How to write a scientific article?

Dr. A.A.M. Gerritsen, Epi Result

Overview lecture

» Why a scientific format?
» Sections of the article
» Process of writing
» Structure of a paragraph
» Reporting and journal guidelines
» Tips to improve your writing & useful resources
Why a scientific format?

» Rigid structure facilitates:
  - Efficiently communicating scientific findings (uniform manner)
  - Allows the article to be read at different levels (skim titles, read title with abstract, look further to tables and figures etc.)

Sections of the article (in general)

» Title Page
» Abstract
» Introduction
» Methods
» Results
» Discussion
» Acknowledgements
» References
» Appendices
Title page

» Title: describes the content/main result of the paper, time and place (key words for search)

» Authors’ names, institutional affiliations, details for correspondence
Abstract (1)

» Summarizes major aspects entire article
» In general 200-300 words
» Use 2-3 sentences per section of the paper (introduction, methods, results, discussion)
» Information that is essential to the reader (to find the article, to decide what to read further or what is needed in case full article can’t be accessed)
» First section, but last to be written

Abstract (2)

» The question(s) you investigated (or purpose) (from Introduction)
» The basic design and methods used (from Methods)
» The major findings that answer the question(s) (from Results)
» A brief summary of your interpretations and conclusions, including implications of the findings (from Discussion)

No references, figures/tables.
Provision and need of HIV/AIDS services in the City of Tshwane Metropolitan Municipality, 2010

Annette A M Gerritsen, Janine S Mitchell, Brenda White

Objectives. To determine the need for HIV/AIDS service provision in the City of Tshwane Metropolitan Municipality (CTMM), especially in municipal areas.

Methods. The Foundation for Professional Development initiated the Compass Project. Using a questionnaire, data were collected during May – June 2010 from organisations providing HIV/AIDS services in the CTMM (organisational information and types of HIV/AIDS services). The need for HIV counselling and testing (HCT), antiretroviral treatment (ART), prevention of mother-to-child transmission (PMTCT), and care for orphans and vulnerable children (OVC) was estimated using data from various sources.

Results. A total of 447 service providers was included in the study: 72.3% non-governmental organisations (NGOs); 18.1% in the public sector; 5.1% in the private sector; and 4.5% faith-based organisations. The majority of the prevention- (70.2%) and support-related services (24.4%) were provided by NGOs, while the majority of treatment-related services originated from the public sector (57.3%). Service need estimates included: HCT – 1 435 438 adults aged 15 - 49 years (11 127/service provider); total ART – 75 211 adults aged 15+ years (1 213/service provider); ART initiation – 30 713 adults aged 15+ years (495/service provider); PMTCT: HCT – 30 092 pregnant women (510/service provider); PMTCT-ART – 7 734 HIV+ pregnant women (221/service provider); and OVC care – 54 590 children (258/service provider).

Conclusions. Service gaps remain in the provision of HCT, PMTCT-ART and OVC care. ART provision must be increased, in light of new treatment guidelines from the Department of Health.


Introduction (1)

Should contain the following sections:

» Problem statement

» Rationale for the study

» Significance of the study

» Objectives of the study

[From general (literature, practice) to specific (your study).]
Introduction (2)

Problem statement:

This is a concise description of the nature of the problem (the discrepancy between what is and what should be) and of the size, distribution and severity of the problem (who is affected, where, since when, and what are the consequences).

Use a literature review (original research papers/review articles) to provide this context: general to specific

e.g. HIV/AIDS, treatment, adherence ... and/or World, Africa, South Africa ...

[Extensive descriptions of similar studies come in the Discussion.]

Introduction (3)

Example problem statement:

Gender differences in health and health care utilisation are well documented. Women generally experience poorer health than men. [1-5] Although some studies have shown that the direction and magnitude of gender differences in health may vary according to the particular health outcome. [6,7] Determinants of gender differences in health include biological (e.g. genetic and hormonal factors), psychological (e.g. gender identity and roles, chronic stressors), behavioural (smoking, drinking, eating, physical exercise) and social factors (e.g. social support, socio-eco-...

