

## „PubMed Commons“: Beschreibung und Analyse von PubMeds neuer Kommentarfunktion

Köhler, Norbert

Abteilung für Medizinische Psychologie und Medizinische Soziologie – Universitätsmedizin Leipzig

### Abstract

Mit „PubMed Commons“ hat die weltweit größte biomedizinische Literaturdatenbank seit einigen Monaten eine Kommentarfunktion, die von Wissenschaftlern genutzt werden kann, deren Publikationen bei PubMed gelistet sind. In diesem Beitrag erfolgt eine Beschreibung und formale Analyse der bisher veröffentlichten Kommentare.

Der analysierte Datensatz enthält alle Kommentare, die im Zeitraum Juni 2013 bis Mai 2014 verfasst wurden, das Geschlecht der Kommentar-Autoren sowie das Publikationsjahr der kommentierten Zeitschriftenartikel und die Namen der Zeitschriften. Außerdem wurde überprüft, wie viele Kommentare eine Frage bzw. einen Verweis auf weiterführende Literatur enthalten.

Von 460 verschiedenen Autoren wurden insgesamt 1.358 Kommentare verfasst, wobei gut 90% der Beiträge einmal kommentiert wurden. Knapp 30% aller Kommentare wurden von 10 Kommentar-Autoren verfasst und die Hälfte aller kommentierten Beiträge wurde in den Jahren 2013 und 2014 publiziert. Seit der öffentlichen Freigabe von PubMed Commons ist die Anzahl der monatlich verfassten Kommentare rückläufig. Ein Viertel aller Kommentare enthielt eine Frage und in zwei Dritteln fanden sich Verweise auf weiterführende Literatur. Knapp 90% der Kommentar-Autoren waren männlich.

Da in PubMed nahezu alle in der Biomedizin relevanten Zeitschriften gelistet sind, hat PubMed Commons das Potential, verlags- und zeitschriftenübergreifend eine stärkere Vernetzung der „scientific community“ zu bewirken. Bisher hat PubMed Commons jedoch kaum einen interaktiven Charakter.



„PubMed Commons“:  
Description and Analysis of PubMed's new commenting system

Köhler, Norbert

Abteilung für Medizinische Psychologie und Medizinische Soziologie – Universitätsmedizin Leipzig

**Abstract**

With „PubMed Commons“, the world's largest biomedical literature database has recently launched a forum for commenting on publications. PubMed Commons is open to any authors in PubMed. The purpose of this investigation is to describe and analyse all comments that have been published so far.

The analysed dataset comprises all comments written between June 2013 und May 2014, the forum users' gender, the year the commented papers were published and the journal names. Furthermore, the number of comments containing a question or further references was assessed.

In total, 1.358 comments were published by 460 forum users. 90% of all papers were commented once, 30% of all comments were written by 10 users and half of the commented papers were published in 2013 and 2014. Since PubMed Commons was officially launched, the number of monthly published comments has continuously declined. 1/4 of comments contained a question and about 2/3 of comments contained further references. 90% of the forum users were male.

Since nearly any relevant biomedical journal is listed in PubMed, PubMed Commons has the potential to improve the networking of the worldwide scientific community regardless of publishing house or journal. However, PubMed Commons has hardly become interactive so far.



## Einleitung

Mit etwa 23 Millionen Zitationen ist PubMed die umfangreichste biomedizinische Literaturdatenbank der Welt [1]. PubMed wird durch das „National Center for Biotechnology Information“ (NCBI) betrieben und bietet einen kostenfreien Zugang zu den Datenbanken MEDLINE und PubMed Central.

Vor einigen Monaten hat das NCBI PubMed um eine Kommentarfunktion erweitert, die den Namen „PubMed Commons“ trägt. Nachdem die Kommentarfunktion in den Monaten Juni bis September 2013 in einer Alpha-Version von knapp 100 Wissenschaftlern getestet wurde, veröffentlichte die NCBI im Oktober 2013 eine Pilot- bzw. Beta-Version von PubMed Commons. In dieser Phase wurde eine größere Zahl Wissenschaftler einiger US-amerikanischer Universitäten und Forschungseinrichtungen eingeladen, die Kommentarfunktion zu nutzen und zu testen. Die Kommentare waren in der Pilotphase nur für registrierte Teilnehmer sichtbar [2,3]. Wie eine Analyse der ersten 200 Kommentare zeigte, wurde in etwa einem Drittel der Kommentare Kritik geäußert oder auf andere Studien verwiesen. Außerdem wurde die Kommentarfunktion genutzt, um auf Kontextinformationen sowie auf neue, relevante Studien aufmerksam zu machen [4].

