Nomenclature: coronavirus and the 2019 novel coronavirus

Zhiwen Hu[1], Zhongliang Yang[2], Qi Li[2], Yongfeng Huang[2,†]

[1]School of Computer and Information Engineering, Zhejiang Gongshang University, Hangzhou 310018, China.
[2]Beijing National Research Center for Information Science and Technology (BNRist), Tsinghua University, Beijing 100084, China.
†Correspondence and requests for materials should be addressed to yfhuang@tsinghua.edu.cn.

Author contributions

Z.H., Z.L. and Q.L. contributed equally to this work. Y.H. and Z.H. were involved in the conceptual design of the study. Z.H., Z.L. and Q.L. performed the metadata analyses. All authors wrote the manuscript.

Competing interests

The authors declare no competing interests.

Acknowledgments

We hereby desire to express our indebtedness to Prof. Kenneth McIntosh for personal communication and insightful comments. This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFC1604002 and the National Natural Science Foundation of China under Grant U1705261, Grant U1836204, Grant U1536201 and Grant U1636113.

Abstract

Less aligned emphasis has been given to naming the 2019 novel coronavirus and pandemic disease. Global profusion of squab names has found their ways in daily communication, and our survey promises to articulate that many of them may have contributed to backlash against Chinese people. Here, based on brief critical reviews on the naming of coronavirus and human coronaviruses, we scrutinize a clear sense of pros and cons of previous multifarious names and punctuate heuristic introspection of naming practices. Our findings suggest that full-fledged official names are duly contribute to the resilience of healthy collective usages in current infodemic scenario.

Keywords: human coronaviruses (HCoVs); infodemic; PHEIC; 2019-nCoV; COVID-19; SARS-CoV-2

Introduction

On the occasion of the Chinese Lunar New Year of 2020, the 2019 novel coronavirus disease (now known as COVID-19 [1]) outbreak was first reported from Wuhan City of China, home to 11 million people. However, less aligned emphasis has been given to naming the 2019 novel coronavirus and pandemic disease. So far, there is not a universally accepted name yet, either for academic-industrial usage or consistency with international virus taxonomy. The past few weeks has witnessed an explosive growth of tangled monikers and hashtags of the virus, which have found their ways in daily communication. There is concern that an appropriate nomenclature should be duly assigned to this newly identified coronavirus and caused respiratory tract disease in humans, which has potential public health impact.

Nomenclature: coronavirus and human coronaviruses

Before discussion of naming the 2019 novel coronavirus, it is necessary to take a glimpse into the history of scientific taxonomy and nomenclature. In 1966, an International Committee on Nomenclature of Viruses (ICNV) was established with the mission of introducing some degree of order and consistency into the naming of viruses. In 1973, the ICNV became the International Committee on Virus Taxonomy (ICTV), a global authority on the designation and naming of viruses. Another international authoritative body, the World Health Organization (WHO) is responsible for naming new human infectious diseases.

However, as an earlier nomenclature practice, the naming history of coronaviruses (CoV) is always misjudged in scientific community. In retrospect, on 16 November 1968, eight virologists proposed the term "coronaviruses" in a brief annotation of *Nature* [2]. In fact, Anthony Peter Waterson (1923 – 1983) and his colleagues should be credited with the coinage of the neologism "coronavirus" [3,4][Personal communication

with Prof. Kenneth McIntosh, the only survivor of the authors of the annotation in *Nature* that proposed this name [2].].

In humans, there are 7 spectrums of human coronaviruses (HCoVs) known to cause the common cold as well as more severe respiratory disease. Of those, human coronaviruses HCoV-229E, HCoV-NL63, HCoV-OC43 and HCoV-HKU1 are routinely responsible for mild respiratory illnesses like the common cold but can cause severe infections in immunocompromised individuals. But three members have caused deadly outbreaks: SARS-CoV, MERS-CoV, and the newly identified coronaviruses (now known as SARS-CoV-2 [5]).

