CWTS

Citation analysis may severely underestimate the impact of clinical research as compared to basic research

Nees Jan van Eck, Ludo Waltman, Anthony F.J. van Raan, Robert J.M. Klautz, and Wilco C. Peul

CWTS Working Paper Series

Paper number	CWTS-WP-2012-012
Publication date	October 4, 2012
Number of pages	13
Email address corresponding author	ecknjpvan@cwts.leidenuniv.nl
Address CWTS	Centre for Science and Technology Studies (CWTS) Leiden University P.O. Box 905 2300 AX Leiden The Netherlands www.cwts.leidenuniv.nl

Universiteit Leiden

Centrum voor Wetenschaps- en Technologie Studies

Citation analysis may severely underestimate the impact of clinical research as compared to basic research

Nees Jan van Eck[1], Ludo Waltman[1], Anthony F.J. van Raan[1], Robert J.M. Klautz[2], and Wilco C. Peul[3]

[1] Leiden University, Centre for Science and Technology Studies, Leiden, The Netherlands
{ecknjpvan, waltmanlr, vanraan}@cwts.leidenuniv.nl

[2] Leiden University Medical Center, Department of Cardiothoracic Surgery, Leiden, The Netherlands
r.j.m.klautz@lumc.nl

[3] Leiden University Medical Center, Department of Neurosurgery, Leiden, The Netherlands
w.c.peul@lumc.nl

Background

Citation analysis has become an important tool for research performance assessment in the medical sciences. However, different areas of medical research may have considerably different citation practices, even within the same medical field. Because of this, it is unclear to what extent citation-based bibliometric indicators allow for valid comparisons between research units active in different areas of medical research.

Methodology

A visualization methodology is introduced that reveals differences in citation practices between medical research areas. The methodology extracts terms from the titles and abstracts of a large collection of publications and uses these terms to visualize the structure of a medical field and to indicate how research areas within this field differ from each other in their average citation impact.

Results

Visualizations are provided for 32 medical fields, defined based on journal subject categories in the Web of Science database. The analysis focuses on three fields: *Cardiac & cardiovascular systems*, *Clinical neurology*, and *Surgery*. In each of these fields, there turn out to be large differences in citation practices between research areas. In particular, it is found that clinical research is generally cited less frequently than basic and non-interventional research.

Conclusions

Popular bibliometric indicators, such as the *h*-index and the impact factor, do not correct for differences in citation practices between medical fields. These indicators therefore cannot be used to make accurate between-field comparisons. More sophisticated bibliometric indicators do correct for field differences but still fail to take into account within-field heterogeneity in citation practices. As a consequence, the impact of clinical research may be substantially underestimated in comparison with basic research.

1. Introduction

Citation analysis is widely used in the assessment of research performance in the medical sciences (Patel et al., 2011). Especially the *h*-index (Hirsch, 2005) and the impact factor (Chew et al., 2007; Garfield, 1996, 2006) are extremely popular bibliometric indicators. However, the use of these indicators for performance assessment has important limitations. In particular, both the *h*-index and the impact factor fail to take into account the enormous differences in citation practices between fields of science (e.g., Radicchi, Fortunato, & Castellano, 2008). For instance, the average length of the reference list of a publication is much larger in molecular biology than in mathematics. As a consequence, publications in molecular biology on average are cited much more frequently than publications in mathematics. This difference can be more than an order of magnitude (Waltman, Van Eck, Van Leeuwen, Visser, & Van Raan, 2011a).

More sophisticated bibliometric indicators used by professional bibliometric centers perform a normalization to correct for differences in citation practices between fields of science (e.g., Glänzel, Thijs, Schubert, & Debackere, 2009; Waltman, Van Eck, Van Leeuwen, Visser, & Van Raan, 2011b). These field-normalized indicators typically rely on a field classification system in which the boundaries of fields are explicitly defined (e.g., the journal subject categories in the Web of Science database). Unfortunately, however, practical applications of field-normalized indicators often suggest the existence of differences in citation practices not only between but also within fields of science. As shown in this paper, this phenomenon can be observed especially clearly in medical fields, in which clinical research generally turns out to be cited substantially less frequently than basic and non-interventional research. Within-field heterogeneity in citation practices is not corrected for by field-normalized bibliometric indicators and therefore poses a serious threat to the accuracy of these indicators.

