Lancet Infectious Diseases. Personal View.

The global epidemiology of non-influenza RNA respiratory viruses – data gaps and a growing need

Tang JW\(^1,2\), Koopmans MP\(^3\), Lam TT\(^4\), Zaraket H\(^5\), Lipkin WI\(^6\), Jean-Michel Heraud\(^7\), and the INSPIRE investigators

\(^1\)University Hospitals Leicester NHS Trust, Leicester, UK
\(^2\)University of Leicester, Leicester, UK
\(^3\)Erasmus Medical Centre, Rotterdam, The Netherlands
\(^4\)The University of Hong Kong, Hong Kong SAR, China
\(^5\)American University of Beirut, Beirut, Lebanon
\(^6\)Columbia University, New York, NYC, USA
\(^7\)Institut Pasteur of Madagascar, Antananarivo, Madagascar.

Correspondence to: Dr Julian W Tang
Clinical Microbiology, University Hospitals of Leicester NHS Trust
Level 5 Sandringham Building, Leicester Royal Infirmary
Infirmary Square, Leicester LE1 5WW, UK.
Email: julian.tang@uhl-tr.nhs.uk; jw tang49@hotmail.com
Tel: 0116 258 6516. Fax: 0116 255 1949
Abstract
Together with influenza, the non-influenza RNA respiratory viruses (NIRVs), including respiratory syncytial-, parainfluenza-, corona-, rhino- and human metapneumoviruses, contribute a significant global health burden, as recognised by the recent World Health Organization (WHO)’s BRaVe (Battle against Respiratory Viruses) initiative. However, in contrast to influenza viruses, very little is known about the contemporaneous global diversity of these viruses, and the relevance for development of pharmaceutical interventions. Although far less advanced than for influenza, antiviral drugs and vaccines are in different stages of development for several of these viruses, though none are yet licensed. This lack of global genetic data is a significant gap and impediment to the eventual licensing of new antiviral drugs and vaccines for these NIRVs. Enhanced genetic surveillance will assist and boost research and development (R&D) into new antiviral drugs and vaccines for these viruses. In addition, understanding the global diversity of respiratory viruses is also part of emerging disease preparedness, as non-human coronaviruses and paramyxoviruses have been listed as priority concerns through a recent WHO R&D blueprint initiative for emerging infectious diseases. In this Personal View, we explain further the rationale and emphasise the need for greater investment into expanding the genetic database for these NIRVs.

Word count: abstract: 199; main text: 2652
Keywords: respiratory viruses; antiviral; vaccine; influenza; RSV; parainfluenza; human metapneumovirus; rhinovirus; coronavirus; sequencing; molecular epidemiology

Authors’ contributions: JWT, MK, WIL developed the concept and drafted the article text; MK, TL and HZ conceptualised and produced the Figure and Table; all authors critically reviewed/commented on the final manuscript text.

Conflicts of interest: All authors are currently completing the authors declaration and Conflict of Interest Forms

Funding source: None

Ethics approval: Not required
According to World Health Organization (WHO) figures, lower respiratory tract infections are listed in the top four causes of death, globally, after ischaemic heart disease, chronic obstructive pulmonary disease and stroke, all of which are non-infectious chronic conditions.¹ Even in these latter, non-infective conditions, respiratory infections often trigger life-threatening exacerbations of these diseases. At the very least, this makes respiratory infections, worldwide, the leading cause of death due to infection.

The rationale for investment in the global surveillance of influenza viruses is driven by the need to ensure the efficacy of seasonal influenza vaccines, and to monitor circulating strains for pandemic potential or resistance against antiviral drugs. Less well established is the importance and feasibility of surveillance for non-influenza respiratory viruses (NIRVs), including respiratory syncytial virus (RSV), parainfluenza viruses (PIV), human metapneumovirus (HMPV), rhinoviruses (RV), and coronaviruses (CoV), despite scientific consensus that the burden of disease attributable to these infections is considerable. Like influenza, these NIRVs are also RNA viruses which have a relatively high mutation rate due to the lack of proof-reading that is inherent in the replication of RNA genomes. In combination, these viruses are responsible for a greater annual morbidity and mortality than influenza viruses, across all age groups.²⁻⁵ This is particularly so when the full range of mild to severe respiratory illness is taken into account, from common colds (mostly due to rhinovirus and coronavirus infections) that affect people of all ages, to more severe respiratory illness requiring hospitalization.⁶

