

#EMUG2015

When Authors Aren't to Blame for Bad References

or

How Publishers Inadvertently Make Their Own Content Less Discoverable, and How to Avoid This

Bruce D. Rosenblum, CEO

Inera Incorporated

EMUG, June 19, 2015



INERA
eXtyle | editX

© Inera, Inc. 2003–2015. All Rights Reserved.

What Makes a Reference “Bad”?

- Common author errors include:
 - Incomplete or incorrect reference data
 - Problematic formatting or punctuation
 - Non-standard journal names or abbreviations
 - Violations of editorial style
- Inera’s reference-processing is designed to address all of these problems
- However...

Correct but Non-Standard Data Can Also Lead to “Bad” References

- Tools and processes that have been developed to work with references often falter when faced with idiosyncratic data
- Publishers who adopt non-standard metadata policies can end up inadvertently sweeping their own content under the rug



How Can Accurate Reference Data Be “Bad”?

- Problematic reference data includes non-standard:
 - Journal names
 - Pagination
 - Publication dates
 - DOIs
- Poor curation or distribution of metadata

NFAIS Best Practices

- Best Practices for Publishing Journal Articles
- Published: February 2009
- <https://nfais.memberclicks.net/assets/docs/BestPractices/article%20publishing.pdf>



Best Practices for Journal Titles

- NFAIS Report
 - The metadata for a journal title should be consistent
 - Each journal title should have a unique ISSN
- Additional recommendations
 - Register common aliases
 - Avoid “cute” titles
 - Avoid potentially confusing titles

If You Change a Journal Name, Change the ISSN

- 1977–1992
 - International Journal of Obesity (ISSN 0307-0565)
- 1992–2002:
 - International Journal of Obesity and Related Metabolic Disorders (ISSN 0307-0565)
- 2005–present
 - International Journal of Obesity (ISSN 0307-0565)

Register Common Aliases

- Transportation Research Part F: Traffic Psychology and Behaviour (ISSN 1369-8478)
 - Matches on CrossRef
 - Transportation Research Part F: Traffic Psychology and Behaviour
 - Transp. Res., Part F Traffic Psychol. Behav
 - Fails to link on CrossRef
 - Transportation Research F
 - Transportation Research Part F
- Subtitles are especially troublesome
- Similar name issues confound OpenURL link servers

Don't Be Cute!

International Journal of Radiation Oncology, Biology,
Physics (ISSN 0360-3016)

ELSEVIER

[Home](#) > [Books & Journals](#) > [International Journal o...](#)

International Journal of Radiation
Oncology • Biology • Physics

An official journal of the [American Society for Radiation Oncology](#)



Home Browse My Settings Alerts Help



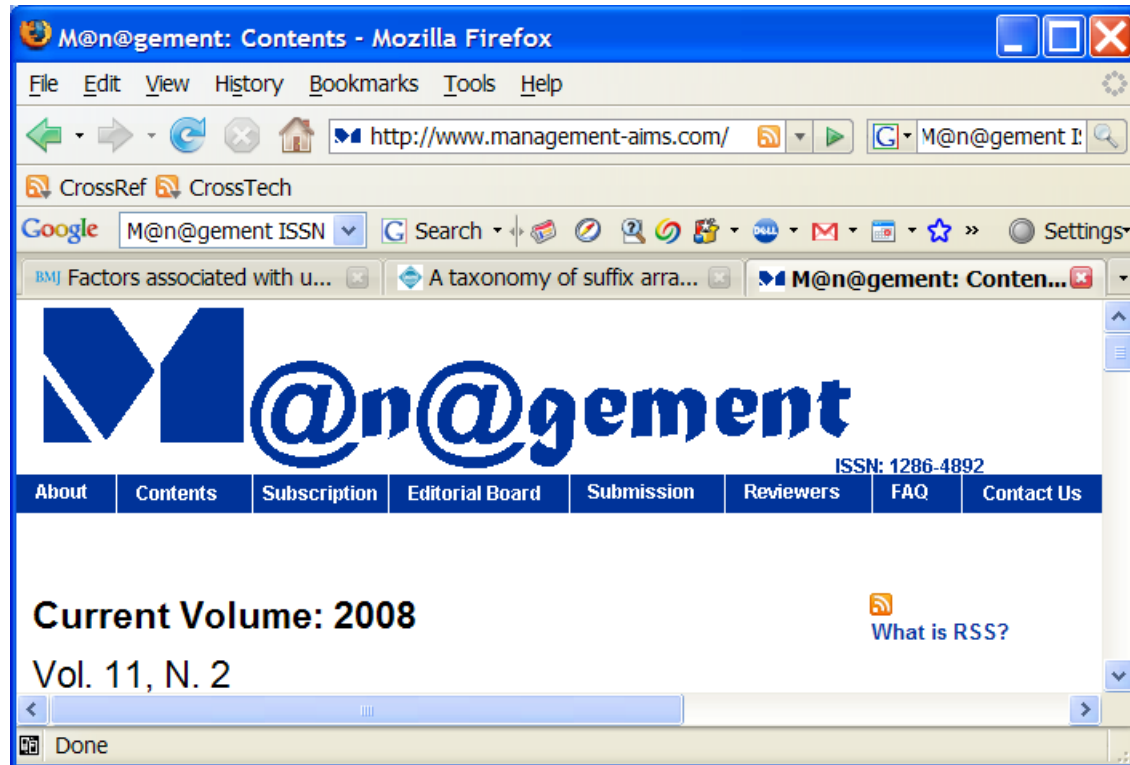
International Journal of Radiation Oncology*Biology*Physics
Copyright © 2008 Elsevier Inc. All rights reserved

