



## AOA Critical Issues in Education

# Doximity Orthopaedic Surgery Program Rankings Are Associated with Academic Productivity

Patrick J. Meade, BA, Sheena J. Amin, MD, Michaela A. Stamm, MS, and Mary K. Mulcahey, MD, FAOA

**Background:** Doximity residency rankings are used to compare training programs. Reputation is a component of Doximity rankings and is based on physician surveys. Outcomes-based measures allow for objective program comparison and may be more appropriate when assessing programs. This study evaluated how Doximity rankings of orthopaedic surgery residency programs compare with an outcomes-based ranking of programs based on academic productivity. This study also evaluated whether program size and type were associated with academic productivity.

**Methods:** Orthopaedic program rankings, size, type, and research productivity were recorded from Doximity. An academic productivity score was calculated by averaging the mean percentage of alumni clinical trials and alumni publications. Analysis of variance and post hoc analyses were performed to determine whether academic productivity was associated with program reputation, size, and type.

**Results:** One hundred seventy-five orthopaedic residency programs were included. Program rankings, size, and type had a statistically significant association with academic productivity ( $p < 0.0001$ ). The 44 orthopaedic programs in the top quartile had an academic productivity score of  $79.1 \pm 13$  in comparison with programs in the bottom quartile, who had a score of  $38.3 \pm 12.1$  ( $p < 0.0001$ ). Of the programs in the top quartile of reputation rankings, 32 (73%) were also in the top quartile for academic productivity. Programs with more residents demonstrated greater academic productivity than smaller programs. Twenty-three programs (13%) had more than 30 residents and a mean academic productivity of  $76.2 \pm 14.8$ . In comparison, 13 programs (7.5%) had 10 or fewer residents with an academic productivity score of  $40.6 \pm 16.3$  ( $p < 0.0001$ ). An academic productivity score differed significantly based on program type: university ( $63.9 \pm 14.8$ ), military ( $54.7 \pm 16.9$ ), university/community ( $46.3 \pm 15.6$ ), and community ( $38.7 \pm 15$ ) ( $p < 0.0001$ ).

**Conclusions:** Academic productivity among orthopaedic residency programs is associated with Doximity ranking, program size, and type. This information can be used by medical students, residents, and physicians to understand the Doximity rankings and how they correlate with this objective measure.

**Level of Evidence:** III.

IRB approval for this study was not required because no patient data were used. All data included in this study were available on the Doximity website and through other public sources.

**Disclosure:** The **Disclosure of Potential Conflicts of Interest** forms are provided with the online version of the article (<http://links.lww.com/JBJSOA/A453>).

Copyright © 2023 The Authors. Published by The Journal of Bone and Joint Surgery, Incorporated. All rights reserved. This is an open access article distributed under the terms of the [Creative Commons Attribution-Non Commercial-No Derivatives License 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/) (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

## Introduction

Doximity is a social network that includes a comprehensive database of more than 1 million US physicians<sup>1</sup>. Since 2014, the website has published residency program rankings for 28 different specialties, including orthopaedic surgery. Residency program rankings provide medical students, residents, and attendings a quantitative reference point with which to compare training programs<sup>2</sup>. The highly visible Doximity rankings are especially influential among medical students who are applying for residency. In 2017, 62% of all applicants to 24 Mayo Clinic residency programs reported that they referenced Doximity rankings during the application process, and 79% reported that the rankings affected their residency applications, including the decision to accept an interview invitation, and match lists<sup>3</sup>.

Doximity program rankings are formulated from 3 components: (1) reputation surveys completed by board-certified physicians in the corresponding specialty, (2) resident and recent alumni satisfaction surveys, and (3) an objective measure of alumni research output, which includes alumni publications and alumni clinical trial participation<sup>2,4</sup>. Because of the subjectivity of some of these components, the Doximity ranking system is not a completely objective evaluation of residency programs. Even the objective components may not always be accurate. One ophthalmology program published a letter to the editor to report that the information on their Doximity program page did not correlate with their internal data, which raised concern about the validity of Doximity data<sup>5</sup>. Furthermore, a recent study by Feinstein et al. found that larger program size was predictive of a higher Doximity rank across 16 specialties, which suggests that larger programs may rank higher by virtue of having more alumni to fill out the surveys<sup>5,6</sup>. Wilson et al. found that Doximity rankings only loosely correlated with board pass rate and alumni publications among general surgery programs, raising concerns about the validity of these subjective reputation surveys as a measure of program quality<sup>7</sup>. A similar result was found when comparing the Doximity program rankings in dermatology with an objective ranking that uses a different set of academic measures to generate a ranking of dermatology residency programs<sup>7</sup>. Only half of the dermatology programs appeared in both the top 20 of the Doximity rankings and the top 20 of the objective dermatology ranking<sup>7</sup>.

