

**Impact of Medical Academic Genealogy on Publication Patterns:
An Analysis of the Literature for Surgical Resection in Brain Tumor Patients**

Running Head: **Impact of Medical Academic Genealogy**

Brian R. Hirshman M.D. M.S.^{1,2,3}, bhirshman@ucsd.edu
Jessica A. Tang B.S.¹, jat030@ucsd.edu
Laurie A. Jones M.S.³, lajones@cmu.edu
James A. Proudfoot M.S.⁴, jproudfoot@ucsd.edu
Kathleen M. Carley Ph.D.^{2,3}, kathleen.carley@cs.cmu.edu
Lawrence Marshall M.D.⁵, lmarshall@ucsd.edu
Bob S. Carter M.D. Ph.D.^{1,5}, bobcarter@ucsd.edu
Clark C. Chen* M.D. Ph.D.^{1,5} clarkchen@ucsd.edu

Body word count: **2694** (without abstract)

Abstract word count: **247**

Intro word count: **385**

Discussion word count: **443**

¹ Center for Translational and Applied Neuro-Oncology, Division of Neurosurgery, University of California, San Diego

² Center for Computational Analysis of Social and Organizational Systems, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA

³ Computation, Organizations & Society Program, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA

⁴ Clinical and Translational Research Institute, University of California San Diego

⁵ Department of Neurosurgery, University of California, San Diego

*Corresponding Author:

Clark C. Chen, M.D., Ph.D.

3855 Health Science Drive #0987

La Jolla, CA 92093-0987

Phone: 858 246 0674

Fax: 858 822 4715

E-mail: clarkchen@ucsd.edu

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version record](#). Please cite this article as [doi:10.1002/ana.24569](https://doi.org/10.1002/ana.24569).

Abstract

“Academic genealogy” refers to the linking of scientists and scholars based on their dissertation supervisors. We propose that this concept can be applied to medical training and that this “medical academic genealogy” may influence the landscape of the peer-reviewed literature. We performed a comprehensive PubMed search to identify U.S. authors who have contributed peer-reviewed articles on a neurosurgery topic that remains controversial: the value of maximal resection for high grade gliomas (HGGs). Training information for each key author (defined as the first or last author of an article) was collected (e.g. author’s medical school, residency, and fellowship training). Authors were recursively linked to faculty mentors to form genealogies. Correlations between genealogy and publication results were examined. Our search identified 108 articles with 160 unique key authors. Authors who were members of two genealogies (14% of key authors) contributed to 38% of all articles. If an article contained an authorship contribution from the first genealogy, its results were more likely to support maximal resection (Log odds ratio ~~OR~~ 2.74, $p < 0.028$) relative to articles without such contribution. In contrast, if an article contained an authorship contribution from the second genealogy, it was less likely to support maximal resection (Log odds ratio ~~OR~~ -1.74, $p < 0.026$). We conclude that the literature on surgical resection for HGGs is influenced by medical academic genealogies, and that articles contributed by authors of select genealogies share common results. These findings have important implications for the interpretation of scientific literature, design of medical training, and health care policy.

Keywords: medical academic genealogy, medical education, extent of resection, high grade glioma, brain cancer, network analysis

Accepted Article

INTRODUCTION

A central tenet of science is that investigators should strive to be objective – to be free from the undue influence of past experience, social context, and the opinions of peers and mentors. Basic intuition, however, suggests that perceptions and conclusions are likely affected by the beliefs of those around us, particularly our mentors. This intuition is supported by a well-established literature in both the social sciences(1–9) and the physical sciences(8–13). In particular, an emerging interdisciplinary literature suggests that mentors and mentoring environments have a strong influence on researcher attitudes, methods of investigation, and career development(14–17).

To date, studies investigating social factors that influence scientific investigations tend to use qualitative methodologies. The notion of “academic genealogy”, in which scientists are linked based on their dissertation supervisors, is one technique designed to qualitatively characterize the influence of mentors(18–21). Genealogies have also been constructed to analyze other creative fields such as philosophy(22), music(23), and art(24, 25) to follow the influences of teachers on their students. The emergence of dynamic network models(26, 27) and social network analysis(28–30) now allow rigorous quantitation of genealogical influences. Here, we apply the concept of academic genealogy to medical training and use network analysis to quantitatively assess the impact of “medical academic genealogy” on medical investigations.