[Sample Issue Online](#) | [About this Journal](#) | [Submit your Article](#) | [Shortcut link to this Title](#)

Alert me about new Volumes / Issues

Add to Favorites

Management - Centre de recherche DMSP (ISSN 1286-4692)



Confusing Titles

- Consider this reference:

Benson V, Marano MA. Current estimates from the National Health Interview Survey, 1992. Vital Health Stat 10 1994 Jan;(189):1-269

- The copy editing is correct!
- ***Quick... What is "10"?***



More Confusing Titles

- Consider this reference
 - Levi, P. Cinema by Other Means.
October. Winter 2010, No. 131: 51–68
- The copy editing is correct
- Quick... What is the journal name?

Serial Titles for Reconsideration?

Title	ISSN
January Magazine	1493-2334
March	0702-5149
April	0282-4205
May	2103-7051
June 12	1117-6814
August	1102-7932
October	0162-2870
Spring	0362-0522
Summer	1513-3672



Best Practices for Article Retrieval

- NFAIS Report
 - Assign a DOI to each individual article
 - Follow industry conventions
- Additional recommendations
 - Deposit all requested metadata
 - Follow the intent of the metadata recipient



Follow CrossRef Conventions

- Character problems:
 - doi:10.1002/(SICI)1097-0142(19970215)79:4<835::AID-CNCR21>3.0.CO;2-#
 - [http://dx.doi.org/10.1002/\(SICI\)1097-0142\(19970215\)79:4%3C835::AID-CNCR21%3E3.0.CO;2-%23](http://dx.doi.org/10.1002/(SICI)1097-0142(19970215)79:4%3C835::AID-CNCR21%3E3.0.CO;2-%23)
 - Mis-coding URLs causes failures
- Allowed DOI suffix characters
 - "a-z", "A-Z", "0-9" or "-. _;()/"
 - Published to CrossRef TWG list, July 2, 2008



Deposit All Requested Metadata

- Required for linking:
 - First author, journal title, year, volume, first page
- Even better for linking:
 - Common journal title aliases
 - All authors
 - Article title
 - Issue number
- For extended uses (like eXtyles Reference Cleanup)
 - Organizational authors
 - Translated title
 - Last page

Version Management


- NFAIS Report
 - Exact date of release is a key metadata component
 - Beware unintended consequences




Exact Release Date?

- Consider this reference:
 - Marmor MF, Zrenner E. Standard for clinical electroretinography (1999 update): International Society for Clinical Electrophysiology of Vision. Doc Ophthalmol. **1998-1999**;97:143-156
- What prompted the year range?
 - <MedlineDate>1998-1999</MedlineDate>
 - Download of PubMed data to citation manager

Correct Release Date Is Important

	Standard for clinical electroretinography (1999 update)
Journal	Documenta Ophthalmologica
Publisher	Springer Netherlands
ISSN	0012-4486 (Print) 1573-2022 (Online)
Issue	Volume 97, Number 2 / November, 1998
DOI	10.1023/A:1002016531591
Pages	143-156
Subject Collection	Medicine
SpringerLink Date	Monday, November 29, 2004