Ende Dezember 2013 wurde die Pilotphase abgeschlossen. Seitdem haben alle Wissenschaftler, die Autor mindestens einer bei PubMed gelisteten Publikation sind und die sich bei PubMed als Nutzer registriert haben, die Möglichkeit, sowohl die eigenen als auch die Publikationen anderer Autoren zu kommentieren. Die Kommentare sowie die Namen der Kommentar-Autoren sind für alle Besucher der PubMed-Seiten sichtbar, auch wenn diese nicht bei PubMed registriert sind.

Da soziale Medien in den letzten Jahren auch in der Wissenschaft an Bedeutung gewonnen haben und sich dieser Trend auch in Zukunft fortsetzen wird [5], erscheint es notwendig, genauere Kenntnisse über die Nutzer und die Art der Nutzung dieser Medien zu gewinnen.

Im Rahmen dieser Arbeit erfolgt eine Beschreibung und formale Analyse der Kommentare sowie der Nutzer von PubMed Commons. Eine Inhaltsanalyse der Kommentare kann im Rahmen dieser Arbeit nicht geleistet werden. Vielmehr stehen die folgenden Fragestellungen im Vordergrund:

1. Welche Trends gibt es hinsichtlich der Anzahl der Kommentare im Zeitverlauf?
2. Welche formalen Merkmale weisen die Kommentare auf?
3. Gibt es Unterschiede im Kommentarverhalten weiblicher und männlicher Nutzer der Kommentarfunktion?

## Material und Methoden

### Erstellung des Datensatzes

Der dieser Arbeit zugrundeliegende Datensatz enthält alle Kommentare, die zwischen Juni 2013 und dem 20. Mai 2014 geschrieben wurden. Zur Erstellung des Datensatzes wurde zunächst über das PubMed Suchformular (<http://www.ncbi.nlm.nih.gov/pubmed/>) mit Hilfe des Suchstrings `has_user_comments[sb]` eine CSV-Datei mit allen Zitationen, die mindestens einmal kommentiert wurden, erstellt. Diese Datei enthält neben der PubMed ID (PMID) den Namen der Publikation und der Zeitschrift, die Namen der Autoren sowie das Jahr der Veröffentlichung. In einer zweiten Datei (MS Excel) wurden die folgenden Angaben gespeichert: die Anzahl der Kommentare, die Namen der Kommentar-Autoren sowie das Datum, an dem der Beitrag erstellt wurde. Beide Dateien wurden mit Hilfe des Statistikprogramms GNU R importiert und anhand der PMID verknüpft. Das Geschlecht der kommentierenden Personen wurde anhand der Namen ermittelt.

Der so erstellte Datensatz liegt im Format RData vor und kann zusammen mit einem R-Syntax-File auf der Startseite des Artikels<sup>1</sup> herunter geladen werden.

### Analyse

Statistische Analysen wurden mit den Softwarepaketen Microsoft Excel 2003 und GNU R 3.02 [6] durchgeführt. Für die Auszählung der Worthäufigkeiten wurde das R-Paket „qdap“ benutzt [7]. Mittelwertsunterschiede in Hinblick auf den Umfang der Kommentare (Anzahl der Wörter) in Abhängigkeit vom Geschlecht des Nutzers sowie von der Art des Kommentars (Frage, Verweis auf Literatur und andere Quellen) wurden mit dem t-Test für unabhängige Stichproben berechnet. Zusammenhänge zwischen dem Geschlecht der Nutzer und der Art der Kommentare wurden mit dem Chi<sup>2</sup>-Test überprüft. Als statistisch signifikant gelten Zusammenhänge mit einem p-Wert < 0.05.

<sup>1</sup> <http://www.bibliometrie-pf.de/article/view/204>





## Ergebnisse

### Anzahl der Artikel, Kommentare und Kommentar-Autoren

Insgesamt wurden im Zeitraum Juni 2013 bis 20.05.2014 1358 Kommentare geschrieben, die sich auf 1170 verschiedene Artikel bezogen. Dabei wurden 1068 Artikel (91%) einmal, 69 Artikel (6%) zweimal, 15 Artikel (1%) dreimal und 18 Artikel (2%) vier bis dreizehn Mal kommentiert.

