

# KJA Experience on Similarity Index *including data mining*

Younsuk Lee

ylee@dongguk.edu

<sup>1</sup>Korean Journal of Anesthesiology

<sup>2</sup>Dept. of Anesthesiology and Pain Medicine  
DUMC, Goyang, Dongguk University

<sup>3</sup>DM Statistical Institute  
Goyang, Dongguk University

1-Jul-2014

KAMJE 2014 Workshop



## Table of Contents

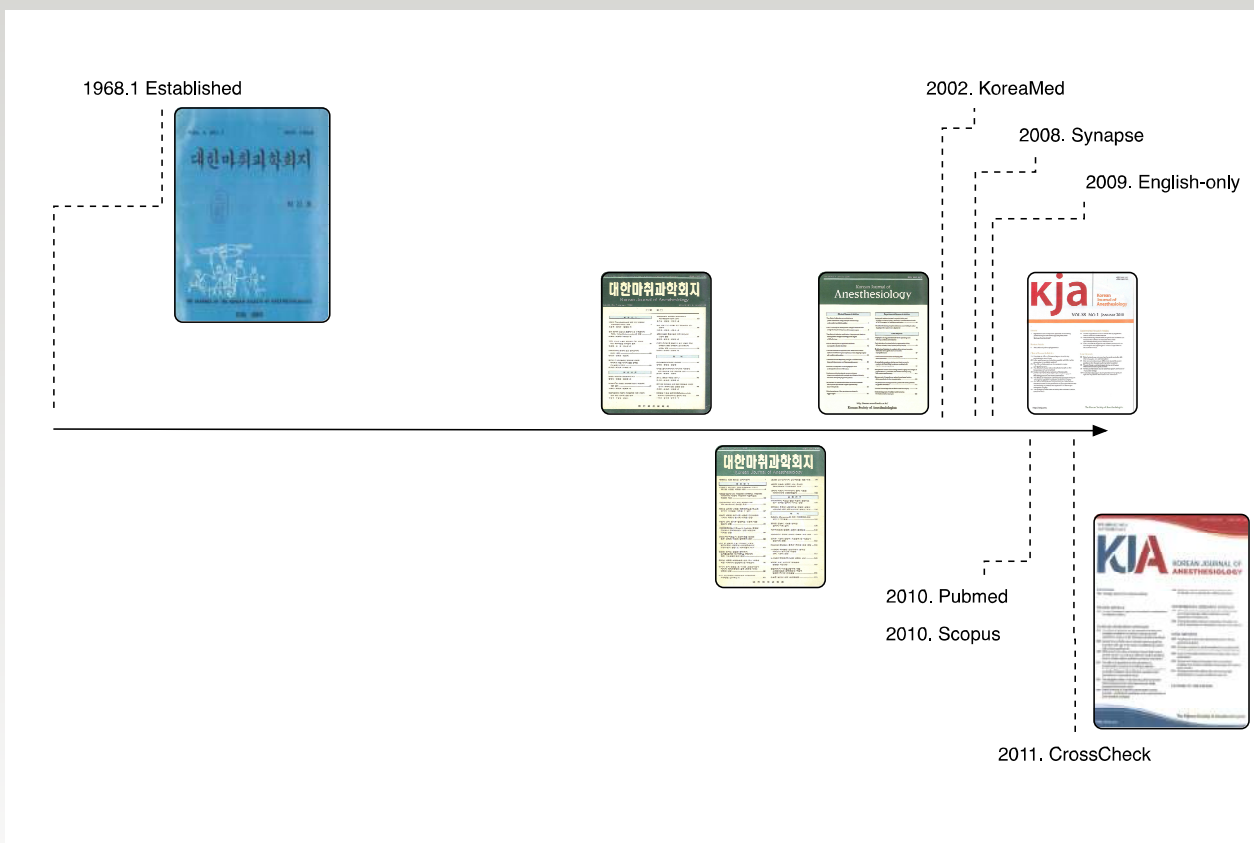
At glance

Data mining

Summary



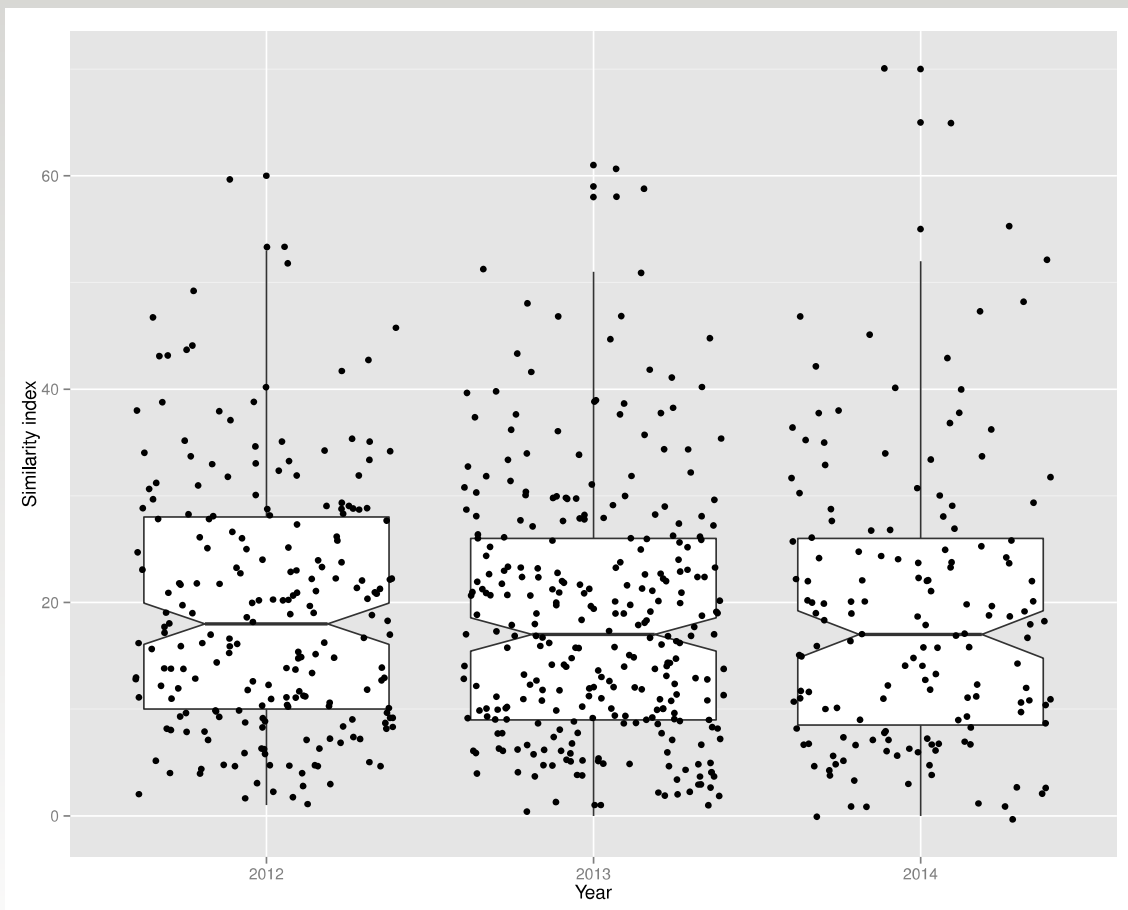