This paper presents an empirical analysis of the above problem, with a focus on the medical sciences. An advanced visualization methodology is used to show how citation practices may differ between research areas within a medical field. In particular, large differences between basic and clinical research areas are revealed. Implications of the analysis for the use of bibliometric indicators in the medical sciences are discussed.

2. Methodology

The approach taken in this paper to analyze differences in citation practices between areas of medical research is based on the idea of visualizing scientific fields using term maps (e.g., Van Eck & Waltman, 2011; Waaijer, Van Bochove, & Van Eck, 2010, 2011). A term map is a two-dimensional representation of a field in which strongly related terms are located close to each other and less strongly related terms are located further away from each other. A term map provides an overview of the structure of a field. Different areas in a map correspond with different subfields or research areas. The color of a term in a map reflects the average citation impact of the publications in which the term occurs. The use of visualizations to analyze the structure and development of scientific fields has a long history (e.g., Börner, 2010), but this approach has not been used before to study differences in citation practices between research areas.

The first step is the definition of scientific fields. Because of their frequent use in field-normalized bibliometric indicators, the journal subject categories in the Web of Science (WoS) bibliographic database are employed to define fields. There are about 250 subject categories in the WoS database, covering all scientific disciplines. The analyses reported in this paper are based on all publications in a particular subject category that are classified as article or review and that were published between 2006 and 2010. For each publication, citations are counted until the end of 2011.

Using natural language processing techniques, the titles and abstracts of the publications in a field are parsed. This yields a list of all noun phrases (i.e., sequences of nouns and adjectives) that occur in these publications. An additional algorithm (Van Eck & Waltman, 2011) selects the 2000 noun phrases that can be regarded as the most characteristic terms of the field. This algorithm aims to filter out trivial noun phrases such as *conclusion* and *result*. All further steps are based on the 2000 selected terms.

The next step is to determine the number of publications in which each pair of terms co-occurs. Two terms are said to co-occur in a publication if they both occur at least once in the title or abstract of the publication. The larger the number of publications in which two terms co-occur, the stronger the terms are considered to be related to each other. In neuroscience, for instance, *Alzheimer* and *short-term memory* may be expected to co-occur a lot, indicating a strong relation between these two terms. The matrix of term co-occurrence frequencies serves as input for the VOS mapping technique (Van Eck, Waltman, Dekker, & Van den Berg, 2010). This technique determines for each term a location in a two-dimensional space. Strongly related terms tend to be located close to each other in the two-dimensional space, while terms that do not have a strong relation are located further away from each other. It is important to note that in the interpretation of a term map only the distances between terms are relevant. A map can be freely rotated, because this does not affect the inter-term distances. This also implies that the horizontal and vertical axes have no special meaning.

In the final step, the color of each term is determined. First, in order to correct for the age of a publication, each publication's number of citations is divided by the average number of citations of all publications that appeared in the same year. This yields a publication's normalized citation score. A score of 1 means that the number of citations of a publication equals the average of all publications that appeared in the same field and in the same year. Next, for each of the 2000 terms, the normalized citation scores of all publications in which the term occurs (in the title or abstract) are averaged. The color of a term is determined based on the resulting average score. Colors range from blue (average score of 0) to green (average score of 1) to red (average score of 2 or higher). Hence, a blue term indicates that the publications in which a term occurs have a low average citation impact, while a red term indicates that the underlying publications have a high average citation impact. The VOSviewer software (Van Eck & Waltman, 2010; freely available at www.vosviewer.com) is used to visualize the term maps resulting from the above steps.