To this end, we study the medical academic genealogy of authors who contributed peer-reviewed articles on a controversial subject in neurosurgery, the utility of maximal surgical resection in patients afflicted with high grade gliomas (HGGs)(31–39). The infiltrative nature of HGGs, the most common forms of adult brain cancer, renders complete surgical resection impossible. The unresolved issue, however, is whether maximal resection leads to increased

survival. Supporters of maximal resection believe that reducing the tumor burden enhances the efficacy of subsequent chemo-radiation(34). Opponents argue that this is of no benefit given the inherent resistance of HGGs to chemotherapy and radiation(37, 38). While the number of retrospective studies exploring this issue has greatly increased in recent years(35, 36), the controversy has not been resolved through a well-designed randomized clinical trial.

The goal of this study is not to resolve this question. Instead our study aims to examine whether medical training influences results. Utilizing quantitative network analysis, we find a statistical association between membership in a genealogy and results published in this field.

METHODS

Identification of key articles. A comprehensive PubMed search was performed in December of 2014 using broad medical subject heading (MeSH) terms relating to this controversy. We required that all articles have a MeSH term related to high grade gliomas (“high-grade glioma”, “astrocytoma”, “anaplastic astrocytoma”, “oligodendroglioma”, “oligoastrocytoma”, “glioblastoma*”, or “intracerebral tumor”) and a MeSH term related to tumor resection (“extent of resection,” “surgical resection”, “gross total resection”, “tumor resection”, “partial resection”, or “resection and/or “extent” in the title or abstract). This process identified 4047 articles for review. Articles were selected for this study if they were i) written in English, ii) published before December 2014, iii) presented original research results on human patients, iv) focused on adult intracranial HGGs, v) discussed maximal resection, vi) were written as a clinical study and not a case report, vii) used mortality as an outcome, viii) performed a univariate or multivariate statistical analysis ix) considered maximal resection as a separate comparison group in their statistical analysis and x) listed a primary address at an American institution for either the first or

last authors. The last criterion was necessary because it was often not possible to ascertain and verify the training history of authors trained at foreign institutions. This process produced 108 articles for analysis.

Article classification. We classified the identified articles into two groups based on “publication result”. Those that found a statistically significant correlation ($p < 0.05$) between maximal resection and survival as a primary or secondary result were labeled “supportive”.

Those that reported no statistical association between maximal surgical resection and overall survival were labeled “not supportive”. This determination was performed by two independent readers (B.R.H. and J.A.T) and discrepant articles were discussed with senior readership (C.C.C).

Author classification. We defined the first and last author of each article as “key authors”, with the rationale that these authors play key roles in shaping the conclusion of the manuscript(40).

To avoid oversampling from the small number of articles with joint first or last authorship ($n=9$ and $n=2$, respectively), we selected only the first and the last of the co-authors for our study.

There were 160 unique key authors of our 108 articles because many investigators were key authors of more than one article. Internet searches were performed to determine their medical subspecialty (if any) as well as the timing and location of their medical school, residency, and fellowship training. All information was compiled (by B.R.H. and J.A.T.) using publicly available academic or institutional resumes and verified against publicly accessible documents such as academic directory listings, alumni pages, and press briefings. Each training history was verified using at least two independent sources by both reviewers.

Construction of medical academic genealogies. We adopted a “top-down” approach to identifying medical academic genealogies (**Fig 1**). We first identified potential “founders” of genealogies in the following manner. We identified the 21 authors who served as department chairmen because these individuals oversee the training of multiple trainees. Next, we linked these potential founders to their trainees, trainees of trainees, and so on in a recursive manner to create what social network analysis calls the “ego network” of the chairmen(26, 28, 29). Links were drawn between authors if one was a faculty member at an institution while another was a trainee (medical student, resident, or fellow) in the same discipline. To be connected, mentor and trainee had to be located at the same institution during the same calendar year. Finally, we excluded networks that did not span at least two generations. This process produced eleven medical academic genealogies for analysis.

Association of genealogy and articles. Many of the key authors were also middle authors of manuscripts for where they did not serve as first or last author. To capture the influence of these authorships, we associated an article with a genealogy if the article had an author who was a member of that genealogy, including middle authors. In analyzing the articles produced by a genealogy, we only counted an article once no matter how many authors were members of that genealogy. If an article included authors from multiple independent genealogies, the article was assigned to each genealogy during analysis. In this way, a single article could be associated with multiple genealogies. We applied the same criteria in classifying articles by medical subspecialty. For example, we considered an article to be written by a member of a specialty if it

had one of our key authors from that specialty on the article, even if he or she was a middle author.