# Brief History



YL

KJA

# Constant perennially I

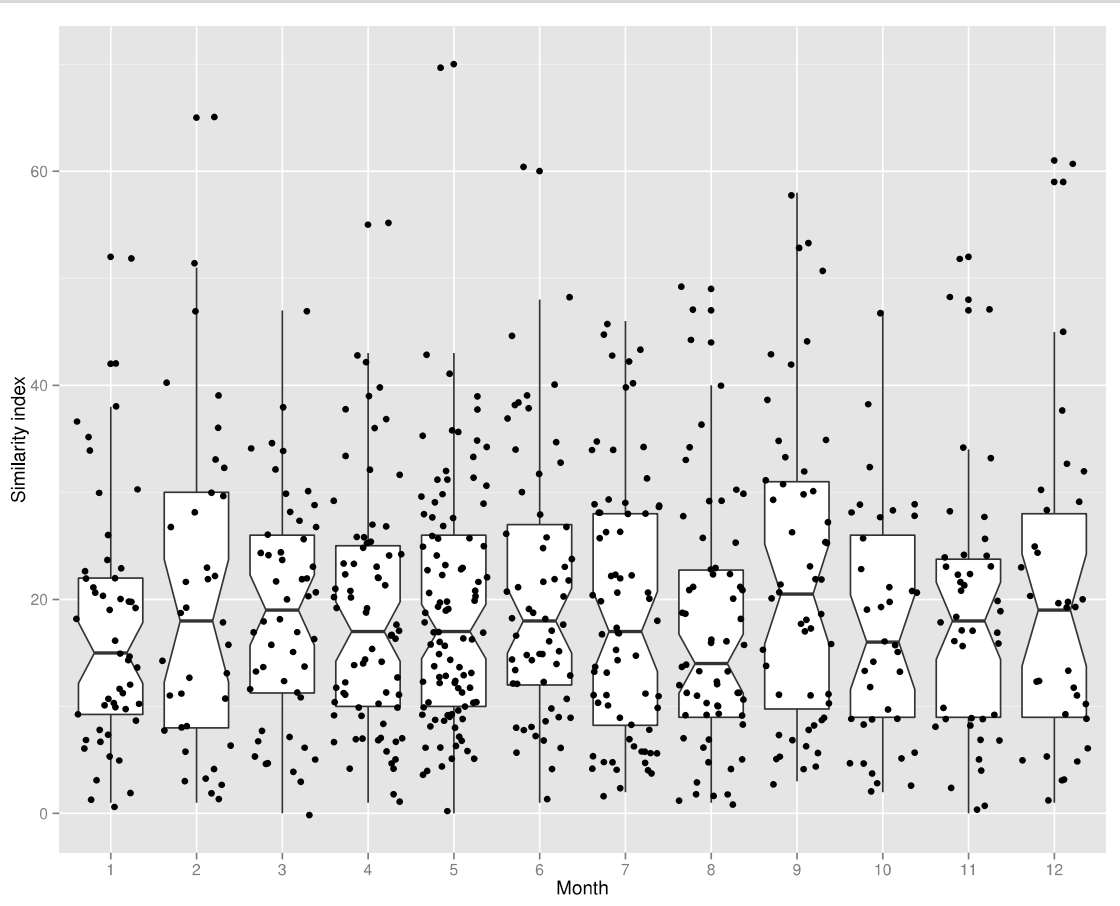


YL

KJA



# Constant perennially II



YL

KJA

At glance

mining

Summary

## Paper Source

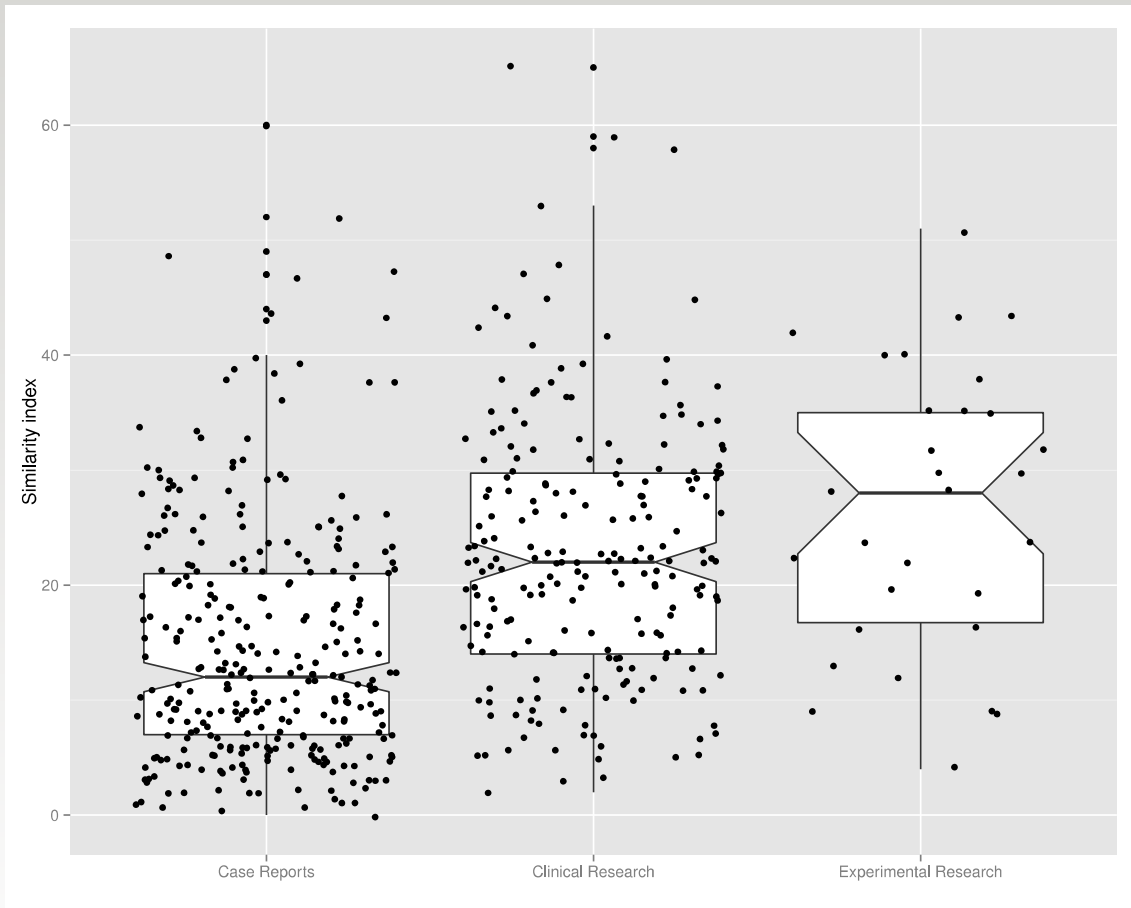
2010.7–2014.6

Paper Source	Submitted (n)
Albania, Brazil, Chile, France, Hong Kong, Indonesia, Italy, Jordan, Pakistan, Poland, Saudi Arabia, Thailand	1
Croatia, Germany, Israel, Sweden	2
Lao PDR, Nepal, Singapore	3
Belarus, Canada, Greece, Malaysia, Spain, Tunisia	4
China	5
Egypt	6
US	9
Japan	28
India	33
Turkey	34
Iran	43
Korea	1471

YL

KJA

# Paper Type



YL

KJA

# PLOS ALM (article-level metrics)

**PLOS** ARTICLE-LEVEL METRICS

[Home](#) [Overview](#) [Altmetrics](#) [Developers](#)

## PLOS Article-Level Metrics (ALMs): measuring the impact of research

ALMs provide a suite of established metrics that measure the overall performance and reach of published research articles.