Besides gender differences in health and health care utilisation, ethnic differences have also been the subject of research. In general, migrant groups experience a poorer

Introduction (4)

**Rationale for the study:**

This relates to the origin/source of the topic and the importance of the problem. A brief description of any solutions to the problem that have been tried in the past should be given, how well they have worked, and why further research is needed.

Again, use literature to support this.

Introduction (5)

**Example rationale:**

Studies on ethnic differences in health and health care utilisation often adjust for gender and/or report results for women and men separately, without systematically reviewing gender differences in health and health care utilisation within and between ethnic groups [10, 11, 13, 15, 18, 19]. However, some studies highlight the gender differences found, despite the fact that this was not the principal aim of the study [12, 14, 17].

Little is known about whether gender disparities persist across different ethnic groups or whether they differ due to cultural differences, for example regarding the position of women in the society. Therefore the purpose of the current study is to determine gender differences in health and
Introduction (6)

Significance of the study:

This is a description of the type of information expected to result from the project and a clarification of how this information will be used to help solve the problem (contribution to existing knowledge).

Example:

- health care utilisation in various ethnic groups in the Netherlands. This information might be helpful to develop policy to focus on the health status and accessibility of the health care system of specific groups according to gender and ethnicity.

Introduction (7)

Objectives of the study:

» General objective (general aim or purpose of the study which is derived from the research topic).

» Specific objectives which are based on your general objective.

Example:

- to cultural differences, for example regarding the position of women in the society. Therefore, the purpose of the current study is to determine gender differences in health and health care utilisation in various ethnic groups in the Netherlands. This information might be helpful to develop policy to focus on the health status and accessibility.
Methods (1)

» Study design
» Study setting, population and sampling
» Measurement instrument(s)
» Data collection
» Data-analysis
» Ethics

Methods (2)

Study design:

Explain the design of your study e.g. cross-sectional, case-control, cohort study, randomized controlled trial. In case of an intervention study describe, if relevant, the method of randomization/concealment of treatment allocation.

Example:

A randomized controlled trial conducted from October 1998 to April 2000 at 13 neurological outpatient clinics in the Netherlands.

**Methods (3)**

**Study setting, population and sampling:**

Define the study setting (e.g. hospital), population (e.g. age, sex, place, condition, etc.), the sampling or selection method/criteria, and justification of sample size (power calculation).

Intervention study: And add a section in which you describe the interventions that the treatment and control groups received.

**Methods (4)**

**Example study setting, population and sampling:**

All patients with *clinically suspected CTS* had been referred to one of the participating neurologists and were examined for eligibility to participate in the study. *Inclusion criteria were* (1) pain, paresthesia, and/or hypoesthesia in the hand in the area innervated by the median nerve; ... *Exclusion criteria were* (1) previous treatment with splinting or surgery; ... *Sample-size calculation* was based on the ability to detect a clinically important difference in success rates of 20% or more 3 months after randomization. A total sample size of 190 patients was required (2-sided $\alpha = .05$, $\beta = .20$).
Methods (5)

Measurement instrument(s):

Describe the instrument(s) used for data collection (e.g. interview guide, questionnaire, checklist or data collection form). Describe the instrument(s) in detail e.g. questions/answers (including validity, reliability).

For intervention studies: Is blinding applied (for caregivers, participants, outcome assessors) where relevant?

Methods (6)

Example measurement instrument(s):

» Patients completed questionnaires: General improvement was scored by the patient on a 6-point ordinal transition scale, ranging from "completely recovered" to "much worse."12 ... Mean (SD) scores were collected using the Symptom Severity Scale (11 questions about symptoms experienced during the past 2 weeks with 1 equaling mildest and 5 equaling most severe) ...

» A research physiotherapist examined the patients: After a standardized history-taking and a physical examination, the overall severity of CTS complaints was scored by a research physiotherapist on an 11-point numerical rating scale with zero equaling no complaints and 10 equaling very severe complaints.12 ....
Methods (7)

**Data collection:**

Who collected what data, when?

Example:

Patients completed questionnaires and were examined by a trained research physiotherapist in the hospital at baseline and 3, 6, and 12 months after randomization. Although different research physiotherapists assessed the outcomes, most patients were seen by the same therapist each time they visited the hospital. In the remaining months, and 18 months after randomization, questionnaires were mailed.