Statistical Analysis. We considered the possibility that each article may not represent an independent investigative unit. For instance, articles contributed by the same first author may harbor similar results. We further considered the possibility that articles that originating from the same data set (e.g. the Surveillance, Epidemiology, and End Result (SEER) database) or academic institution may share comparable results. We therefore tested the association between genealogy and publication result using a mixed logistic regression model. Publication result was treated as the dependent variable, genealogy membership was treated as a fixed effect, and first author of the article and the dataset of the publication-article were treated as random effects. False discovery rate correction was performed for multiple comparisons(41).

Because others have suggested that the sample size of the study, ~~the medical specialty of the author, and~~ the time of publication, and the medical specialty of the author potentially influences whether a study supports maximal resection, we examined these variables using the same univariate mixed logistic model(35, 36, 39, 42). As article sample sizes reported in our identified literature ranged from 19 to 40,137 patients, with a skew toward the larger numbers, we performed a logarithmic conversion in order to meaningfully account for this distribution. Additionally, ~~articles with any key author neurosurgeons were identified to differentiate from any overall effect of author field. Lastly,~~ as approximately half of our articles were published prior to 2010, we used this date as our division point. Lastly, we examined whether the specialties of the first or the last author (neurosurgeons, medical oncologists,

radiation oncologists, or other) or whether the presence of at least one neurosurgeon key author were associated with publication result.

Variables that were significantly associated with publication ~~result outcomes~~ in the univariate analyses were then incorporated in to a multivariate mixed logistic regression model. Specifically, the final mixed logistic regression model treated publication results as the dependent variable. Genealogy (A or B), time of publication, and the sample size of the study were treated as fixed effects. The first author of the manuscript and the dataset of the manuscript were treated as random effects. To take into consideration the potential effect of the senior authorship, additional models were performed incorporating the senior author of the manuscript as random effects.

All statistical analysis were performed using R version 3.1.2 (R Foundation for Statistical Computing) and p or q values less than 0.05 were considered significant. Mixed models were performed using the lme4 package version 1.1-9 and bobyqa optimizer. Visual representation of genealogies were created using *ORA version 3.0.9 (CASOS center, Carnegie Mellon University).

RESULTS

Univariate association between genealogy and publication result. Our search identified 108 original articles, 160 key authors, and 11 genealogies (**Table 1, Tables S2-S5**). A majority of the articles reported results in support of maximal HGG resection (70%). Mixed model logistic regressions were performed to determine whether articles published by genealogy members were more likely to support HGG resection as compared to articles published by authors not belonging to that genealogy (first column, **Table 2**) using first author and data set as random effects. ~~In this~~

analysis, we considered the possibility that articles contributed by the same first author may harbor similar results and may therefore not represent independent units of analysis. Similarly, articles that originating from the same data set (e.g. the Surveillance, Epidemiology, and End Result (SEER) database) or academic institution may share comparable results. In this context, we controlled for these potential confounders by treating them as random effects in the mixed regression model. As there were eleven genealogies, eleven comparisons were performed. After False discovery rate correction for eleven pairwise comparisons, we identified two genealogies where members were more likely to publish article in support (or not in support) of surgical resection (**Fig 2**). ~~One genealogy was founded by A1 (“Genealogy A”), a neurosurgical chair at a prominent American medical school.~~ The presence of an author from ~~this genealogy~~ “Genealogy A”, a genealogy of 14 neurosurgeons, increased the log odds ratio (LOR) that an article would support of support maximal resection by 3.50 ($q=0.043$). ~~The second genealogy was founded by B1 (“Genealogy B”), a radiation oncology chair incidentally at the same medical school.~~ The presence of an author from ~~this genealogy~~ “Genealogy B”, a genealogy of 8 radiation oncologists, decreased the log odds of support for maximal resection by -2.08 ($q=0.043$). The presence of an author from the remaining nine genealogies did not have a statistically significant association with article results.

Notably, there was no overlap between the authors or articles that comprised the two genealogies (**Table 3**). No author was a member of both genealogies, and none of the articles contained authors from both sets of genealogies. Furthermore, the 22 members of Genealogies A and B contributed to 38% of all articles studied (25% Genealogy A, 13% Genealogy B, **Table 3**) while accounting for 14% of all key authors (9% Genealogy A, 5% Genealogy B).