Because Doximity rankings often impact medical student, resident, and physician perception of training programs, it is worthwhile to evaluate how these rankings compare against other objective standards. Subjective physician surveys are a major component of Doximity rankings, so it is uncertain whether they accurately gauge program quality. Outcomes-based measures may provide a more objective and accurate means of program assessment. Therefore, the purpose of this study was to evaluate how Doximity rankings of orthopaedic surgery residency programs compare with an outcomes-based ranking of programs based on academic productivity. This study also evaluated whether program size and type were associated with academic productivity.

## Materials and Methods

An observational study was performed in August 2021 to evaluate how Doximity orthopaedic program rankings, size, and type correlated to academic productivity. Publicly available Doximity data were used to collect program rank, size, mean percent of alumni clinical trials, and mean percent of alumni publications. Data on program type (university, university/community, community, or military) were collected from each program's page in the Fellowship and Residency Electronic Interactive Database Access System.

All 201 orthopaedic residency programs in the United States were assessed. Twenty-six programs were excluded because of missing data, leaving 175 programs for inclusion in the study. Programs were excluded from the study if they were missing one of the metrics used to calculate academic productivity. The mean academic productivity score and the SD from the mean were calculated for each program by averaging the mean percentage of alumni clinical trials and mean percentage of alumni publications as reported by Doximity. The mean academic productivity score was used as an objective measure to assess one aspect of program quality: research output. These data were taken directly from each program's Doximity page, which relies on alumni Doximity profiles to calculate alumni publication percentage and alumni clinical trial percentage.

## Statistical Analysis

Analysis of variance (ANOVA) was performed in Microsoft Excel (Microsoft) to determine whether academic productivity was associated with program reputation, size, and type. Variables were considered significant if  $p < 0.05$ . If a significant association was found, post hoc analysis was performed using a simple  $t$  test.

## Results

One hundred seventy-five (87.1%) of 201 orthopaedic residency programs in the United States had Doximity pages that included all relevant information and were therefore included in the study. The mean academic productivity score of all programs was  $56.4 \pm 19.6$  (range = 26.5-96.5). The mean program size was  $24 \pm 9.7$  (range = 10-70).

## Program Ranking

Programs were divided into quartiles based on Doximity rankings, with 44 programs in quartiles 1, 2, and 3 and 43 programs in quartile 4. Programs in the top quartile had an academic productivity score of  $79.1 \pm 13$  (range = 51-98.5), while programs in the bottom quartile had a score of  $38.8 \pm 12.1$  (range = 26.5-62) ( $p < 0.0001$ ) (Fig. 1). Post hoc analysis demonstrated that there was a statistically significant difference between academic productivity in programs in the first and second quartile ( $p = 0.0001$ ), first and third quartile ( $p = 0.0001$ ), and first and fourth quartile ( $p = 0.0001$ ). There was also a statistically significant difference in academic productivity in programs ranked in the second and fourth quartile ( $p = 0.0001$ ) as