The earliest usage track-down of the terms "coronavirus" and "coronaviruses" could provide an insightful and compelling argument for rigorous historical story. Meta-analysis on their diachronic discourse promises to articulate the unfolding chronological picture since its debut on a historical time scale. The diachronic discourse of "coronavirus" and "coronaviruses" in English corpus from 1960 to 2008 unveils that there was a mild increase in the numbers of printed books dealing with them after the initial description of coronaviruses in 1968. Then, each human coronavirus epidemic – SARS-CoV in 2002-2003, HCoV-NL63 in 2004 and HCoV-HKU1 in 2005 – leads to a new wave of hot research (**Figure 1**).

Furtherly, analysis results of metadata of Web of Science and PubMed indicate that known knowledge still remains off-limit in the field of combating emerging HCoVs (**Figure 2**). The SARS-CoV-2 is the seventh identified coronavirus that can cause diseases of the respiratory tract via human-to-human transmission. It caused mysterious pneumonia outbreak is spreading far more quickly than the SARS-CoV and MERS-CoV diseases [1,6,7], although the epicenter of the outbreak was locked down to curb the pandemic spread. Presently, its clinical severity is yet to be determined, although many fatal cases have occurred. Meanwhile, much remains unclear about the pandemics caused by HCoVs.
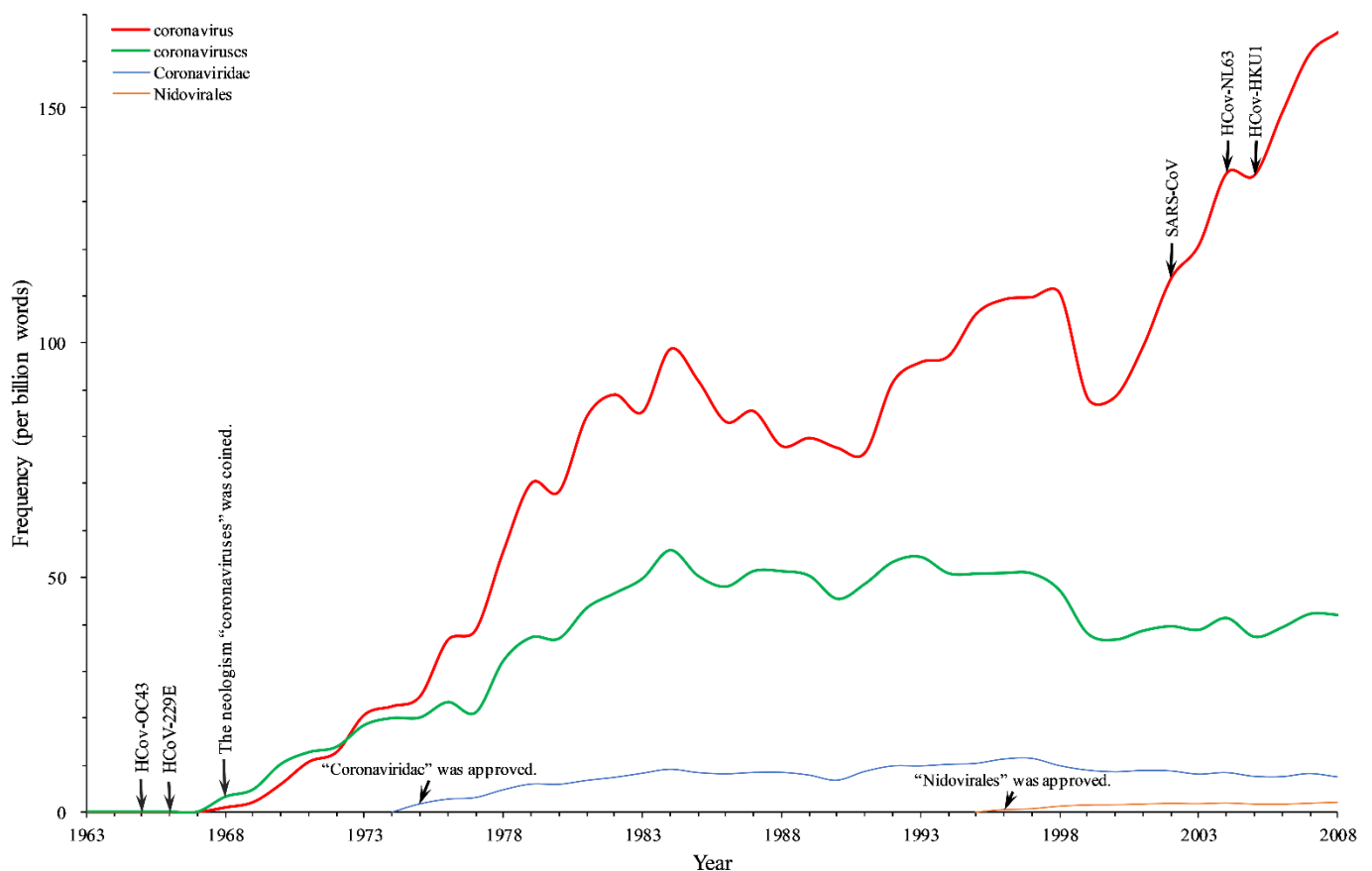
Fig. 1. Diachronic discourse of "coronavirus", "coronaviruses", "Coronaviridae" and "Nidovirales" in English corpus from 1963 to 2008. Google Books Ngram Corpus (GBNC) facsimiles the word frequency of "coronavirus", "coronaviruses", "Coronaviridae" and "Nidovirales" in English corpus from 1963 to 2008, respectively. The Google Books Ngram Corpus (GBNC) [8] is a unique linguistic landscape that benefits from centuries of development of rich grammatical and lexical resources as well as cultural contexts. After the initial description of coronaviruses in 1968 [2]

