Check for updates

RESEARCH ARTICLE

Particle physics facing a pandemic

[version 1; peer review: 1 approved, 1 approved with reservations]

Adam KARDOS[1], Sven-Olaf MOCH[2], German RODRIGO[3]

[1]Department of Experimental Physics, Institute of Physics, Faculty of Science and Technology, University of Debrecen, 4010 Debrecen, PO Box 105, Hungary
[2]II. Institut fur Theoretische Physik, University of Hamburg, Hamburg, D-22761, Germany
[3]Instituto de Fisica Corpuscular, Universitat de Valencia - Consejo Superior de Investigaciones Científicas, Parc Cientific, E-46980 Paterna, Valencia, Spain

v1 **First published:** 19 Jan 2023, **3**:13
https://doi.org/10.12688/openreseurope.15199.1
Latest published: 19 Jan 2023, **3**:13
https://doi.org/10.12688/openreseurope.15199.1

Abstract

Background: Ordinary life as we knew it changed in March of 2020 due to the global COVID-19 pandemic. While springtime in general was well awaited and regarded as a synonym for rejuvenation, March of 2020 on the other hand brought lock-down, curfews, home working and digital education to the lives of many. The particle physics community was not an exception: research institutes and universities introduced remote working and digital lecturing and all workshops, conferences and summer schools were cancelled, got postponed or took place online.
Methods: Using publicly available data from the INSPIRE and arXiv databases we investigate the effects of this dramatic change of life to the publishing trends of the high-energy physics community with an emphasis on particle phenomenology, theory and CERN's two major LHC experiments, ATLAS and CMS. To get insights, we gather information about publishing trends in the last 20 years, ending by December 2021, and analyse it in detail.
Results: Our analysis revealed that the publishing trend in particle physics was only been affected in a minor way.
Conclusions: Publication data show that difficult times were successfully overcome and that the community even increased scientific output.

Keywords

covid, pandemic, particle physics, publishing trends

This article is included in the COST Actions gateway.

Open Peer Review

Approval Status ✓ ?

	1	2
version 1 19 Jan 2023	✓ view	? view

1. **Janusz Gluza**, University of Silesia, Katowice, Poland

2. **Avik Roy**, University of Illinois at Urbana-Champaign, Champaign, USA

Any reports and responses or comments on the article can be found at the end of the article.



This article is included in the Marie-Sklodowska-Curie Actions (MSCA) gateway.



This article is included in the Research on Research gateway.



This article is included in the Horizon 2020 gateway.

Corresponding author: Adam KARDOS (kardos.adam@science.unideb.hu)

Author roles: KARDOS A: Formal Analysis, Investigation, Resources, Software, Validation, Visualization, Writing – Original Draft Preparation; **MOCH SO**: Conceptualization, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; **RODRIGO G**: Conceptualization, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This research was financially supported by the European Union's Horizon 2020 research and innovation programme, under the Marie Sklodowska-Curie grant agreement No 764850 (Scattering Amplitudes: from Geometry to Experiment [SAGEX]). This work is supported by the COST Action CA16201 PARTICLEFACE, by the Spanish Government (MCIN/AEI/10.13039/501100011033) Grant No. PID2020-114473GB-I00 and Generalitat Valenciana (Grant No. PROMETEO/2021/071), by grant K 125105 of the National Research, Development and Innovation Fund in Hungary and by the ÚNKP-21-Bolyai+ New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund. Kardos kindly acknowledges financial support from the Bolyai Fellowship programme of the Hungarian Academy of Sciences
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2023 KARDOS A *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: KARDOS A, MOCH SO and RODRIGO G. **Particle physics facing a pandemic [version 1; peer review: 1 approved, 1 approved with reservations]** Open Research Europe 2023, **3**:13 https://doi.org/10.12688/openreseurope.15199.1

First published: 19 Jan 2023, **3**:13 https://doi.org/10.12688/openreseurope.15199.1

Introduction

In January 2020, the high-energy physics community was looking forward to another busy year full of workshops, schools and conferences. The situation dramatically worsened in Europe by the end of February and the final and lethal blow to all these live events happened in the middle of March. The quick escalation of the situation was personally experienced by the authors during the organization of the *PRecision Effective FIeld Theory School (PREFIT20)* from the 2nd to the 13th of March[1] and during a COST short-term scientific mission to Hamburg. The *PREFIT20* school had to be switched from lecturers being on-site to online presentations as events were running and turned out to be the last major live event in our field. The short-term scientific mission looked like an ordinary trip during the out-bound journey (9th of March), but one and a half weeks later (18th of March) almost all flights were cancelled and the airport at Frankfurt became completely deserted, with all shops closed.

By the end of March, lock-down, working from home and online education became the standard. This also meant all conferences and workshops were cancelled or postponed until the end of summer. Within a couple of weeks the dominant platform to interact with colleagues became Zoom. The system got so wide-spread that term "Zoom meeting" became the 15th most frequently used term in 2020[1]. In the particle physics community, collaborations spanning over several countries or continents are quite common, having members from different countries in Europe or even from the Americas and Asia. Hence people were more-or-less familiar with virtual meeting platforms and the technical transition had no major difficulties or problems. This was also witnessed by the authors participating in several virtual collaboration meetings, seminars, online workshops and by giving lectures.

Our experience suggests that although distant collaborations had been set up and were working flawlessly well before the pandemic, the situation changed because several institutions and universities in Europe introduced travelling restrictions, and completely prohibited short visits and brain-storming sessions during conference coffee breaks. Also, it is very common that new collaborations form spontaneously, a process for which in-person meetings, workshops and structured programs at centers for scientific exchange like those offered by, e.g., Aspen, GGI, INT, KITP, MIAPP and MITP, play an important role[2]. During a pandemic virtual presence and online platforms can keep an already working collaboration going, but new collaborations become much harder to start due to the lack of personal contacts between potential members. In the particle physics community well before the pandemic the usual habit was to approach people personally during a workshop or conference and start an informal collaboration growing out of stimulating discussions over coffee breaks.