Methods (8)

**Data analysis:**

This should cover: primary and secondary outcomes, how the data were summarized (e.g. means, %) and measures of variability (e.g. SD, 95% CI), the statistical techniques used; the programme used.
Methods (9)

Example data analysis:

To determine the long-term effects, the assessment time of 18 months after randomization was chosen. Differences in success rates between the treatment groups along with 95% confidence intervals were calculated using $\chi^2$ tests. Continuous outcomes were analyzed as change scores (difference between baseline assessment and each follow-up assessment). Subsequently, differences in improvement between the groups (mean change score in surgery group minus mean change score in splint group) along with 95% confidence intervals were calculated using t tests. Multivariate analyses (logistic or linear regression) were performed to adjust for the influence of eventual differences between the groups at baseline in prognostic indicators ...

Data were analyzed using SPSS statistical software (Version 10.1; SPSS Inc, Chicago Ill).

Methods (10)

Ethics:

Review by ethics committee (number).

Use of informed consent.

Example:

The medical ethics committees of the 13 participating hospitals approved the study protocol of this multicenter RCT. After they had provided written informed consent, patients were included in the study.
## Results (1)

» **Function:** Objectively (and concisely) present key results, without interpretation (see Discussion).

» **Organisation:**
  - Develop Tables/Figures based on your analysis; place these in a logical sequence following your objectives.
  - Add text that highlights one or two findings (e.g. trends, differences, similarities, correlations ...) that can be found in each Table/Figure (don't repeat the information).

## Results (2)

» **Details:**
  - Report your results with name of the statistical test and result (e.g. 95% CI or difference, and p-value, and n)
  - Report negative results as well
  - Don't feel obliged to report on all the results in one article (make a decision based on objectives addressed in this article)
  - Use appendices if needed
The prevalence of diarrhoea in this study was 24.5%. The mean number of diarrhoea episodes was 0.38 per person in 2 weeks. For participants with diarrhoea, the most recent episode had lasted 2 days (9 cases, 33.3%), 3 days (8, 29.6%), 1 day (6, 22.2%), 4 days (2, 7.4%) and 7 days (2, 7.4%). The relationships between diarrhoea and labour force status (Fisher’s exact test $p=0.023$), living conditions (Fisher’s exact test $p=0.050$) and monthly expenditure (Fisher’s exact test $p=0.037$) were statistically significant. Other variables evaluated and their relationship with diarrhoea are set out in Table I.

### Table I. Variables evaluated and their relationship with diarrhoea

<table>
<thead>
<tr>
<th>Variable</th>
<th>Diarrhoea</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>20/89 (22.5%)</td>
<td>0.298</td>
</tr>
<tr>
<td>Female</td>
<td>7/21 (33.3%)</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illiterate</td>
<td>1/2 (50.0%)</td>
<td>0.223</td>
</tr>
<tr>
<td>Primary</td>
<td>3/13 (23.1%)</td>
<td></td>
</tr>
<tr>
<td>Secondary</td>
<td>18/84 (21.4%)</td>
<td></td>
</tr>
<tr>
<td>Tertiary</td>
<td>5/11 (45.5%)</td>
<td></td>
</tr>
</tbody>
</table>

Discussion (1)

> Organisation:

- Give an answer to the objectives of the study. Interpret all important results (refer to other studies if needed). Also unexpected findings.
- Compare results with other studies done on the topic: similarities/differences (explain the latter)
- Highlight limitations (and consequences thereof) and strengths of your study
- Implications of the results and way forward (research, practice, policy)
- Concluding sentence

Discussion (2)

> Details:

- Information needed in general on other studies in order to compare with current study: Year conducted (or published), study design, location/setting, population (age, gender, number), most important findings.
- Do not present any new results here (need to go into the Results section)
- From specific (your study) to general (literature, practice) [Note: Introduction the other way around.]
Discussion (3)

Example answer objectives, interpret results:

The main purpose of this study was to compare the use of health care services between refugees and asylum seekers in The Netherlands. It was expected that, in the unadjusted analyses, asylum seekers would have a higher use of (mental) health care services, because, in general, asylum seekers report more (mental) health problems than refugees. However, the unadjusted analyses did not show any statistically significant differences between refugees and asylum seekers, although asylum seekers seemed to make more use of mental health services (but the difference was just statistically insignificant). A reason for not finding a difference in contact with a general practitioner could be lying in the fact that asylum seekers do not have direct access to a general practitioner, but have to be referred by a nurse of the MOA team. When comparing the