there was a mild increase in the numbers of printed books dealing with them, followed by several peaks, after several human coronavirus epidemics: SARS-CoV in 2002-2003, HCoV-NL63 in 2004 and HCoV-HKU1 in 2005. In 1971, "coronaviruses" was officially approved by the International Committee on Nomenclature of Viruses (ICNV) [9,10]. The terms "Coronaviridae", "Nidovirales" and "Coronavirinae" were officially approved by the International Committee on the Taxonomy of Viruses (ICTV) in 1975, 1996, 2009, respectively.
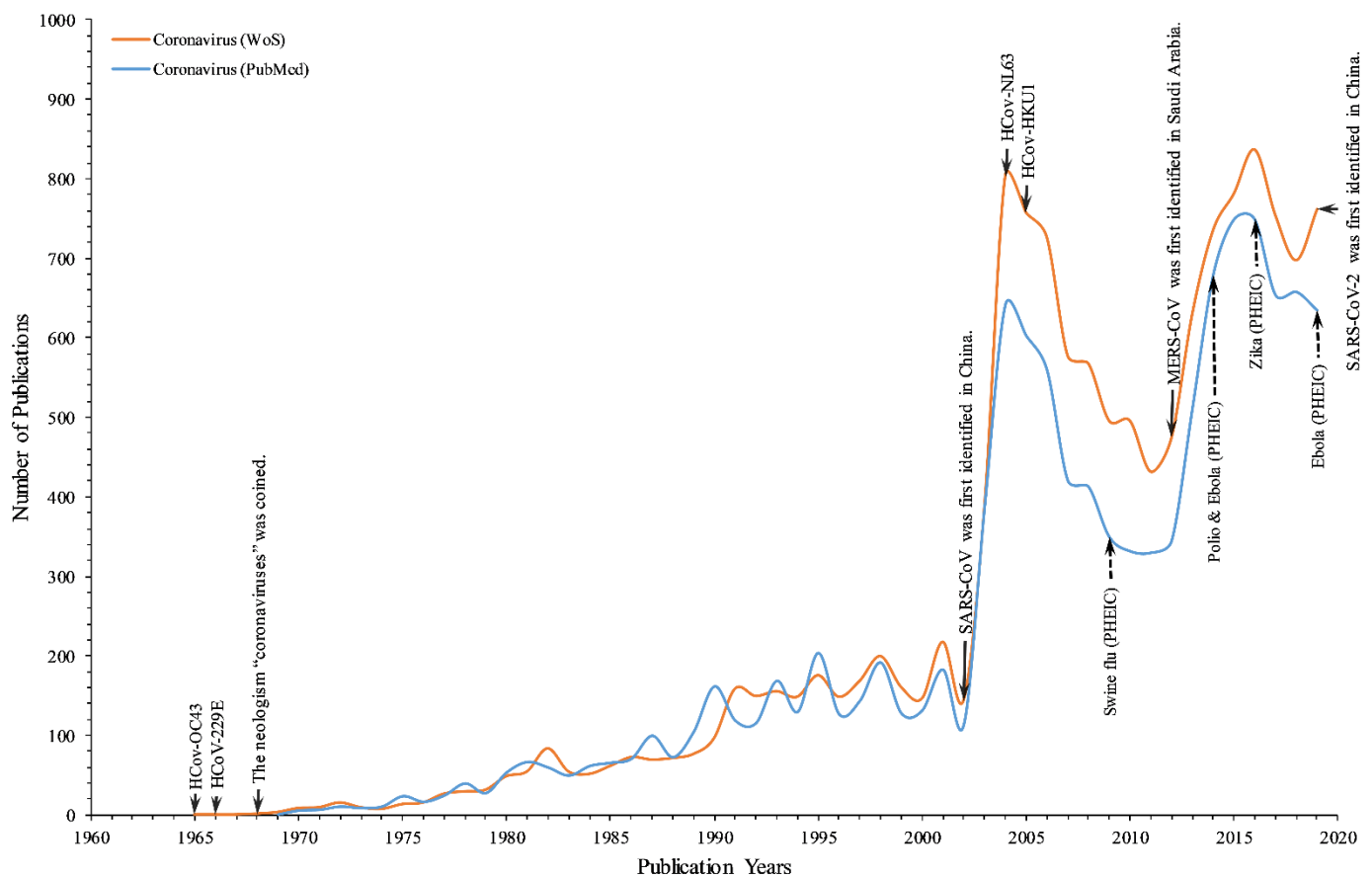
Fig. 2. Number of publications with "coronavirus" in Web of Science and PubMed. Coronaviruses are positive-sense single-stranded RNA viruses found in a wide variety of animals, as well as in humans. Seven strains of respiratory coronaviruses known to infect human-to-human: HCoV-OC43 (1965), HCoV-229E (1966), SARS-CoV (2002), HCoV-NL63 (2004), HCoV-HKU1 (2005), MERS-CoV (2012) and SARS-CoV-2 (2019). WHO declared the 2019-nCoV outbreak a public health emergency of international concern (PHEIC) on 30 January 2020. This is the 6th time WHO has declared a PHEIC since the International Health Regulations (IHR) came into force in 2005. Before that, there have been five global health emergencies since such declaration was formalized: swine flu (2009), polio (2014), Ebola (2014 then again in 2019), and Zika (2016).