The pandemic not only changed the professional lives of researchers but also their personal lives. This can also have consequences to scientific output and should not be disregarded. With lockdown and online education introduced in schools researchers not only had to cope with new ways to keep their current collaborations going and to establish new ones despite the lack of all personal contacts but also had to prepare for lecturing using virtual classrooms. The latter aspect is of particular significance as the community generally regards chalk board lecturing as the best way to convey course material and scientific ideas[11,12]. All of these factors have an effect on scientific research and it is not at all clear how the community could adopt to new circumstances.

In this paper we analyze the research output in the form of papers and proceedings of the high-energy particle physics phenomenology community by using various open search engines to retrieve data on published papers during the pandemic and in the past twenty years to get a better understanding of the true effects of the pandemic and to see what kind of changes it induced in the community. To have a broader perspective we also collected data related to the high-energy particle physics theory community and the two major experiments, ATLAS and CMS, conducted at CERN's Large Hadron Collider (LHC) as well. We have selected these two experiments since these are the biggest ones in our community founded in the middle of the 1990s, hence their publishing trends extends well before the pandemic.

Methods

In order to extract publication data we used the search engines of two publicly available databases, which cover basically all relevant scientific publications (these include both peer-reviewed and non-peer-reviewed publications as well) in the field of high-energy physics: INSPIRE[13] and arXiv[14]. These engines allow for sophisticated searches by restricting hits through the use of many criteria, like paper categories[3] (hep-ph, hep-th, hep-ex, etc.), by requiring papers to have a certain number of authors or by listing papers published or uploaded within a given period of time. As these highly specialized tools have become quite mature it is possible to perform automatic queries by invoking the site's search engine directly with an API (Application Programming Interface). By using a special URL, the search engine can be instructed to perform a search according to the parameters specified through the URL and to return the result in some machine readable format. For example, if publications that appeared in 2021 and categorized as hep-ph are extracted

[1] The *PREFIT20* school[2] was organized as a joint event of the COST Actions *PARTICLEFACE*[3] and *VBSCan*[4] at DESY in Hamburg.

[2] Aspen Center for Physics[5], Galileo Galilei Institute (GGI) for Theoretical Physics[6], Institute for Nuclear Theory (INT)[7], Kavli Institute for Theoretical Physics (KITP)[8], Munich Institute for Astro- and Particle Physics (MIAPP)[9], Mainz Institute for Theoretical Physics (MITP)[10] provided.

[3] The paper category represents a subfield in the high-energy physics community: hep-ph is destined for papers by the high-energy physics phenomenology community, hep-th for high-energy theoretical investigations and findings and hep-ex is exclusively used by the high-energy experiments publishing their results.

from INSPIRE the following URL should be invoked from the browser:

https://inspirehep.net/api/literature?fields=hits.total&format=json&q=publication_info.year:2021%20AND%20arxiv_eprints.categories:hep-ph

The result is delivered in the `JSON` format, which is tolerably readable by the human eye but easily digestible by the many commonly available interpreting languages, like `Python`. In our case only the total number of papers fulfilling certain criteria was needed, hence the method we have employed is:

1. URL is constructed with all needed keywords.
2. URL is invoked from `Python`.
3. Result is delivered in `JSON` format.
4. Total number of papers is extracted and written into data file.
5. Histograms are created with `GNUplot`.

In the case of the arXiv database we did not use the arXiv API. Instead, by selecting the advanced search option, the total number of records fulfilling the search criteria (for details see the corresponding section) were gathered manually.

In recent times a strong increase in conference proceedings can be seen in all the categories we have investigated. In order to avoid this proliferation as much as possible and to get a clearer view of publishing trends we have filtered out all records without a DOI number. This method does not fully eliminate proceedings from the search results but restricts their number to a subset which is indeed published in one form or another. The filtering according to the DOI number was only possible with the INSPIRE API. Hence, we have validated our arXiv results by taking a look at the annual statistics reported from both search engines to see if the presence of proceedings in the arXiv search results affects the trends obtained with the INSPIRE API.

With the INSPIRE API we have collected annual statistics starting in the year 2001 and up to 2021. We have not considered 2022 because entries in INSPIRE do not acquire a DOI until they are published in a journal a few months later. To get the annual number of papers we have requested that the field `publication_info.year` includes the year for which total number of publications are collected. Moreover, the document type had to be `article`, all the records were supposed to have a DOI number and to belong to the target category (hep-ph, hep-th or hep-ex) for which we collected the data. In case of the hep-ex category we have only accepted papers by the ATLAS and CMS collaborations and, hence, rejected any other. By varying the number of authors we were able to get statistics data for various collaboration sizes.

In the case of the arXiv search engine our set of selection criteria was very modest. We have performed searches in the same three categories as in case of the INSPIRE API, filtered records according to their submission year and, for a finer resolution, the month in which the submission happened.

Annual statistics using inspire API

For the long-term annual trend in publishing the INSPIRE API was used in the three different categories: particle phenomenology (hep-ph), theory (hep-th) and experiment (hep-ex), the latter being restricted to the ATLAS and CMS collaborations only, and the results were filtered by requiring a valid DOI number attached to each record in order to minimize bias due to the proliferation of proceedings. Specifically we are interested in the total number of papers on an annual basis and in the number of authors.

For the interpretation of the data collected in this way, it is also important to identify scientific events with potential impact on the publishing behavior of the high-energy physics community, such as new discoveries, the schedule of LHC operations, or other important developments. For the categories we have analyzed events with potential impact are listed in Table 1.

