ORIGINAL ARTICLE

Visualising Digital Pathology Research : A Bibliometric Analysis from 1991-2021

Rafidah Hod[1], Siti Khadijah Adam[1], Faridah Idris[2]

[1] Department of Human Anatomy, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia. 43400 Serdang, Selangor, Malaysia.

[2] Department of Pathology, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia. 43400 Serdang, Selangor, Malaysia

ABSTRACT

Introduction: Digital pathology encompasses the acquisition, management, sharing and interpretation of pathology information in a digital environment. Bibliometric analysis is a quantitative method to examine scholarly publications including the number of publications, citations, co-authorships, and collaboration network. Aim of this study is to provide a bibliometric analysis of academic documents on digital pathology (DP) from 1991-2021. **Methods:** The literature on digital pathology were obtained from the Scopus database. Frequency, percentage, data visualisation and citation metric were analysed using Microsoft Excel 365 and VOSviewer. **Results:** A total of 1848 documents from the Scopus database were analysed. There is a continuous growth of publications on DP with a total of 28330 citations. The United States was the most productive contributor to the publications followed by the United Kingdom and European countries, whilst University of Pittsburgh Medical Center, US produced the most publications. Progress in Biomedical Optics and Imaging Proceedings of SPIE was the largest source title while the Medical Image Analysis was the most prestigious journal. The keyword analysis suggests that DP research is mainly a medical imaging and engineering research domain with application in the histopathology subject. **Conclusion:** Digital Pathology research and publications continue to grow and concentrated in the Western countries. The publications focused on the image analysis, machine learning and engineering research domain in histopathology subject. Potential research areas include the implementation, validation of use and impact of DP to the pathology services and health care with exploration in other pathology subjects such as haematology.
Malaysian Journal of Medicine and Health Sciences (2022) 18(SUPP21): 43-53. doi:10.47836/mjmhs18.s21.8

Keywords: Digital pathology, Bibliometric analysis, Scopus, VOSviewer, Top articles

Corresponding Author:
Faridah binti Idris, MPath
Email: i_faridah@upm.edu.my
Tel: +6019-6548073

INTRODUCTION

Digital pathology (DP) is defined as a dynamic, image-based environment that allows pathology information to be acquired, managed, and interpreted from a digitised glass slide (1). This digital environment had changed the landscape of pathology services. Digital slides are easily stored, accessed and shared among pathologists for digital consultation and used for training and education for laboratory personnel. In the era of pandemic and remote working, digital pathology workflow serves as a useful and flexible working environment to ensure continuity of laboratory services and teaching while efficiently utilising the human resources involved (2).

Since the emergence of whole slide imaging (WSI) in the early 1990s, DP had transformed not only in the diagnostic services, but also had accelerated the advancement in research, quality management system, training and education (3). The diagnosis made with WSI are excellently correlated with the diagnosis made using traditional slides in different subspecialties in anatomical pathology including cytopathology (4).

With the advancement of machine learning and artificial intelligence, the digital collection of slides augments the development of algorithms for diagnosis and supports the diagnostic process (5). With this intelligent system, it aids pathologists to identify imaging markers that are related to specific diseases leading to faster detection, fine-tuning the prognosis and deciding the most effective treatment modality. This digital revolution of pathology, contributed to the progress of personalised medicine resulting in improved patient care, patient safety and overall health services (6).

Even though the research on WSI started more than twenty years ago, only a handful of its commercial platforms were endorsed recently by regulators, and this development accelerated the adoption of DP in laboratory services (7). There are still many more aspects to explore and investigate especially related to technical validation, safe implementation, cost-analysis, quality framework, adoption and medical education (8).

Bibliometric analysis refers to quantitative statistical techniques used to examine the properties of books and other forms of media communication (11). It is a useful tool to observe publication trends, impact and performance of articles such as citations per year and h-index. Based on the aim of the study, our focus is to analyse all research on DP related to diagnostic and health practices that are available in Scopus database as in April 2021.

The bibliometric analysis is a practical method to have an overall view of research and developments in this field. It measures the research impacts and trends, major key players, networking and interconnection of various literature documents, done by quantitatively analysing the scientific publications indexed in big bibliographic databases. It also provides various information including the progress in the area of research, the emerging topics of interest and future research directions (9). The search into the Scopus repository only found one bibliometric analysis on telepathology, a service component of digital pathology, hence lacking the coverage of other aspects of digital pathology research (10). This study aims to provide an overview of bibliometric analysis covering the development, key areas, main researchers and collaboration of DP research in the clinical practice areas. This information provides the latest trends of the current research, therefore give useful insight for researchers, research funders and policymakers to navigate the future research in the field of digital pathology.

MATERIALS AND METHODS

Data source and search strategy

We used "Digital Pathology" as the search term and all documents written in any language which contained this term in the title, abstract and keyword were selected. We included the abstract and keyword in the search to capture comprehensive literature in this field of study in all available languages. The search strategy is summarized in Fig.1.

Information extraction

Bibliometric analysis was performed on all the documents to give various information related to DP research. We used Microsoft Excel to analyse all selected publications to extract data on the frequencies, percentages (overall and by subjects), publication growth, most productive countries and total citations.