Mit 896 Artikeln (77%) wurde die übergroße Anzahl aller kommentierten Artikel in den Jahren 2010 bis 2014 geschrieben, wobei 608 Artikel (52%) aus den Jahren 2013 und 2014 stammten. Weitere 220 Artikel (19%) wurden

in den Jahren 2000 bis 2009 verfasst und noch einmal 54 Artikel (5%) stammen aus den Jahren 1956 bis 1999.

Die veröffentlichten Kommentare wurden von insgesamt 460 verschiedenen Autoren verfasst. Knapp zwei Drittel der Nutzer (N=286) schrieben nicht mehr als einen Kommentar, ein weiteres Drittel veröffentlichte zwischen zwei und neun Kommentare (N=152) und knapp fünf Prozent der Nutzer (N=22) verfassten mehr als zehn Kommentare. Knapp 30% aller Kommentare (N=395) wurden von insgesamt zehn Nutzern verfasst. Diese „Vielschreiber“ veröffentlichten zwischen 18 und 112 Kommentare.

Abb. 1: Anzahl der Kommentare im Zeitverlauf

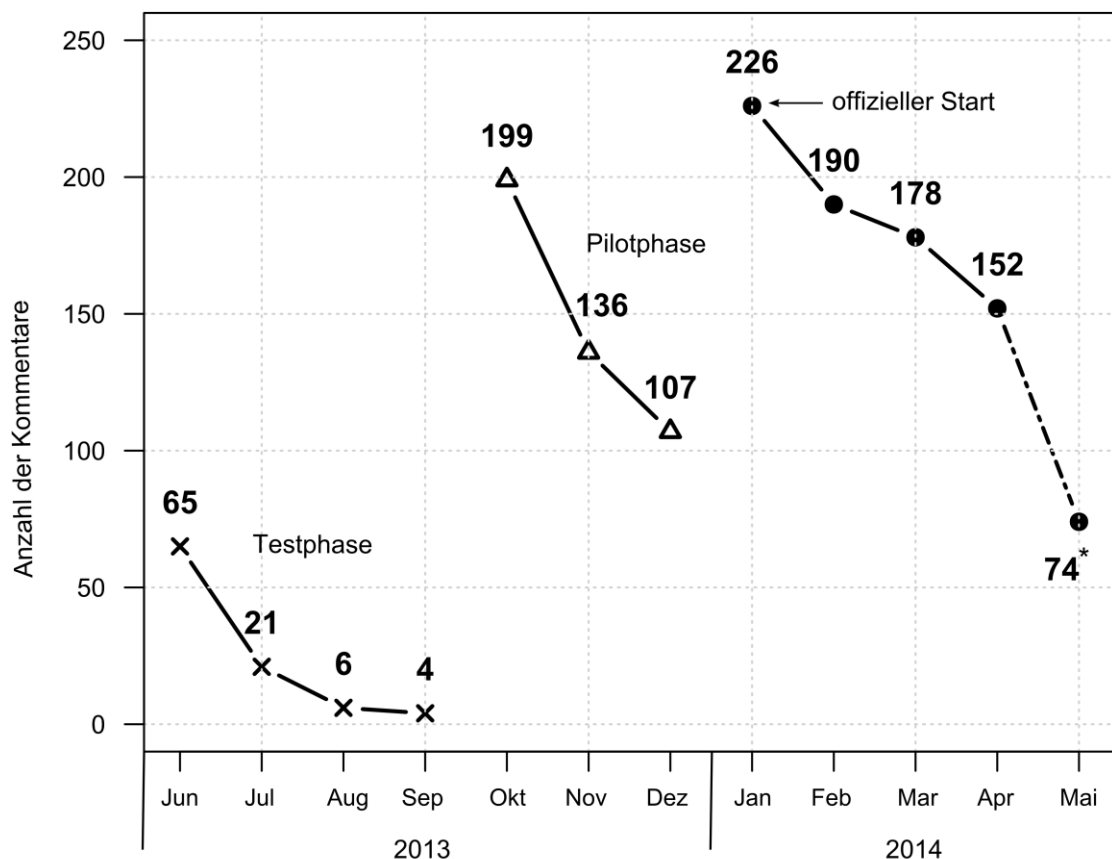


Abbildung 1: \*nur für den Zeitraum 01.05.2014 - 20.06.2014

Wie Abb. 1 zeigt, wurden 96 Kommentare (7%) in der Testphase und 442 Kommentare (33%) in der Pilotphase verfasst. Nach der offiziellen Freigabe der Kommentarfunktion kamen weitere 820 Kommentare (60%) hinzu.