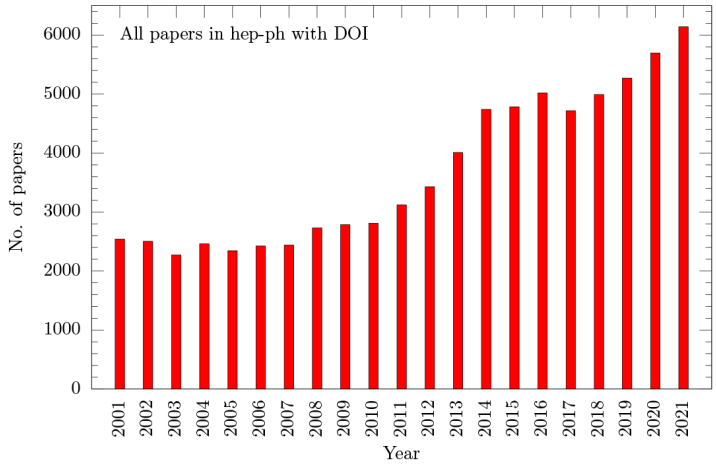
Results

Total number of papers

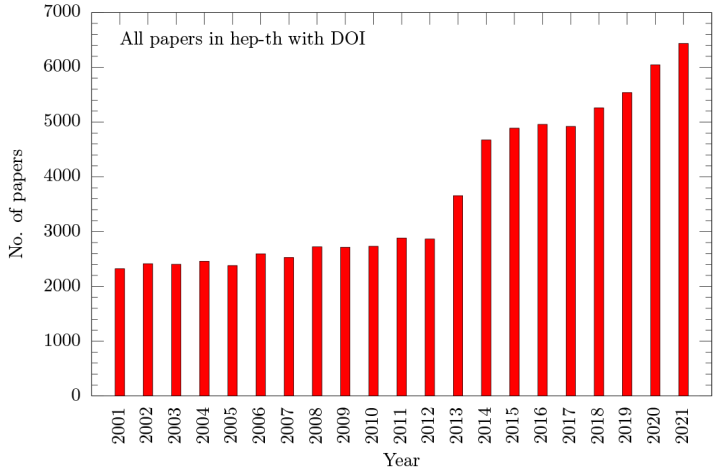
The total number of published papers in the period 2001 to 2021 are depicted in Figure 1. It is evident that the start of LHC operations with Run I had a tremendous effect on publication activity in high-energy physics. As it can be seen in Figure 1a, before 2010 the number of papers produced in the phenomenology community remained at a steady level with a slight increase hardly discernable due to fluctuations. However, as first LHC results started to pour in the number of papers increased rapidly. A certain fraction of the increasing number of papers in 2011 can be also attributed to the OPERA announcement on the measurement of superluminal neutrinos[15]. In 2014 and with the end of Run I the increase flattens out, the annual publication rate becomes roughly constant

Table 1. Noteworthy events with potential impact on publishing trends and tendencies in the high-energy physics community since the start of LHC operations.

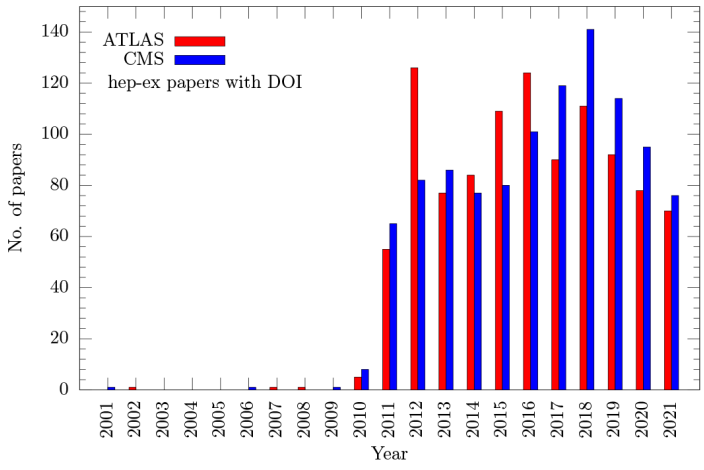
Event	Date
LHC Run I	late 2009 - early 2013
OPERA announcement of measurement of superluminal neutrinos[15]	23 September 2011
CERN announcement of Higgs boson discovery[16]	04 July 2012
ATLAS and CMS presentation of diphoton excess at 750 GeV[17,18]	15 December 2015
LIGO announcement of first measurement of gravitational waves[19]	11 February 2016
LHC Run II	2016 – 2018



(a) Papers published with category label hep-ph.



(b) Papers published with category label hep-th.



(c) Papers published with category label hep-ex by the ATLAS and CMS experiments.

Figure 1. Total number of published papers according to INSPIRE filtered using various arXiv category labels and requesting a valid DOI number.

although with a different normalization, roughly a factor of two compared to pre-LHC days. A small peak can be seen in Figure 1a in 2016, which can be attributed to the 750 GeV diphoton excess[17,18]. The Run II of LHC has a less dramatic effect on the number of phenomenology papers and it is also shifted by a year or so starting to ramp up only in 2018. The increase in the number of publications at the end of the LHC Run II did not finish in 2019, instead it is still visible in data taken for 2020 and 2021. Knowing that the COVID-19 pandemic had its first real effects in March 2020 this is a remarkable observation. The publishing trend did not change by the measures enforced on the phenomenology community but on the contrary even with all difficulties induced by the pandemic the community flourished.

A look at the high-energy theory category in Figure 1b shows that the number of papers published grows only moderately until the year 2013, where a sudden increase happened in the normalization. This change of normalization coincides with new achievements in multi-loop amplitude calculations where the first publications appeared on arXiv in 2013[20,21] and completely changed the way loop integrals are being calculated nowadays. The real change in the trend appeared after the Run I starting in 2015 and 2016. Then, a second, less dramatic increase starts in 2018, a trend in agreement with the previous LHC related ramp-up. The number of theory papers increases not during but just after the end of the LHC Run II. This could be a consequence of the lack of new physics found in both runs, leading the community to speculate about the lack of new signals.

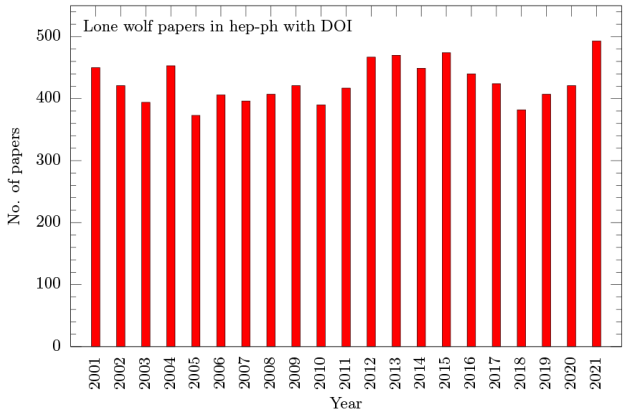
The annual number of published papers produced by the ATLAS and CMS collaborations are shown in Figure 1c and it is evident that the publication activity really started around 2011 and 2012 once real data was taken and subsequently analyzed. As for Run I, ATLAS outnumbered CMS which started to catch up during Run II later on. Also between the two runs the continuous analysis of recorded data produced a steady flow of published papers. As more and more data has been analyzed the rate of papers decreases, which is already visible in the numbers recorded after 2019.