Analysis of sample size, ~~author specialty, and time of publication, and author specialty~~ with result. Previous investigators have proposed that study results depend, in part, on the sample size of the study (with larger studies more likely to support resection due to greater statistical power(35, 36)), ~~the author specialty (with neurosurgeons more likely to support resection than non-neurosurgeons(38, 39))~~, and the time of publication (with the more “recent” ~~publications articles~~ more likely to support resection(35, 36)). We wished to determine the validity of these proposed associations in our dataset. Univariate analysis showed that studies with larger cohorts (LOR 1.60, $p < 0.01$; first column, Table 2) and those published after 2010 (LOR 1.49, $p < 0.01$) were more likely to support surgical resection.

Previous investigators have also proposed that author specialty may be associated with publication result, with neurosurgeons more likely to support maximal resection than non-neurosurgeons (38, 39). However, our analysis indicated that medical specialties of either the first author (all $p > 0.05$; first column, Table 2) or the last author (all $p > 0.05$) were not associated with publication result when examined individually. We also did not find evidence of an association between publications written by at least one neurosurgeon key author as compared to articles without a neurosurgeon key author ($p > 0.05$). However, ~~articles including a key author who is neurosurgeon were no more likely to support resection when compared to articles without a neurosurgeon key author ($p > 0.184$; first column, Table 2).~~

Multivariate association between genealogy and publication result. To determine whether the association between genealogy A and B with publication results remain robust after controlling for all potential confounding variables, we analyzed our results using a multivariate mixed model logistic regression model that incorporated time of publication, sample size, and

medical academic genealogy as fixed effects, in addition to data set and first author as random effects. We found that articles authored by members of Genealogy A were more likely to support maximal resection than articles not written by members of Genealogy A (LOR 2.74, $p < 0.028$; second column, **Table 2**). On the other hand, articles authored by members of Genealogy B were less likely to support maximal resection than those not written by members of Genealogy B (LOR -1.74, $p < 0.026$, last column, **Table 2**). Similar results were observed when the mixed model logistic regression was repeated using last author as a random effect variable. The association between genealogy and publication result remained significant in a mixed model logistic regression for genealogy A (LOR 2.24, $p < 0.047$; fourth column, **Table 2**) and in a mixed model logistic regression for genealogy B (LOR -1.53, $p < 0.024$; last column, **Table 2**).

DISCUSSION

In this article, we introduce the concept of “medical academic genealogy” that links authors by common medical mentors. Using a quantitative method developed in social network analysis, we demonstrate that articles published by authors in a medical academic genealogy are more likely to share similar conclusions in the HGG literature. While our study examines a specific neurosurgical question, we propose that the issues raised are pertinent to critical evaluation of other medical literature. Based on this finding, we suggest that medical academic genealogy plays a previously unrecognized role in shaping medical literature. Recognition and reconciliation of these effects should improve our ability to evaluate medical literature.

Our findings have particular significance in the era of health reform. Increasingly, the effectiveness of medical practice will be evaluated by central panels who review the published literature. Care should be taken in the evaluation of medical literature disproportionately shaped

by members of medical academic genealogies. Further, the potential influence of medical academic genealogy on publication ~~result outcome~~ challenges a fundamental premise of meta-analyses, since each individual ~~article publication~~ may not represent an independent investigative unit(43). If so, the development of statistical tools that adjust for the influence of genealogy will be needed for future quantitative reviews.

Our results also highlight an inherent tension in medical mentorship. While in clinical care it often necessary to have hierarchical interactions, in research we should strive to foster independence. The challenge of medical training lies in fostering appropriate mentor-trainee relationships while minimizing the unconscious adaptation of mentor biases. It is therefore necessary to consciously structure the educational experience to reflect these goals. In this context, an integrated educational approach involving thoughtful curriculum design, mentor self-awareness, and training individualized to the tendencies of the trainee will be necessary to minimize genealogical bias.

There are several limitations inherent in our study design. As in all retrospective studies, our conclusions were based on correlative associations with causation inferred. We further recognize that dividing complex variables into discrete groups may have potential impact on statistical analysis. Additionally, the genealogies we created are distillations of the complex medical communities and training environments in academic medical centers, and the exclusion of non-U.S. authors potentially limits the generalizability of our conclusion. In addition, although our analysis weights all education links equally, the literature suggests that mentor influence vary during training. Finally the interpretation of our results is limited by publication bias, as we do not know what was not published(43–45). Future work will be necessary to determine if certain genealogies disproportionately pursued and/or abandoned non-supportive

results at different rates. Despite this, we believe that our data compellingly demonstrate the effects of medical academic genealogy on published literature.