Does a virus' name really matter?

As mentioned above, in recent years, humans have witnessed several outbreaks of infectious diseases caused by viruses, with common names given by stakeholders. However, each round of naming practice is not always successful. As a case in point, some strongly-held names such as "Middle Eastern Respiratory Syndrome" [11] and "Swine flu" were accused of unintentional social impacts and negative economic by stigmatizing certain industries or communities (**Figure 2**). "Swine flu," an influenza strain known to have originated in pigs, resulted in causing great financial damage to farmers, despite there being no evidence that it could be spread via pork consumption. Since these incidents, in May 2015, WHO released some naming conventions for the naming of new human diseases [12].

Unfortunately, accompanying with the 2019-nCoV outbreak, another massive epidemic spread virally over the world. WHO has to declare this massive infodemic as the "2019-nCoV 'infodemic'" on 2 February [13]. On 8 February, the *Lancet* published a statement in solidarity with Chinese professionals in combating the novel coronavirus outbreak and called upon fighting against the army of infodemics [14,15]. Later, more and

more public health scientists have endorsed this statement. Previous evidence suggests that the Internet, by its very nature, could amplify and relay such infodemics swiftly worldwide, furtherly cause global panic and worsen stigmatization of the coronavirus outbreak epicenter [16,17].

Recognizing the current case in such infodemic scenario, can we learn from history how not to make mistakes? To find the answer, we examine the popularity of top search queries of the 2019 novel coronavirus via the Google Trends Index (**Figure 3**). Those dynamic shares are faithful indicators of collective behaviours across various regions over time. The findings reveal that the 2019 novel coronavirus is thought to have originated in China, which led to it being frequently named the "Wuhan coronavirus" or "Chinese coronavirus" and our survey pinpoint that those unofficial names might have contributed to recent backlash against Chinese people.
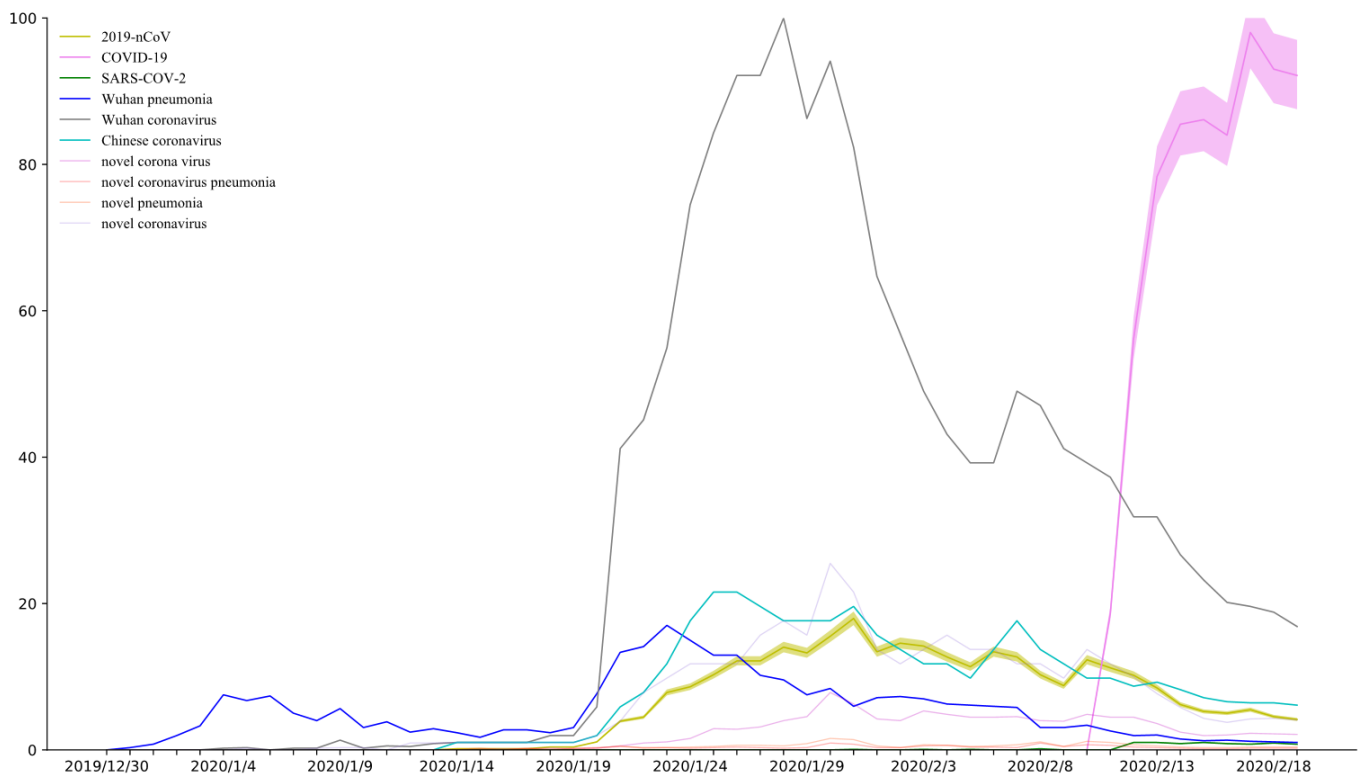
Fig. 3. Illustration on the top search query volumes of the 2019 novel coronavirus in infodemic scenario. The Google Trends facsimiles the popularity of "2019-nCoV" (yellow), "COVID-19" (violet), "SARS-CoV-2" (green), "Wuhan pneumonia" (blue), "Wuhan coronavirus" (grey), "Chinese coronavirus" (cyan), "novel corona virus" (magenta), "novel coronavirus pneumonia" (red), "novel pneumonia" (orange), and "novel coronavirus" (purple) from 30 January 2019 to 19 February, 2020, respectively. Comparatively, some stigmatized monikers enjoy high frequencies of collective consumption. The portfolio of full-fledged official names would be beneficial to overdue correctives to such stigmatization. Notably, substantial pattern shifts are observed after the announcements of "2019-nCoV" and "COVID-19", while "SARS-CoV-2" had failed to keep abreast of "COVID-19".