A. Number of authors

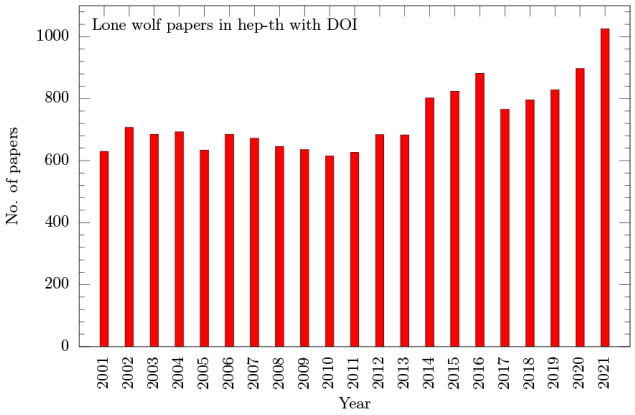
In high-energy physics collaborations of many sizes exist. A pandemic can have different effects on different collaboration sizes hence it is very informative to dissect the annual number of papers according to different collaboration size ranges. This partitioning can only affect phenomenology and theory papers because in the hep-ex category we only consider ATLAS and CMS publications. The latter two collaborations have constant size since the start of LHC operations in 2009 with approximately 3.000 authors in case of ATLAS and around 2.500 authors in case of CMS.

For the hep-ph and hep-th categories, the annual statistics for single-author papers is shown in Figure 2. Restricting ourselves to the phenomenology papers only, in Figure 2a it can be seen that the number of papers displays a rather steady trend with somewhat larger fluctuations due to the smaller amount of papers produced on an annual basis. Due to higher fluctuations, possible trends are harder to identify but a slight increase can be observed at the start of LHC Run I in 2010 and ending in 2012. Recent years, starting with 2016, brought a steady decrease of number of papers with a slight increase starting in 2018 in accord with the end of LHC Run II. This slight increase has continued until 2020 and has ramped up in 2021.

In Figure 2b, single-author hep-th papers are listed as a function of publication year. Before LHC Run I it follows a more-or-less constant trend with some fluctuations, even some decrease can be seen ending in 2010. A ramp-up period



(a) Papers published by a single author with category label hep-ph.



(b) Papers published by a single author with category label hep-th.

Figure 2. Total number of published papers according to INSPIRE filtered using hep-ph (left) and hep-th (right) arXiv category labels, requesting a valid DOI number and with a single author.

can be identified starting in 2010 in line with the start of LHC operations resulting in a peak in 2016, part of which can be attributed to the diphoton excess announcement in December 2015 and the LIGO announcement of first measurement of gravitational waves in February 2016[19]. Another period of increasing publication activity started in 2017 in line with the end of the LHC Run II. The year 2020 witnessed an increasing publication activity that continued in 2021 for single-author papers in hep-th, so it seems that the pandemic was not a significant factor in this respect.

For small-size collaborations, which we define to consist of two to five authors, the number of published papers on an annual basis is plotted in Figure 3 for both communities, phenomenology and theory. In Figure 3a, it is visible that before the LHC Run I the phenomenology community more-or-less followed a steady publishing tendency with the same amount of papers produced annually with some fluctuations. The start of Run I resulted in a steady increase in published papers with a slight slow-down tendency as nearing the Run II. The year of 2016 witnessed a remarkably good year for publishing most probably by the diphoton anomaly announced in late 2015[22,23]. Run II ignited one more increase which seems to continue.

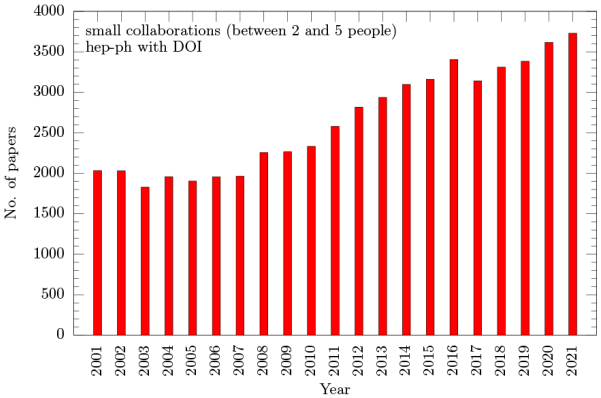
On the other hand, in Figure 3b, publishing trends of small hep-th collaborations can be seen with a very similar behavior as in hep-ph, although with a slightly elongated slow increase. The big jump in publishing happens not through the Run I period but slightly after with no clear sign of superluminal neutrino or the diphoton excess announcements. The same trend is visible in recent years, as the numbers of publications grow again with the end of Run II.

Medium-size collaboration[4] are neither very common in phenomenology nor in theory as can be seen from the annual statistics in Figure 4. It seems evident that the start of LHC operations during its Run I was a crucial factor in forging collaborations of this size since both in phenomenology and in theory the number of papers published drastically increased throughout and after this period. In the phenomenology statistics a clear peak is visible in 2016, correlated in time with the diphoton excess. In recent years, trends continue to be positive.