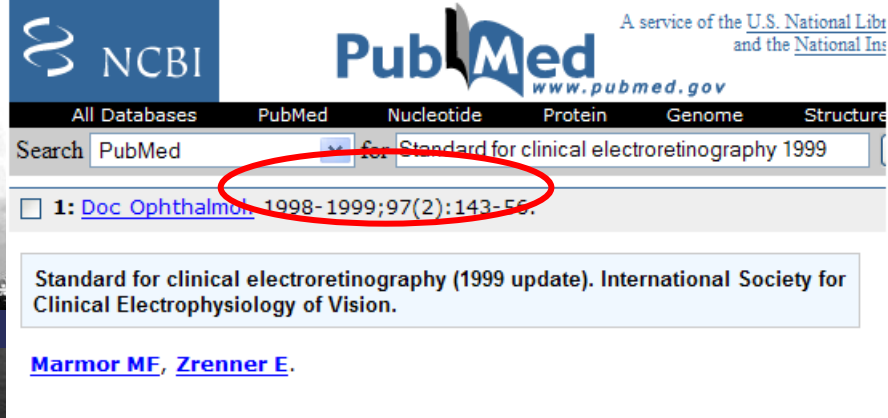
Michael F. Marmor¹ and Eberhart Zrenner²

 *Documenta Ophthalmologica* **97**: 143–156, 1999.
 © 1999 Kluwer Academic Publishers. Printed in the Netherlands.

Standard for clinical electroretinography (1999 update)

MICHAEL F. MARMOR & EBERHART ZRENNER

Accepted 11 June 1999



NCBI PubMed A service of the U.S. National Library of Medicine and the National Institutes of Health
 www.pubmed.gov

All Databases PubMed Nucleotide Protein Genome Structure

Search PubMed for Standard for clinical electroretinography 1999

1: [Doc Ophthalmol.](#) 1998-1999;97(2):143-56.

Standard for clinical electroretinography (1999 update). International Society for Clinical Electrophysiology of Vision.

[Marmor MF](#), [Zrenner E](#).



Beware Unintended Consequences

- Or how about this publication date:
 - Ullrich A, Coussens L, Hayflick JS, Dull TJ, Gray A, Tam AW, Lee J, Yarden Y, Libermann TA, Schlessinger J, et al. Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. Nature **1984 May 31-Jun 6**; 309:418-25
- What prompted the date range?
 - <MedlineDate>1984 May 31-Jun 6</MedlineDate>
 - Download of PubMed data to citation manager

Cover Date Consequences

- “It is our policy to cite the date published on the cover of the journal issue in which an article appears. I have the hardcopy of *Nature*, volume 309, issue 5967 in hand; the published cover date shows May 31-Jun 6 1984”
- *Fran Spina (NLM)*
- “Nature's policy has always ever since its first issue in 1869 been that the article is published on that one day, ie the 1201 hrs on the Thursday of the week of publication”
- *Maxine Clarke (Nature)*

Retractions

- NFAIS Report
 - Issue a prominent notice
 - If article is removed, some notice should be left
 - Register CrossMark record, even for older retractions
- Additional recommendations
 - Electronic publication does not mean ephemeral publication
 - Bits are harder to burn than paper



No, You Can't Make It Disappear

- “Question: Have you ever had an editor that decides that even though an article is online ahead of print he/she does not want to compile it in an issue, ever? And better still, he/she wants the paper removed from the Internet all together and hopes the world with think it was never published with the journal.”
 - *Query to CrossRef Technical Working Group, 24 October 2008*
- One Answer: “Unless there are legal reasons (court order), you probably should not completely remove the paper. Librarians (and others) frown on this very much, especially if the paper has been out there for five months.”
 - *Reply to CrossRef Technical Working Group, 24 October 2008*



Bits Propagate

- From a real contract:
 - “Customer hereby consents to allow Service Provider to include the work product in Service Provider's normal backup processes, and understands that this may produce non-erasable copies of the work”
- Consider *any* electronic document to be a permanent record



Indication of Length

- NFAIS Report
 - Use page ranges when articles have sequential pagination
 - Use total number of pages when article IDs are used



It's Four Pages

- Consider this publication:

BMJ medical publication of the year

Research **FREE** Education News Comment Topics Archive About BMJ

Home > Research > BMJ 2008;336(7647):754 (5 April), doi:10.1136/bmj.39489.590671.25

This Article
Abstract **FREE**
Abridged PDF
PDF
All Versions of this Article:
336/7647/754 *most recent*
bmj.39489.590671.25v1
Respond to this article
Read responses to this article

BMJ 2008;336:754-757 (5 April), doi:10.1136/bmj.39489.590671.25
(published 28 February 2008)

Research

Factors associated with uptake of measles, mumps, and rubella vaccine (MMR) and use of single antigen vaccines in a contemporary UK cohort: prospective cohort study