In den fünf Monaten nach der offiziellen Freigabe der Kommentarfunktion sank die Anzahl der monatlich veröffentlichten Kommentare kontinuierlich. Wurden im Januar 2014 pro Tag im Durchschnitt 7,3 Kommentare veröffentlicht, waren es im Mai 2014 mit täglich 3,7 Kommentaren nur noch etwa halb so viele.



### Formale Merkmale der Kommentare

Die Kommentare umfassten im Durchschnitt 183 Wörter, was in etwa einer halben A4-Seite entspricht (Schriftgröße 12). Der längste Kommentar beinhaltete 1.235 Wörter und der kürzeste Kommentar bestand aus nur einem Wort (Median=84, SD=225).

Insgesamt 329 Kommentare (24%) enthielten eine Frage und in etwa zwei Drittel der Kommentare (N=846) fanden sich Verweise auf weiterführende Literatur.

### Geschlechtsunterschiede

Von den insgesamt 460 Kommentar-Autoren konnten 405 (88%) als männlich und 50 (11%) als weiblich identifiziert werden. Bei weiteren 5 Autoren (1%) war eine Zuordnung des Geschlechts nicht möglich. Betrachtet man nur den Zeitraum nach der offiziellen Freigabe von PubMed Commons, zeigt sich eine nahezu identische Geschlechterverteilung (12% Frauen, 88% Männer).

Hinsichtlich der Anzahl verfasster Kommentare gab es zwischen männlichen und weiblichen Nutzern keine statistisch signifikanten Unterschiede (t-Test für unabhängige Stichproben,  $p=0,703$ ). Sowohl männliche als auch weibliche Nutzer schrieben im Mittel etwa drei Kommentare (Median: jeweils ein Kommentar). Mit durchschnittlich 188 Wörtern umfassten die Kommentare weiblicher Nutzer 30 Wörter weniger als die Kommentare männlicher Nutzer. Zwar ist dieser Unterschied statistisch signifikant (t-Test für unabhängige Stichproben,  $p=0,016$ ), aber kaum bedeutsam (40 Wörter entsprechen gut drei Zeilen auf einer A4-Seite). Hinsichtlich des Anteils von Kommentaren, die eine Frage enthielten, gab es keine statistisch signifikanten Unterschiede zwischen weiblichen und männlichen Nutzern (Chi-squared=2,79, df=1,  $p=0,095$ ).

### Zeitschriften und Themengebiete

Die insgesamt 1.170 kommentierten wissenschaftlichen Zeitschriftenbeiträge wurden in 613 verschiedenen Zeitschriften publiziert. Aus 438 Zeitschriften (72%) wurde jeweils ein Beitrag kommentiert, aus 97 Zeitschriften (16%) jeweils zwei Beiträge, aus 70 Zeitschriften (11%) drei bis neun Beiträge und aus acht Zeitschriften (1%) wurden mehr als 10 Beiträge kommentiert (Plos One = 53, New England Journal of Medicine = 40, Proceedings of the National Academy of Sciences of the USA = 34, Science = 31, Nature = 24, Journal of the American Medical Association = 17, Lancet = 15, Journal of the American Medical Association Internal Medicine = 13).

Darüber hinaus zeigte eine Auszählung der Worthäufigkeiten der Zeitschriftentitel, dass schwerpunktmäßig Beiträge aus Zeitschriften mit den folgenden fachlichen Ausrichtungen kommentiert wurden: Biologie (inkl. Molekular- und Zellbiologie; N=128 kommentierte Beiträge), Neurologie (N=70), Innere Medizin (N=52), Psychologie und Psychiatrie (N=48), Kardiologie (N=46) und Onkologie (N=44).

### Diskussion

Mit PubMed Commons gibt es seit Januar 2014 die Möglichkeit, in der biomedizinischen Literaturdatenbank PubMed gelistete Einträge zu kommentieren. Allerdings steht die Kommentarfunktion nur Wissenschaftlern offen, die Autor mindestens einer bei PubMed gelisteten Publikation sind. Der interessierten Öffentlichkeit, z.B. Studenten, Patientenvertretern und Wissenschaftsjournalisten, steht die Kommentarfunktion hingegen nicht zur Verfügung [2].

Ähnlich wie verschiedene Wissenschaftsverlage (z.B. Springer, PLOS, BioMed Central), zahlreiche Wissenschaftsblogs und andere soziale Medien, bietet PubMed Commons Wissenschaftlern weltweit die Möglichkeit eines „post-publication peer review“ [8], in dessen Rahmen aktuelle Publikationen ohne zeitlichen Verzug diskutiert werden können [5]. Anders als die eben genannten Medien, ermöglicht PubMed Commons jedoch einen schnellen Zugriff auf die Zitationen (fast) aller im Bereich der Biomedizin relevanten Fachzeitschriften, was zumindest potentiell eine stärkere Vernetzung der „scientific Community“ begünstigt.