Recently, the World Health Organization’s (WHO’s) BRaVe (Battle against Respiratory Viruses) initiative has highlighted the need for enhanced clinical and epidemiological surveillance for respiratory viruses,⁷ with a focus on the development of a vaccine for RSV.⁸ This acknowledges the significant impact and health burden of RSV,³⁻⁹⁻¹² which is now one of the highest priorities for intervention out of all of these NIRVs. Under
one of the document’s sections on improving the diagnosis of severe acute respiratory illness (SARI) and diagnostic testing, viral sequencing by ‘deep sequencing’ (where all viral populations present in the sample, including majority and minority, are sequenced) is explicitly listed as a research priority. We would like to emphasise and expand upon this point.

A key reason for the lack of global data on the burden of NIRVs is the relatively fewer cost-effective, sensitive assays that can be used, routinely, in everyday diagnostic settings. Many healthcare facilities will have some capability for influenza virus and RSV (particularly where paediatric services are provided) testing, whereas the other common respiratory viruses may not be routinely screened for at all. Transmission of respiratory viruses have caused outbreaks among patients in intensive care units and oncology wards, and timely testing may decrease the unnecessary use of antimicrobials and importantly, limit the transmission by effective isolation of infected patients.

In some regions, specialized centers can offer multiplex PCR-based testing, but the cost of the diagnostic testing compared to the (perceived) lack of impact on clinical decision-making precludes widespread use. Nevertheless, these multiplex molecular assays and the high throughput sequencing protocols that are being piloted, now allow the detection of a wider spectrum of respiratory viruses from a single sample. This has enabled an unprecedented opportunity to focus on the burden of these NIRVs.

How then would one organize surveillance for these other respiratory pathogens? One important starting point is whom and how to sample. Groups at higher risk of disease associated with seasonal influenza virus and NIRV infections are similar, and include: children; the elderly; the immunocompromised; and individuals with chronic co-morbidities (asthma, chronic obstructive pulmonary disease, cardiac and renal failure, and/or
diabetes). Both influenza and NIRV infections may result in hospitalisation due to exacerbation of these conditions. In these at risk populations, co-infections with more than one of these respiratory viruses are not uncommon, though the clinical significance of multiple co-infections is still unclear.

Many studies have demonstrated that in young children, NIRVs, especially RSV, are the predominant viral cause of respiratory morbidity and mortality, with accumulating evidence to suggest that infection in early childhood with NIRVs, such as RSV and RV, in predisposed children, may result in the development of increased airways sensitivity and asthma later in life. This is one example of the impact of both the acute and more chronic respiratory virus-associated healthcare burden. Taken together, we conclude that there is a case to be made for enhanced surveillance for these other NIRVs.

One key factor questioning the usefulness of such surveillance is that there are no agreed interventions. Nevertheless, current research and development (R&D) interest into NIRV therapeutics (as evidenced by PubMed searches for each of these viruses by name AND ‘vaccine’ or ‘antiviral’) will benefit from the availability of more large-scale, full-genome NIRV sequences. Although a limited survey of the United States (US) and European Union (EU) clinical trials websites shows a substantially greater R&D investment into influenza viruses, quite a few clinical trials targeting the other NIRVs are currently in progress (Table 1).

Would understanding the global genomic diversity of these viruses become increasingly important with the progression of these product pipelines? Yes, we believe so, because not all antiviral targets are subject to the same mutational pressure, with some being highly conserved, and some highly variable, such as the fusion (F) and attachment (G) proteins of the paramyxoviruses (e.g. RSV, HMPV), respectively.
and PIV vaccines are in various stages of clinical evaluation, and assessing the impact of vaccine-induced immune responses in the context of continued viral evolution and the subsequent potential for vaccine escape will be an essential consideration when deciding on the annual vaccine composition. For antiviral agent development, since modern rational drug design allows for the development of new antivirals targeting specific viral proteins, it is vital to assess the mutation rates and genetic diversity of these therapeutic targets, to ensure the long-lasting efficacy of these agents. Thus, this baseline data will allow an estimate of the naturally occurring mutation rate for the individual genes in each of these viruses, which may then affect the choice of target for the development of any future specific antiviral drugs or vaccines against them.