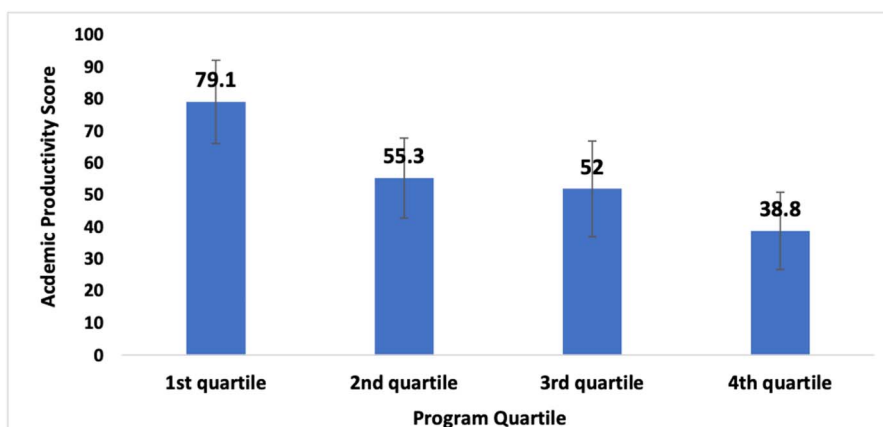


Fig. 1 Mean academic productivity score of orthopaedic surgery residency programs ranked by Doximity.

well as the third and fourth quartile ( $p = 0.0001$ ). There was, however, no difference in academic productivity between programs in the second and third quartile ( $p = 0.4683$ ) (Table I).

Each program was ranked based on the academic productivity score calculated during analysis. Of the 44 programs in the top Doximity quartile, 32 (72.7%) were also ranked in the top quartile for academic productivity; 10 programs (22.7%) were ranked in the second quartile, and 2 programs (4.6%) were ranked in the third quartile. None were ranked in the bottom quartile. Of the 43 programs in the bottom Doximity quartile, 26 (60.5%) were also in the bottom quartile for academic productivity; 12 programs (27.9%) were in the third quartile for academic productivity, and 5 programs (11.6%) were in the second quartile on Doximity. None were ranked in the top quartile.

**Program Size**

Programs were also divided into 4 groups based on the number of residents. Twenty-three programs (13%) had >30 residents, 67 (38%) had 21 to 30 residents, 72 (41%) had 11 to 20 residents, and 13 (7%) had ≤10 residents. Programs ranged in size from 10 to 70 residents, with an average size of 23.9 residents. Programs with more than 30 residents had an academic productivity score of  $76.2 \pm 14.8$  (range = 47.5-97.5), whereas programs with 10 or fewer residents had a score of  $40.6 \pm 16.3$  ( $p = 0.0001$ ) (range = 26.5-76) (Fig. 2). A statistically significant association was found between program size and academic productivity score during ANOVA ( $p < 0.01$ ). The mean size of programs ranked in the first quartile by Doximity is  $33.8 \pm 25.6$  residents, whereas the mean size of the bottom quartile of programs on Doximity is  $16.9 \pm 5.1$  ( $p = 0.0001$ ).

During post hoc analysis, there was a statistically significant difference in academic productivity between programs with >30 residents and programs with 21 to 30 residents ( $p = 0.005$ ), 11 to 20 residents ( $p = 0.0001$ ), and <10 residents ( $p = 0.001$ ). Programs with 21 to 30 residents also had significantly higher mean academic productivity ( $64.5 \pm 17.2$ ) than programs with 11 to 20 residents ( $45.4 \pm 13.8$ ) ( $p = 0.0001$ ) and programs with less than 10 residents ( $40.6 \pm 16.3$ ) ( $p = 0.0001$ ). However, there was no significant difference in academic productivity for programs with 11 to 20 residents ( $45.4 \pm 13.8$ ) and programs with less than 10 residents ( $40.6 \pm 16.3$ ) ( $p = 0.2663$ ).

**Program Type**

Programs were classified based on the type of program. There were 103 (59%) university programs, 49 (28%) university/community, 15 (9%) community, and 8 (5%) military affiliated. In the top quartile of programs ranked by Doximity, 39 (88.6%) were university programs and 5 (11.4%) were university/community programs. There was a significant difference in the academic productivity score based on program type: university ( $63.9 \pm 14.8$ ), military ( $54.7 \pm 16.9$ ), university/community ( $46.3 \pm 15.6$ ), and community ( $38.7 \pm 15$ ) (all  $p < 0.01$ ) (Fig. 3). University programs had significantly greater academic productivity than university/community programs ( $p = 0.0001$ ) and community programs ( $p = 0.0001$ ). There was no significant difference in academic productivity between university and military programs ( $p = 0.1733$ ) or community and community/university programs ( $p = 0.1053$ ). Military programs had significantly greater academic productivity than community programs ( $p = 0.0375$ ), but not university/community programs ( $p = 0.1794$ ).

