Centre of the EMWA website). Thank you and we hope to see more of you in Brussels … we are not alone!

In this edition of the OOOO we continue along those lines …

Marco Torregrossa, the Secretary General of the European Forum of Independent Professionals (EFIP; www.efip.org) gives us an introduction to this collaboration of various European national groups, and the activities EFIP are involved in to support and promote the freelance way of working. We are also excited to inform you that in future OOOO editions, Marco will be covering various topics related to freelancers that we'll definitely find interesting, pertinent, and enlightening!

In the last few years, a novel concept of co-working has made rapid ground among freelancers and small-business owners worldwide. The central idea behind this concept is innovation and networking, allowing FLs to rent spaces in 'workhubs' along with others on a part-time basis and use the support services provided by such facilities. In this edition, the Association of Independent Professionals and the Self Employed (IPSE) gives us a glimpse into the world of co-working and useful links for those wishing for more information.

We also bring you the third and last part of Marion Alzer's article on her career as a medical translator. In this part, Marion talks about her established business and the approaches she has made to diversify and add value to her services, and reflects on her career choice. To those of us interested in becoming medical translators, Marion's article is certainly helpful.

Finally, if you have an article that you wish to contribute to the OOOO and share with your fellow FLs, please feel free to send it to us. In the meantime, happy reading and best wishes.

Julie Charlesworth
and Satyen Shenoy

The coworking revolution

"I can't believe how much more productive I am since working from home," said Andrea, who had just started working as a freelance medical writer after years of doing the nine to five (or more often nine to nine) routine in an office.

Without the distractions of office life it can be much easier to focus. None of those constant interruptions from colleagues. No more emails advising you of the latest parking policy, or asking you to attend another all-staff meeting. Gone are the weekly fire alarm tests.

You're free to devote your energy to producing excellent work, which is why clients hire you.

Working from home also has an added benefit – you can take a 15 minute break to go and hang out the washing. Short breaks have been shown to boost concentration and productivity, so you can have an ordered house without feeling guilty!

However, there is a downside to working on your own. Over time it can start to feel a little bit … quiet.

You also miss out on those 'watercooler moments', as Silicon Valley executives like to call them, in other words, the casual encounters with colleagues, such as when making coffee in the staff kitchen. These moments can be a great source of new ideas, or an exchange of important information you wouldn't otherwise have come across.

A study[1] by workhubs.com found that nearly half of homeworkers (44%) struggled with having their work disturbed by other household members and 37% said that the lack of mental stimulation or interaction was a problem.

This is why the concept of coworking is becoming so popular. It offers the best of both worlds.

What is coworking?

Coworking spaces have been appearing in cities all over the world, and increasingly in smaller towns as well. Sometimes also known as 'workhubs', they are shared office spaces, providing a cost-effective solution for freelancers and other small business owners to work from, often on a part-time basis.

Most of them offer flexible packages, allowing you to book a workspace by the hour or day, or for a certain number of days per month. For example, some people work two days a week from a coworking hub, and for the remaining three days they work from home.

The key thing that distinguishes a 'coworking' space from a more traditional rented workspace is the idea of collaboration. Many of them have this as their core philosophy, and make it an integral part of their offering.

For example, THECUBE in London (thecubelondon.com) says:

"When you join THECUBE, you join more than just a workspace, you join a curated, diverse and smart community of scientists, engineers, designers, technologists, artists, futurists and anthropologists. Our workspace helps our members innovate through our events, innovation labs and one to one mentoring. As a community we actively collaborate to create innovative sol-

utions via our consultancy and independent projects."

The concept of a collaborative workspace for independent professionals is widely thought to have started with the iconic C-base, which was founded in Berlin in 1995 as a 'hackerspace', where people with common interests could meet, share knowledge, and experiment.

By 2007 the term 'coworking' had become one of the most widely searched terms on Google, and in 2012 it even got its own conference to celebrate "a worldwide explosion of coworking awesomeness". In true digital style it was named the Global Coworking Unconference Conference (gcuc.co).