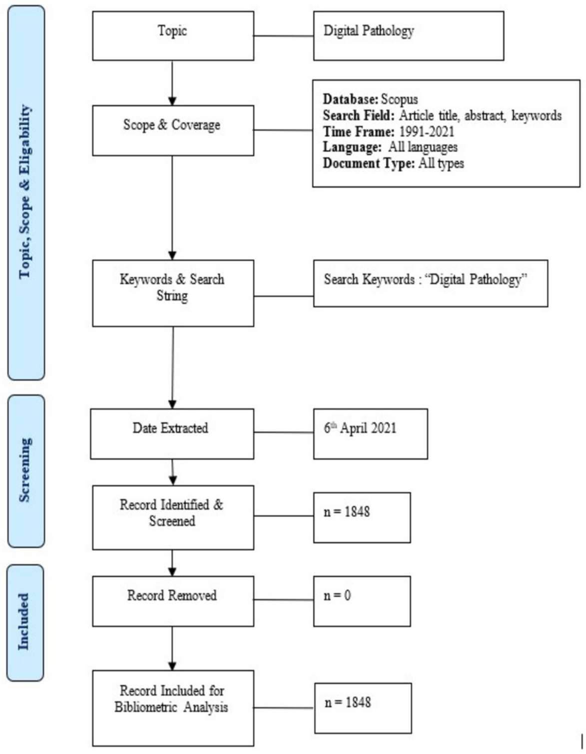

Figure 1: Flow Diagram of the search strategy

Subsequently, we refined our analysis to publications categorized under these selected subject areas: i) medicine, ii) biochemistry, genetics, molecular biology (BGMB), iii) health professions, iv) neuroscience, v) immunology and microbiology (IM), vi) dentistry and vii) nursing to focus on publications associated to the practice of medicine and its related clinical areas, to align the findings with our study aim. We used VOSviewer to visualise the publication networks and analysed the Scopus Metrics available in the Scopus database.

i. Visualisation of publication network

In this analysis, visualisation or bibliometric mapping is shown for indicators such as co-authorship among authors, co-authorship between countries and co-occurrences of author keywords. Visualisation mapping is presented with circles, labels and links between items, based on which item we want to explore, for example, co-authorship by countries, researchers or author keywords. The size of the label and circles are based on the weight of the items, the larger the size showed the stronger weight of the items. For example, the larger the circle of a country in the co-authorship by countries showing higher number of publications has been co-authored by that country. Similarly, the strength of the link between two items is indicated by the strength of co-authorship. In the case of co-authorship links between research, the number of publications that two researchers have co-authored determines the strength of the co-authorship links. For co-authorship mapping by countries and authors, documents co-authored by many countries (more than 20 countries per document) and authors (maximum of 25 authors) are excluded.

ii. Scopus Metrics

These metrics aim at assisting in the evaluations of authors, journals, and articles and provide a bird-eye view of research areas and valuable insights on the impact of the research output.

a. CiteScore 2019
This indicator measured the number of citations by a journal in 2019. It was obtained by dividing the number of citations for each of the five peer-reviewed document categories (articles, reviews, conference papers, data papers and book chapters) in 2016-2019, by the number of peer-reviewed papers indexed in Scopus and published during the same period (12).

b. SJR (SCImago Journal Rank) 2019
This indicator is weighted according to the journal's prestige. The subject area, quality, and reputation of the journal have an important influence on the citation value. SJR assigns relative scores to all the sources in a citation network on the basis that not all citations are equal (13). A citation from a source with a high SJR is more valuable than one from a source with a low SJR.

c. SNIP (Source Normalized Impact per Paper) 2019
This indicator measures a source's contextual impact. This is done by weighting citations according to the total number of citations in a subject field. It enables direct comparisons of sources in a variety of subject fields. SNIP is determined by how often authors cite publications in their reference list, how quickly citation impact matures, and how well the database used in the evaluation covered the field's literature. It is computed by the ratio between the average number of citations per paper for a source and the citation potential of its subject field. The citation potential of a source's subject field is calculated as the average number of references in each document that cites that source. It indicates the possibility of a document being cited in a particular field. A high citated source in a field typically has a high impact per paper (14).

RESULTS

We reviewed 1848 academic publications on DP during 1991-2021, published in 160 sources and written by a total of 159 authors. The majority of the documents were original articles, conference papers and review types with 53.03%, 26.84% and 11.53% of all publications respectively. The other 8.6% document types were book chapters, editorial, letters, conference reviews, notes, short surveys, erratum, books, data papers including one retracted publication. Majority of the publications originated from journals (68.24%), conference proceedings (21.1%), book series (8.06%), books (2.16%) and trade journals (0.43%) and were written in 13 different languages with English as the main language used (97.78%), while seven documents are available in more than one language.

Publication growth and citation trends

Presented in Figure 2 is the number of publications and citations as time progresses. The growth of DP publication can be seen starting in 2009 which rises steadily and exponentially from 2014 onwards. The total citations were 28330 with citations per year ranging from 0 to 6974 as shown by Fig 2. There was a sharp increase in the citation of the DP articles in 2009 with 1664 citations. The citation trend continues to increase and again led to a prominent citation peak in 2017 with 6974 citations. Following that is a drop in citation numbers for the years 2018 – 2021.

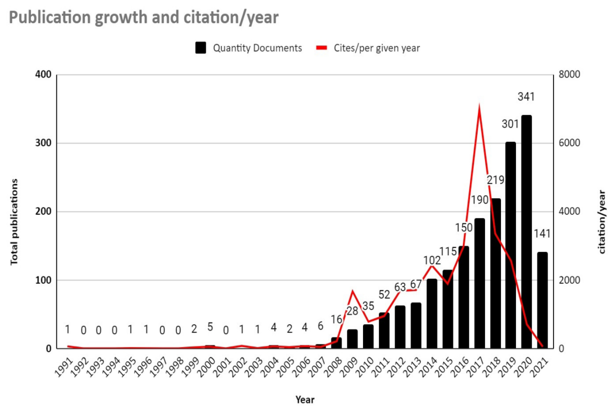

Figure 2: Publication growth and total citations of scientific publications in DP from 1991 – 2021

Subject Area

Publications on DP are found in various subject areas and one document may belong to more than one area. Top subject areas are Medicine (total publications (TP) = 1229, 34% of all publications) Computer Science (TP= 576, 16%), Biochemistry, Genetics and Molecular Biology (BGMB). (TP=388. 10.8%) and engineering (TP=323, 9%). The rest of the subject areas such as physics, astronomy, materials sciences and mathematics constituted less than 6% of the publications.

Most productive countries

Publications in DP involved 80 countries. The top 20 countries on DP publications in the selected subject areas are retrieved based on their total link strength. Most of the productive countries are western countries (US, UK, and European countries). Brazil and Columbia are from the South America continent and China, India and South Korea are countries from Asia. The highest volume of DP publications in the area was dominated by the United States (US) (45.91%) followed by the United Kingdom (UK) (13.69%), Canada (8.25%) and German (7.68%). The network visualisation among countries was also analysed. The minimum number of documents and citation of a country is set to 5. For each country, total strength of the co-authorship links with other countries is determined. We selected the countries with the greatest total link strength. Expectedly, United States has the most collaborations in publications with many countries followed by the United Kingdom. US was observed to build collaboration with Colombia beside European countries and Australia. Malaysia collaborated with Australia, China and the US.