[Learn More!](#) [Get the App](#) [Get the Data](#)

**Researchers**

Maximize the Impact of your research

**Publishers**

Enhance publication value through real-time views of reach and influence

**Institutions**

Capture researcher impact for hiring, tenure, and promotion decisions

**Funders**

Track the performance and impact of research funding decisions

### Videos and Presentations

- Finding Insights in ALMs for Research Evaluation**  
Charleston Conference, 2013, Richard Cave
- Article-Level Metrics at PLOS & Beyond**  
SPARC Webcast, Apr 2012, Peter Binfield, Publisher of PLoS ONE, PLoS
- Metrics: the New Black?**  
NFAIS 2012, Kristin Ratan, Director of Products, PLoS

[MORE](#)

### Latest News

- 30 July, 2013  
The Summer issue of Information Standards Quarterly (ISQ), NISO's print and electronic magazine for communicating standards-based technology and...
- 2 July, 2013  
We are pleased to announce that ALM Workshop 2013 will be held on October 10-12 in San...
- 18 June, 2013  
Today, we are delighted to announce the release of ALM Reports which allow you to view and download ALMs...
- 11 Jun, 2013  
New paper just published: PlosOpenR Exploring FP7 Funded PLOS Publications PLOS's Martin Fenner in collaborating with Najko Jahn and Jochen...
- 16 May, 2013  
The San Francisco Declaration on Research Assessment (DORA): Join PLOS and a long list of publishers, funders,...

[MORE](#)