CONCLUSION

Analysis of the literature on the utility of surgical resection for HGG reveals that members of medical academic genealogies make significant contributions to the peer-reviewed literature. Articles written by authors belonging to select genealogies are more likely to support (or not support) surgical resection relative to articles written by non-genealogy members.

Accepted Article

Acknowledgements

This work was supported by the Czech Duck research fellowship in Neurosurgery at UCSD which was awarded to Dr. Brian Hirshman. Statistics were discussed with Mr. James Proudfoot (Clinical & Translational Research Institute, UCSD) as well as Ms. Laurie Jones and Dr. Kathleen Carley (Computation, Organizations, and Society Program, Carnegie Mellon University).

Author contributions

BRH and CCC were responsible for study conception and design. KMC was responsible for the design and supervision of the network analysis. Acquisition of data was performed by BRH, JAT and CCC. Statistical analysis and data analysis was performed by BRH, LAJ, JAP, BSC, LM, and KMC. The manuscript and figures were drafted by BRH and CCC. The manuscript was critically reviewed by all authors.

Accepted Article

References

1. Strang D (1997) Cheap Talk: managerial discourse on quality circles as an organizational innovation. *Proceedings of the American Sociological Association* (Toronto, Canada).
2. Mahoney M (1977) Publication Prejudices: an experimental study of confirmatory bias in the peer review system. *Cogn Ther Res* 1(2):161–175.
3. Young SN (2009) Bias in the research literature and conflict of interest: an issue for publishers, editors, reviewers and authors, and it is not just about the money. *J Psychiatry Neurosci* 34(6):412–417.
4. Crane D (1972) *Invisible colleges; Diffusion of knowledge in scientific communities* (University of Chicago Press, Chicago, IL).
5. Anderson MS, Ronning EA, De Vries R, Martinson BC (2007) The perverse effects of competition on scientists' work and relationship. *Sci Eng Ethics* 13(4):437–461.
6. Case DO, Higgins GM (2000) How can we investigate citation behavior? A study of reasons for citing literature in communication. *J Am Soc Inf Sci* 51(7):635–645.
7. Pickering A (1992) *Science as Practice and Culture* (University of Chicago Press, Chicago, IL).
8. MacCoun RJ (1998) Biases in the Interpretation and Use of Research Results. *Annu Rev Psychol* 49(1):259–287.
9. Kuhn T (1970) *The Structure of Scientific Revolutions* (University of Chicago Press, Chicago, IL). 2nd Ed.
10. Kim DD, Tang JY, Ioannidis J (2014) Network geometry shows evidence sequestration for medical vs. surgical practices: treatments for basal cell carcinoma. *J Clin Epidemiol* 67(4):391–400.
11. Newman M (2003) The Structure and Function of Complex Networks. *SIAM Rev* 45(2):167–256.
12. Latour B, Woolgar S (1986) *Laboratory life: the construction of scientific facts* (Princeton University Press, Princeton, NJ). 2nd Ed.
13. Park I-U, Peacey MW, Munafo MR (2014) Modelling the effects of subjective and objective decision making in scientific peer review. *Nature* 506(7486):93–96.
14. Sambunjak D, Straus SE, Marusic A (2010) Mentoring in Academic Medicine. *J Am Med Assoc* 296(9):1103–1115.
15. Malmgren RD, Ottino JM, Amaral LAN (2010) The role of mentorship in protégé performance. *Nature* 465:622–626.
16. Eby LT, Allan TD, Evans SC, Ng T, DuBois DL (2008) Does Mentoring Matter? a multidisciplinary meta-analysis comparing mentored and non-mentored individuals. *J Vocat Behav* 72(2):254–267.