Most productive institutions

Institutions with more than 25 DP publications are considered as the most productive and shown in Table I. Institutions are listed based on their total number of publications. These 15 institutions produce 35% of the total DP publications in the selected subject areas with an h-index ranging from 10-19. Nine of the institutions are in the United States. University of Pittsburgh Medical Center has the highest number of publications (TP=52), but the highest average citation per cited publication (C/CP) of 66.85 is achieved by Case Western Reserve University with lesser number of publications (TP=41). Most of the institutions have higher average citation per cited publication (C/CP) than average citation per publication. Most of the publications from these institutions are cited. Number of publications that are not cited (TP – NCP) for each institution ranges from 2 – 8.

Most active source titles

Data presented in Table II shows source titles, total publications, total citations, publisher, cite score,

Table I: Most productive institutions

Affiliation	Country	TP	NCP	TC	C/P	C/CP	*h*
University of Pittsburgh Medical Center	United States	52	46	794	15.27	17.26	17
Case Western Reserve University	United States	41	33	2206	53.80	66.85	19
University of Toronto	Canada	39	31	1088	27.90	35.10	13
National Cancer Institute NCI	United States	34	28	334	9.82	11.93	13
Harvard Medical School	United States	32	27	659	20.59	24.41	10
University of Leeds	United Kingdom	32	27	337	10.53	12.48	11
University of Pittsburgh	United States	31	24	751	24.23	31.29	12
University of Pennsylvania	United States	29	27	621	21.41	23.00	13
Inserm	France	28	25	1129	40.32	45.16	11
Emory University	United States	27	21	782	28.96	37.24	12
The Ohio State University	United States	27	23	383	14.19	16.65	11
Leeds Teaching Hospitals NHS Trust	United Kingdom	27	23	330	12.22	14.35	10
University Health Network University of Toronto	Canada	27	24	1334	49.41	55.58	13
University of Warwick	United Kingdom	27	24	825	30.56	34.38	14
Massachusetts General Hospital	United States	27	23	652	24.15	28.35	11

Notes: TP=total number of publications; NCP=number of cited publications; TC=total citations; C/P=average citations per publication (TC/TP); C/CP=average citations per cited publication (TC/NCP); h=h-index

Table II: Most Active Source Titles

Source Title	TP	TC	Publisher	Cite Score	SJR 2019	SNIP 2019
Progress In Biomedical Optics And Imaging Proceedings Of SPIE	134	647	The Society of Photo-Optical Instrumentation Engineers (SPIE)	1.1	0.269	0.298
Journal Of Pathology Informatics	76	1172	Wolters Kluwer Medknow Publications	5.5	1.005	1.26
Journal Of Clinical Pathology	38	297	BMJ Publishing Group	4.8	0.971	0.942
Proceedings International Symposium On Biomedical Imaging	36	378	IEEE Computer Society	n/a	n/a	n/a
Archives Of Pathology and Laboratory Medicine	34	851	College of American Pathologists	7.3	1.763	1.999
Toxicologic Pathology	30	155	SAGE Publications Inc.	2.9	0.63	0.849
Histopathology	27	663	Blackwell Publishing Ltd	6.3	1.424	1.466
Diagnostic Pathology	25	638	BioMed Central	3.6	0.827	0.971
Medical Image Analysis	23	4465	Elsevier B.V.	17.2	3.877	5.351
Cancers	22	82	MDPI AG	3.4	1.938	1.445
Computerized Medical Imaging And Graphics	21	531	Elsevier Ltd	6.6	1.035	2.056
IEEE Transactions on Medical Imaging	20	1005	Institute of Electrical and Electronics Engineers Inc.	16.6	3.276	4.569
Journal of Medical Imaging	20	303	The Society of Photo-Optical Instrumentation Engineers (SPIE)	3.4	0.798	1.008

Notes: TP= total number of publication; TC= total citations
seletected subjects = medicine, BGMB, health proferssions, neruoscience, immunology and microbiology (IM), denstistry and nursing

SJR2019 and SNIP 2019. Listed are the source titles with 20 and more publications in DP from the selected subject areas. Total publications from each source title range from 20 – 134. The source with the highest number of publications is Progress in Biomedical Optics and Imaging Proceedings of SPIE. Articles from Medical Image Analysis received the highest total citation of 4465 giving it the highest cite score of 17.2 from 23 of its total publications since its first DP publication in 2010.

Most influential articles

Top 20 publications in DP arranged according to the highest number of citations are shown in Table III. The citations range from 116 – 3313 total citations while cites per year range from 14.5 – 828.25. The most influential article - *A survey on deep learning in medical image analysis* received an average of 828.25 cites per year with a total of 3313 citations since it was first published in 2017.

Authorships and keywords

Table IV listed the most productive authors based on the total number of publications with more than 15 DP publications. The top three authors Pantanowitz,L., Madabhushi, A., and Treanor, D. contributed 10.5% of the total DP publications in the selected subject areas since 1991. Madabhushi, A is the most cited author mainly contributed by the article entitled: *Deep Learning for digital pathology Image Analysis: A Comprehensive tutorial with selected use cases* (3) with 419 citations as shown in Table III.

Figure 3 showed the co-authorship network and the authors' keyword co-occurrence mapping. Figure 3A displayed the co-authorship network from 1991-2021 in which author names are represented by the labels and circles, whereas the link between two authors is represented by a line between them. The size of the circles and labels indicate the numbers of the publications that had been co-authored by that researcher. The colour of an item is determined by the cluster to which the item belongs. This network shows prominent authors in digital pathology such as Pantanowitz L., Madabhushi A., Treanor D., and Hewitt S.M.