YL

KJA



# PLOS ALM (article-level metrics)

10-year Dataset Retrieved at 2013.5.17

R Data Editor

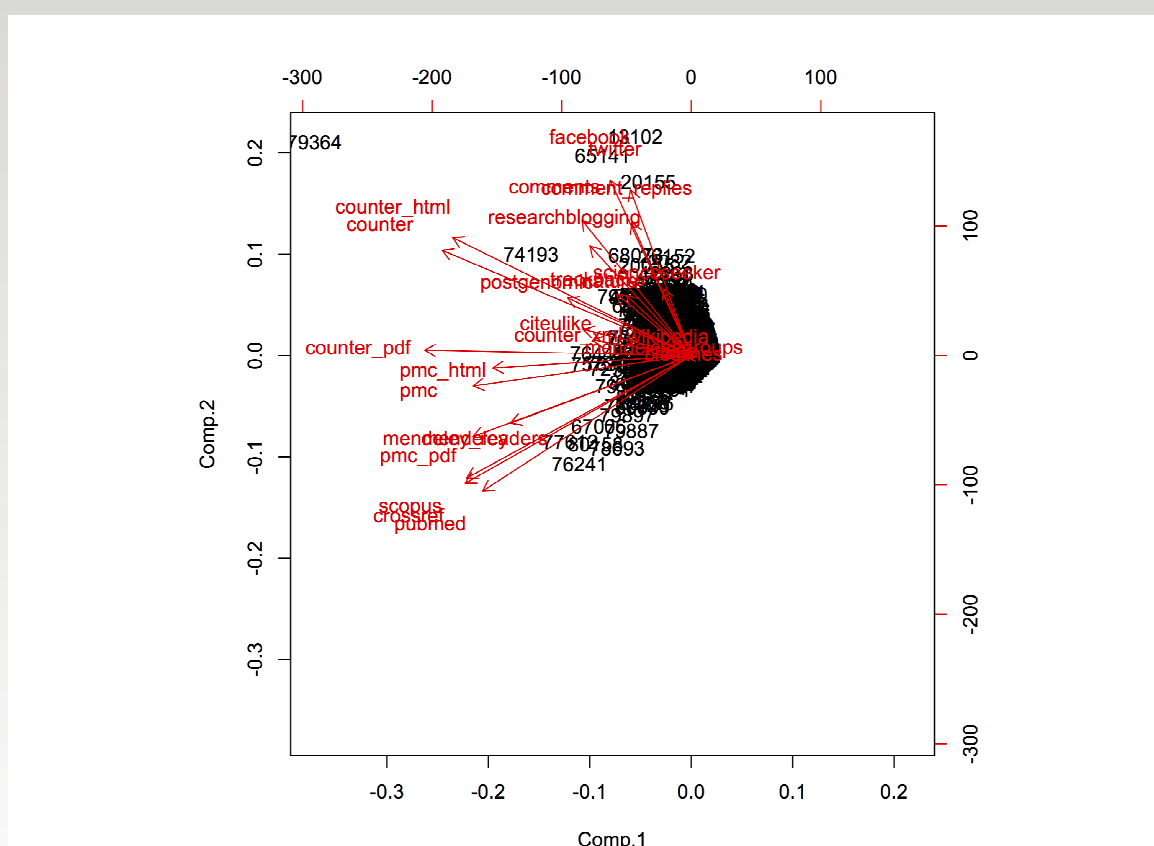
title	bioid	citeulike	crossref	scopus	pubmed	researchblogging	nature	bioglines	postgenomic	facebook	twitter	wikipedia	scienceseeker	counter_html
s of Patients with Multiple Sclerosis...	0	0	0	0	0	0	0	0	0	0	0	0	0	184
Delays DNA Damage-Induced Casp...	0	0	0	0	0	0	0	0	0	0	0	0	0	80
edly Rational Stock Markets and I...	0	0	0	0	0	0	0	0	0	0	0	0	0	171
Recognition Approach to Distinguis...	0	0	0	0	0	0	0	0	0	1	0	0	0	210
ing (CBS): A Novel, Efficient and Cas...	0	0	0	0	0	0	0	0	0	0	0	0	0	220
Free Mass Standard Deviation Score...	0	0	0	0	0	0	0	0	0	0	0	0	0	198
MT of Gastric Cancer Cells through...	0	0	0	0	0	0	0	0	0	0	0	0	0	213
terventions on Appropriateness of...	0	0	0	0	0	0	0	0	0	0	0	0	0	202
iChIP $\alpha$ , an Explorative Phylogenet...	0	0	0	0	0	0	0	0	0	0	0	0	0	216
cal Antenna Regulating Human Adi...	0	0	0	0	0	0	0	0	0	0	0	0	0	187
on of SH2 Peptide Interactions Fro...	0	0	0	0	0	0	0	0	0	0	0	0	0	183
peals on Fear Arousal and Driver Be...	0	0	0	0	0	0	0	0	0	0	5	0	0	185
l Cancer Cell Growth and Metastasi...	0	0	0	0	0	0	0	0	0	0	0	0	0	120
g and Fruiting Trails in Olive	0	0	0	0	0	0	0	0	0	0	0	0	0	112
lowing in Vitro Stimulation with <it...	0	0	0	0	0	0	0	0	0	0	0	0	0	61
7 pH Changes on Intracellular pH an...	0	0	0	0	0	0	0	0	0	0	0	0	0	68
und Xanthohrysin Defines a New Gr...	0	0	0	0	0	0	0	0	0	0	0	0	0	98
T Xon 11 Deletions, and Gastroint...	0	0	0	0	0	0	0	0	0	0	0	0	0	107
spontaneous Calcium Release in Car...	0	0	0	0	0	0	0	0	0	0	0	0	0	83
</italic> Methylation Patterns in So...	0	0	0	0	0	0	0	0	0	0	0	0	0	114
and Intrauterine and Perinatal Envi...	0	0	0	0	0	0	0	0	0	0	0	0	0	64
roRNAs Regulating Proliferation an...	0	0	0	0	0	0	0	0	0	0	0	0	0	219
Satellite Telemetry Locations Using...	0	0	0	0	0	0	0	0	0	0	0	0	0	178
rmine Invasion Success? A Compari...	0	0	0	0	0	0	0	0	0	5	0	0	0	170
ceptor Tyrosine Kinase Pathway by T...	0	0	0	0	0	0	0	0	0	0	0	0	0	155
ritis: Characteristics and Outcomes...	0	0	0	0	0	0	0	0	0	0	0	0	0	153
3 Treatment by Gene Expression Pr...	0	0	0	0	0	0	0	0	0	0	0	0	0	176
Hydroxy-2-Hexenal-Induced Infla...	0	0	0	0	0	0	0	0	0	0	0	0	0	172
dicts the Time Depth of the World's...	0	0	0	0	0	0	0	0	0	31	0	0	0	287
aggy Flux through Lysosomes	0	0	0	0	0	0	0	0	0	0	0	0	0	172
Engineered with Amniotic Epitheli...	0	0	0	0	0	0	0	0	0	0	0	0	0	293
n of COPD: A Population Based Cas...	0	0	0	0	0	0	0	0	0	0	0	0	0	131
Diffusible Resource Prevents Coexi...	0	0	0	0	0	0	0	0	0	1	0	0	0	195
ance: Enhancing Drug Game	0	0	0	0	0	0	0	0	0	0	3	0	0	175
<sup>7</sup></sup>Sca-1</sup>+ </sup>+ </sup>...	0	0	0	0	0	0	0	0	0	0	0	0	0	127
del of Tuberculosis	0	0	0	0	0	0	0	0	0	0	0	0	0	143
Fetal Lung and Skin/Amnion (but...	0	0	0	0	0	0	0	0	0	0	0	0	0	175
MAP3B&C Homologous Epitopes are...	0	0	0	0	0	0	0	0	0	0	0	0	0	133
Yall Lipids from the Pathogenic Pha...	0	0	0	0	0	0	0	0	0	0	0	0	0	138
ination Analysis of Tomato Spotte...	0	0	0	0	0	0	0	0	0	0	0	0	0	174
y Pathways Inhibition in <italic>Rhi...	0	0	0	0	0	0	0	0	0	0	0	0	0	112
cs of Fin Whales <italic>Balaenopt...	0	0	0	0	0	0	0	0	0	0	0	0	0	113
5-Rox Gene from Rirch <italic>Ret...	0	0	0	0	0	0	0	0	0	0	0	0	0	171
sion Membrane Proteins in Inclusio...	0	0	0	0	0	0	0	0	0	0	0	0	0	169
t HIV Infection and Late HIV Diagn...	0	0	0	0	0	0	0	0	0	23	1	0	0	149
Contributes to Podocyte Damage i...	0	0	0	0	0	0	0	0	0	0	0	0	0	148
a Scaling Based Function	0	0	0	0	0	0	0	0	0	0	1	0	0	156
Immune Reconstitution Inflammat...	0	0	0	0	0	0	0	0	0	0	0	0	0	129
Response to HIV-1 Core-24 by In...	0	0	0	0	0	0	0	0	0	0	0	0	0	222

