

# SM2 - Supplemental figures, tables and text for

## *Citizen Science for Health:*

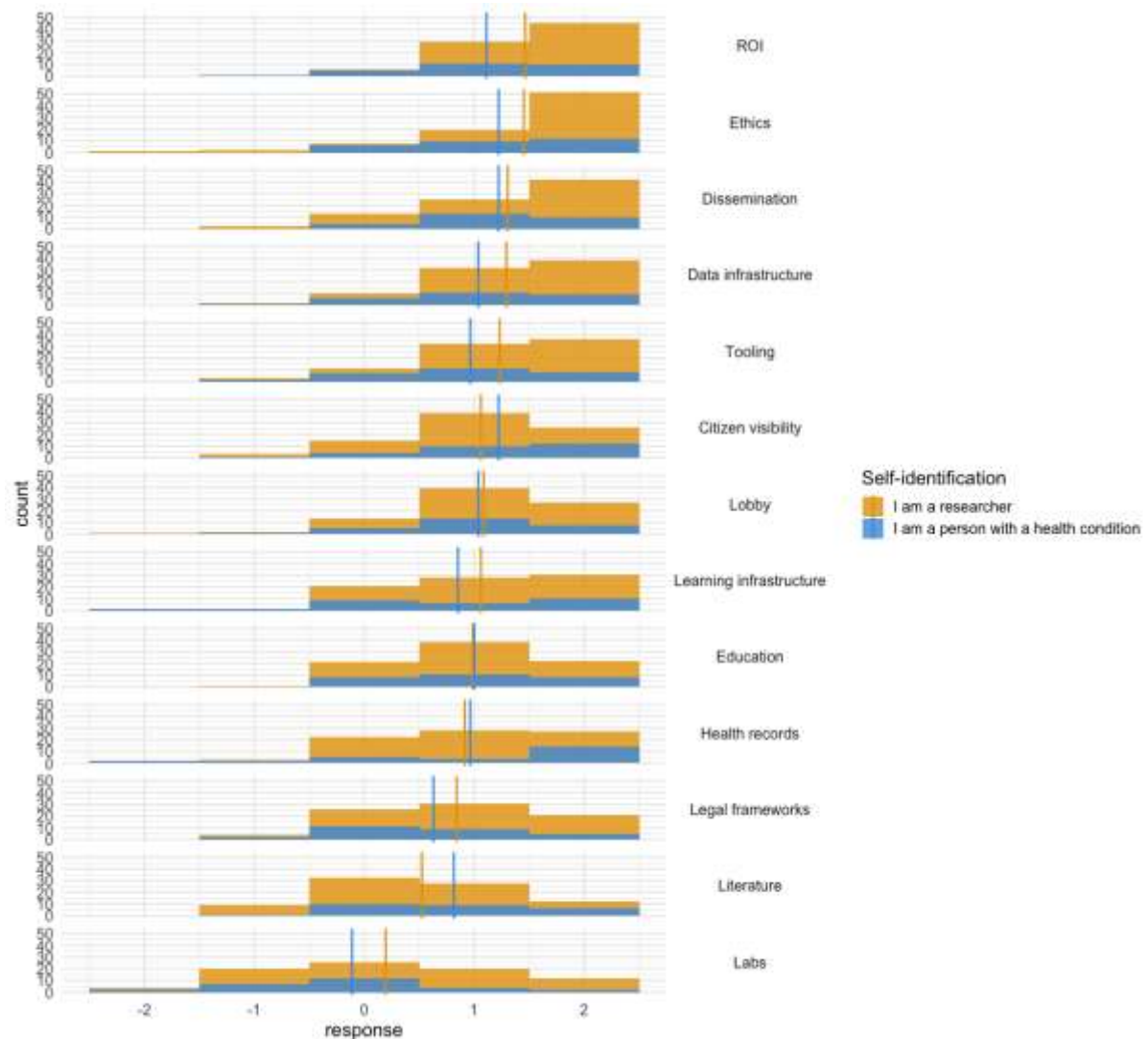
### *an international survey on its characteristics and enabling factors*