There are 2755 author keywords where 140 meet the threshold set for mapping the co-occurrence (minimum of 5 occurrences). For each of the 140 keywords, total strength of the co-occurrence links with other keywords were calculated. The top five keywords used in DP publications are "digital pathology" (total publication (TP) = 727), "deep learning" (TP = 124), "whole slide imaging" (TP=107), "image analysis" (TP=92) and "machine learning" (TP= 79). The keywords with the greatest total link strength were selected to create a map of the co-occurrence of author keywords as shown in Figure 3B.

DISCUSSION

In this era of Industry Revolution 4.0 (IR4.0) digital technology plays a vital role in the transformation of healthcare and has a powerful impact on how health services are delivered and managed (35). The application of digital solutions in healthcare is made timelier and more relevant in this era of pandemic and remote working to reduce risk of exposure and infection (36). In the field of pathology, the methodological concept of digital pathology has started ever since the usage of image from the microscopes onto photographic plates, followed by the practice of telepathology in the 1960s. However, the wider acceptance of digital pathology only gained momentum when the whole slide imaging (WSI) technology entered the commercial market in the 1990s (37). The improvement of the WSI technology and the accessibility to an enormous number of digital images in pathology for the past 20 years had facilitated the development of artificial intelligence and machine learning in the field of DP (4).

The publications in DP started in 1991 but the growth was relatively slow for the first 15 years. The first publication on DP appeared in the Journal of Pathology, discussing image analysis. For the following 10-15 years, the related publications remained minimal. This might be due to the duration needed to secure the patent and relatively slow progress of the technology advancement and digital economy in the late 1990s and 2000s (4). However, the publications started to pick up in 2008 and kept increasing until the beginning of 2021. In 2020, more than 300 publications had been published and until April 2021, more than one hundred publications had been indexed in the Scopus. This indicates increasing interest and rapid development in this field particularly from the Western countries, as shown by data of the most productive countries, and English being the most common language of publication. Only a handful of other countries outside the US and Europe contributed to a significant publication output in DP and this included China, Japan and Australia.

This is in congruence with a report that listed mainly cities in US and Europe as the global innovation hub according to the Global Innovation Index 2020(38). Even though the data is still preliminary and not specific to the development of DP technology, it still gives us valuable insight into the technological development of the Western countries in general. Expectedly, the US and the UK were the main countries that were involved in the co- authorship with other countries. Authors from the US were observed to have close collaboration with Colombia, few European countries and Australia. Collaboration from Malaysia can be seen with few countries like Australia, China and the US. With the occurrence of Covid-19 and the need of flexible working environment during pandemic, it is highly likely that DP

Table III: Highly Cited Articles from Selected Subject Areas

Authors	Title/Reference	Year	Cites	Cites per Year
Litjens, G., Kooi, T., Bejnordi, B.E., (...), van Ginneken, B., S6nchez, C.I.	A survey on deep learning in medical image analysis (16)	2017	3313	828.25
Pagès, F., Mlecnik, B., Marliot, F., (...), Fox, B.A., Galon, J.	International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study (17)	2018	573	191
Janowczyk, A., Madabhushi, A.	Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases (18)	2016	419	83.8
Xu, J., Xiang, L., Liu, Q., (...), Tang, J., Madabhushi, A.	Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images (19)	2016	404	80.8
Madabhushi, A., Lee, G.	Image analysis and machine learning in digital pathology: Challenges and opportunities (20)	2016	276	55.2
Pantanowitz, L., Sinard, J.H., Henricks, W.H., (...), Lal, A., Parwani, A.V.	Validating whole slide imaging for diagnostic purposes in Pathology: Guideline from the College of American pathologists Pathology and Laboratory Quality Center (21)	2013	267	33.38
Ghaznavi, F., Evans, A., Madabhushi, A., Feldman, M.	Digital imaging in pathology: Whole-slide imaging and beyond (22)	2013	226	28.25
Rizzardi, A.E., Johnson, A.T., Vogel, R.I., (...), Metzger, G.J., Schmechel, S.C.	Quantitative comparison of immunohistochemical staining measured by digital image analysis versus pathologist visual scoring (23)	2012	224	24.89
Veta, M., van Diest, P.J., Willems, S.M., (...), Viergever, M.A., Pluim, J.P.W.	Assessment of algorithms for mitosis detection in breast cancer histopathology images (24)	2015	222	34.5
Weinstein, R.S., Graham, A.R., Richter, L.C., (...), Yagi, Y., Gilbertson, J.R.	Overview of telepathology, virtual microscopy, and whole slide imaging: prospects for the future (25)	2009	218	17.25
Cruz-Roa, A., Basavanhally, A., Gonz6lez, F., (...), Tomaszewski, J., Madabhushi, A.	Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks (26)	2014	207	29.57
Saltz, J., Gupta, R., Hou, L., (...), Sharma, A., Thorsson, V.	Spatial Organization and Molecular Correlation of Tumor-Infiltrating Lymphocytes Using Deep Learning on Pathology Images (27)	2018	193	64.33
Sirinukunwattana, K., Pluim, J.P.W., Chen, H., (...), Snead, D.R.J., Rajpoot, N.M.	Gland segmentation in colon histology images: The glas challenge contest (28)	2017	188	47
Al-Janabi, S., Huisman, A., Van Diest, P.J.	Digital pathology: Current status and future perspectives (29)	2012	186	20.67
Tsujikawa, T., Kumar, S., Borkar, R.N., (...), Flint, P.W., Coussens, L.M.	Quantitative Multiplex Immunohistochemistry Reveals Myeloid-Inflamed Tumor-Immune Complexity Associated with Poor Prognosis (30)	2017	174	43.5
Wang, H., Cruz-Roa, A., Basavanhally, A., (...), Gonzalez, F., Madabhushi, A.	Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features (31)	2014	162	23.14
Ali, S., Madabhushi, A.	An integrated region-, boundary-, shape-based active contour for multiple object overlap resolution in histological imagery (32)	2012	159	17.67
Xu, Y., Jia, Z., Wang, L.-B., (...), Lai, M., Chang, E.I.-C.	Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features (33)	2017	143	35.75
Clapper, J.R., Hendricks, M.D., Gu, G., (...), Lowe, C., Roth, J.D.	Diet-induced mouse model of fatty liver disease and nonalcoholic steatohepatitis reflecting clinical disease progression and methods of assessment (34)	2013	120	15
Assayag, O., Grieve, K., Devaux, B., (...), Boccara, C., Varlet, P.	Imaging of non-tumorous and tumorous human brain tissues with full-field optical coherence tomography (35)	2013	116	14.5