TABLE I Post hoc analysis: comparing mean academic productivity of Q1, Q2, Q3, and Q4 programs					
Q1 vs. Q2	Q1 vs. Q3	Q1 vs. Q4	Q2 vs. Q4	Q3 vs. Q4	Q2 vs. Q3
$p < 0.0001$	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$	$p < 0.0001$	$p = 0.4683$

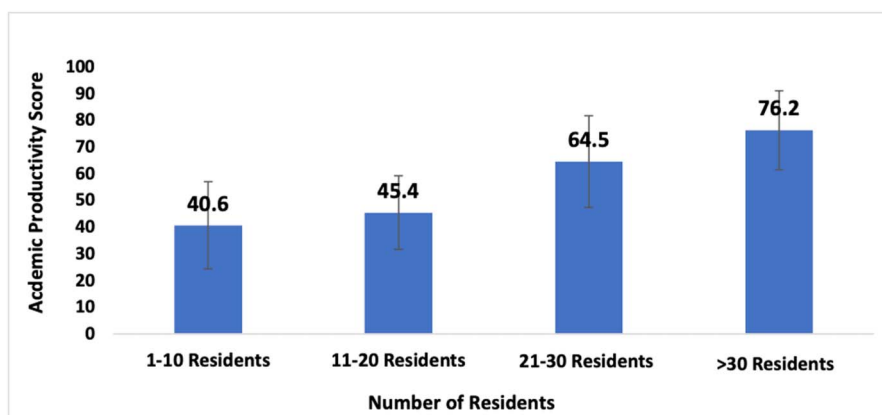


Fig. 2

Mean academic productivity score by orthopaedic surgery residency program size.

### Discussion

This study demonstrates that academic productivity of orthopaedic surgery residency programs was significantly associated with Doximity ranking, program size, and program type. In addition, there was a significant difference between the academic productivity of programs in the first quartile and every other quartile. Although reputation and satisfaction surveys are a part of the Doximity ranking methodology, the objective measure of academic productivity is also a significant factor. Notably, there was no difference in academic productivity between programs in the second and third quartile. These 50% of programs in the middle of the rankings are not significantly different in academic output, suggesting that reputation surveys may play a more decisive role in determining how these programs are ordered.

Medical students frequently reference Doximity to obtain information about orthopaedic residency programs, and survey results suggest that these rankings are taken into account by most applicants during the residency application process<sup>3,8,9</sup>. However, in a 2018 survey of 2,152 residency applicants nationwide, most students expressed concerns about the accuracy of Doximity rankings, with 58% responding that the rankings may be either “slightly accurate” or “not accurate”<sup>3</sup>. In another study from 2016, 65% of emergency medicine applicants responded

with negative comments when asked for their assessment of the accuracy of Doximity rankings<sup>8</sup>. These findings suggest that, although program reputation and Doximity rankings are highly influential among medical students, students are aware of the limitations. This study provides more context for Doximity rankings by demonstrating that academic productivity was significantly associated with Doximity ranking, program size, and program type.

When Doximity rankings were compared with academic productivity, this study found that programs ranked in the first quartile on Doximity had significantly greater academic productivity than every other quartile. These findings show that although reputation and satisfaction surveys are major components of the ranking methodology on Doximity, the rankings still correlate with an objective standard. Most programs ranked in the top quartile on Doximity (72.7%) were also in the top quartile for academic productivity. Likewise, 60.5% of programs in the bottom quartile on Doximity were also in the bottom quartile for academic productivity.

Notably, there was no significant difference in academic productivity between programs in the second and third quartile. This suggests that subjective measures may play a more decisive role in determining how these middle programs are ranked by

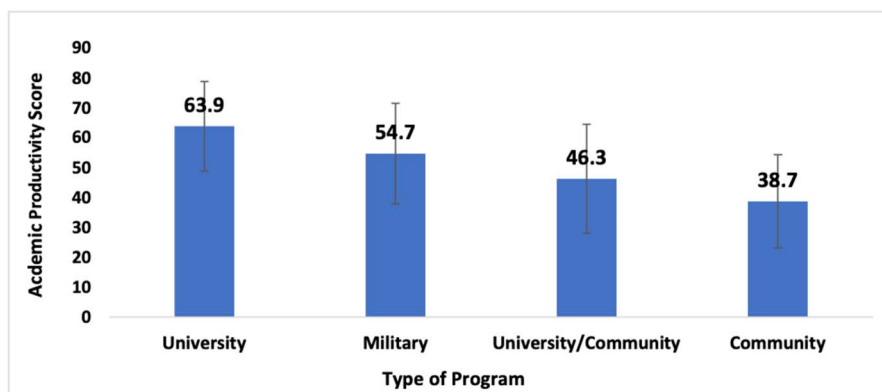


Fig. 3

Mean academic productivity score by orthopaedic surgery residency program type.