YL

KJA

# PLOS ALM (article-level metrics)

10-year Dataset Retrieved at 2013.5.17

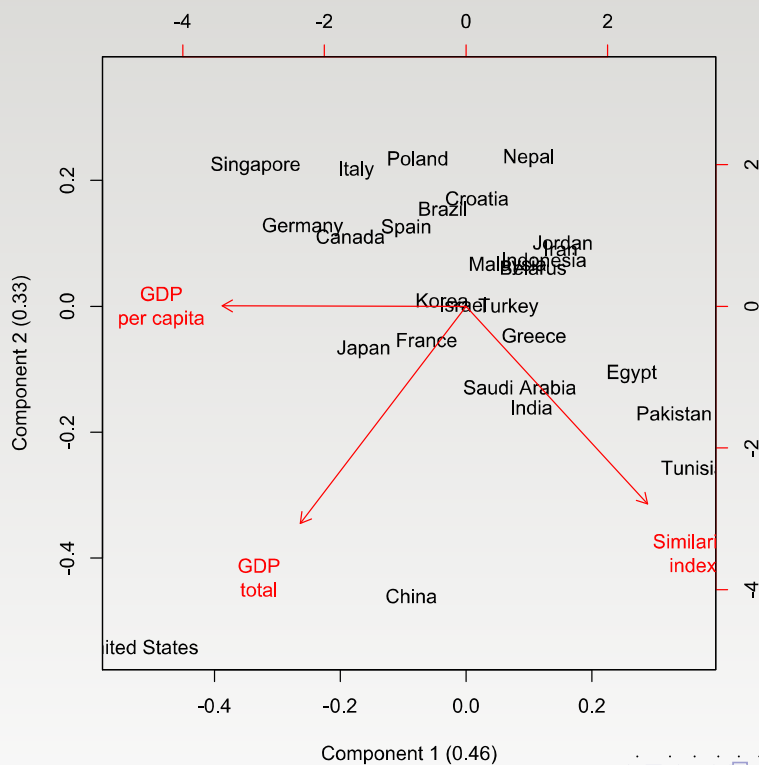


YL

KJA

# Paper Source Country

Principal component analysis with those GDP

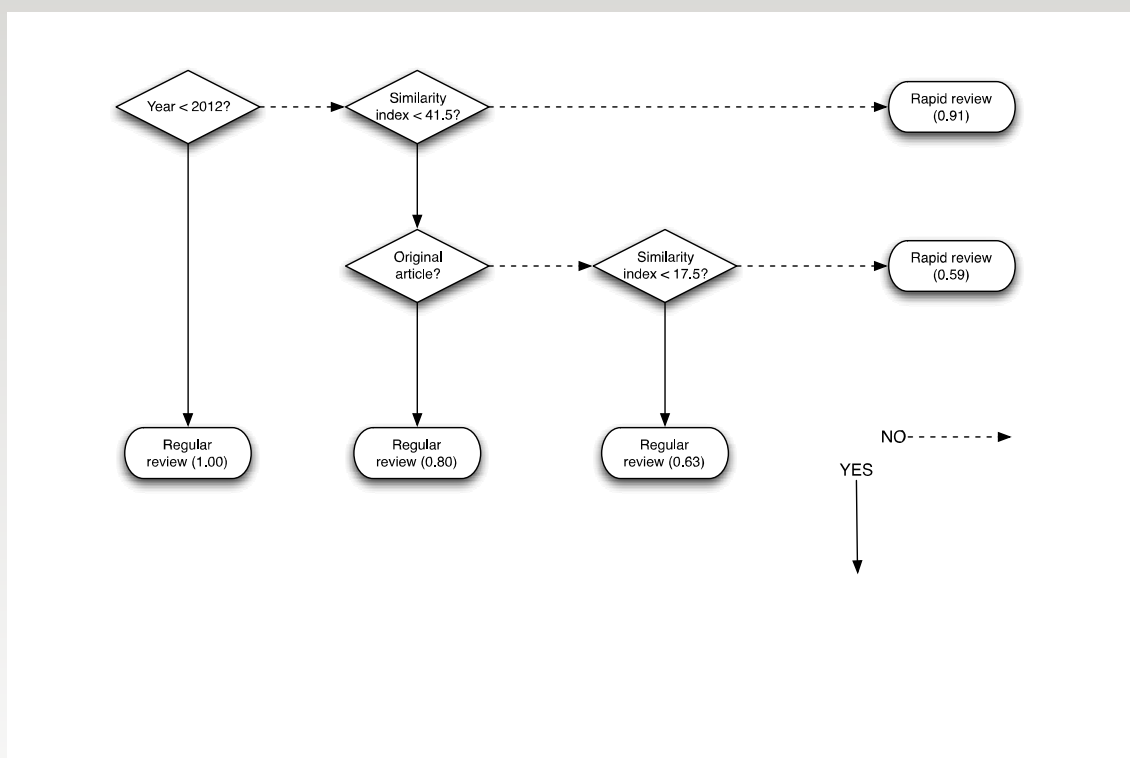


YL

KJA

# Rapid or Regular

Recursive partitioning modeling

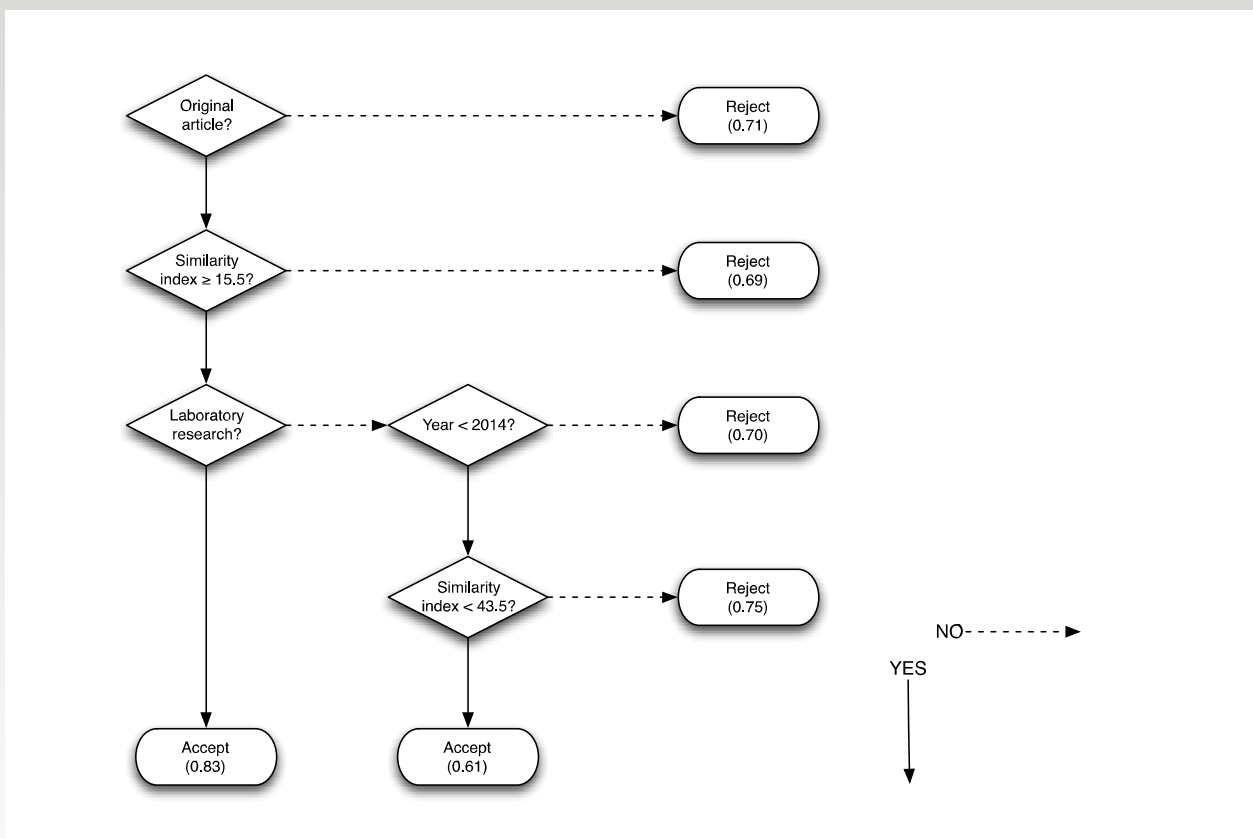


YL

KJA

# Accept or Reject

## Recursive partitioning modeling



YL



KJA

At glance

mining

Summary

# Long titles are associate with higher citations

Farrokh and Yadollahie (2010). Are Shorter Article Titles More Attractive for Citations? Cross-sectional Study of 22 Scientific Journals. *Croat Med J* 51: 165-70.

## Are Shorter Article Titles More Attractive for Citations? Cross-sectional Study of 22 Scientific Journals

Farrokh Habibzadeh,  
Mahboobeh Yadollahie  
Shiraz NOC Medical Education and  
Research Center, Shiraz, Iran

**Aim** To investigate the correlation between the length of the title of a scientific article and the number of citations it receives, in view of the common editorial call for shorter titles.

**Methods** Title and the number of citations to all articles published in 2005 in 22 arbitrarily chosen English-language journals ( $n=9031$ ) were retrieved from citation database Scopus. The 2008 journal impact factors of these 22 journals were also retrieved from Thomson Reuters' Journal Citation Report (JCR). Assuming the article title length as the independent variable, and the number of citations to the article as the dependent variable, a linear regression model was applied.

**Results** The slope of the regression line for some journals ( $n=4$ , when titles were measured in characters but 7 when titles were measured in words) was negative – none was significantly different from 0. The overall slope for all journals was 0.140 (when titles were measured in characters) and 0.778 (when titles were measured in words), which is significantly different from 0 ( $P < 0.001$ ,  $t$  test). Overall, articles with longer titles received more citations – Spearman  $\rho = 0.266$  – when titles were measured in characters, and  $\rho = 0.244$  when titles were measured in words ( $P < 0.001$ ). This association was found for 7 of 8 journals with impact factor  $> 10$  and for 2 out of 14 journals with impact factor  $< 10$  ( $P < 0.001$ , Fisher exact test).

**Conclusion** Longer titles seem to be associated with higher citation rates. This association is more pronounced for journals with high impact factors. Editors who insist on brief and concise titles should perhaps update the guidelines for authors of their journals and have more flexibility regarding the length of the title.