Table IV: Most Productive Authors from selected subject areas

Author's Name	No. of Documents	Percentage (%)	h-index	TC	Publication Year (No. of articles)
Pantanowitz, L	62	4.49	18	1037	2021(4), 2020(10), 2019(9), 2018(12), 2017(10), 2016(7), 2015(3), 2014(2), 2013(3), 2012(2)
Madabhushi, A	51	3.69	24	3071	2021(2), 2020(2), 2019(4), 2018(3), 2017(7), 2016(6), 2015(6), 2014(4), 2013(3), 2012(2), 2011(8), 2010(4) 2009(1)
Treanor, D.	31	2.24	10	378	2021(1), 2020(5), 2019(7), 2018(3), 2017(6), 2016(2), 2015(3) 2014(2), 2013(1), 2012(1)
Parwani, A.V	18	1.30	11	683	2020(1), 2019(4) 2018(2), 2016(2), 2015(1), 2014(2), 2013(2), 2012(1), 2011(1), 2008(2)
Hewitt, S.M.	17	1.23	8	198	2021(2), 2020(1), 2019(3), 2018(1), 2017(2), 2016(2), 2015(2), 2014(3), 2013(1)
Salto-Tellez, M.	17	1.23	11	336	2021(1), 2020(4), 2019(3), 2018(2), 2017(1), 2016(1), 2015(3), 2014(1), 2013(1)
Yagi, Y.	17	1.23	9	434	2020(1), 2019(4), 2018(2), 2015(1), 2012(6), 2011(1), 2009(1), 2008(1)
Kurc, T.	16	1.16	7	385	2020(4), 2019(4), 2018(2), 2017(1), 2016(3), 2013(1), 2011(1)

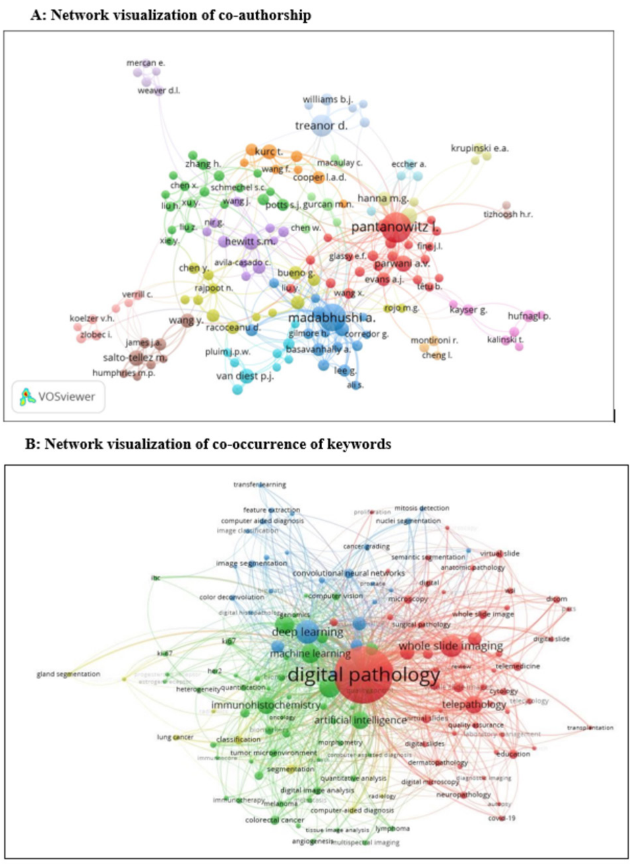

Figure 3: Network visualisation map
Note: Unit of analysis = Authors(A), Keywords (B)
Counting method: Fractional counting
Minimum number of i) documents of an author (A), ii) occurrence of keywords(B) = 5
Minimum number of citations of an author (A) = 5

will play a more important role in the pathology services and boost publication growth and collaboration in the near future (39,40).

Most of the publications were found in the journal and conference proceedings and commonly occurred in the area of medicine, computer science, BGMB and engineering. The field of digital pathology mainly related to medicine and computer science hence that explains why these two fields are among the common subject areas in this analysis. Out of a total of 160 sources of publication, the top 20 sources published around 27% of the total number of publications. Unsurprisingly, most of the active sources are related to the field of medical imaging and only a few appeared under sources directly related to pathology and laboratory services. First on the list is Progress in Biomedical Optics and Imaging - Proceedings of SPIE, which fall under the subject area of engineering focusing, among others, on radiology, nuclear medicine and imaging. SPIE (The Society of Photo-Optical Instrumentation Engineers) organised many conferences each year to showcase and share the findings related to photo-optical research advancement including medical imaging. Digital pathology is one of the dedicated categories during the conference and produced 20-40 conference papers each year (41). At that rate, it is understandable this publication became the most active source of publication in DP. SPIE is also the publisher of the Journal of Medical Imaging which contributed 20 publications in the field during the period of analysis.

More than half of the publications in the active sources belong to medical imaging and computer science compared to the source in the clinical pathology field. Analysis of frequency and clustering of keywords also confined mainly within the domain of computer science, imaging and engineering. There are four main clusters, and two major clusters have keywords mostly related to computer science and imaging subject areas. The top five keywords which hold the bulk of the total publications are i) digital pathology ii) deep learning iii) whole slide imaging iv) image analysis and v) machine learning, which is evident in the visualisation map. In the map, the nodes in the same colour belong to the same clusters and share similar keywords. Digital pathology, being the core keyword, is linked to almost all the main keywords listed, particularly has a strong link to whole image analysis, machine learning and deep learning. Keyword linkage analysis also suggested that the application of DP research mainly occurred in the area of immunohistochemistry and cancer.