Coworking has seen astonishing growth over the last five years – the number of coworking hubs around the world has almost doubled every year. This has been fuelled in large part by the overall trend towards independent working and the modern innovation-driven economy, which relies heavily on freelance talent (see www.ipse.co.uk/research).

Embracing innovation

Coworking hubs tend to place a great deal of emphasis on creating a distinctive look and feel. As a result they have a lot of character, which attracts forward-thinking people and gives them an air of excitement – a certain 'buzz'.

As well as being an inspiring place to work, they can be a great place to meet clients – many coworking hubs offer affordable rates for meeting rooms. Often they also offer business address and mail forwarding services, allowing you to present a professional, innovative image to the outside world.

The chance to mingle with diverse professionals is another key factor. Although the majority of people who use coworking hubs are freelancers working on their own independent projects, it's not unusual to see groups huddled together swapping ideas.

Of course you're under no obligation to chat if you don't want to, but many freelancers value the chance to share information freely about each other's businesses, which can trigger thoughts for a different way of doing things, and sometimes even lead to new projects.

Mandy Taylor, who works from Coachwerks, a UK coworking hub, says that the mix of people and skills makes it "amazingly fecund." She says, "you're encouraged to try running workshops because everyone assumes you can do it. You're pushed to try things you wouldn't normally. There's also the benefit of a critical second opinion."[2]

Freelancer James Holloway says: "Even though I've only been here a few weeks I've noticed a significant increase in my productivity. I read a study that showed that mixing in diverse social groups produces more innovative thinking and I feel that's an important aspect of coworking. On a more practical level, the email group has proved helpful – people will chip in with advice. I'm also starting a personal branding exercise, so those different perspectives will be helpful."

An iconic coworking hub near you

There are now over 1600 coworking hubs across Europe. For a list of prominent hubs in each country, visit www.jobfluent.com/the-ultimate-list-of-coworking-spaces-in-europe. You can also find a comprehensive map of hubs all over the world at www.coworkingvisamap.com.

If you're a member of IPSE (www.ipse.co.uk), the UK Association of Independent Professionals and the Self Employed, you can use a London coworking hub for up to four days a month, completely free of charge. You can choose from any one of 80 coworking hubs across the city, courtesy of www.workspace.co.uk/co-working. The IPSE membership card also allows you to touch in and touch out at over 250 workspaces around the UK – just pay for the time you use.

So if you love the experience of working for yourself, but want to add that extra buzz to your week, spending some time in a coworking hub could give you a much needed dose of energy and inspiration!

References

1. Dwelly T, Lake A, Thompson L. Workhubs: smart workspace for the low carbon economy. Workhubs Network 2010 (cited 2016 May 4). Available from: http://www.flexibility.co.uk/downloads/Workhubs-finalreport.pdf
2. The Werks, Hove. Workhubs Network 2010 (cited 2016 May 4). Available from: http://www. workhubs.com/sites/default/files/ pagedocs/the_werks_hove.pdf

Michelle Storm Lane and Tim Bradburn
Association of Independent Professionals and the Self-Employed, London, UK
michelle.lane@ipse.co.uk

Medical translation – a dead-end job or a gateway to opportunity? Part 3

An established business

Today, my translation business is well-established. Clients range from large pharmaceutical and biotech companies and CROs to small and medium-sized organisations which include regulatory consultancies, medical centres, patient organisations, translation agencies, academic institutions, and even private individuals.

My clients typically require translation from English into German of one of the following types of text:

Typical texts I translate

- Press releases / newsletters / websites
- Training / educational materials
- Product monographs
- Patient brochures
- Protocol synopses
- Patient informed consent documents
- Patient diaries / questionnaires
- Original articles
- Posters
- Abstracts
- Presentations
- Summary of Product Characteristics (SmPC), patient information leaflets, labelling
- Dear Health Care Professional letters
- Medical reports
- Benefit dossiers for health technology assessment
- Inspection reports (GCP, GMP)
- Standard operating procedures
- Guidelines
- Job descriptions

Diversifying

To offer my clients additional value, I applied to the higher regional court for authorisation to certify translations. This is an official procedure common in Germany which grants the applicant the right to certify translations needed for official and legal purposes. I am now publicly appointed and sworn in by a Bavarian court and have the right to officially certify translations. Box 2 provides some examples.