7. Chen C, Petterson S, Phillips R, Bazemore A, Mullan F (2014) Spending Patterns in Region of Residency Training and Subsequent Expenditures for Care Provided by Practicing Physicians for Medicare Beneficiaries. *J Am Med Assoc* 312(22):2385–2393.
8. Academic Family Tree (2014) Available at: <http://academictree.org/>.
9. The PhD Tree Project (2014) Academic Genealogy Wiki. *PhD Tree Acad Geneal Fam Tree*. Available at: <http://phdree.org/>.
0. Mathematics Genealogy Project (2014) (North Dakota State University, Fargo, North Dakota) Available at: <http://genealogy.math.ndsu.nodak.edu/>.
1. Jackson A (2007) A Labor of Love: The Mathematics Genealogy Project. *Not Am Math Soc* 54(8):1002–1003.
2. Scharp K (2014) History of Western Philosophy. Available at: <http://kevinscharp.com/Sociology%20of%20Philosophy%20%28Western%29%203.1%20%20%28part%201%29.jpg>.
3. Piano Genealogy Database Available at: <http://www.pianogenealogy.com.au/> [Accessed May 5, 2014].
4. Williamson RK (1991) *American architects and the mechanics of fame* (University of Texas Press, Austin, Texas).
5. Barr A (1936) *Cubism and Abstract Art* (Belknap Press, Cambridge, MA).
6. Carley K (2003) Dynamic Network Analysis. *NRC Workshop on Social Network Modeling and Analysis*: (Washington DC), pp 133–145.
7. Carley K (2014) ORA: A Toolkit for Dynamic Network Analysis and Visualization. *Encycl Soc Netw Anal Min*:1219–1228.
8. Wasserman S, Faust K (1994) *Social Network Analysis* (Cambridge University Press, Cambridge, MA).
9. Scott J (2000) *Social network analysis: a handbook* (Sage, London, England). Second.
0. Kas M, Carley K, Carley LR (2012) Trends in Scientific Networks: understanding structures and statistics in scientific networks. *Soc Netw Anal Min* 2(2):169–187.
1. Wen PY, Kesari S (2008) Malignant Gliomas in Adults. *N Engl J Med* 395(5):492–507.
2. Ng K, Kim R, Kesari S, Carter BS, Chen CC (2012) Genomic profiling of glioblastoma: convergence of fundamental biologic tenets and novel insights. *J Neurooncol* 107(1):1–12.
3. Bartek Jr. J, et al. (2012) Key concepts in glioblastoma therapy. *J Neurol Neurosurg Psychiatry* 83(7):753–760.

4. Lacroix M, et al. (2001) A multivariate analysis of 416 patients with glioblastoma multiforme: prognosis; extent of resection; and survival. *J Neurosurg* 95(2):190–198.
5. Sanai N, Berger MS (2008) Glioma Extent of Resection and its Impact on Patient Outcome. *Neurosurgery* 62(4):753–764.
6. Hardesty D, Sanai N (2012) The value of glioma extent of resection in the modern neurosurgical era. *Front Neurol* 3(140):1–8.
7. Gonda DD, et al. (2013) The Value of Extended Glioblastoma Resection: insights from randomized controlled trials. *Surg Neurol* 4:110–114.
8. Hess KR (1999) Extent of resection as a prognostic variable in the treatment of gliomas. *Neuro-Oncol* 4(3):227–231.
9. Quigley MR, Maroon JC (1991) The Relationship between Survival and the Extent of the Resection in Patients with Supratentorial Malignant Gliomas. *J Neurosurg* 29(3):385–389.
0. Tscharrntke T, Hochberg ME, Rand TA, Resh VH, Kraus J (2007) Author sequence and credit for contributions in multiauthored publications. *PLoS Biol* 5.1:e18.
1. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B* 57(1):289–300.
2. Beiko J, et al. (2014) IDH1 mutant malignant astrocytomas are more amenable to surgical resection and have a survival benefit associated with maximal surgical resection. *Neuro-Oncol* 16(1):81–91.
3. Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F (2000) *Methods for Meta-analysis in Medical Research* (John Wiley & Sons, Ltd, New York, NY).
4. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR (1991) Publication Bias in Clinical Research. *The Lancet* 337(8746):867–872.
5. Yoshimoto Y (2003) Publication Bias in Neurosurgery: lessons from a series of unruptured aneurysms. *Acta Neurochir (Wien)* 145(1):45–48.

Acknowledgements

This work was supported by the Czech Duck research fellowship in Neurosurgery at UCSD which was awarded to Dr. Brian Hirshman. Statistics were discussed with Mr. James Proudfoot (Clinical & Translational Research Institute, UCSD) as well as Ms. Laurie Jones and Dr. Kathleen Carley (Computation, Organizations, and Society Program, Carnegie Mellon University).

Accepted Article

Table & Figure Legends

Table 1: Descriptive features of articles analyzed.

Table 2: Analysis of article results by medical academic genealogy

Table 3: Impact of key medical academic genealogies on publication results

Figure 1: Genealogy generation. Genealogies were created using by linking key authors (first and last authors) of the identified literature based on who trained whom. Once genealogies were identified, all articles written by genealogy members were compiled and analyzed.