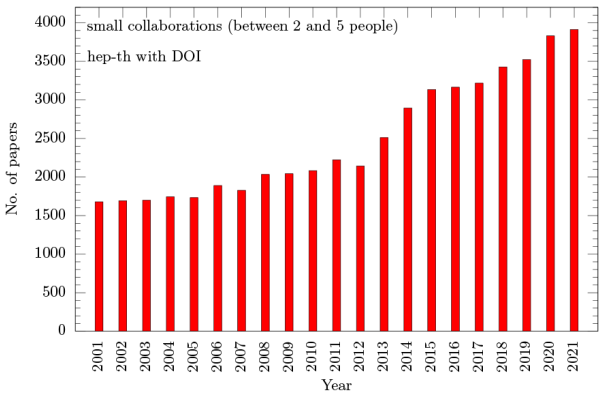
Large-size collaboration[5] of more than ten authors involved do not have a long history in the analyzed two communities as can be seen in Figure 5. In the theory community even nowadays these can be regarded as scarce with less than a dozen papers produced annually. As Figure 5a shows the advent of LHC physics and the first results from LHC Run I catalyzed the formation of these kind of collaborations. In comparison, less than twenty such papers were published on an annual basis before the LHC era. The past recent years however do witness a rising trend of publications from large-size collaborations in the hep-ph category, culminating in an enormous peak in 2020 (more than 20% increase

[4] There is no solid definition what collaboration size can be considered medium-size. In this paper drawing from our experience have settled for the number of authors in the range of six to ten in order to fit into this category.

[5] As with medium-size collaborations there is no firm definition of what could be considered a large-size collaboration in these areas of academic research. Hence the lower bound for these collaborations of more than ten people reflects the authors' judgment of what can be regarded as big in our communities.

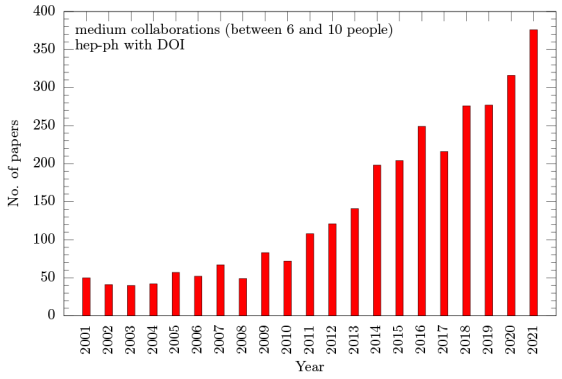


(a) Papers published by small-size collaborations of two to five authors with category label hep-ph.

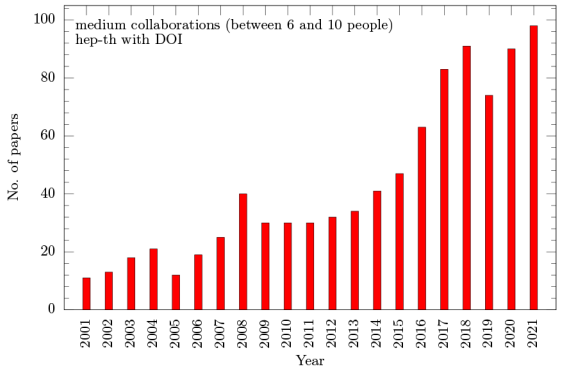


(b) Papers published by small-size collaborations of two to five authors with category label hep-th.

Figure 3. Total number of published papers according to INSPIRE filtered using hep-ph (left) and hep-th (right) arXiv category labels, requesting a valid DOI number and by collaborations with between two to five authors.

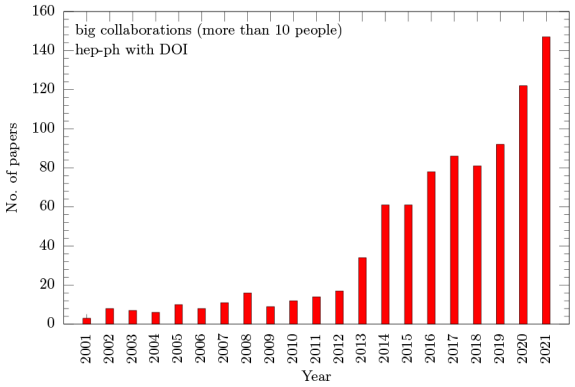


(a) Papers published by medium-size collaborations of six to ten authors with category label hep-ph.

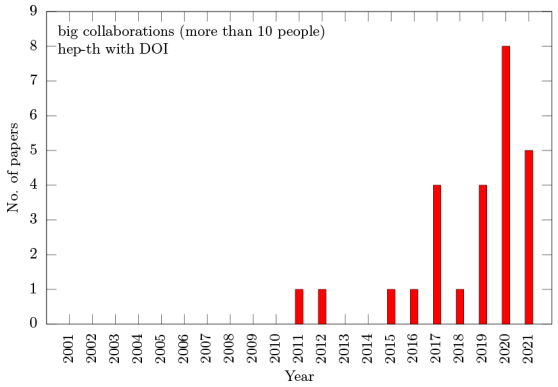


(b) Papers published by medium-size collaborations of six to ten authors with category label hep-th.

Figure 4. Total number of published papers according to INSPIRE filtered using hep-ph (left) and hep-th (right) arXiv category labels, requesting a valid DOI number and by collaborations with between six to ten authors.



(a) Papers published by large-size collaborations of more than ten authors with category label hep-ph.



(b) Papers published by large-size collaborations of more than ten authors with category label hep-th.

Figure 5. Total number of published papers according to INSPIRE filtered using hep-ph (left) and hep-th (right) arXiv category labels, requesting a valid DOI number by collaborations with more than ten authors.

compared to 2019) and 2021. This dramatic increase in the number of publications by large-size collaborations in phenomenology can be explained by the extraordinary performance of LHC during and since Run I. The high-quality data created a strong driving force for precise predictions, which, on the other hand, require a significant amount of collaborators. Consequently, the demand for state-of-the-art predictions can often only be fulfilled within a reasonable time-frame by joining efforts, hence large(r) collaborations have been and are being formed. The peaking behavior can also be seen in the publication trend in the hep-th community but since the number of publications there is smaller by a factor of two the increase in 2020 compared to 2019, and the decrease in 2021, can be partially attributed to fluctuations. However, it is clear from the trend (at least for the hep-ph community) that the pandemic did not have any effect on the publishing trend. On the contrary, the past two years were fruitful for producing papers for this size of collaborations.

Judging from these annual statistics, it seems that the effect of the pandemic on the three communities in high-energy particle physics (hep-ph, hep-th and ATLAS, CMS in hep-ex) is at most moderate and in the worst case resulted in a

stagnation in the publishing trends or a very slight decrease. Collaborations of larger sizes seem to be immune to the COVID-19 restrictions, being able to publish with the same or even higher frequency. In the case of the ATLAS and CMS collaborations the overall trend for the recent years shows a decreasing rate of publications and the data for 2020 and 2021 also mirror this well-established trend, which is likely to change after the start of the LHC Run III in 2022..