3. Results

Figures 1, 2, and 3 show the term maps obtained for the WoS fields *Cardiac & cardiovascular systems*, *Clinical neurology*, and *Surgery*. These fields were selected because they match well with our own areas of expertise. The maps are based on, respectively, 75,314, 105,405, and 141,155 publications from the period 2006–2010. Only a limited level of detail is offered in Figures 1, 2, and 3. To explore the term maps in full detail, the reader is referred to the interactive versions of the maps that are available at www.neesjanvaneck.nl/basic_vs_clinical/. This webpage also provides maps of 29 other medical fields as well as of all medical fields taken together.

The term maps shown in Figures 1, 2, and 3 all indicate a clear distinction between basic and clinical research, with clinical research generally located in the left part of a map and basic and non-interventional research in the right part. The distinction is best visible in the *Cardiac & cardiovascular systems* and *Clinical neurology* maps (Figures 1 and 2), in which the left part indicates clinical research areas (e.g., cardiopulmonary and neurological surgery) while the right part corresponds with basic research areas (e.g., cardiology and neurology). The *Surgery* map (Figure 3) gives a somewhat different picture, probably because of the more clinical focus of surgical research. Yet, also in this map, clinical research areas (e.g., orthopedic surgery, oncological surgery, and cardiovascular surgery) are concentrated in the left and middle part, while research areas with a more basic focus (e.g., oncology and transplantation) can be found in the right part.

Connections between basic research areas on the one hand and clinical research areas on the other hand are well visible in the term maps, especially for the *Cardiac & cardiovascular systems* and *Clinical neurology* fields (Figures 1 and 2). For these fields, the term maps display 'bridges' that seem to represent translational research, that is, research aimed at translating basic research results into clinical practice. In the *Cardiac & cardiovascular systems* map (Figure 1), for instance, two bridges are visible, one in the upper part of the map and one in the lower part. In the upper part, the topic of atherosclerosis can be found, starting in the upper-right part of the map with basic research on vascular damage, continuing in the middle part with research on cholesterol and cholesterol lowering drugs, and extending in the upper-left part with interventional therapies such as coronary bypass surgery and percutaneous interventions (PCI) and its modifications (BMS and DES). In the lower part of the map, the topic of arrhythmias can be identified. It starts in the lower-right part of the map with basic research on electrophysiological phenomena, it continues in the middle part with diagnostic tools, and it ends in the lower-left part with the clinical application of ablation therapy for arrhythmias.

Looking at Figures 1, 2, and 3, a crucial observation is that the distinction between basic and clinical research areas is visible not only in the structure of the maps but also in the colors of the terms. In general, in the right part of each map, in which the more basic research areas are located, there are many yellow and red terms, which clearly indicates an above-average citation impact. On the other hand, in the left part of each map, research areas can be found with mainly blue and green terms, implying a below-average citation impact. This pattern is most strongly visible in the *Clinical neurology* map (Figure 2) and can also be clearly observed in the *Surgery* map (Figure 3). In the *Cardiac & cardiovascular systems* map (Figure 1), a clear distinction between high- and low-impact research areas is visible as well, but it does not completely coincide with the distinction between basic and clinical areas. The general picture emerging from Figures 1, 2, and 3, and supported by term maps for other medical fields provided online, is that within medical fields there is often a considerable heterogeneity in citation impact, with some research areas on average receiving two or three times more citations per publication than others. In general, high-impact research areas tend to have a more basic focus, while low-impact research areas tend to be more clinically oriented.

4. Discussion and conclusion

The citation impact of a publication can be influenced by many factors. In the medical sciences, previous studies have for instance analyzed the effect of study design (e.g., case report, randomized controlled trial, or meta-analysis; Patsopoulos, Analatos, & Ioannidis, 2005) and article type (i.e., brief report or full-size article; Mavros, Bardakas, Rafailidis, Sardi, Demetriou, & Falagas, in press). In this paper, the effect of differences in citation practices between medical research areas has been investigated.