Debate goes on: naming the 2019 novel coronavirus

Prior to a final standard name for the 2019 novel coronavirus and pandemic disease, debate on interim solutions has been going on. On 12 January 2020, WHO provisionally named the 2019 novel coronavirus disease "2019-nCoV". Later, an emergency ICD-10 code of U07.1 is assigned to the disease diagnosis of 2019-nCoV by the International Classification of Diseases (ICD).

On 7 February 2020, China's National Health Commission (CNHC) decided to temporarily call the disease "Novel Coronavirus Pneumonia" or "NCP". This official name has invoked intensive arguments outside as well as inside the scientific community. Firstly, Chinese scientists are divided on that official name. Supporters say the descriptive name follows typical classification practices, whereas opponents claim that it could be easily misunderstood and abused to cause unnecessary fears. Secondly, the word 'novel' is

confusing in the way that neither the disease nor the host range can be used to reliably determine the virus novelty. Arguably, high mutation and gene recombination rates make this type of virus ideal for pathogen evolution. Once viral mutation happened, it won't be "novel" any more.

Before that, the 2019 novel coronavirus was designated as "WH-Human-1 coronavirus" or "Wuhan-Human-1 coronavirus" by a group of scientists in *Nature* on 3 February [18]. In the same vein, on 11 February, another name "HARS-CoV", with 'Han' standing for 'Wuhan in Chinese', was proposed in *The Lancet* [19]. Obviously, such practices are against the naming principles of WHO [20] – geographic locations should be avoided in disease names, and the name should be short and easy to pronounce. Such names might provoke unintended negative impacts by stigmatizing Wuhan citizens and even Chinese people. Those dire messages take hold should be duly corrected, as well as other similar paradigms [21–26].

In response to such concerns, on 11 February, WHO officially renamed "2019-nCoV" as "COVID-19", with 'CO' meaning 'corona', 'VI' for 'virus', 'D' for 'disease', and '19' referring to 2019. This generic descriptive reassignment offers an overdue corrective to those strongly-held but flawed notions, with the hope of minimizing stigma.

Coinciding with the WHO's latest announcement, in a bioRxiv preprint [5], a new name "Severe Acute Respiratory Syndrome coronavirus 2" or "SARS-CoV-2" was penned by the Coronavirus Study Group of the International Committee on Taxonomy of Viruses (ICTV-CSG) on the same day. ICTV-CSG explains that this designation highlights the new strain's similarity to the SARS-CoV. It is unclear whether the ICTV will take this suggestion into consideration.

In the real dilemma, "SARS-CoV-2" is not a nomenclature endorsed by WHO and some prominent virologists yet [27,28]. Outside the academic-industrial sphere, it also was argued against. Although "SARS-CoV-2" seems to be natural for ICTV-CSG to add a numeral '2' behind "SARS-CoV" to signify their relation, many prominent scientists scramble to refute the latest claim. To the untrained eye, the hasty designation may mislead the public to perceive a severer strain virus as a direct descendant of SARS-CoV, rather than a close affinity for the causative agent of China's another major viral outbreak in 2002-03. Before that, on 5 February, Prof. Shibo Jiang and his colleagues proposed another name, "Pneumonia Acute Respiratory Syndrome Coronavirus" or "PARS-CoV" in *Cellular & Molecular Immunology* [29]. In the same token, this assignment also intends to retain equivalent terminology of SARS-CoV.