Monthly statistics using arxiv

The search engine of arXiv allows us to extract publishing data not only on an annual but also on a monthly basis, i.e., we can get some insights on the monthly distribution of publications. The engine does not just restrict the search to published papers, but even with conference proceedings contained in the collected samples, it is interesting to look at publication and arXiv submission trends at a finer resolution. In order to convince ourselves that statistics gathered on a monthly basis including conference proceedings still can be used to identify trends in publishing, the total annual statistics collected via arXiv can be compared with the histograms in the previous section based on the data of the INSPIRE API. As an example of this, cross-check annual submission statistics were collected in the hep-ph category on arXiv and the resulting histogram is plotted in Figure 6. Comparing this figure to Figure 1a shows that the trends are the same, although the overall normalization is different. Before the LHC Run I period both statistics follow a more-or-less constant trend with an increase in the number of publications due to the LHC Run I. Even the publishing peak in 2016 is visible with arXiv results and the increasing frequency of recent years can be identified. These findings give us enough confidence to take a look at the monthly charts obtained from arXiv and to search for tendencies in them.

In Figure 7 we have plotted the monthly statistics for seven consecutive years, from 2015 to 2021, for both hep-ph and hep-th. As it can be seen from both plots the publishing trend is relatively constant for phenomenology papers but for theory papers a clear pattern emerges. The publishing frequency is systematically lower in the first two months of the year and increases to a local maximum in the summer but just to turn back and reach a second minimum in August and September and then it ramps up to reach the maximum frequency as the year ends. The histograms contain both 2020 and 2021 results and these data fit perfectly well into the general trend of recent years. In order to create these histograms we have used the original submission date for the records to filter results. In this way it is easy to see for example that the 750 GeV diphoton excess really created a swell of publication[6], both in phenomenology and theory papers, resulting in an excess of 15–20 % of publishing in December 2015 with a significant surplus carried over to January 2016 (more than 30% for phenomenology and 17% for theory compared to the same period of previous year).

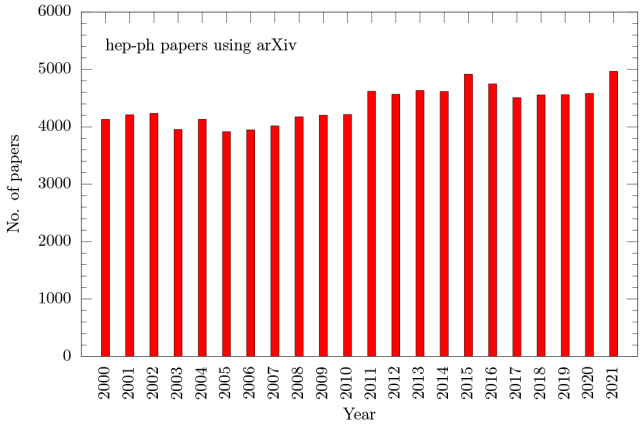


Figure 6. Annual submission statistics in the hep-ph category using the arXiv search engine.

From these statistics it seems that the pandemic did little to change the already settled publishing trends and at times it helped overshoot the same month from previous years. We cannot identify any dramatic increase nor dramatic decrease in the publication rates.

Conclusions

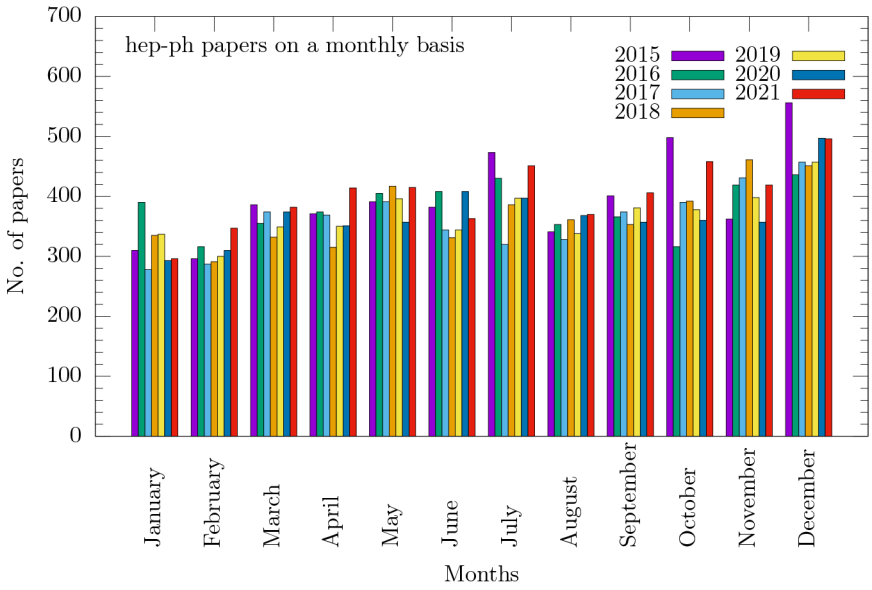
Despite the pandemic, the high-energy physics community continued with the increasing trend in the number of publications that started with the operation of the LHC. At the same time the face-to-face interaction remains a vital component in the process of exchanging scientific ideas, for work in existing collaborations and for the start of new ones. Here, workshops, conferences and topical programs at centers for scientific exchange play an important role by creating suitable environments and atmospheres. The complete lack of such interactions could have a negative impact. While the various virtual presence platforms seem suitable and efficient for administrative purposes and supersede in-person meetings, they cannot serve as a substitute for the true melting pots of new ideas during social events. Travel and workshop attendance can also help to focus on research in the absence of the usual other errands, lectures and administrative burdens.

In summary, the particle physics community successfully adapted to the COVID-19 pandemic and the associated restrictions for the time being, but to avoid complete stagnation and to really progress towards a bright future the high-energy physics community needs to restart in-person events, scientific visits and black board discussions.