Doximity. This is corroborated by the results of a recent study by Esmaeeli et al., who found significant differences in the academic characteristics and academic achievements of anesthesiology program directors in the first quartile of Doximity rankings compared with quartiles 2 to 4<sup>10</sup>. The study found fewer differences between quartiles 2, 3, and 4<sup>10</sup>. Similarly, in 2018, Zhang et al. found that faculty members at plastic surgery programs in the first quartile of the Doximity rankings had a significantly higher mean number of publications than faculty members in quartiles 2 to 4<sup>11</sup>. These objective measures demonstrate that residency programs among various subspecialties ranked in the first quartile on Doximity have greater academic productivity than the other quartiles and suggest that the differences between the other quartiles may be less significant<sup>10,11</sup>. The top quartile programs are also more likely to have more residents and may have more resources in general to devote to research.

In contrast to the findings of this study, in 2014, Wilson et al. found only a “moderate correlation” between Doximity reputation rankings and an outcomes-based ranking of general surgery programs<sup>2</sup>. These data suggest that Doximity rankings may better correlate to objective standards in orthopaedic surgery than general surgery, or that the Doximity ranking methodology has changed in the past 7 years. Since the precise ranking formula has never been published by Doximity, it is not possible to know whether there were any alterations. Greater transparency regarding the ranking methodology used by Doximity would allow medical students and others who rely on them to have a better understanding on how the system ranks programs.

Orthopaedic surgery training programs vary widely in number of residents (range = 10-70). The findings of this study showed that programs ranked in the top quartile by Doximity had a larger mean number of residents than those in the bottom quartile. This result is in accordance with a 2019 study by Feinstein et al., which found that larger program size was predictive of a higher Doximity ranking<sup>6</sup>. Results from Esmaeeli et al. in 2021 also demonstrated that program size was strongly correlated with program rank<sup>10</sup>. The correlation between program size and Doximity ranking raises concerns that the reputation surveys may be more advantageous to larger programs, who have more alumni to survey. The rankings may then favor these larger programs, regardless of program quality or objective outcomes. In response to these concerns, this study did demonstrate that when comparing with one objective measure, larger program size is associated with a higher academic productivity score, and thus, Doximity rankings may be objectively correct by favoring larger programs.

As residency programs evolve and affiliations change, the difference between program types may sometimes be difficult to determine<sup>12</sup>. The Accreditation Council for Graduate Medical Education requires programs to identify as either university, university/community, community, or military<sup>13</sup>. The criteria to decide a program “type” are subjective, and this label may change over time<sup>12</sup>. Although the distinction between different program types is not absolute, this study did find a significant difference between the academic productivity of university and university/community programs. This is also reflected in the

Doximity rankings, where 89% of orthopaedic surgery programs in the top quartile were identified as university programs. The results of this study demonstrate that type of program does have concrete implications for academic productivity.

The consideration of Doximity residency rankings raises the question of how a medical student can best evaluate an orthopaedic surgery residency program. The limited number of publicly available resources for comparing different programs helps contribute to the prominence of the Doximity rankings, despite their inherent flaws. In the absence of other resources, medical students will use what is available to them. However, in addition to their subjective nature, the Doximity rankings fail to consider important factors to medical students such as the quality of instruction in each subspecialty, the demographic makeup of the program and its efforts regarding diversity/inclusion, and the fellowship/career outcomes of its graduates. These measures are difficult to quantify, but most would agree are more important to medical students than alumni research output.