In fact, the looming worry is that the public are susceptible to SARS-CoV [16], which evokes the memory of higher case fatality ratio. On 9 February, Chen Huan-chun, Chinese academician and virologist, made a public apology for mistakenly saying 2019-nCoV is SARS-CoV, which had stung a sensitive nerve and aroused great consternation in the Chinese public.

It is necessary to punctuate heuristic cautions and continuous introspection of previous multifarious names [11,16,20,27], which is a bedrock of such scientific efforts. Recently, global profusion of squab candidates has been discussed inside the scientific community, as well as on social media. For example, HCoV-19 (Human coronavirus 2019) [28]; TARS-CoV [30] and CARS-CoV, with 'ARS' standing for 'acute respiratory syndrome', 'T' for 'transmissible', and 'C' for 'contagious'. Whatever merits and demerits, some of them with plausible reasons should be fairly recognized. Before the pull of academic accession, authority should take an open mind to appreciate modest introspections and rededications to such collective efforts and then make informed and judicious choice. On 22 February, CNHC officially renamed the temporary English name "NCP" as "COVID-19", with the hope of sitting well with the reference of WHO and further discouraging the use of stigmatized titles [31].

Conclusions and implications

Previous evidences strongly suggest that full-fledged official names are duly contribute to the resilience of healthy collective usages in infodemic scenario. Understanding of the way the army of unsung naming rules has strengthened and enriched the integrity and quality of naming practices under the umbrella of original mission of stakeholders remains nominal rather than substantial [11,16,20,27]. It is necessary to punctuate heuristic cautions and continuous introspection of previous multifarious names, which is a bedrock of such scientific efforts. In retrospect, according to the traditional naming conventions [12], names that should be avoided in disease names include people's names (e.g. Creutzfeldt-Jakob disease, Chagas disease), geographic locations (e.g. Middle East Respiratory Syndrome, Rift Valley fever, Spanish Flu, Lyme disease,

Crimean Congo hemorrhagic fever, Japanese encephalitis), species of animal or food (e.g. Swine flu, bird flu, monkey pox, equine encephalitis, paralytic shellfish poisoning), cultural, populational, industrial or occupational references (e.g. legionnaires, miners, butchers, cooks, nurses), and names that incite undue fear (e.g. unknown, death, fatal, epidemic). Obviously, some naming practices have deviated from course with regard to the ways it has conflicted with the advancement of both scientific doctrine and social progress.

As the precaution, noteworthily, the word "novel" was recommended by WHO [12] for "indicating a new pathogen of a previously known type, recognizing that this term will become obsolete if other new pathogens of that type are identified." Therefore, stakeholders frequently reserve 'novel' for striking new type of virus, lest the word lose fundamentally its impact without regular amendments.

Learning lessons of naming practices could surely be seen as a pressing necessity of yielding up some guidelines for the adoption of practical principles intended to enhance the possibility for the lessening of stigmatization and for the avoidance of discrimination. At this critical moment, an epoch-making name is expected to be scientifically pithy and socially acceptable, with the faith of minimizing unintentional negative impacts on nations, economies and people. This is a positivist doctrine, not merely for naming a virus but for the vitality of science and the promotion of social progress.