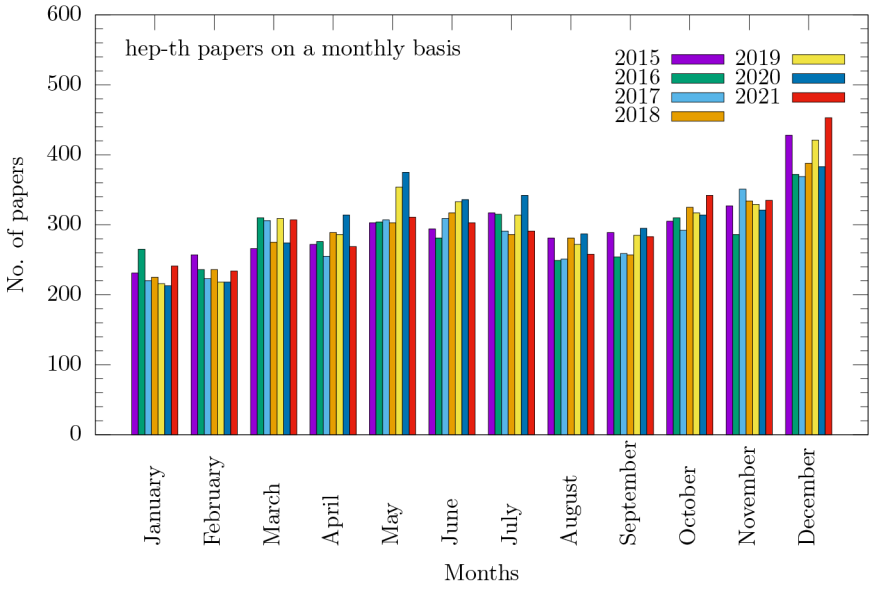
Ethics and consent

No ethical approval or consent was required.

[6] In the community this effect is sometimes referred to as “chasing the ambulance”.



(a) Submissions to arXiv on a monthly basis in the category label hep-ph.



(b) Submissions to arXiv on a monthly basis in the category label hep-th.

Figure 7. Submission to arXiv in the hep-ph (up) and hep-th (down) categories between 2015 and 2021. For further details see the main text.

Disclaimer

Author German Rodrigo is currently on leave at the European Research Council Executive Agency, European Commission, BE-1049 Brussels, Belgium. The views expressed are purely those of the writer and may not in any circumstances be regarded as stating an official position of the European Commission.

Data availability

Repository: Hepstat: Particle physics facing a pandemic - dataset https://doi.org/10.5281/zenodo.7392919 This project contains the following underlying data:

- `arxiv` (containing annual and monthly statistics extracted from arXiv)

- `gnuplot` (contains all the `GNUplot` scripts used to create all the figures of the paper in two subfolders: one for arXiv and one for the INSPIRE statistics)
- `inspire` (contains the annual statistics obtained from INSPIRE)
- `.py` files:
  - `arxiv-annual.py`: this script is used to extract annual statistics from arXiv
  - `arxiv.py`: this script is used to obtain the monthly statistics from arXiv
  - `inspire.py`: this script is used to get the various annual statistics from INSPIRE

All data has been extracted from open sources and underlying results and all computer code are available on `GitHub`[24]

Data are available under the terms of a CC0 licence and no additional source data are required.

Acknowledgements

A preprint version of this paper is also made available on arXiv[25].

References

1. Irish Times: **What's the top word of 2020? Oh you guessed that one easily**. **Reference Source**
2. **PREFIT20: PRecision Effective FIeld Theory School**. **Reference Source**
3. COST Action CA16201: **PARTICLEFACE - Unraveling new physics at the LHC through the precision frontier**. **Reference Source**
4. COST Action CA16108: **VBSCan - Vector Boson Scattering Coordination and Action Network**. **Reference Source**
5. **The Aspen Center for Physics**. **Reference Source**
6. **The Galileo Galilei Institute (GGI) for Theoretical Physics**. **Reference Source**
7. **The Institute for Nuclear Theory (INT)**. **Reference Source**
8. **The Kavli Institute for Theoretical Physics (KITP)**. **Reference Source**
9. **The Munich Institute for Astro- and Particle Physics (MIAPP)**. **Reference Source**
10. **The Mainz Institute for Theoretical Physics (MITP)**. **Reference Source**
11. Finc JE, Rudow SR: **Pointing with power or creating with chalk, Contemporary Issues in Education Research (CIER)**. 2015. **Reference Source**
12. Waters CM: **Rock the chalk: A five-year comparative analysis of a large microbiology lecture course reveals improved outcomes of chalk-talk compared to powerpoint.** *bioRxiv.* 2019. **Publisher Full Text**
13. INSPIRE Collaboration: **INSPIRE**. **Reference Source**
14. Cornell University: **arXiv**. **Reference Source**
15. OPERA collaboration, Adam T, Agafonova N, *et al.*: **Measurement of the neutrino velocity with the OPERA detector in the CNGS beam.** *JHEP.* 2012; **10**: 93 [arXiv: 1109.4897]. **Publisher Full Text**
16. ATLAS and CMS collaboration, Gianotti F, Incandela J: **Higgs announcement seminar on 4 July 2012**. Jul, 2012. **Reference Source**
17. ATLAS collaboration: **Search for resonances decaying to photon pairs in 3.2 fb^{-1} of *pp* collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector**. preprint ATLAS-CONF-2015-081. 2015. **Reference Source**
18. CMS collaboration: **Search for new physics in high mass diphoton events in proton-proton collisions at $\sqrt{s}$ = 13 TeV**. preprint CMS-PAS-EXO-15-004. 2015. **Reference Source**
19. LIGO Scientific, Virgo collaboration, Abbott BP, *et al.*: **Observation of Gravitational Waves from a Binary Black Hole Merger.** *Phys Rev Lett.* 2016; **116**: 061102. [arXiv: 1602.03837]. **Publisher Full Text**
20. Henn JM: **Multiloop integrals in dimensional regularization made simple.** *Phys Rev Lett.* 2013; **110**: 251601. [arXiv: 1304.1806]. **Publisher Full Text**
21. Henn JM, Smirnov AV, Smirnov VA: **Evaluating single-scale and/or non-planar diagrams by differential equations.** *JHEP.* 2014; **03**: 088. [arXiv: 1312.2588]. **Publisher Full Text**
22. ATLAS collaboration, Aaboud M, Aad G, *et al.*: **Search for resonances in diphoton events at $\sqrt{s}$ = 13 TeV with the ATLAS detector.** *JHEP.* 2016; **09**: 001. [arXiv: 1606.03833]. **Publisher Full Text**
23. CMS collaboration, Khachatryan V, Sirunyan AM, *et al.*: **Search for Resonant Production of High-Mass Photon Pairs in Proton-Proton Collisions at $\sqrt{s}$ = 8 and 13 TeV.** *Phys Rev Lett.* 2016; **117**: 051802. [arXiv: 1606.04093]. **Publisher Full Text**
24. Kardos A, Moch SO, Rodrigo G: **Hepstat: Particle physics facing a pandemic - dataset.** **http://www.doi.org/10.5281/zenodo.7392919**
25. Kardos A, Moch SO, Rodrigo G: **Particle physics facing a pandemic**. arXiv: 2111.05585. **Publisher Full Text**