Die Analyse der Kommentare zeigte zum einen, dass in den ersten Monaten ein großes Interesse bestand, die Kommentarfunktion zu nutzen. Es bleibt jedoch abzuwarten, ob sich der kontinuierliche Rückgang der veröffentlichten Kommentare zwischen Januar und Mai 2014 fortsetzen wird. Außerdem konnte gezeigt werden, dass das Gros der kommentierten wissenschaftlichen Beiträge (>90%) nicht mehr als einmal kommentiert wurde. PubMed Commons hat demnach bisher kaum einen interaktiven Charakter entwickeln können. Der hohe Anteil von Kommentaren, die eine Frage beinhalten, zeigt jedoch, dass viele Kommentare zumindest auf Interaktivität ausgerichtet sind.

Auffällig ist weiterhin, dass zwar fast 90% aller Nutzer männlich waren, es aber hinsichtlich der Anzahl und der Art der geschriebenen Kommentare (Fragen, Verweise) keine nennenswerten Unterschiede zwischen weiblichen und männlichen Nutzern gab. Relevante Ge-





schlechterunterschiede in der Art des wissenschaftlichen Diskurses scheint es demzufolge in diesem Zusammenhang nicht zu geben.

Der große Anteil von Kommentaren, in denen auf weiterführende Fachliteratur oder Datenbanken verwiesen wird, deutet an, dass es lohnenswert sein kann, PubMed Commons zu nutzen. In diesem Zusammenhang kann es nützlich sein, eine Benachrichtigungsfunktion einzurichten [9]. Des Weiteren könnte PubMed Commons dazu genutzt werden, kritische Reflexionen, die im Rahmen von Journal Clubs entstehen, einer breiteren fachlichen Öffentlichkeit zur Verfügung zu stellen.

### Literaturverzeichnis

- <sup>1</sup> U.S. Department of Health & Human Services. Meet PubMed Commons: The new comments forum in PubMed (19.12.2013). Im Internet: <http://pubmedcommonsblog.ncbi.nlm.nih.gov/2013/12/19/meet-pubmed-commons-the-new-comments-forum-in-pubmed>; Stand: 22.06.2014
- <sup>2</sup> Tibshirani R. PubMed Commons: A system for commenting on articles in PubMed (23.10.2013). Im Internet: <http://andrewgelman.com/2013/10/23/pubmed-commons-system-commenting-articles-pubmed>; Stand: 22.06.2014
- <sup>3</sup> U.S. Department of Health & Human Services. PubMed Commons: A New Forum for Scientific Discourse (22.10.2013). Im Internet: <http://ncbiinsights.ncbi.nlm.nih.gov/2013/10/22/pubmed-commons-a-new-forum-for-scientific-discourse>; Stand: 22.06.2014
- <sup>4</sup> U.S. Department of Health & Human Services. Early Developments in the PubMed Commons Pilot (01.11.2013). Im Internet: <http://ncbiinsights.ncbi.nlm.nih.gov/2013/11/01/early-developments-in-the-pubmed-commons-pilot>; Stand: 22.06.2014
- <sup>5</sup> Bik HM, Goldstein MC. An Introduction to Social Media for Scientists. PLoS Biol 2013; 11: e1001535
- <sup>6</sup> R Core Team. R: A Language and Environment for Statistical Computing. Wien: The R Foundation for Statistical Computing; 2013
- <sup>7</sup> Rinker T. qdap: Bridging the gap between qualitative data and quantitative analysis (15.06.2013). Im Internet: <http://cran.r-project.org/web/packages/qdap/index.html>; Stand: 22.06.2014

<sup>8</sup> Oransky I. PubMed now allows comments on abstracts — but only by a select few (22.10.2013). Im Internet: <http://retractionwatch.com/2013/10/22/pubmed-now-allows-comments-on-abstracts-but-only-by-a-select-few>; Stand: 22.06.2014

<sup>9</sup> U.S. Department of Health & Human Services. Comment search and alert: A PubMed Commons guide (13.12.2013). Im Internet: <http://pubmedcommonsblog.ncbi.nlm.nih.gov/2013/12/13/comment-search-and-alert-a-pubmed-commons-guide>; Stand: 22.06.2014