References

1.   Heymann DL, Shindo N (2020) COVID-19: what is next for public health? *Lancet* **395**: 542–545.
2.   Almeida JD, Berry DM, Cunningham CH, Hamre D, Hofstad MS, Mallucci L, Mcintosh K, Tyrrell DAJ (1968) Virology: Coronaviruses. *Nature* **220**: 650–650.
3.   Waterson AP, Wilkinson L (1978) *An Introduction to the History of Virology*. Cambridge University Press, Cambridge, Massachusetts.
4.   Tyrrell DAJ (1979) A New Science: What, Why, And How? *Br Med J* **1**: 45.
5.   Gorbalenya AE, Baker SC, Baric RS, Groot RJ De, Gulyaeva AA, Haagmans BL, Lauber C, Leontovich AM (2020) The species and its viruses – a statement of the Coronavirus Study Group. *bioRxiv*.
6.   de Wit E, van Doremalen N, Falzarano D, Munster VJ (2016) SARS and MERS: recent insights into emerging coronaviruses. *Nat Rev Microbiol* **14**: 523–534.
7.   Cui J, Li F, Shi Z-L (2019) Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol* **17**: 181–192.
8.   Michel J-B, Shen YK, Aiden AP, Veres A, Gray MK, Pickett JP, Hoiberg D, Clancy D, Norvig P, Orwant J, et al. (2011) Quantitative Analysis of Culture Using Millions of Digitized Books. *Science* **331**: 176–182.
9.   Fenner F (1971) The nomenclature and classification of viruses the International Committee on Nomenclature of Viruses. *Virology* **46**: 979–980.
10.   Fenner F (1976) The Classification and Nomenclature of Viruses: Summary of Results of Meetings of the International Committee on Taxonomy of Viruses in Madrid, September 1975. *J Gen Virol* **31**: 463–470.
11.   Enserink M (2013) Amid heightened concerns, new name for novel coronavirus emerges. *Science* **340**: 673.
12.   WHO (2015) World Health Organization Best Practices for the Naming of New Human Infectious Diseases. 1–3.
13.   WHO (2020) *Novel Coronavirus (2019-nCoV) Situation Report - 13*. , Geneva.
14.   (2020) COVID-19: fighting panic with information. *Lancet* **395**: 537.
15.   Calisher C, Carroll D, Colwell R, Corley RB, Daszak P, Drosten C, Enjuanes L, Farrar J, Field H, Golding J, et al. (2020) Statement in support of the scientists, public health professionals, and medical professionals of China combatting COVID-19. *Lancet*.
16.   Enserink M (2013) War Stories. *Science* **339**: 1264–1268.
17.   Rothkopf DJ (2003) When the Buzz Bites Back. *Washington Post* B01.
18.   Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, Hu Y, Tao Z-W, Tian J-H, Pei Y-Y, et al. (2020) A new coronavirus associated with human respiratory disease in China. *Nature*.
19.   Wang L-F, Anderson DE, Mackenzie JS, Merson MH (2020) From Hendra to Wuhan: what has been learned in responding to emerging zoonotic viruses. *Lancet* **395**: e33–e34.
20.   Fukuda K, Wang R, Vallat B (2015) Naming diseases: First do no harm. *Science* **348**: 643–643.
21.   (2020) Stop the Wuhan virus. *Nature* **577**: 450–450.
22.   Callaway E, Cyranoski D (2020) Why snakes probably aren't spreading the new China virus. *Nature*.
23.   Callaway E, Cyranoski D (2020) China coronavirus: Six questions scientists are asking. *Nature* **577**: 605–607.
24.   Callaway E (2020) China coronavirus: labs worldwide scramble to analyse live samples. *Nature* **578**: 16–16.
25.   Cyranoski D (2020) Did pangolins spread the China coronavirus to people? *Nature*.
26.   Liu S-L, Saif L (2020) Emerging Viruses without Borders: The Wuhan Coronavirus. *Viruses* **12**: 130.

27. Enserink M (2020) Update: 'A bit chaotic.' Christening of new coronavirus and its disease name create confusion. *Science*.

28. Jiang S, Shi Z, Shu Y, Song J, Gao GF, Tan W, Guo D (2020) A distinct name is needed for the new coronavirus. *Lancet*.

29. Jiang S, Xia S, Ying T, Lu L (2020) A novel coronavirus (2019-nCoV) causing pneumonia-associated respiratory syndrome. *Cell Mol Immunol*.

30. Jiang S, Shi Z-L (2020) The First Disease X is Caused by a Highly Transmissible Acute Respiratory Syndrome Coronavirus. *Virol Sin*.

31. China Daily, China adopts COVID-19 as official English name for disease, Last updated 2020, Accessed on 2020.