Open Peer Review

Current Peer Review Status: ✔ ?

Version 1

Reviewer Report 23 August 2023

https://doi.org/10.21956/openreseurope.16439.r34223

© **2023 Roy A.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Avik Roy
National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Champaign, Illinois, USA

This article studies the impact of the COVID-19 pandemic on publishing trends in HEP. Papers and proceedings are categorized into three major categories- theory (hep-th), phenomenology (hep-ph), and experimental results from ATLAS and CMS (hep-ex). The authors obtain publication records by searching INSPIRE and Arxiv databases and present the results in various categories showing yearly distributions as bar charts. They also take into account the impacts of some of the potential “noteworthy events” like LHC Run I, superluminal neutrino measurements, the diphoton excess, etc. They further categorize the information in terms of the number of authors and the month of the year. The analysis, however, mostly relies on a qualitative assessment of the observed trends in publication records in such categories. They conclude that publication trends are only marginally impacted by the pandemic.

This analysis is an important one, and therefore, requires further sophistication. While I agree with the conclusion that publication record in the pandemic years has seen a marginal impact, this may be a highly “averaged out” conclusion. Even in terms of publication trends, there remains a large scope for investigating more nuanced impacts. I am pointing out some of those points in the following list-

- The authors can easily filter out hep-ex results from Run I, early Run II (2015-16), and full Run II results. The distribution of full Run II publications, when compared with Run I and early Run II results, can be an important aspect to judge whether the pandemic introduced noticeable delays in getting those results ready.

- The analysis can shed light on the geographical distribution of hep-ph and hep-th papers. Were there any noticeable differences among North American, European, or Asian collaborations? While intercontinental collaborations are ubiquitous, an analysis based on which continent/country most of the authors are based on can be performed. Different countries handled the pandemic differently, did those effects become visible?

- Similarly, the impact of publication trends in different age, gender, and race groups can be

studied. However, I realize that it can be a more involved undertaking and correctly categorizing the data in these groups may not be entirely possible based on information available in INSPIRE and Arxiv alone.

- The rise in large pheno collaborations is an important observation. The authors attribute it to the high quality data from LHC since Run I which drove the need to make precise predictions. While I partly agree with this assessment, it does not fully explain why the boost in larger collaborative efforts only starts to show up in 2020 and not earlier. I think the authors should explore the following potential avenues-
  - The shifted norm of communicating over zoom incentivized the formation and efficient management of larger collaborations
  - The pandemic noticeably impacted in-person work on the upgrade and instrumentation efforts during the Long Shutdown 2. It may have freed up FTEs for students and postdocs who might have contributed to simulation and performance studies for those pheno results
- I think overall excitement about the field and the availability of funding should be factored in while truly analyzing publication trends. The year-by-year trend in publication definitely is a function of these parameters and only after disentangling these effects, one can truly fish out the effect of the pandemic. Students, postdocs, and young faculty members constitute the larger part of the army that drives research and publication trends. The excitement about HEP in general must have been much higher right after the Higgs discovery and the associated Nobel Prize awarded in 2012-13. How have the student intake, postdoc hiring, and faculty appointment trends evolved since then? Does that account for any of the fluctuations we see (or don't see)? Has COVID impacted any of these trends in any way?

Finally, I think the paper has taken a restrictive approach to analyzing the impact of the pandemic. The title was suggestive of a comprehensive study. While methods and results indeed clarify the scope of the paper, the conclusion "difficult times were successfully overcome", as stated in the abstract, is too general. It fails to recognize the impact, especially on the younger physicists- graduations delayed, career growth hindered by lack of exposure, and we are yet to determine if virtual physics education is just as effective as its in-person equivalent. Maybe, a more honest assessment would be a rather conservative one- while the community maintained its scientific productivity during the pandemic years, the true range and nature of its impact on the community, even in terms of publication trends, can only be understood by performing a more thorough and categorized study of current and future data.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and does the work have academic merit?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Experimental high energy physics, Explainable AI and machine learning for particle physics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 07 March 2023

https://doi.org/10.21956/openreseurope.16439.r30868

© 2023 Gluza J. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Janusz Gluza
Institute of Physics, University of Silesia, Katowice, Poland

The paper involves a case study in which the Authors discuss if the covid pandemic period changed statistically the publication activity in the high-energy experimental and theoretical society. The analysis is based on papers published in 2001-2021 which are extracted from arxiv and inspires preprint repositories. The study includes data taken from hep-ph (phenomenology), hep-th (theory) and hep-ex (experiment). The Authors analyze papers published by a single author and small-size (2-5), medium-size (6-10), and large-size (>10) collaborators. Though small irregularities in a total number of publications were identified due to specific experimental results (e.g. the OPERA announcement on the measurement of superluminal neutrinos in 2011, the diphoton anomaly in late 2015, increase or decrease of publications due to LHC runs 1 and 2 periods and shutdowns), the overall conclusion of the work is that not any dramatic increase nor decrease in the publication rates was identified due to the covid pandemic situation in years 2020-2021.

The analysis is valuable and interesting, and in my opinion, the analysis can be repeated later, at least for 2022-2023. The covid period can have long-term repercussions after its maximum in 2020-2021.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and does the work have academic merit?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Particle physics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.