#### Content

|                                                                |    |
|----------------------------------------------------------------|----|
| Supplemental Figures and Tables.....                           | 2  |
| Figures S1, S2 and Tables ST1, ST2 .....                       | 2  |
| Supplemental discussion .....                                  | 6  |
| Factors for growing citizen science in the health domain ..... | 6  |
| Dissemination & Access to Health Literature.....               | 7  |
| Data Infrastructure .....                                      | 8  |
| Education and learning .....                                   | 10 |
| References .....                                               | 11 |

# Supplemental Figures and Tables

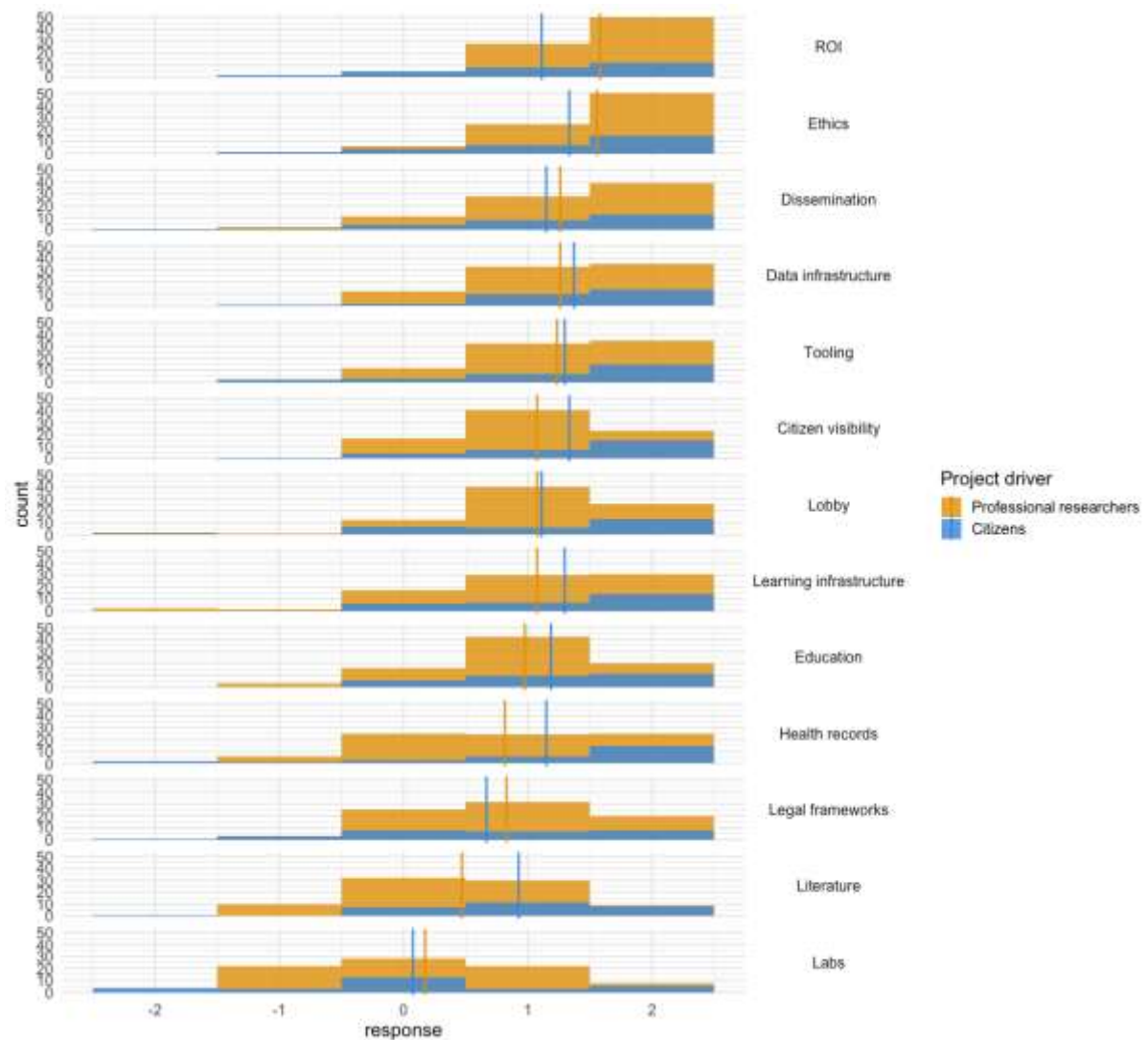
## Figures S1, S2 and Tables ST1, ST2



**Figure S1** Histogram of the responses for perceived importance, split between respondents that self-identified as “*researchers*” and “*person with a health condition*”. Bar plots for both groups are overlapping, both starting at zero. Vertical lines give the average values for each group/condition. Differences between groups were compared with Mann-Whitney-U test, and none of the differences were statistically significant (see Table ST1).