All this data indicates that the progress of research is active at the level of technological and algorithm advancement in the histopathology subject of pathology. There is potential research in the area of the utilization and standardization of DP systems in the laboratory services particularly as part of remote digital diagnosis which is not fully explored (42, 8). In the current situation of pandemic and remote working, it is likely that more laboratory services will integrate DP in their existing workflow, and this opens more potential research opportunities related to the experience of users, impact on the laboratory services and adherence to the quality requirement. It is also observed that research on digital images in pathology was heavily explored in histopathology which deals with solid tissues and very few on other areas of pathology, for example in haematology that also use glass slides for diagnosis of blood disorders (43). This also creates a huge potential opportunity of DP research for haematology diagnosis in the future.

Analysis showed that most productive institutions are mainly from the US, Canada, UK and France, which is in concordance with the most active countries involved in the DP research and publications. The listed institutions are quite different from the previous bibliometric analysis on telepathology even though there are a few overlapping institutions from our list. This is most likely because of the different focus on the keyword during the search strategy and the database used. We used Scopus database instead of Web of Science used by Senel & Bas (10)

University of Pittsburgh Medical Centre (UPMC), an affiliate with the School of the Health Sciences, University of Pittsburgh leads the list with the highest number of publications. UPMC has actively utilised DP technology in their clinical services and has a dedicated DP portal for pathology consultation for histopathology diagnosis worldwide (44). Second institution in the list is Case Western Reserve University (CRWU), a university in Cleveland, US. School of Engineering, CRWU, through its Centre of Computational Imaging and Personalised Diagnostic (CCIPD) has DP as one of their main research areas which focuses on the image analytic methods for diagnosis and prognosis in common solid organ histopathology (45). With this active research and wide application in histopathology clinical services, made both institutions being the top two contributors to publications related to DP.

Liron Pantanowitz (Pantanowitz, L), Anant Madabhushi (Madabhushi, A) and Darrean Treanor (Treanor, D) are among the top three most prolific authors with high citations and h-index. Pantanowitz, L had published a total of 62 publications, followed by Madabhushi, A with 51 and Treanor D had 31 publications. Pantanowitz L is a renowned pathologist with interest in pathology informatics meanwhile Madabhushi, A, is a faculty in Biomedical Engineering, CRWU. Madabhushi has the highest total citations and h-index even though he published eleven papers less than Pantanowitz, L. Six articles co-authored by him were among the highly cited journal articles hence explaining his high citations and h-index. Most of the top productive authors in our list are not listed in the previous analysis on telepathology, except a few (10). This showed that the main research groups for DP and telepathology were from different researchers even though telepathology is a part of DP framework.

In the co-authorship visualisation map, the colour represents the cluster and the authors who commonly work together will be in the same cluster as indicated by the same colour. The map suggests that the top productive authors had their own unique research group and collaboration. Pantanotwitz, L and Madabhushi, A apparently did not co-author with each other and their co-authors hardly overlapped. The collaboration is mostly from institutions within North America.

Citation pattern on digital pathology publications

Though the number of publications increases yearly, citation numbers seem to drop for 2017/18 onwards. However, this is expected as studies published in recent years may take time to make an impact. This is supported by Adams (2005) who commented that citation statistics produced may not be stable for studies in the past three years (46). Recent studies are still building up their citations as it takes time for research. Therefore, the low citation numbers may still increase given more time.

As expected, The US has the highest citations compared to other countries as researchers from the US are actively doing DP research and have published almost half from all the publications related to DP. UPMC, the institution which produced the highest number of publications, had received citation to 88% out of its 52 publications compared to CWRU which received citations to 80% of its publication. However, Case Western Reserve University shows greater citation power with highest total citations, h- index and more than 3-fold average citations per cited publication which indicates its impactful research on DP compared to other institutions.

Analysing the most influential titles in DP publications based on the number of citations shows that most research related to machine learning, assessment of image analysis and algorithms are the main articles being cited. Two of the top cited articles are review papers (15, 28) which are expected to have high citations. The most cited article (15) is a comprehensive review article which gives a summary to the deep learning algorithms utilised in medical image analysis and points out the successful ones by analysing hundreds of papers that used deep learning in various fields. There is also one conference paper (25) and one editorial (19) being highly cited. In general, editorials seldomly get cited, however in this study an editorial is among the top 5 highly cited publications. That editorial was published in a special issue, in conjunction with the journal's 20th anniversary that collected articles from its editorial board members. The editorial write-up related to DP in this issue was written by one of the prominent authors in DP and most likely contributed to its high citation.

Bibliometric, which is a quantitative analysis of scholarly publications, seeks to demonstrate their impact on academic and public conversion, one of which is to analyse the citation pattern of the publications. However, we should interpret it with care. Articles with high citations do not always mean high quality research. An article with low quality may still be highly cited when other researchers are trying to challenge or refute the findings. Furthermore, building up citation of a publication needs time even though most publications reach the peak of its citation within 2 years of publication. Other research metrics can be used to give a more meaningful overview of the impact and influence of a research output. (47)

Scopus metrics such as CiteScore, SJR and SNIP are useful to measure the performance and prestige of a journal and give a guide to researchers to choose the most appropriate journal for their publications. CiteScore provides more accurate indication of a journal's impact, whilst SNIP measures citation impact between-field-comparison more rigorously (12, 13). SJR used weight of prestige to quantify citations of a journal and helpful to compare journals within the same discipline (14). As presented in table IV, Medical Image Analysis journal has the highest score of all three metrics making it the most influential and impactful journal on DP research followed by IEEE Transaction of Medical Imaging. This indicates that DP is mainly a medical imaging research domain which applications fall under pathology subject area, particularly histopathology. It is also interesting to note that these two titles did not produce as many publications as other top source titles but outrank others in terms of impact and prestige. The sources with more publications in an area of research do not necessarily mean that they are more influential and produce more impactful publications. However, using bibliometric coupled with Scopus metric in this analysis provided a more holistic outlook of publication sources and research output related to digital pathology.