Different fields of science have different citation practices. In some fields, publications have much longer reference lists than in others. Also, in some fields researchers mainly refer to recent work, while in other fields it is more common to cite older work. Because of such differences between fields, publications in one field may on average receive many more citations than publications in another field. Popular bibliometric indicators, such as the *h*-index and the impact factor, do not correct for this. The use of these indicators to make comparisons between fields may therefore easily lead to invalid conclusions.

The results obtained using the visualization methodology introduced in this paper go one step further and show that even within a single field of science there may be large differences in citation practices. Similar findings have been reported in some earlier studies (Neuhaus & Daniel, 2009; Smolinsky & Lercher, 2012; Van Leeuwen & Calero Medina, 2012), but based on smaller analyses and not within the medical domain. The present results show that in medical fields clinical research often has a substantially lower citation impact than basic or non-interventional research. In some cases, considerable differences in citation impact can also be observed between different basic or different clinical research areas.

Although differences in citation impact between basic and clinical research have been mentioned in earlier studies (e.g., Seglen, 1997), almost no systematic evidence of such differences has been collected. We are aware of only one earlier study in which differences in citation impact between basic and clinical research have been analyzed (Opthof, 2011). Contrary to the present results, this study concludes that clinical research is cited more frequently than basic research. However, the study is limited in scope. It is restricted to a single medical field (i.e., cardiovascular research), and it only considers publications from a small set of journals.

The present results lead to the conclusion that one should be rather careful with citation-based comparisons between medical research areas, even if the areas are part of the same field. Field-normalized bibliometric indicators, which are typically used by professional bibliometric centers, correct for differences in citation practices between fields, but at present they fail to correct for within-field differences. The use of bibliometric indicators, either the *h*-index and the impact factor or more sophisticated field-normalized indicators, may therefore lead to an underestimation of the impact of certain types of research compared with others. In particular, the impact of clinical research is likely to be underestimated, while the impact of basic and non-interventional research may be overestimated.

There is an urgent need for more accurately normalized bibliometric indicators. These indicators should correct not only for differences in citation practices between fields of science, but also for differences between research areas within the same field. Research areas could for instance be defined algorithmically based on citation patterns (e.g., Klavans & Boyack, 2010; Waltman & Van Eck, in press). Alternatively, a normalization could be performed at the side of the citing publications by giving a lower weight to citations from publications with long reference lists and a higher weight to citations from publications that cite only a few works. A number of steps towards such citing-side normalization procedures have already been taken (e.g., Glänzel, Schubert, Thijs, & Debackere, 2011; Leydesdorff & Opthof, 2010; Moed, 2010; Waltman & Van Eck, 2012; Zitt & Small, 2008), but more research in this direction is needed. Using the presently available bibliometric indicators, one should be aware of biases caused by differences in citation practices between areas of medical research, especially between basic and clinical areas.

References

Börner, K. (2010). *Atlas of science: Visualizing what we know*. MIT Press.

Chew, M., Villanueva, E.V., & Van der Weyden, M.B. (2007). Life and times of the impact factor: Retrospective analysis of trends for seven medical journals (1994–2005) and their editors' views. *Journal of the Royal Society of Medicine*, *100*(3), 142–150.

Garfield, E. (1996). How can impact factors be improved? *British Medical Journal*, *313*(7054), 411–413.

Garfield, E. (2006). The history and meaning of the journal impact factor. *JAMA*, *295*(1), 90–93.

Glänzel, W., Schubert, A., Thijs, B., & Debackere, K. (2011). A priori vs. a posteriori normalisation of citation indicators. The case of journal ranking. *Scientometrics*, *87*(2), 415–424.

Glänzel, W., Thijs, B., Schubert, A., & Debackere, K. (2009). Subfield-specific normalized relative indicators and a new generation of relational charts: Methodological foundations illustrated on the assessment of institutional research performance. *Scientometrics*, *78*(1), 165–188.

Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, *102*(46), 16569–16572.

Klavans, R., & Boyack, K.W. (2010). Toward an objective, reliable and accurate method for measuring research leadership. *Scientometrics*, *82*(3), 539–553.

Leydesdorff, L., & Opthof, T. (2010). Scopus's source normalized impact per paper (SNIP) versus a journal impact factor based on fractional counting of citations. *Journal of the American Society for Information Science and Technology*, *61*(11), 2365–2369.

Mavros, M.N., Bardakas, V., Rafailidis, P.I., Sardi, T.A., Demetriou, E., & Falagas, M.E. (in press). Comparison of number of citations to full original articles versus brief reports. *Scientometrics*.

Moed, H.F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, *4*(3), 265–277.

Neuhaus, C., & Daniel, H.-D. (2009). A new reference standard for citation analysis in chemistry and related fields based on the sections of Chemical Abstracts. *Scientometrics*, *78*(2), 219–229.

Opthof, T. (2011). Differences in citation frequency of clinical and basic science papers in cardiovascular research. *Medical and Biological Engineering and Computing*, *49*(6), 613–621.

Patel, V.M., Ashrafian H., Ahmed, K., Arora, S., Jiwan, S., Nicholson, J.K., Darzi, A., & Athanasiou, T. (2011). How has healthcare research performance been assessed? A systematic review. *Journal of the Royal Society of Medicine*, *104*(6), 251–261.

Patsopoulos, N.A., Analatos, A.A., & Ioannidis, J.P.A. (2005). Relative citation impact of various study designs in the health sciences. *JAMA*, *293*(19), 2362–2366.

Radicchi, F., Fortunato, S., & Castellano, C. (2008). Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences*, *105*(45), 17268–17272.

Seglen, P.O. (1997). Why the impact factor of journals should not be used for evaluating research. *British Medical Journal*, *314*(7079), 498–502.

Smolinsky, L., & Lercher, A. (2012). Citation rates in mathematics: A study of variation by subdiscipline. *Scientometrics*, *91*(3), 911–924.

Van Eck, N.J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538.

Van Eck, N.J., & Waltman, L. (2011). Text mining and visualization using VOSviewer. *ISSI Newsletter*, *7*(3), 50–54.

Van Eck, N.J., Waltman, L., Dekker, R., & Van den Berg, J. (2010). A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS. *Journal of the American Society for Information Science and Technology*, *61*(12), 2405–2416.

Van Leeuwen, T.N., & Calero Medina, C. (2012). Redefining the field of economics: Improving field normalization for the application of bibliometric techniques in the field of economics. *Research Evaluation*, *21*(1), 61–70.

Waaijer, C.J.F., Van Bochove, C.A., & Van Eck, N.J. (2010). Journal editorials give indication of driving science issues. *Nature*, *463*, 157.

Waaijer, C.J.F., Van Bochove, C.A., & Van Eck, N.J. (2011). On the map: Nature and Science editorials. *Scientometrics*, *86*(1), 99–112.

Waltman, L., Van Eck, N.J., Van Leeuwen, T.N., Visser, M.S., & Van Raan, A.F.J. (2011a). Towards a new crown indicator: An empirical analysis. *Scientometrics*, *87*(3), 467–481.

Waltman, L., Van Eck, N.J., Van Leeuwen, T.N., Visser, M.S., & Van Raan, A.F.J. (2011b). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, *5*(1), 37–47.

Waltman, L., & Van Eck, N.J. (2012). *Source normalized indicators of citation impact: An overview of different approaches and an empirical comparison*. arXiv:1208.6122.

Waltman, L., & Van Eck, N.J. (in press). A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology*.

Zitt, M., & Small, H. (2008). Modifying the journal impact factor by fractional citation weighting: The audience factor. *Journal of the American Society for Information Science and Technology*, *59*(11), 1856–1860.