**Table ST1:** Statistical results for the comparison of responses of professional researchers compared to people with health conditions in ranking the importance of conditions.

| Condition               | P-value (uncorrected) | P-value (False Discovery Rate) |
|-------------------------|-----------------------|--------------------------------|
| ROI                     | 0.04349087537         | 0.4229801344                   |
| Ethics                  | 0.09798384329         | 0.4229801344                   |
| Dissemination           | 0.4227574037          | 0.5211423137                   |
| Data infrastructure     | 0.1525134262          | 0.4229801344                   |
| Tooling                 | 0.1252926666          | 0.4229801344                   |
| Citizen visibility      | 0.3123953203          | 0.5076423955                   |
| Lobby                   | 0.7182252879          | 0.7780773952                   |
| Learning infrastructure | 0.4360964894          | 0.5211423137                   |
| Education               | 0.9789807344          | 0.9789807344                   |
| Health records          | 0.4409665731          | 0.5211423137                   |
| Legal frameworks        | 0.2712785562          | 0.5038030329                   |
| Literature              | 0.1626846671          | 0.4229801344                   |
| Labs                    | 0.219902281           | 0.4764549421                   |



**Figure S2** Histogram of the responses for perceived importance, split between respondents that engage with CS4H projects driven by professional researchers or citizens. Bar plots for both groups are overlapping, both starting at zero. Vertical lines give the average values for each group/condition. Differences between groups were compared with Mann-Whitney-U test, and none of the differences were statistically significant (see Table ST2).

**Table ST2:** Statistical results for the comparison of responses by people's experience with projects driven by professional researchers compared to those by citizens in ranking the importance of conditions

| Condition               | P-value (uncorrected) | P-value (False Discovery Rate) |
|-------------------------|-----------------------|--------------------------------|
| ROI                     | 0.02517167011         | 0.1636158557                   |
| Ethics                  | 0.3147326127          | 0.5845034236                   |
| Dissemination           | 0.7993809189          | 0.7993809189                   |
| Data infrastructure     | 0.4189169036          | 0.6750606775                   |
| Tooling                 | 0.4673496998          | 0.6750606775                   |
| Citizen visibility      | 0.06107658123         | 0.198498889                    |
| Lobby                   | 0.6264314912          | 0.740328126                    |
| Learning infrastructure | 0.2995335917          | 0.5845034236                   |
| Education               | 0.240768123           | 0.5845034236                   |
| Health records          | 0.05224809968         | 0.198498889                    |
| Legal frameworks        | 0.6019955583          | 0.740328126                    |
| Literature              | 0.01411454879         | 0.1636158557                   |
| Labs                    | 0.7676414282          | 0.7993809189                   |

# Supplemental discussion

The paper ‘Citizen Science for Health: an international survey on its characteristics and enabling factors’ reports on a survey developed by the European Citizen Science Association (ECSA) Working Group “Citizen Science for Health”. The paper aims to lay bare the characteristics and enabling factors of citizen science in the health domain, and to formulate a direction for future work and research. In the main text of the paper we focused on the two main conditions to support and grow citizen science for health that were prioritised by the majority survey respondents. In this supplementary document, we also briefly address the other conditions which merit their own discussion, but were not as highly prioritized by the survey respondents.

## Factors for growing citizen science in the health domain

The survey asked respondents to rate factors which could contribute to support or grow health-related citizen science (see figure 5 in the paper). Out of the 13 factors listed, two stood out as particularly important to the respondents: (1) A balanced return on investment (ROI) for both researchers and participants (patients or citizens), as well as (2) adequate ethical frameworks. See figure 5, copied in the Supplemental File from the paper. While these were discussed in the main paper, we also offer reflections on the other conditions below.