Translations I have certified

- Certificates of Pharmaceutical Products and SmPCs for importing medicinal products into countries in the Middle East
- License to practice dental medicine for a dentist emigrating to Dubai
- Scientific opinion issued by the European Commission on the safety of breast implants for a lawsuit in Germany

Being a sworn translator unlocks additional job opportunities. In one case, the client explicitly requested the presence of a sworn translator to assist communication during an FDA inspection at a clinical site. My tasks were to interpret the inspectors' interactions with the investigator and to read entries in medical files and convert them from German into English for the inspectors. This enabled them to verify data submitted for a New Drug Application against source. My previous work experience as a CRA and familiarity with patient files proved invaluable for this job. Similarly, in specific settings, I offer liaison interpreting which is the oral transfer of short spoken passages. These settings also include inspections by supervisory or regulatory authorities, or local Ethics Committee meetings where non-German speaking sponsors defend their clinical trials.

Another area that translators may consider is foreign language teaching. I have taught German and English as a foreign language to health-care professionals and led medical translation workshops. Although teaching is enjoyable, the preparation required can be time consuming.

Finally, medical writing should not be underestimated in providing translators with a whole range of opportunities. Some years ago, I attended a medical translation workshop where I met Susanne Geercken. Susanne, an experienced EMWA workshop leader, explained to me that translators could benefit from the association's professional development programme and encouraged me to become an EMWA member. Thereafter, I completed my foundation certificate in medical writing and have subsequently secured my first freelance medical writing jobs.

Reflections of a translator

More recently my work experience has included more emphasis on medical writing and even monitoring clinical trials which reflects the diversity of my work. The work never stops changing. Far from being a dead end, medical translation provided me with a solid basis to take on so many new roles – certainly not the boring career I feared when I started!

Marion Alzer
Pharma Communications GmbH,
Erding, Germany
alzer@pharmacoms.com

What's happening at the European level:

The European Forum of Independent Professionals (EFIP) and why you should know about it

The European Forum of Independent Professionals (EFIP; www.efip.org) is furthering a commitment to build a better economy – one that puts people at the centre of it, empowering anyone to work for themselves and succeed on their own terms. We are a European not-for-profit collaboration of national associations *(see box right)* which represents over 10 million independent professionals at EU level through targeted research, advocacy and campaigning. Independent professionals (often referred to as freelancers or contractors) are highly-skilled self-employed workers without employers or employees. They offer specialised services of a knowledge-based nature and work on a flexible basis in a range of creative, managerial, scientific and technical occupations, primarily in B2B. They are the smallest of small businesses and, with a 45% increase since 2004, they are the fastest growing segment of the EU labour market.[1]

Freelance workers are an important, but hidden, part of the small business population. While we know the creative and information sectors combined (where medical writers belong) is the second largest group of freelancers, we lack EU wide data for the freelance medical profession. In the UK it has been estimated that independent healthcare professionals in 2015 have grown at a rate of 67.7% compared to 2008.[2] The main driver behind working as an independent professional is a strong commitment and reliance on skills and professional development. The problems independent healthcare professionals may experience are related to keeping up to date on developments within the medical industry, which is subject to considerable regulation and operates within strict codes of practice and standards that frequently change. Keeping up with these matters is difficult which is why EFIP has for a long time advocated that access to affordable training is of paramount importance to independent professionals which needs to be at the same levels and standards offered to employees.