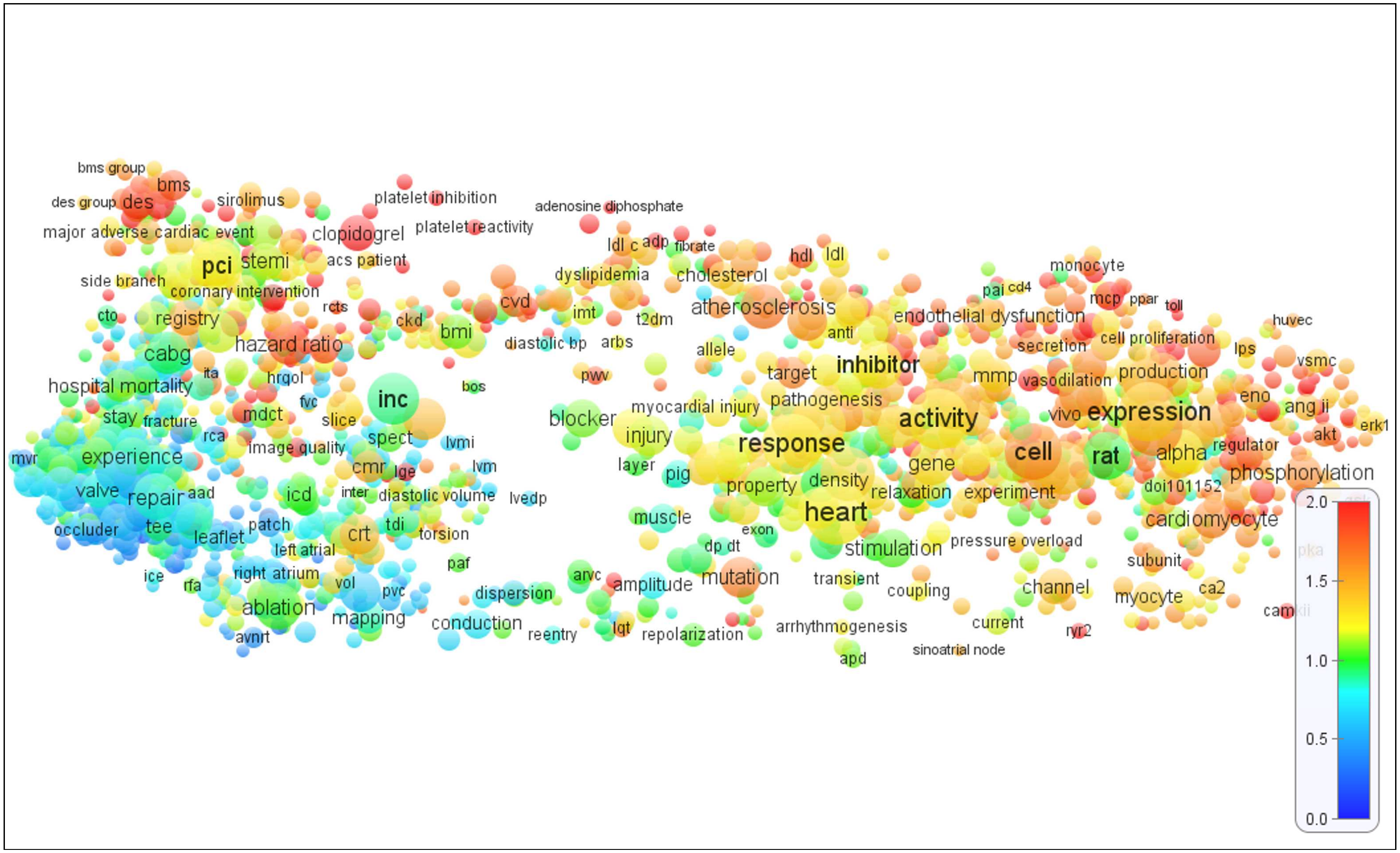

Figure 1. Term map of the *Cardiac & cardiovascular systems* field.