## Dissemination & Access to Health Literature

An apparent contradiction can be observed between the relative importance of publishing and disseminating results and access to health literature; Overall, publishing & dissemination comes out to be the third-highest ranked factor. In contrast, “access to health literature for citizens” is the second least important factor according to the respondents of our survey. Despite recurring calls to abolish the “publish or perish” regime of academia (Kiai, 2019) and the San Francisco declaration for research assessment (DORA); scientific publishing remains key to the career progression of academic researchers (Zhang and Yu, 2020, Cuker et al 2017). Given these external pressures, and the strong presence of researchers among the respondents, it might be unsurprising that respondents put a strong emphasis on the publishing aspects. Similarly, patients/ “Do-It-Yourself” (DIY) researchers might want to publish to make themselves heard, and to show their progress. At the same time, researchers might be less familiar with the struggles of access to literature than non-institutional actors, such as patients and citizens. While the Open Access movement has highlighted the importance of patient-access to literature for many years (Eysenbach 2006, Grouse 2014, Davis 2011), subscription fees remain a barrier that informal researchers need to overcome (Greshake 2016, Greshake 2017, Himmelstein et al., 2018). We see some evidence of this in our survey, as people with health conditions on average do rank access to literature higher than academic researchers (Figure S2 in this Supplemental File). Beyond barriers to access in terms of reading the literature, and even though that some publishers are currently promoting the inclusion of patients as co-authors and also involved in the peer-review process in journals such as BMC Medicine or the Lancet (Boyce 2018), patients face an even higher barrier when it comes to trying to publish

their own original research (Lewis 2022). This is reflected in the other condition for growth of citizen science practice in the health domain, ‘citizen visibility’. It might also explain why ‘lobby’ was ranked just under ‘visibility’.

In a recent work on the factors that give rise to successful DIY Medicine, Wexler (2022) identified four key factors: 1) the presence of frustrated patients, 2) a lack of access to an effective treatment, 3) availability of online social media outlets, and 4) the availability of a therapeutic that is relatively easy to create or access. While our categories do not easily map onto those four factors, we can find some hints that those frustrations and a broader lack of access apply here as well, as seen by the “access to literature”-aspects discussed above, but also in the relative importance of the factor ‘patient’s access to health records’, which people with a health condition score as slightly more relevant than researchers (Figure S2) and which are a frequent topic of patient advocacy (Blease 2022, Esch et al 2016). ‘Tooling’, which emerged as another factor in need of development, is likely to relate to ‘patient access to health records’. The exact nature of this ‘tooling’, should be a matter for future research, as respondents may have resonated with any of the pre-given items: the availability of apps, e-health devices and adequate research procedures.

## Data Infrastructure

Data-infrastructure was another topic of high interest amongst survey respondents, and potentially comprises many underlying themes. At minimum, it includes issues such as data quality, ownership and security, while findability, accessibility, interoperability and reusability become increasingly important (Wilkinson et al, 2016). In the era of personalised medicine, the assemblage of different types and sources of data (such as Real-World Data, clinical data, Citizen Generated Data, etc) is of

utmost relevance (Rudrapatna, V. A. & Butte, 2020). This holds true not only for professional researchers, but also for patients and citizens, who often struggle in pulling their data together for the sake of their treatments and their own research. In various countries large efforts are being invested to create a data infrastructure that overcomes these hurdles; the European Commission aims to bring the European efforts together in one European Health Data Space (EU Commission, 2022). The challenge is very much recognized in ‘normal’ health research, even though the focus is basically on connecting (academic or health care) data silos, rather than on connecting citizen-generated data, despite the relevance the European Commission attaches to it (Corcho et al, 2022).

Connecting and correlating different datasets is complex and time-consuming, resulting in formal researchers most often taking the lead and responsibility for the topic, such as when creating health commons (Jensen et al 2017, All of Us 2019). At the same time, this introduces a dependency for patients, which is why some patients have started to invest in their own data-infrastructure (Lewis & Price Ball, 2017, Greshake Tzovaras et al 2019; MMV, 2023). These patient-driven platforms also help to solve some of the ethical issues around consent mechanisms and governance (cf. Micheli et al., 2023; Remmers et al 2021). At the same time, such patient-infrastructure solutions are often limited in their impact, due to a lack of visibility in the broader discourse, as for example found in the Netherlands, where major patient organisations composed an agenda on ‘Research by and for Patients’ and found a lack of awareness and appreciation for the investigative efforts of patients among institutional partners (Transitieteam GROZ, 2019).