Discussion (4)

Example compare results with other studies:

The results of the current study regarding gender differences within ethnic groups are in agreement with the study conducted in the UK, showing that the general health status of women is worse compared to that of men in some groups (e.g. the Moroccan, Turkish and Surinamese groups) and that the gender differences are small or absent in other groups (e.g. the indigenous group and the Antillean group, respectively). Regarding health care utilization, the results differ somewhat from the study conducted among older Americans, as gender differences in health care use vary among the different ethnic groups,
Discussion (5)

Example strengths and limitations:

The non-response rate in the migrant groups was higher than in the indigenous population. This was mainly due to difficulties in reaching the sampled persons, a well-known problem in population-based studies among migrants. \[10,18\] It is not clear whether the results of the study were consequently biased. However, no differences in age, gender, socio-economic position and general health status (obtained from the census data) between respondents and non-respondents were found.

The strengths of the current study include the relatively large sample sizes of the migrant groups compared to other studies on health and health care utilisation among these groups conducted in the Netherlands,\[10,19\] and

Discussion (6)

Examples implications of the results/way forward:

Future research could focus on, for example, the nature of problems the MOA nurse does or does not refer to a general practitioner; differences in health services use between various ethnic groups; possible limited access to primary care for both asylum seekers and refugees compared to the general population; and the relatively low use of mental health services by both asylum seekers and refugees.

In general the self-reported health of women is worse compared to that of men, although the size of the gender differences may vary according to the particular health outcome and among the ethnic groups. This information might be helpful to develop policy to improve the health status of specific groups according to gender and ethnicity.
Discussion (7)

Example concluding sentence:

In conclusion, asylum seekers and refugees seem to have equal access to the Dutch health care system in general. However, there are differences between the various ethnic groups in their self-reported use of the health care services.


Acknowledgements

» Name persons that are not an author but which assisted you when conducting the research, writing up the article etc.

» Mention sources of funding that supported the research.

Example:

ACKNOWLEDGEMENTS

We wish to thank those who co-operated in the participating hospitals: BovenIJ Hospital (Amsterdam), Flevo Hospital (Almere), Hofpoort Hospital (Woerden), Lorentz Hospital (Zeist), Onze Lieve Vrouwe Gasthuis (Amsterdam), Rode Kruis Hospital (Beverwijk), Spaarne Hospital (Haarlem), Sint Lucas Andreas Hospital (Amsterdam), Waterland Hospital (Purmerend), Westfries Gasthuis (Hoorn), Amstelveen Hospital (Amstelveen), Bronovo Hospital (Den Haag), Hilversum Hospital (Hilversum). Furthermore, we wish to thank the firm’s Loft, Meijer Orthopaedics and Welzorg for providing wrist splints.

Funding: The Health Care Insurance Council and the Foundation “Amazons” provided financial support.
References

» Use an accepted reference format e.g.:
   - Vancouver (use of sequential numbers in the text that link with endnotes)
     e.g. Women generally experience poorer health than men¹ ... 
   - Harvard (use of name, year in the text and an alphabetical list at the end)
     e.g. Women generally experience poorer health than men (Fernandez, 1999) ...

» Use a referencing programme (e.g. Endnote, Reference manager).

» Adhere to the guidelines of the journal.

Appendices

Contains information that is not essential to the understanding of the paper, but that e.g. further clarifies a point without burdening the body of the presentation. An appendix is an optional part of the paper, and more and more used for online publications.

Examples of what might be put in an appendix:

» large tables, figures, maps
» raw data
» explanation of ‘new’ statistical/mathematical procedures
**Process of writing these sections/the article**

- Start with an outline (subheadings, paragraphs, key words on what is covered)
- Then write a first draft (change key words into sentences, develop tables/figures).
- Put it aside, revise it, etc.
- Finalize article for internal review (supervisor(s), co-authors).
- Revise according to comments, send out for a second (or final) internal review.
- Finalize journal choice and adapt the article according to requirements. [Can also be done at an earlier stage.]

