Critical Appraisal of the Medical Literature

James A. Hokanson, Ph.D.
Department of Preventive Medicine and Community Health
The University of Texas Medical Branch
2001
Power Reading For Dummies

A Reference for the Rest of Us!

The Fun and Easy Way™
to Figure Out Power Reading
— Completely Updated!

Your First Aid Kit®
for Using Windows 95 with
Internet Explorer

What to Do When
Bad Things Happen
— Explained in
Plain English
United States Crude Death Rates 1900-2000
Survival Manual

Questions to ask of every paper you are willing to read
Quality of Medical Evidence

United States Public Health Task Force
Guide to Clinical Preventive Services
Quality of Medical Evidence

I. Evidence obtained from at least one properly designed randomized controlled trial
Quality of Medical Evidence

II.1 Evidence obtained from well designed controlled trials without randomization.

II.2 Evidence obtained from well designed cohort or case control studies, preferably from more than one center.
II.3 Evidence obtained from multiple time series with or without intervention. Dramatic results in uncontrolled experiments (such as the results of the introduction of penicillin treatments in the 1940s) could also be regarded as this type of evidence.
III. Opinions of respected authorities, based on clinical experience, descriptive studies, case reports, or reports of expert committees
ABSTRACT

Should I Spend My Time Reading This Paper?

States the Purpose of Article, Major Procedures and Methods, Main Findings, and Conclusions

More and More Journals are using Structured Abstracts
Structured Abstract

Objectives

Study Design

Methods

Results

Conclusions
If properly designed and analyzed, is this study, important and worth knowing about?
ABSTRACT, contd

If the results are statistically significant, do they also have clinical significance? If the results are not statistically significant, was the sample size sufficiently large to detect a meaningful difference or effect?
Introduction

Why is this study needed?

What is the purpose of this study?

Was purpose known before the study or a chance finding discovered as part of ‘data dredging?’
Introduction, contd

What has been done before and how does this study differ? (Places study in proper context such as inadequacies of earlier work or next step in an overall research project)

May also be found in DISCUSSION
Introduction, cont'd

Does the location of the study have Relevance (TO ME)?

What is the population to which the study findings apply?
Introduction, contd

Is the time period covered by the study Appropriate (TO ME). Long studies may have informative censoring.

Short studies may not have adequate follow-up time.
Cross Sectional Studies

A snap-shot in time for the study population

Was the sample selected in an appropriate manner (random, convenience, etc)?
Cross Sectional Studies, contd

Were efforts made to ensure a good response rate or to minimize the occurrence of missing data?

Were reliability (reproducibility) and validity reported?
Cohort Studies

Prospective, expensive (Framingham)

Are the subjects representative of the population to which the findings are applied?

Is there evidence of volunteer bias?

Was there adequate follow-up time?

What was the drop-out rate?
Case Control Studies

Retrospective, often few cases, cheap

Were records of cases and controls reviewed blindly?

How were possible selection biases controlled (Prevalence bias, Admission Rate bias, Volunteer bias, Recall bias, Lead Time bias, Detection bias, etc)?
Clinical Trials

Steps to Drug Development

IND (Investigational New Drug License)
Phase I (toxicity)
Phase II (efficacy)

Phase III (comparability)
NDA (New Drug Application-’on-label’)
Phase IV (after market research)
Clinical Trials

Phase I - Does it hurt the Patient?

Phase II - Does it help the Patient?

Phase III - Is it any better?

Phase IV - Does it work in the community?
Phase III Clinical Trials

Are the number of therapy arms appropriate?

Is the choice of controls appropriate (Placebo Arm)?
Phase III Clinical Trials

Were the patients randomized? How? If not, how were patients chosen to avoid selection bias?
Phase III Clinical Trials

If HISTORICAL CONTROLS were used, were the methods and criteria the same for the new experimental group, and were cases and controls compared on prognostic factors?
Phase III Clinical Trials

Was a power analysis performed to estimate required sample sizes?

Were there multiple endpoints? (Data Dredging)

Were subgroups reported and analyzed?
Phase III Clinical Trials

Were repeated measures made over time? If so, were they analyzed properly?

Were there censored (lost to f/u) observations? How was this determined?
Meta Analysis
(Secondary Analysis)

Do the authors specify how the literature review was conducted? Did they make any effort to overcome publication bias? (File Drawer Effect)?
Meta Analysis (Secondary Analysis)

Were the criteria for inclusion and exclusion of studies clearly stated?