Received: January 7, 2010  
Accepted: March 24, 2010

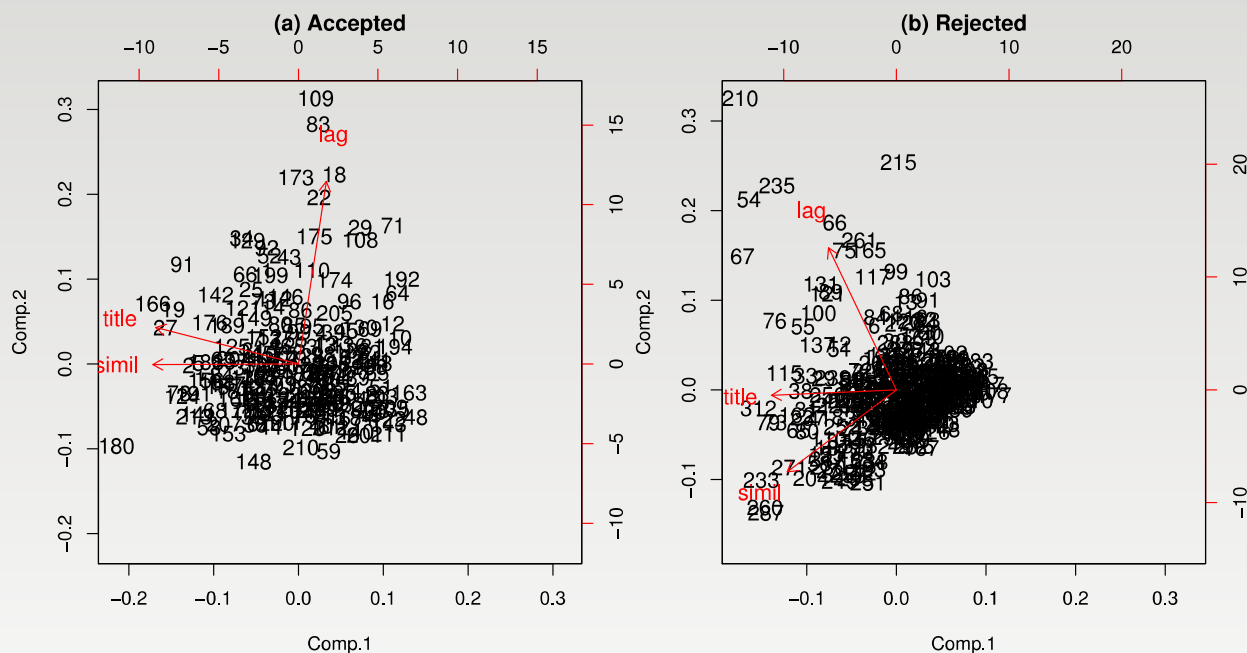
YL



KJA

# Length of Title and Acceptance

## Principial component analysis



YL

KJA

At glance

mining

Summary

# Title with colon?

Jamali and Nikzad (2011). Article title type and its relation with the number of downloads and citations. *Scientometrics* 88: 653-61.

### Colon, downloads and citations

There were 325 titles with and 1847 titles without colon. Statistics showed that titles with colon tend to be longer ( $Mdn = 10$ ) than titles without colon ( $Mdn = 9$ ;  $U = 222427$ ,  $P < .001$ ;  $r = .16$ ). Titles with colon also received fewer citations ( $Mdn = 9$ ) compared to titles without colon ( $Mdn = 12$ ;  $U = 267253$ ,  $P = .012$ ;  $r = .054$ ).

Table 6 Man-Whitney U tests results for presence of colon paired with other variables

	Median		$U$	$P$	$r$
	Without Colon	With Colon			
Title Length	9	10	222427	.000	.16
Times Cited	12	9	267253	.012	.047
Downloads	2705	2777	283916	.120	.033



YL

KJA

## Titles containing Question mark and Acceptance

$P = 0.4251$  using *Fisher's exact test*

	Question mark	
	Without	With
Accept	209	4
Reject	323	11



YL

KJA

## Titles containing Colon and Acceptance

$P=0.9279$  using  $\chi^2$  test

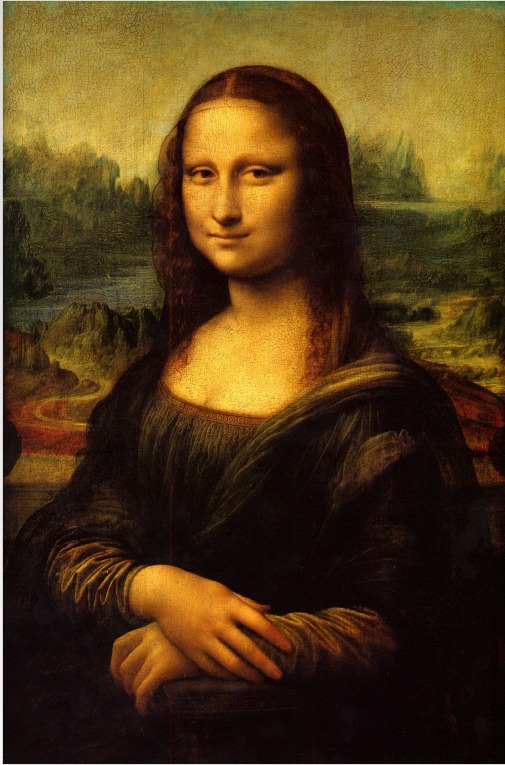
	Colon	
	Without	With
Accept	180	33
Reject	280	54



YL

KJA

# How similar?



At glance

mining

Summary

## Summary

Anyone who wants to submit an article into KJA *SHOULD*:

- ▶ avoid a long title.
- ▶ come from the OECD counry.
- ▶ control your similarity index tightly within 15.5 – 41.5%.