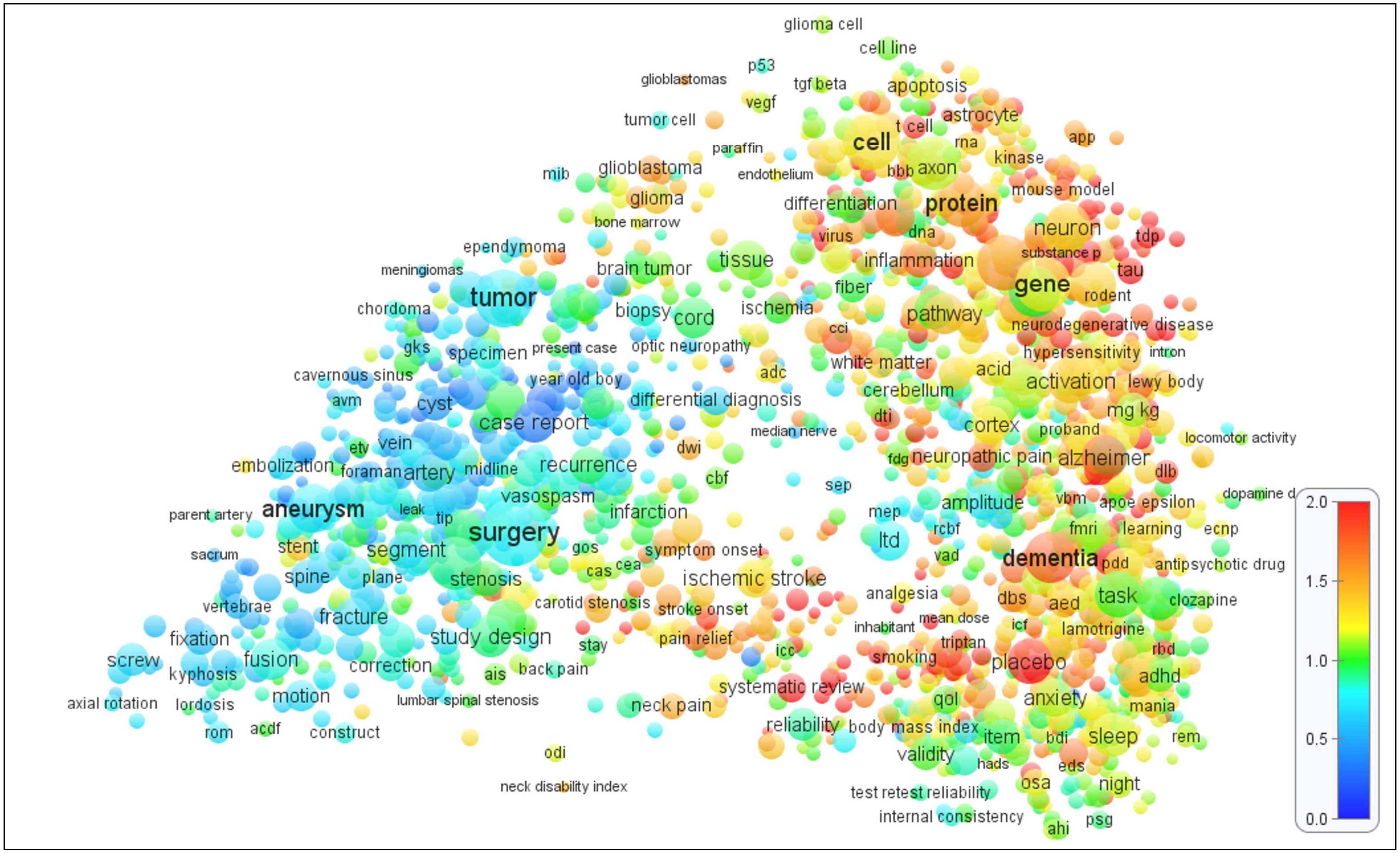

Figure 2. Term map of the *Clinical neurology* field.