Organisation	Country of origin
Association of Independent Professionals and the Self Employed (IPSE)	United Kingdom
Verband der Gründer und Selbstständigen e.V (VGSD)	Germany
Associazone Consulenti Terziario Avanzato (ACTA)	Italy
FEDIPRO vzw	Belgium
Platform Zelfstandige Ondernemers (PZO)	Netherlands
Fédération des Auto-Entrepreneurs (FEDAE)	France
Syndicat des Consultants Formateurs Indépendants (SYCFI)	France
Asociatia Freelancerilor (AF)	Romania
Stowarzyszenie Samozatrudnieni	Poland
Croatian Independent Professionals Association (CIPA)	Croatia
Unión de Profesionales y Trabajadores Autónomos (UPTA)	Spain
Swedish Umbrella Companies Trade Association *(Affiliate Member)*	Sweden

EFIP
European Forum of Independent Professionals

References

1. Leighton P. Future working: The rise of Europe's independent professionals (iPROs). 2016 [cited Apr 2016]. Available from: http://www.efip.org/sites/default/files/Future_Working_Full_Report.pdf.
2. Kitchling J. Exploring the UK freelance workforce in 2015. 2016 [cited May 2016]. Available from: https://www.ipse.co.uk/sites/default/ files/documents/research/Exploring-uk-freelance-workforce-2015-reportv1.pdf.

Marco Torregrossa
European Forum of Independent Professionals, Brussels, Belgium
marco.torregrossa@efip.org

Clarity and Openness in Reporting: E3-based (CORE) Reference

An Open Access Resource to Support Authoring of Clinical Study Reports for Interventional Studies

DOWNLOAD THE LAUNCH PUBLICATION: http://dx.doi.org/10.1186/s41073-016-0009-4

DOWNLOAD THE USER MANUAL: http://www.core-reference.org

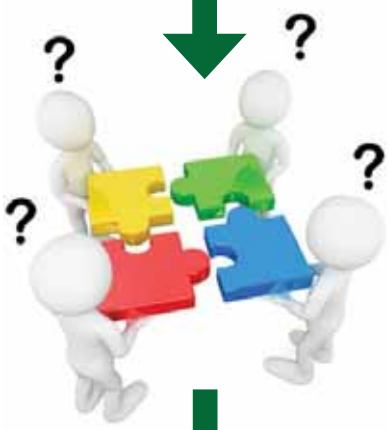
WRITE OR REVIEW CLINICAL STUDY REPORTS (CSRs)?

WRITE OR REVIEW STATISTICAL ANALYSIS PLANS (SAPs)?

YES

NEED HELP INTERPRETING ICH CSR AUTHORING REQUIREMENTS?

WHAT IS 'RESPONSIBLE CLINICAL TRIAL DATA SHARING'?



NEED HELP UNDERSTANDING PUBLIC DISCLOSURE REQUIREMENTS FOR CSRs?

HOW DOES PUBLIC DISCLOSURE AFFECT CSRs AND PRESENTATION OF DATA?

SHARING KNOWLEDGE TO HELP YOU WRITE FIT-FOR-PURPOSE CSRs

Working in these areas?

- Medical Writing
- Regulatory Affairs
- Statistics
- Clinical Research
- Publication Planning
- Medical Communications
- Clinical-Regulatory Document Public Disclosure
- Regulatory Document Publishing

You should know about: http://www.core-reference.org

Please inform your senior colleagues

Consider CORE Reference a 'User Manual' that may be used in conjunction with company Standard Operating Procedures to support the authoring of Clinical Study Reports fit for today's modern drug development environment.

Themes of forthcoming issues of **Medical Writing**

December 2016:
'Medical Education'

This will include articles on running advisory boards and preparing slide kits, conference presentations, and other learning resources.

This issue is closed to new feature articles.

June 2017:
'Medical Devices'

This will include articles on the regulatory approval process for medical devices, preparing related documents, writing publications on clinical studies about medical devices, and other aspects of the medical device field relevant to medical writers.

The deadline for feature articles is March 13, 2017.

March 2017:
'Writing Better'

This will include articles and exercises to help medical writers write better in English.

The deadline for feature articles is December 12, 2016.

CONTACT US

If you have ideas for themes or would like to discuss any other issues, please write to **editor@emwa.org**.