CONCLUSION

Our analysis resulted in 1848 publications on DP obtained from the Scopus database for the period of 1991-2021. Out of those, 1381 documents were from selected subjects to focus on the practice of medicine and its related clinical areas. The research and publication on DP are concentrated in the Western countries and mainly in the medical imaging and engineering research domain which has its applications in histopathology. Most of the research was focused on the technological, development, machine learning and assessment of algorithm of image analysis. With more laboratories likely to adopt DP in their existing practice due to pandemic and acceptance of remote working, there will be potential research opportunities in other areas such as user experience, validation and standardization of usage and its impact on the laboratory and health services.

REFERENCES

1. DPA. About Digital Pathology: Digital Pathology Association; 2020 [cited 2021. Available from: https://digitalpathologyassociation.org/about-digital-pathology.
2. Stathonikos N, van Varsseveld NC, Vink A, van Dijk MR, Nguyen TQ, Leng WWJ, et al. Digital pathology in the time of corona. J Clin Pathol. 2020;73(11):706-12. doi: 10.1136/jclinpath-2020-206845.
3. Pantanowitz L. Digital images and the future of digital pathology. Journal of Pathology Informatics. 2010;1(1):15. doi: 10.4103/2153-3539.68332.
4. Pantanowitz L, Sharma A, Carter AB, Kurc T, Sussman A, Saltz J. Twenty Years of Digital Pathology: An Overview of the Road Travelled, What is on the Horizon, and the Emergence of Vendor-Neutral Archives. J Pathol Inform. 2018;9:40. doi: 10.4103/jpi.jpi_69_18
5. Parwani AV. Next generation diagnostic pathology: use of digital pathology and artificial intelligence tools to augment a pathological diagnosis. Diagn Pathol. 2019;14(1):138. doi: 10.1186/s13000-019-0921-2.
6. Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. Lancet Oncol. 2019;20(5):e253-e61. doi: 10.1016/S1470-2045(19)30154-8.
7. Administration USFD. U.S. Food and Drug Administration. 2021. Digital Pathology Program: Research on Digital Pathology Medical Devices. 2021.
8. Jahn SW, Plass M, Moinfar F. Digital Pathology: Advantages, Limitations and Emerging Perspectives. J Clin Med. 2020;9(11). doi: 10.3390/jcm9113697.
9. Gutiérrez-Salcedo M, Martínez M, Moral-Munoz J, Herrera-Viedma E, Cobo M. Some bibliometric procedures for analysing and evaluating research fields. Applied Intelligence. 2018;48:1275-87. doi: 10.1007/s10489-017-1105-y
10. Senel E, Bas Y. Evolution of Telepathology: A Comprehensive Analysis of Global Telepathology Literature Between 1986 and 2017. Turk Patoloji Derg. 2020;36(3):218-26. doi: 10.5146/tjpath.2019.01484
11. Pritchard A. Statistical bibliography or bibliometrics. Journal of Documentation. 1969;25(4):348–9.
12. CiteScore Journal Metric - FAQs - Scopus: Access and use Support Center [Internet]. Service.elsevier.com. 2021 [cited 4 May 2021]. Available from: https://service.elsevier.com/app/answers/detail/a_id/30562/supporthub/scopus/
13. How is SJR (SCImago Journal Rank) used in Scopus? - Scopus: Access and use Support Center [Internet]. Service.elsevier.com. 2021 [cited 4 May 2021]. Available from: https://service.elsevier.com/app/answers/detail/a_id/14883/c/10546/supporthub/scopus/related/1/
14. How is SNIP (Source Normalized Impact per Paper) used in Scopus? - Scopus: Access and use Support Center [Internet]. Service.elsevier.com. 2021 [cited 4 May 2021]. Available from: https://service.elsevier.com/app/answers/detail/a_id/14884/c/10546/supporthub/scopus/related/1/
15. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017;42:60-88. doi: 10.1016/j.media.2017.07.005.
16. Pages F, Mlecnik B, Marliot F, Bindea G, Ou FS, Bifulco C, et al. International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. Lancet. 2018;391(10135):2128-39. doi: 10.1016/S0140-6736(18)30789-X.
17. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. J Pathol Inform. 2016;7:29. doi: 10.4103/2153-3539.186902
18. Xu J, Xiang L, Liu Q, Gilmore H, Wu J, Tang J, et al. Stacked Sparse Autoencoder (SSAE) for Nuclei Detection on Breast Cancer Histopathology