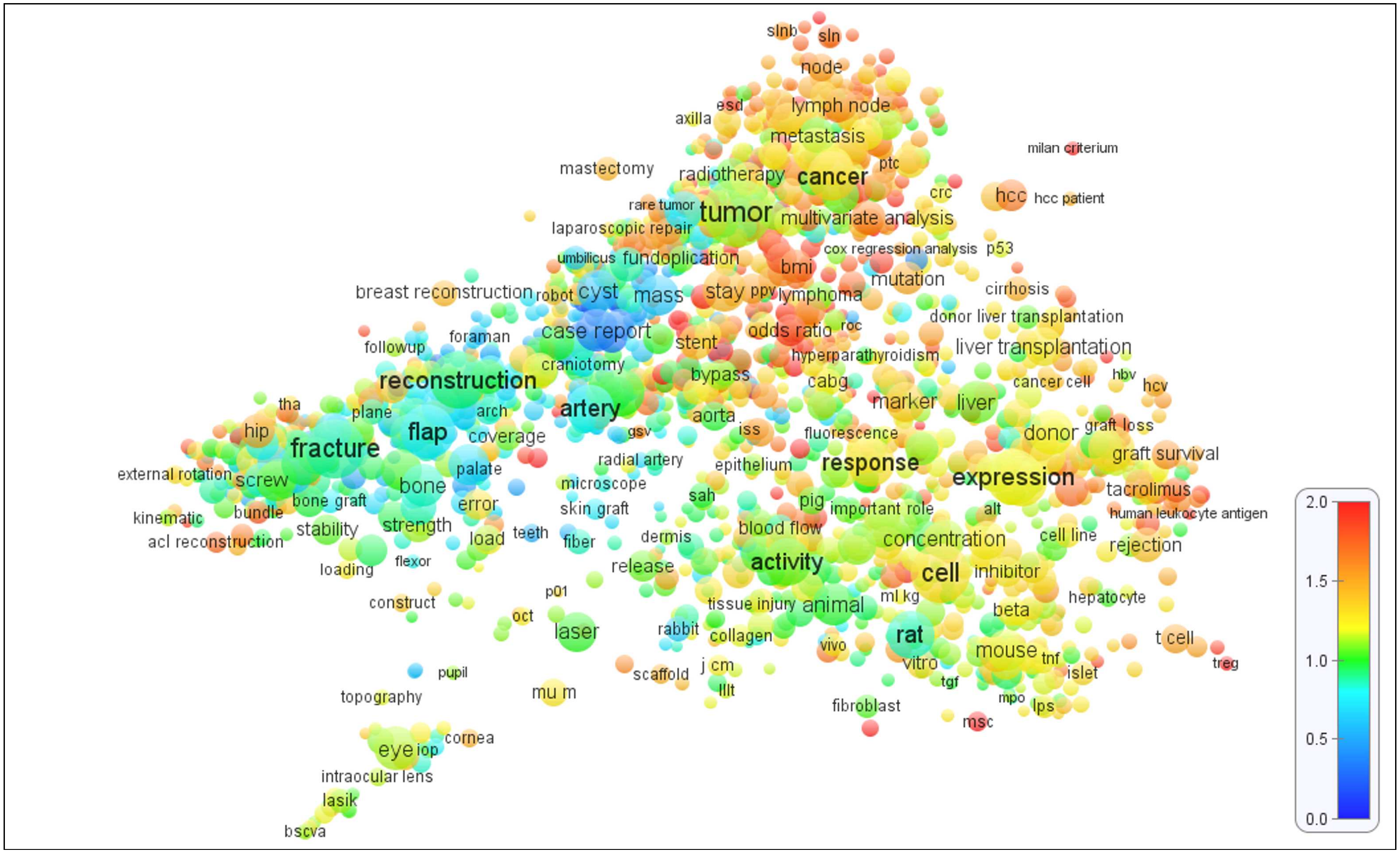

Figure 3. Term map of the *Surgery* field.