**Structure of a paragraph within a section**

- Each paragraph should have a coherent topic sentence, most often as the lead sentence.
- In each paragraph the other sentences should support the topic sentence.
- The transitions between paragraphs should be logical and smooth.

In general, limited the number of words, but say what you need to say.
Reporting guidelines (1)

**CONSORT 2010 checklist of information to include when reporting a randomized trial**

<table>
<thead>
<tr>
<th>Section/Topic</th>
<th>Item No.</th>
<th>Checklist Item</th>
<th>Reported on page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title and abstract</strong></td>
<td>1</td>
<td>a) Identification as a randomized trial in the title</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>b) Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for elements)</td>
<td></td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td>2a</td>
<td>c) Specific objectives or hypotheses</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2b</td>
<td>d) Scientific background and explanation of rationale</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2c</td>
<td>e) Description of trial design such as parallel, factorial, including allocation ratio</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2d</td>
<td>f) Important changes to methods after trial commencement (such as eligibility criteria), with reasons</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2e</td>
<td>g) Settings and locations where the data were collected</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2f</td>
<td>h) How and when they were actually administered</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2g</td>
<td>i) The interventions for each group with sufficient details to allow replication, including how and when they were actually administered</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2h</td>
<td>j) Completely defined pre-specified primary and secondary outcome measures, including how and when they were measured</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2i</td>
<td>k) Any changes to trial outcomes after the trial commenced, with reasons</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2j</td>
<td>l) When applicable, explanation of any interim analyses and stopping guidelines</td>
<td></td>
</tr>
<tr>
<td><strong>Sample size</strong></td>
<td>3a</td>
<td>m) How sample size was determined</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3b</td>
<td>n) If sample size was determined</td>
<td></td>
</tr>
<tr>
<td><strong>Randomization</strong></td>
<td>4a</td>
<td>o) Method used to generate the random allocation sequence</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4b</td>
<td>p) Type of randomization; details of any restriction such as blocking and block size</td>
<td></td>
</tr>
<tr>
<td><strong>Allocation concealment mechanism</strong></td>
<td>5b</td>
<td>q) Mechanism used to implement the random allocation sequence (such as sequentially numbered envelopes)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5c</td>
<td>r) Any steps taken to conceal the sequence until interventions were assigned</td>
<td></td>
</tr>
<tr>
<td><strong>Blinding</strong></td>
<td>6a</td>
<td>s) If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes of outcomes</td>
<td></td>
</tr>
</tbody>
</table>

**Reporting guidelines (2)**

**STROBE Statement—checklist of items that should be included in reports of observational studies**

<table>
<thead>
<tr>
<th>Item No.</th>
<th>Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(c) Indicate the study's design with a commonly used term in the title or the abstract</td>
</tr>
<tr>
<td></td>
<td>(d) Provide in the abstract an informative and balanced summary of what was done and what was found</td>
</tr>
<tr>
<td>2</td>
<td>Explain the scientific background and rationale for the investigation being reported</td>
</tr>
<tr>
<td>3</td>
<td>State specific objectives, including any prespecified hypotheses</td>
</tr>
<tr>
<td>4</td>
<td>Present key elements of study design early in the paper</td>
</tr>
<tr>
<td>5</td>
<td>Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection</td>
</tr>
<tr>
<td>6</td>
<td>(a) Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up</td>
</tr>
</tbody>
</table>
Journal guidelines

Tips on how to improve your writing

» Use guidelines (general, specific research design, journal)
» “Copy and paste” format from existing articles on a similar topic
» Read a lot of articles
» Have your article checked by an experienced writer
» Have your article checked by a language editor
### Useful resources

- Uniform Requirements for Manuscripts Submitted to Biomedical Journals  [http://www.icmje.org/manuscript_1prepare.html](http://www.icmje.org/manuscript_1prepare.html)
- Newsletters San Francisco Edit  [http://www.sfedit.net/newsletters.htm](http://www.sfedit.net/newsletters.htm)
- How to write a paper in scientific journal style and format  [http://abacus.bates.edu/~ganderso/biology/resources/writing/HTWtoc.html](http://abacus.bates.edu/~ganderso/biology/resources/writing/HTWtoc.html)