Figure 2: Authors from Genealogy A (left) and Genealogy B (right). Authors are colored white if all articles by that author support maximal resection, light gray if over half support maximal resection, dark gray if less than or equal to half support maximal resection, and black if none support maximal resection. Node size is proportional to the number of articles written.

Accepted Article

Table 1: Descriptive features of articles analyzed

	Results supportive of maximal resection		Results not supportive of maximal resection		All articles	
	n (% row)	mean (sd)	n (% row)	mean (sd)	n (% all)	mean (sd)
Articles (105 total)						
Total number of articles	76 (70%)		32 (30%)		108 (100%)	
Article sample size (range 19-40,137)						
Number of patients studied	76 (70%)	1875 (6279)	32 (30%)	121 (80)		1356 (5317)
Base 10 log, number of patients	76 (70%)	2.39 (0.71)	32 (30%)	1.98 (0.32)		2.27 (0.65)
Year of publication (range 1983-2014)						
Year of publication	76 (70%)	2008 (6.9)	32 (30%)	1998 (8.5)	108 (100%)	2005 (8.5)
Published before 2010	36 (58%)	2002 (6.4)	26 (42%)	1995 (6.6)	62 (57%)	1999 (7.2)
Published during or after 2010	40 (87%)	2013 (1.3)	6 (13%)	2011 (0.8)	46 (43%)	2012 (1.4)
<u>Medical specialty of first author</u>						
Medical oncologist	<u>16 (21%)</u>		<u>6 (19%)</u>		<u>22 (20%)</u>	
Neurosurgeon	<u>33 (43%)</u>		<u>8 (25%)</u>		<u>41 (38%)</u>	
Radiation oncologist	<u>16 (21%)</u>		<u>14 (44%)</u>		<u>40 (28%)</u>	
Other	<u>11 (15%)</u>		<u>4 (12%)</u>		<u>15 (14%)</u>	
<u>Medical specialty of last author</u>						
<u>Medical oncologist</u>	<u>16 (73%)</u>		<u>6 (27%)</u>		<u>22 (20%)</u>	
<u>Neurosurgeon</u>	<u>37 (76%)</u>		<u>12 (24%)</u>		<u>49 (46%)</u>	
<u>Radiation oncologist</u>	<u>14 (61%)</u>		<u>9 (39%)</u>		<u>23 (21%)</u>	
<u>Other</u>	<u>9 (64%)</u>		<u>5 (36%)</u>		<u>14 (13%)</u>	
<u>Medical specialty of either author</u>						
One or more neurosurgeons‡	52 (74%)		18 (26%)		70 (65%)	
<u>No neurosurgeons</u>	<u>24 (63%)</u>		<u>14 (37%)</u>		<u>38 (35%)</u>	
Co-authors per article who are key authors†						

Mean number of authors per article	76 (70%)	3.23 (1.39)	32 (30%)	3.63 (1.91)	108 (100%)	3.35 (1.57)
Joint first or last authorship						
Articles with joint first authorship	8 (89%)		1 (11%)		9 (8%)	
Joint first or last authorship						
Articles with joint first authorship	8 (89%)		1 (11%)		9 (8%)	
Articles with joint last authorship	2 (100%)		0 (0%)		2 (2%)	

* Aggregated as “non neurosurgeons” for analysis

‡Key authors are authors who appear in a first or last author on any article studied

‡Key authors from multiple specialties collaborate on the same article, therefore values do not sum to 100%

Accepted Article

Table 2: Analysis of article results by medical academic genealogy

Parameter	Univariate Parameters LOR ¹	Mixed Logistic Regression Model ¹ +		Mixed Logistic Regression Model ² +	
		Genealogy A Authorship LOR	Genealogy B Authorship LOR	Genealogy A Authorship LOR	Genealogy B Authorship LOR
Medical academic genealogy					
Genealogy A (14 neurosurgeons)	3.50 **†	2.74 **	--	2.24 **	--
Genealogy B (8 radiation oncologists)	-2.08 **†	--	-1.74 **	--	-1.53 **
Nine other genealogies	NS	--	--	--	--
Known literature covariates					
Article's sample size (log transformed)	1.60 ***	1.33 ***	1.49 **	1.23 **	1.41 **
Published during / after 2010	1.49 ***	NS	1.10 **	NS	1.10 **
<u>Author specialty</u>					
<u>First author‡</u>	<u>NS</u>	<u>==</u>	<u>==</u>	<u>==</u>	<u>==</u>
<u>Last author‡</u>	<u>NS</u>	<u>==</u>	<u>==</u>	<u>==</u>	<u>==</u>
<u>Presence of >=1 neurosurgeon key author‡</u>	NS	--	--	--	--
<p>LOR: log odds ratio, NS: not significant, ***: p/q<0.01, **: p/q<0.05, ○: p/q<0.10 1. Accounting for data set and first author using random effects 2. Accounting for last author using random effects †After Benjamini-Hochberg false discovery rate correction for 11 comparisons ‡Grouped into medical oncologists (ref group), neurosurgeons, radiation oncologists, or other lvs absence of neurosurgery key author</p>					