Anna Pearce, *research fellow*¹, Catherine Law, *professor*¹, David Elliman, *consultant*², Tim J Cole, *professor*¹, Helen Bedford, *senior lecturer*¹, the Millennium Cohort Study Child Health Group



No, It's Seven Pages

#EMUG2015

Downloaded from bmj.com on 24 November 2008

BMJ

RESEARCH

Factors associated with uptake of measles, mumps, and rubella vaccine (MMR) and use of single antigen vaccines in a contemporary UK cohort: prospective cohort study

Anna Pearce, research fellow,¹ Catherine Law, professor,¹ David Elliman, consultant,² Tim J Cole, professor,¹ Helen Bedford, senior lecturer,¹ the Millennium Cohort Study Child Health Group

13 Sharland M, Atkinson P, Maguire H, Begg N. Lone parent families are an independent risk factor for lower rates of childhood immunisation in London. *Commun Dis Rep CDR Rev* 1997;7:R196-72.

England and Wales 1995-2006* by quarter. www.hpa.org.uk/infections/topics_az/measles/data_not_confirmed.htm.

Accepted: 28 January 2008

BMJ | ONLINE FIRST | bmj.com

page 7 of 7

INERA
eXtyle | edify

© Inera, Inc. 2003–2015. All Rights Reserved.

Problems

- How to cite the full vs. abbreviated PDF?
 - Specifying version?
 - Referring to pages?

<p>This article is an abridged version of a paper that was published on bmj.com on 28 February 2008. Cite this article as: <i>BMJ</i> 28 February 2008, doi: 10.1136/bmj.39489.590671.25</p> <p>754</p>	<p>with MIMR made Conclusions Alth substantial prop avoidable infecti decide not to im could be used to uptake.</p>
--	---



One Publisher's Response: Drop Last Page From Citations

“From the article page, note that the citation with page info refers to the print edition (confirm by calling up the abridged PDF). The online version, which is the definitive version, is longer; the PDF is numbered starting from page 1. Our practice... is to consider the full (online) version as our source, but the PubMed citation does not reflect this. Hence our decision to provide only the first page (we also would provide the DOI in a parenthetical at the end of the citation).”

Article Identifiers

- NFAIS Report
 - Must contain 6 or more alpha-numeric characters
 - Avoid ambiguity with volume, issue, pagination
 - Avoid single leading or trailing letter, especially “S”
 - Don’t number each volume from article 1 to n
 - Must not contain punctuation
 - Avoid confusion with ISSN and other identifiers
- Additional recommendations
 - Discourage “unusual” citation styles

Article IDs: Don't Number 1 to n

- Consider this reference:
 - Foulds J, Furberg H. Is low-nicotine Marlboro snus really snus? Harm Reduct J. 2008;5:9
- What is “9”?
 - Issue number, page number, or article ID?
- Fails to link on CrossRef if misidentified
 - May match PubMed; depends on lookup method

Discourage “Unusual” Page Styles

- Don't turn article IDs into page ranges
 - Puglisi, S. J., Smyth, W. F., and Turpin, A. H. 2007. A taxonomy of suffix array construction algorithms. *ACM Comput. Surv.* 39, 2, pp. 4:1--4:31 (June 2007)
 - Yoichi Kizaki, Yuji Noguchi, and Masaru Miyayama, "Defect control for low leakage current in $K_{0.5}Na_{0.5}NbO_3$ single crystals," *Applied Physics Letters*, vol. 89, 142910-1-142910-3, October 2006

"this way lies madness"

Comment to CrossRef Technical Working Group, 10 September 2007

More Wonky “Pagination”

- Jiang L, Lee VJ, Lim WY, Chen MI, Chen Y, Tan L, Lin RT, Leo YS, Barr I, Cook AR. Performance of case definitions for influenza surveillance. *Euro Surveill.* 2015;20(22):pii=21145
- Vandamme, J., Volkel, P., Rosnoblet, C., Le Faou, P. and Angrand, P. O. (2011). Interaction proteomics analysis of polycomb proteins defines distinct PRC1 complexes in mammalian cells. *Mol. Cell Proteomics* 10, M110.002642
- Yamamoto, Y.; Ohno, Y.; Maehashi, K.; Matsumoto, K. Noise reduction of carbon nanotube field-effect transistor biosensors by alternating current measurement. *Jpn. J. Appl. Phys.* 2009, 48, 06FJ01:1–06FJ01:4
- Kalpouzos, G., Persson, J., & Nyberg, L. (2012). Local brain atrophy accounts for functional activity differences in normal aging. *Neurobiol. Aging*, 33, 623.e1-623.e13
- Van der Auwera GA, Carneiro M, Hartl C, et al. From FastQ data to high-confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 2013;43:11.10.1–11.10.33.
- S. Ballandras, A. Reinhardt, A. Khelif, M. Wilm, V. Laude, W. Daniau, V. Blondeau-Pâtissier, “Theoretical analysis of damping effects of guided elastic waves at solid/fluid interfaces”, *Journal of Applied Physics*, Vol.99, pp.054907:1–9, 2006.