This study has several limitations. First, the data used to calculate a program's mean academic score were taken from each program's Doximity page and could not be independently validated. These data are based on the Doximity profiles of program alumni, which may not accurately reflect the entire research output of a residency program. Second, the Doximity profiles of the orthopaedic programs do not publish data regarding the program's mean Orthopaedic In-Training Examination scores or average number of operative cases. These variables, among others, would provide more objective data to rank orthopaedic residency programs. The mean academic progress score is also part of a program's overall Doximity ranking and is not the ideal measure to compare these rankings with an objective measure. However, considering the influence of the Doximity rankings and the lack of data available to compare residency programs, we do believe this study offers some insight into how the rankings work.

Although Doximity rankings are subjective because of the reliance on survey data, academic productivity among orthopaedic residency programs is associated with Doximity ranking, program size, and type. This information can be used by medical students, residents, and physicians to understand the Doximity rankings and how they correlate with this objective measure. ■

Patrick J. Meade, BA<sup>1</sup>  
Sheena J. Amin, MD<sup>2</sup>  
Michaela A. Stamm, MS<sup>3</sup>  
Mary K. Mulcahey, MD, FAOA<sup>3</sup>

<sup>1</sup>Tulane University School of Medicine, New Orleans, Louisiana

<sup>2</sup>Department of Orthopaedic Surgery, University of Missouri-Kansas City, Kansas City, Missouri

<sup>3</sup>Department of Orthopaedics, Tulane University School of Medicine, New Orleans, Louisiana

E-mail address for M.K. Mulcahey: mary.mulcahey.md@gmail.com

---

## References

---

1. Kapoor N, Blumenthal DM, Smith SE, Ip IK, Khorasani R. Gender differences in academic rank of radiologists in US medical schools. *Radiology*. 2017;283(1):140-7.
2. Wilson AB, Torbeck LJ, Dunnington GL. Ranking surgical residency programs: reputation survey or outcomes measures? *J Surg Educ*. 2015;72(6):e243-50.
3. Smith BB, Long TR, Tooley AA, Doherty JA, Billings HA, Dozois EJ. Impact of Doximity residency navigator on graduate medical education recruitment. *Mayo Clinic Proc Innov Qual Outcomes*. 2018;2(2):113-8.
4. Doximity. Doximity Residency Navigator. Available at: [https://s3.amazonaws.com/s3.doximity.com/mediakit/Doximity\\_Residency\\_Navigator\\_Survey\\_Methodology.pdf](https://s3.amazonaws.com/s3.doximity.com/mediakit/Doximity_Residency_Navigator_Survey_Methodology.pdf). Accessed September 2021.
5. Lorch AC, Miller JW, Kloek CE. Accuracy in residency program rankings on social media. *J Grad Med Educ*. 2019;11(2):127-8.
6. Feinstein MM, Niforatos JD, Mosteller L, Chelnick D, Raza S, Otteson T. Association of Doximity ranking and residency program characteristics across 16 specialty training programs. *J Grad Med Educ*. 2019;11(5):580-4.
7. Ashack KA, Burton KA, Dellavalle RP. Dermatology in doximity. *Dermatol Online J*. 2016;22(2).
8. Peterson WJ, Hopson LR, Khandelwal S, White M, Gallahue FE, Burkhardt J, Rolston AM, Santen SA. Impact of Doximity residency rankings on emergency medicine applicant rank lists. *West J Emerg Med*. 2016;17(3):350-4.
9. Rolston AM, Hartley SE, Khandelwal S, Christner JG, Cheng DF, Caty RM, Santen SA. Effect of Doximity residency rankings on residency applicants' program choices. *West J Emerg Med*. 2015;16(6):889-93.
10. Esmaeeli S, Seu M, Akin J, Nejatmahmoodaliloo P, Knezevic NN. Program directors research productivity and other factors of anesthesiology residency programs that relate to program doximity ranking. *J Educ Perioper Med*. 2021;23(2):E662.
11. Zhang JQ, Herman SB, Tepper OM, Garfein ES, Weichman KE. Rank and research: the correlation between integrated plastic surgery program reputation and academic productivity. *Ann Plast Surg*. 2018;80(5):553-60.
12. Chen JG, Saidi A, Rivkees S, Black NP. University-versus community-based residency programs: does the distinction matter? *J Grad Med Educ*. 2017;9(4):426-9.
13. American Medical Association. FREIDA Online. Available at: <http://www.ama-assn.org/ama/pub/education-careers/graduate-medical-education/freida-online.page>. Accessed September 2021.