Table 3: Impact of key medical academic genealogies on publication results

	Number of authors (% of dataset)	Number of unique articles (% of dataset)	Number of articles supporting maximal resection	Number of articles not supporting maximal resection
Identified medical academic genealogies†				
Genealogy A (neurosurgery)	14 (9%)	27 (25%)	26	1
Genealogy B (radiation oncology)	8 (5%)	14 (13%)	5	9
Both genealogies combined	22 (14%)	41 (38%)		
†See Figure 2 for genealogy membership				

Accepted Article

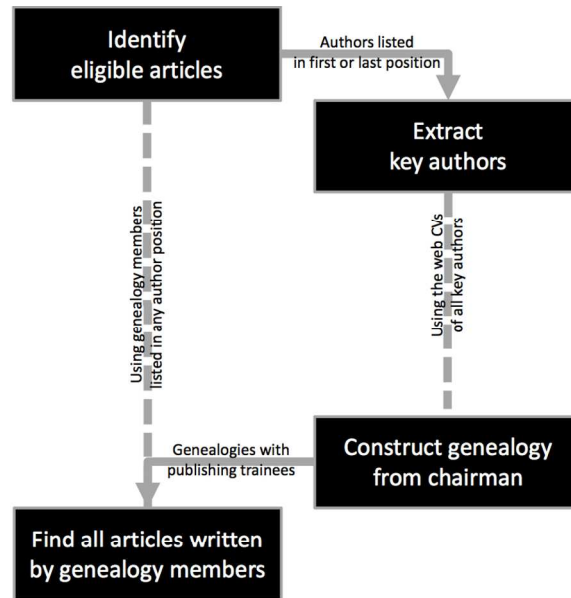


Figure 1: Genealogy generation. Genealogies were created using by linking key authors (first and last authors) of the identified literature based on who trained whom. Once genealogies were identified, all articles written by genealogy members were compiled and analyzed.
254x190mm (150 x 150 DPI)

Accep

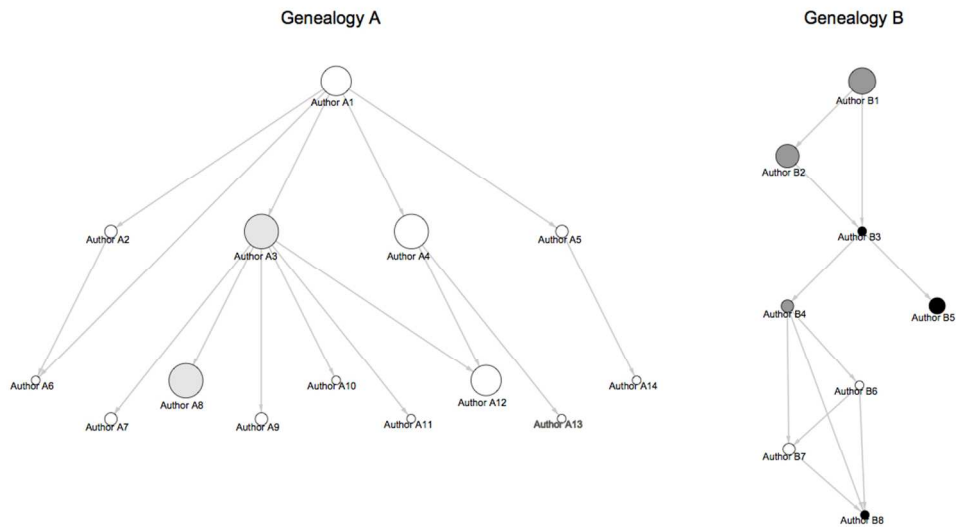


Figure 2: Authors from Genealogy A (left) and Genealogy B (right). Authors are colored white if all articles by that author support maximal resection, light gray if over half support maximal resection, dark gray if less than or equal to half support maximal resection, and black if none support maximal resection. Node size is proportional to the number of articles written.

196x106mm (150 x 150 DPI)

Accepte