Some Closing Thoughts

- Use rational and unambiguous numbers
- Check on how you're indexed
 - Indexers make mistakes
 - Test: Can you find your content on services your readers use?
- Proof your metadata carefully *before* publication
 - You may find unfortunate results



Use Unambiguous Numbers

- Do not use arbitrary volume numbers
- Biochim Biophys Acta
 - Volume 1: 1947
 - Volume 54: 1961
 - Volume 500: 1977
 - Volume 1000: 1989
 - Volume 1786: 2008
- Consider this possible reference in a few years:
 - Tanel A, Averill-Bates DA. The aldehyde acrolein induces apoptosis via activation of the mitochondrial pathway. Biochim Biophys Acta 2018;2019:255-267

Is This Too Complicated?

- <http://www.ohsu.edu/library/ejournals/bba2revised.shtml>

[OHSU Library](#)

BBA Decoder

Use this chart when you have a reference to a known Biochimica et Biophysica Acta (BBA) journal article. You can then know which **section** it's in. The online version of Biochimica et Biophysica Acta (BBA) is divided up between the 25 different section Web pages, not in volume sequence. *document!* Current volumes are at the top. For older volumes you can save sort to:

- ▶ [Volume 1499 \(2000\)](#)
- ▶ [Volume 1000 \(1989\)](#)
- ▶ [Volume 500 \(1977\)](#)

Many thanks to the [UCSD Library](#) for the initial formatting of this page.

Volume no	Year	BBA Section
1835-1836	2013	Reviews on Cancer
1834	2013	Proteins & Proteomics
1833	2013	Molecular Cell Research
1832	2013	Molecular Basis of Disease
1831	2013	Molecular and Cell Biology of Lipids
1830	2013	General Subjects
1829	2013	Gene Regulatory Mechanisms
1828	2013	Biomembranes

1863	BBA (no section)
	Specialized Section on Nucleic Acids and Related Subjects
	BBA (no section)
	Specialized Section on Nucleic Acids and Related Subjects
1861	BBA (no section)



Indexers Make Mistakes

- Published vs. indexed title and group author

JAMA & ARCHIVES Select Journal or Resource

JAMA[®]

The Journal of the American Medical Association

Vol. 282 No. 18, November 10, 1999 [TABLE OF CONTENTS >](#)

Original Contribution

Validation and Utility of a Self-report Version of PRIME-MD

The PHQ Primary Care Study

Robert L. Spitzer, MD; Kurt Kroenke, MD; Janet B. W. Williams, DSW; and the Patient Health Questionnaire Primary Care Study Group

JAMA. 1999;282:1737-1744.

NCBI PubMed A service of the U.S. National Library and the National Institutes of Health www.pubmed.gov

All Databases PubMed Nucleotide Protein Genome Structure

1: [JAMA](#). 1999 Nov 10;282(18):1737-44.

Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire.

[Spitzer RL](#), [Kroenke K](#), [Williams JB](#).



Carefully Proof Your Metadata

- What's wrong with this title?
 - Risk Strataification of Brugada Syndrome Revisited
- Unfortunate result – PubMed's entry:

PubMed A service of the U.S. National Library of Medicine and the National Institutes of Health
www.pubmed.gov

PubMed Nucleotide Protein Genome Structure OMIM PM

1: [Isr Med Assoc J. 2008 Jun;10\(6\):462-4.](#)

IMAJ Free Full Text

Erratum in:

- ♦ [Isr Med Assoc J. 2008 Jul;10\(7\):549.](#)

Risk stratification [corrected] of Brugada syndrome revisited.

[Nof E. Antzelevitch C.](#)

Summary

- Round one of e-pub experimentation is over
 - Some successes; some failures
- Print is dead; long live print!
 - Can't discard 500 years of print-based metaphors
 - Even online-only must retain some vestiges of print
- Essential to find e-content reliably
 - Follow best practices
 - Ensure metadata accuracy



#EMUG2015

Thank You!

Bruce Rosenblum
Inera Incorporated
+1 (617) 932 - 1932
brosenblum@inera.com
www.inera.com