Another potentially underlying topic of 'data infrastructure' category is not only the ability for patients and citizens to access the data but also to participate in the data interpretation process. Most of the time, this interpretation process is done solely by researchers, although patients and citizens have lived experiences that could enrich the results interpretation (Senabre Hidalgo et al, 2021). It could also be a powerful addition to or way to increase the ROI for patients and citizens (Phillips et al, 2019).

## Education and learning

Finally, education and learning was another topic to emerge from the survey respondents' prioritising. The education of health professionals regarding the potentials and pitfalls of citizen science is relevant here. Yet, this condition or topic also refers to a learning environment in which other actors can also learn and benefit from citizen science through the exchange of different perspectives on an issue. Bearing in mind that, as was shown in section 3 of the original paper, the dynamics in the health domain are considered to be more complex than in other domains, *and* that health institutions are considered to be more reluctant to adopt citizen science, the creation of such a learning environment seems of paramount importance.

# References

All of Us. 2019. The “All of Us” Research Program Special Report. *N Engl J Med* 2019; 381:668-676. DOI: 10.1056/NEJMSr1809937.

Blease C, Salmi L, Rexhepi H, et al. 2022. Patients, clinicians and open notes: information blocking as a case of epistemic injustice *Journal of Medical Ethics* 2022;48:785-793.

Boyce N., Marsh J., Wayman C., Pinfold V, Kabir T. 2018. Service user reviewers: extending peer review in *The Lancet Psychiatry*, *The Lancet Psychiatry*, Vol 5, 10, 780-781

Corcho, O. J. Jiménez Blanco, C. Morote and E. Simperl, 2022. Data.europa.eu and Citizen-generated Data: opportunities and challenges associated with the inclusion of citizen-generated data in data.europa.eu. European Commission Directorate-General for Communications Networks, Content and Technology. doi: 10.2830/137589

Cuker A, Ay C, Cushman M. 2017. The role of scientific publishing in the development of early career investigators. *Res Pract Thromb Haemost*. 2017 Dec 27;2(1):6-7. doi: 10.1002/rth2.12072.

Davis PM, Walters WH. 2011. The impact of free access to the scientific literature: a review of recent research. *J Med Libr Assoc*. 2011 Jul;99(3):208-17. doi: 10.3163/1536-5050.99.3.008.

Esch T, Mejilla R, Anselmo M, et al. 2016. Engaging patients through open notes: an evaluation using mixed methods. *BMJ Open* 2016;**6**:e010034. doi: 10.1136/bmjopen-2015-010034.

EU Commission, 2022: Proposal for a Regulation of the European parliament and of the council of Europe on the European Health Data Space. [https://eur-lex.europa.eu/resource.html?uri=cellar:dbfd8974-cb79-11ec-b6f4-01aa75ed71a1.0001.02/DOC\\_1&format=PDF](https://eur-lex.europa.eu/resource.html?uri=cellar:dbfd8974-cb79-11ec-b6f4-01aa75ed71a1.0001.02/DOC_1&format=PDF)

Eysenbach G. The open access advantage. *J Med Internet Res*. 2006 May 15;**8**(2):e8. doi: 10.2196/jmir.8.2.e8.

Greshake, B.. 2016. Correlating the Sci-Hub data with World Bank Indicators and Identifying Academic Use. Authorea. June 02, 2016. DOI: 10.15200/winn.146485.57797

Greshake B. 2017. Looking into Pandora's Box: The Content of *Sci-Hub* and its Usage. *F1000Res*. 2017 Apr 21;**6**:541. doi: 10.12688/f1000research.11366.1.

Greshake Tzovaras, B., Angrist, M., Arvai, K., Dulaney, M., Estrada-Galiñanes, V., Gunderson, B., Head, T., Lewis, D., Nov, O., Shaer, O., Tzovara, A., Bobe, J. and Price Ball, M. 2019. Open Humans: A platform for participant-centered research and personal data exploration, *GigaScience*, Volume 8, Issue 6, June 2019, giz076, <https://doi.org/10.1093/gigascience/giz076>.