If significant findings were determined, did the authors specify the number of additional negative studies that would be needed to eliminate the observed significance?
<table>
<thead>
<tr>
<th>Citation</th>
<th>Year</th>
<th>N</th>
<th>Total</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australian 1</td>
<td>1973</td>
<td>517</td>
<td>.754</td>
<td></td>
</tr>
<tr>
<td>Australian 2</td>
<td>1977</td>
<td>230</td>
<td>.625</td>
<td></td>
</tr>
<tr>
<td>Australian 3</td>
<td>1980</td>
<td>726</td>
<td>.562</td>
<td></td>
</tr>
<tr>
<td>Bassetti</td>
<td>1986</td>
<td>59</td>
<td>.084</td>
<td></td>
</tr>
<tr>
<td>Bassand</td>
<td>1987</td>
<td>107</td>
<td>.571</td>
<td></td>
</tr>
<tr>
<td>Citrius</td>
<td>1986</td>
<td>44</td>
<td>1.100</td>
<td></td>
</tr>
<tr>
<td>Clow</td>
<td>1983</td>
<td>42</td>
<td>.471</td>
<td></td>
</tr>
<tr>
<td>Duxler</td>
<td>1987</td>
<td>94</td>
<td>.566</td>
<td></td>
</tr>
<tr>
<td>European 1</td>
<td>1969</td>
<td>167</td>
<td>1.400</td>
<td></td>
</tr>
<tr>
<td>European 2</td>
<td>1971</td>
<td>730</td>
<td>.635</td>
<td></td>
</tr>
<tr>
<td>European 3</td>
<td>1979</td>
<td>315</td>
<td>.561</td>
<td></td>
</tr>
<tr>
<td>Fletcher</td>
<td>1959</td>
<td>23</td>
<td>.159</td>
<td></td>
</tr>
<tr>
<td>Florida 2</td>
<td>1973</td>
<td>290</td>
<td>.378</td>
<td></td>
</tr>
<tr>
<td>Fitch</td>
<td>1975</td>
<td>189</td>
<td>.959</td>
<td></td>
</tr>
<tr>
<td>GIS S1-1</td>
<td>1986</td>
<td>11772</td>
<td>.807</td>
<td></td>
</tr>
<tr>
<td>Heikinheimo</td>
<td>1971</td>
<td>406</td>
<td>1.248</td>
<td></td>
</tr>
<tr>
<td>ISAM</td>
<td>1986</td>
<td>1741</td>
<td>.872</td>
<td></td>
</tr>
<tr>
<td>IS II</td>
<td>1986</td>
<td>17187</td>
<td>.746</td>
<td></td>
</tr>
<tr>
<td>Italian</td>
<td>1971</td>
<td>321</td>
<td>1.012</td>
<td></td>
</tr>
<tr>
<td>Kennedy</td>
<td>1969</td>
<td>369</td>
<td>.621</td>
<td></td>
</tr>
<tr>
<td>Kiss</td>
<td>1975</td>
<td>23</td>
<td>3.200</td>
<td></td>
</tr>
<tr>
<td>Lusina</td>
<td>1977</td>
<td>24</td>
<td>.222</td>
<td></td>
</tr>
<tr>
<td>N-Gal-Collab</td>
<td>1977</td>
<td>483</td>
<td>1.215</td>
<td></td>
</tr>
<tr>
<td>NHLB-SWIT</td>
<td>1974</td>
<td>107</td>
<td>2.507</td>
<td></td>
</tr>
<tr>
<td>Olson</td>
<td>1986</td>
<td>52</td>
<td>.407</td>
<td></td>
</tr>
<tr>
<td>Senneca</td>
<td>1986</td>
<td>99</td>
<td>.467</td>
<td></td>
</tr>
<tr>
<td>Schmidbauer</td>
<td>1969</td>
<td>36</td>
<td>.296</td>
<td></td>
</tr>
<tr>
<td>S.R-Collab</td>
<td>1976</td>
<td>585</td>
<td>.310</td>
<td></td>
</tr>
<tr>
<td>Valler</td>
<td>1975</td>
<td>91</td>
<td>1.061</td>
<td></td>
</tr>
<tr>
<td>Velay</td>
<td>1986</td>
<td>25</td>
<td>.417</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>1987</td>
<td>219</td>
<td>.159</td>
<td></td>
</tr>
<tr>
<td>Wisselberg</td>
<td>1969</td>
<td>96</td>
<td>.265</td>
<td></td>
</tr>
<tr>
<td>Witchett</td>
<td>1977</td>
<td>58</td>
<td>.778</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fixed</th>
<th>Combined</th>
<th>38674</th>
<th>.768</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>Combined</td>
<td>38674</td>
<td>.762</td>
</tr>
</tbody>
</table>
How were subjects chosen or recruited? If not random, are they representative of the population? (Random selection is not random assignment)
Materials and Methods, contd

Types of Blinding (Masking) Single, Double, Triple.

Is there a control group? How was it chosen?
How are patients followed up? Who are the dropouts? Why and how many are there?

Materials and Methods, contd

Are the independent (predictor) and dependent (outcome) variables in the study clearly identified, defined, and measured?
Materials and Methods, contd

Do the authors explain or reference any unusual methods?

Are statistical methods specified in sufficient detail (If I had access to the raw data, could I reproduce the analysis)?
Materials and Methods, contd

Is there a statement about sample size issues or statistical power (Especially important in negative studies)?

If a multicenter study, what quality assurance measures were employed to obtain consistency across sites?
Materials and Methods, contd

If a study involves human subjects, human tissues, or animals, was approval from appropriate institutional or governmental entities obtained?
Results

Do the results relate to research questions and the purpose of the study?

Do Statistical tests answer the research question?

Are many Statistical tests performed and many comparisons made (Data Dredging)?
Are actual values reported (Means, Standard Deviations, Frequencies, etc) and not just the results of statistical tests?
Results, contd

Are groups similar at baseline? If not, were appropriate adjustments made?

Are informative and appropriate graphics used to present results clearly?
Conclusions/Discussion

Are the questions posed in the study adequately addressed?

Are the conclusions justified by the data?

Do the authors extrapolate beyond the data?

Are shortcomings of the study addressed and constructive suggestions given for future research?
Bibliography/References

Do the citations follow one of the Council of Biological Editors’ (CBE) standard formats?

Several ‘dialects’ exist, but in general, can you find the cited paper or book?
Authors’ Affiliations
(Issues of Scientific Misconduct)

Is the list of contributors reasonable? Thirty authors of a small study is bogus.

Do authors disclose financial relationships for product endorsement, consulting arrangements, etc?