Images. IEEE Trans Med Imaging. 2016;35(1):119-30. doi: 10.1109/TMI.2015.2458702
19. Madabhushi A, Lee G. Image analysis and machine learning in digital pathology: Challenges and opportunities. Med Image Anal. 2016;33:170-5. doi: 10.1016/j.media.2016.06.037.
20. Pantanowitz L, Sinard JH, Henricks WH, Fatheree LA, Carter AB, Contis L, et al. Validating whole slide imaging for diagnostic purposes in pathology: guideline from the College of American Pathologists Pathology and Laboratory Quality Center. Arch Pathol Lab Med. 2013;137(12):1710-22. doi: 10.5858/arpa.2013-0093-CP.
21. Ghaznavi F, Evans A, Madabhushi A, Feldman M. Digital imaging in pathology: whole-slide imaging and beyond. Annu Rev Pathol. 2013;8:331-59. oi: 10.1146/annurev-pathol-011811-120902.
22. Rizzardi AE, Johnson AT, Vogel RI, Pambuccian SE, Henriksen J, Skubitz AP, et al. Quantitative comparison of immunohistochemical staining measured by digital image analysis versus pathologist visual scoring. Diagn Pathol. 2012;7:42. doi: 10.1186/1746-1596-7-42.
23. Veta M, van Diest PJ, Willems SM, Wang H, Madabhushi A, Cruz-Roa A, et al. Assessment of algorithms for mitosis detection in breast cancer histopathology images. Med Image Anal. 2015;20(1):237-48. doi: 10.1016/j.media.2014.11.010.
24. Weinstein RS, Graham AR, Richter LC, Barker GP, Krupinski EA, Lopez AM, et al. Overview of telepathology, virtual microscopy, and whole slide imaging: prospects for the future. Hum Pathol. 2009;40(8):1057-69. doi: 10.1016/j.humpath.2009.04.006
25. Cruz-Roa A, Basavanhally A, Gonzalez F, Gilmore H, Feldman M, Ganesan S, et al., editors. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. Progress in Biomedical Optics and Imaging - Proceedings of SPIE; 2014: SPIE. doi: 10.1117/12.2043872
26. Saltz J, Gupta R, Hou L, Kurc T, Singh P, Nguyen V, et al. Spatial Organization and Molecular Correlation of Tumor-Infiltrating Lymphocytes Using Deep Learning on Pathology Images. Cell Rep. 2018;23(1):181-93 e7. doi: 10.1016/j.celrep.2018.03.086
27. Sirinukunwattana K, Pluim JPW, Chen H, Qi X, Heng PA, Guo YB, et al. Gland segmentation in colon histology images: The glas challenge contest. Med Image Anal. 2017;35:489-502. doi: 10.1016/j.media.2016.08.008.
28. Al-Janabi S, Huisman A, Van Diest PJ. Digital pathology: current status and future perspectives. Histopathology. 2012;61(1):1-9. doi: 10.1111/j.1365-2559.2011.03814.x.
29. Tsujikawa T, Kumar S, Borkar RN, Azimi V, Thibault G, Chang YH, et al. Quantitative Multiplex Immunohistochemistry Reveals Myeloid-Inflamed Tumor-Immune Complexity Associated with Poor Prognosis. Cell Rep. 2017;19(1):203-17. doi: 10.1016/j.celrep.2017.03.037.
30. Wang H, Cruz-Roa A, Basavanhally A, Gilmore H, Shih N, Feldman M, et al. Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features. J Med Imaging (Bellingham). 2014;1(3):034003. doi: 10.1117/1.JMI.1.3.034003.
31. Ali S, Madabhushi A. An integrated region-, boundary-, shape-based active contour for multiple object overlap resolution in histological imagery. IEEE Trans Med Imaging. 2012;31(7):1448-60. An integrated region-, boundary-, shape-based active contour for multiple object overlap resolution in histological imagery
32. Xu Y, Jia Z, Wang LB, Ai Y, Zhang F, Lai M, et al. Large scale tissue histopathology image classification, segmentation, and visualisation via deep convolutional activation features. BMC Bioinformatics. 2017;18(1):281. doi: 10.1186/s12859-017-1685-x
33. Clapper JR, Hendricks MD, Gu G, Wittmer C, Dolman CS, Herich J, et al. Diet-induced mouse model of fatty liver disease and nonalcoholic steatohepatitis reflecting clinical disease progression and methods of assessment. Am J Physiol Gastrointest Liver Physiol. 2013;305(7):G483-95. doi: 10.1152/ajpgi.00079.2013.
34. Assayag O, Grieve K, Devaux B, Harms F, Pallud J, Chretien F, et al. Imaging of non-tumorous and tumorous human brain tissues with full-field optical coherence tomography. Neuroimage Clin. 2013;2:549-57. doi: 10.1016/j.nicl.2013.04.005.
35. Digital technologies: shaping the future of primary health care [Internet]. Geneva: World Health Organizaton; 2018 [cited 22 April 2021]. Available from: https://www.who.int/docs/default-source/primary-health-care-conference/digital-technoloGies.pdf? sfvrsn=3efc47e0_2
36. Javaid M, Haleem A, Vaishya R, Bahl S, Suman R, Vaish A. Industry 4.0 technologies and their applications in fighting COVID-19 pandemic. Diabetes & Metabolic Syndrome: Clinical Research & Reviews. 2020;14(4):419-422. doi: 10.1016/j.dsx.2020.04.032.
37. Sarah Moore M. History of Digital Pathology [Internet]. News-Medical.net. 2021 [cited 22 April 2021]. Available from: https://www.news-medical.net/life-sciences/History-of-Digital-Pathology.aspx
38. Global Innovation Hubs Index 2020 [Internet]. Nature.com. 2021 [cited 23 April 2021]. Available from: https://www.nature.com/articles/d42473-020-00535-9
39. Cimadamore A, Lopez-Beltran A, Scarpelli M, Cheng L, Montironi R. Digital pathology and COVID-19 and future crises: pathologists can safely diagnose cases from home using a consumer

monitor and a mini PC. Journal of Clinical Pathology. 2020;73(11):695-696. doi: 10.1136/jclinpath-2020-206943
40. Browning L, Fryer E, Roskell D, White K, Colling R, Rittscher J et al. Role of digital pathology in diagnostic histopathology in the response to COVID-19: results from a survey of experience in a UK tertiary referral hospital. Journal of Clinical Pathology. 2020;74(2):129-132. doi: 10.1136/jclinpath-2020-206786
41. Browse Proceedings [Internet]. Spiedigitallibrary.org. 2021 [cited 4 May 2021]. Available from: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/browse/SPIE-Medical-Imaging/2021
42. Hanna M, Reuter V, Ardon O, Kim D, Sirintrapun S, Schüffler P et al. Validation of a digital pathology system including remote review during the COVID-19 pandemic. Modern Pathology. 2020;33(11):2115-2127. doi: 10.1038/s41379-020-0601-5
43. VanVranken S, Patterson E, Rudmann S, Waller K. A Survey Study of Benefits and Limitations of using CellaVision™ DM96 for Peripheral Blood Differentials. American Society for Clinical Laboratory Science. 2014;27(1):32-39.
44. UPMC Pathology Consultation Services [Internet]. UPMC | Life Changing Medicine. 2021 [cited 30 April 2021]. Available from: https://www.upmc.com/healthcare-professionals/physicians/pathology-consultation
45. Digital Pathology | Center for Computational Imaging and Personalised Diagnostics [Internet]. Engineering.case.edu. 2021 [cited 30 April 2021]. Available from: https://engineering.case.edu/centers/ccipd/research/digital-pathology
46. Adams J. Early citation counts correlate with accumulated impact. Scientometrics. 2005;63(3):567-81. doi: 10.1177/01455613211042113.
47. Citation analysis & bibliometrics - Library, University of York [Internet]. York.ac.uk. 2021 [cited 4 May 2021]. Available from: https://www.york.ac.uk/library/info-for/researchers/citation/#tab-3