Grouse L. 2014. Open access medical publications. *J Thorac Dis*. Jun;**6**(6):E133-6. doi: 10.3978/j.issn.2072-1439.2014.03.21.

Himmelstein, D., Rodriguez Romero, A., Levernier, J., Munro, T., Reid McLaughlin, S., Greshake Tzovaras, B. and Greene, C. 2018. Research: Sci-Hub provides

access to nearly all scholarly literature eLife 7:e32822. DOI:

<https://doi.org/10.7554/eLife.32822>

Jensen, M., Ferretti, V., Grossman, R. and Staudt, L. 2017. The NCI Genomic Data Commons as an engine for precision medicine. *Blood* 2017; 130 (4): 453–459. doi: <https://doi.org/10.1182/blood-2017-03-735654>.

Kiai, A. 2019. To protect credibility in science, banish “publish or perish”. *Nat Hum Behav* 3, 1017–1018. <https://doi.org/10.1038/s41562-019-0741-0>

Lewis, D., 2022. Barriers to Citizen Science and Dissemination of Knowledge in Healthcare. *Citizen Science: Theory and Practice*, 7(1), p.40. DOI: <http://doi.org/10.5334/cstp.511>

Lewis, D. and Price Ball, M. 2017. OpenAPS Data Commons on Open Humans. figshare. Poster. <https://doi.org/10.6084/m9.figshare.5428498.v1>

Micheli, M., Farrell, E., Carballa-Smichowski, B., Posada-Sánchez, M., Signorelli, S., Vespe, M., 2023. Mapping the landscape of data intermediaries — Emerging models for more inclusive data governance, Publications Office of the European Union, Luxembourg, 2023, doi:10.2760/261724, JRC133988.

MMV, 2023: Zelfonderzoek - werken aan gezond leven.

<https://www.mmv.nl/zelfonderzoek/> Accessed 20 June 2023

Phillips, T. B., Ballard, H. L., Lewenstein, B. V., & Bonney, R. 2019. Engagement in science through citizen science: Moving beyond data collection. *Science Education*, 103(3), 665–690. <https://doi.org/10.1002/sce.21501>

Remmers, GGA, A. Boorsma, H. Duinkerken and M. van Lieshout, 2021. Gezond AKKOORD: een coöperatief beheer van je gezondheidsgegevens. Position paper.

<https://mdog.nl/wp-content/uploads/2021/11/1-GA-Position-paper-Gezond-AKKOORD.pdf>

Rudrapatna, V. A. and Butte, A. J. 2020. Opportunities and challenges in using real-world data for health care. *The Journal of Clinical Investigation*, 130(2), 565-574.

Senabre Hidalgo, E., Perelló, J., Becker, F., Bonhoure, I., Legris, M. and Cigarini, A. 2021. Participation and Co-creation in Citizen Science. In: , et al. *The Science of Citizen Science*. Springer, Cham. [https://doi.org/10.1007/978-3-030-58278-4\\_11](https://doi.org/10.1007/978-3-030-58278-4_11)

Transitieteam Patiententafel GROZ, 2019. *Onderzoek door en voor patiënten: een kennisagenda voor hogere kwaliteit van leven en meer maatschappelijke participatie door patiënt-gedreven onderzoek in gezondheid*. Pp. 33

[https://publicaties.zonmw.nl/fileadmin/zonmw/documenten/LSH/MPNN/Kennisagenda\\_Onderzoek\\_voor\\_en\\_door\\_Patienten\\_11\\_okt\\_2019.pdf](https://publicaties.zonmw.nl/fileadmin/zonmw/documenten/LSH/MPNN/Kennisagenda_Onderzoek_voor_en_door_Patienten_11_okt_2019.pdf)

Wexler, A., 2022. Mapping the Landscape of Do-it-Yourself Medicine. *Citizen Science: Theory and Practice*, 7(1), p.38. DOI: <http://doi.org/10.5334/cstp.553>

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>.

Zhang, Y., Yu, Q. What is the best article publishing strategy for early career scientists?. *Scientometrics* 122, 397–408 (2020). <https://doi.org/10.1007/s11192-019-03297-4>