With influenza treatment, it is recognized that immunocompromised patients on long-term antiviral treatment (or prophylaxis) are predisposed to the rapid evolution and development of drug resistance, either in total or subpopulations of viruses within the host. It is therefore very likely that any antiviral agent developed for the NIRVs may show similar characteristics – unless it can be specifically designed to act on a more stable viral protein.

This has already been reported for palivizumab, a humanized monoclonal antibody directed against the RSV F protein. Despite its relatively rare usage and the highly conserved nature of its viral protein target, resistance to palivizumab (though strictly speaking, not an antiviral drug) has already been reported, due to variations in the F protein binding site.

One study from Italy described a human monoclonal antibody (MPE8) that cross-neutralizes RSV and HMPV by binding to two highly conserved anti-parallel β-strands on the pre-fusion viral F protein. The authors also found naturally occurring antibodies with this same target specificity in some patient sera, and therefore postulated this pre-fusion F protein
as a potential vaccine candidate.\textsuperscript{69} A wider, global genetic survey of this target in RSVs from other countries would help to confirm its suitability for development into a global RSV vaccine.

For rhinoviruses, the expansion of the limited number of complete viral genome data available (currently only around 200 full genomes, compared to over 5000 for influenza A/H1N1pdm09 and over 6000 for influenza A/H3N2) is essential to understand further their natural genetic diversity and underlying evolutionary forces (i.e. viral gene mutation and recombination driven by selection pressure), as a foundation for designing new antiviral drugs and vaccines.\textsuperscript{70-72}

Recent work on coronavirus infections has been dominated by studies on severe acute respiratory syndrome (SARS)- and Middle East respiratory syndrome (MERS)-associated coronaviruses. However, the development of an antiviral drug effective against the coronavirus family as a whole will also be effective against the milder, but much more prevalent common cold viruses (e.g. CoV OC43, 229E, NL63 and HKU1).\textsuperscript{73} Using a mouse hepatitis coronavirus, one study found a relatively limited repertoire of resistance mutations to an experimental compound that inhibits the action of a coronavirus fusion protein. This fusion protein, HR2 (heptad repeat 2) is required to enter the host cell and its inhibition will block infection. The authors found that after multiple passages in vitro, most mutations conferring resistance were found to lie within a limited 19 amino acid region of the related HR1 region in the mouse hepatitis coronavirus spike (S) protein.\textsuperscript{74} Should the development of such an HR2 fusion inhibitor progress further, it will be essential to perform a global surveillance to determine if any of these HR1 mutations conferring resistance already exists, naturally; this will imply a higher potential for the emergence of resistant mutants.
These examples further highlight the need for baseline genetic information about these NIRVs, as well as ongoing monitoring for the emergence of drug resistance, once such antiviral drugs become available.

Thus, from a design viewpoint, it is imperative that an initial antiviral target be selected from a well-conserved genomic region that mutates only very slowly. This careful selection can only be achieved with a large-scale characterization of the natural genetic diversity of these NIRVs. Such an enhanced genetic surveillance approach to the NIRVs, together with comprehensive conventional epidemiological data, will naturally fuel the search for new drugs and vaccines to combat them. However, when reviewing the currently available genomic information in GenBank, the publicly available sequence database, the total number of whole genome influenza sequences outnumbers the combined total for the NIRVs by ten-fold (Figure 1). We believe that this is a large data gap that needs to be filled with some urgency.

In addition, genetic surveillance has many other important public health implications, such as the identification of novel respiratory viruses which can clinically mimic other more common NIRVs. With the recent emergence of several novel respiratory viruses able to cause multiple outbreaks with a varying degree of person-to-person transmission potential (e.g. SARS-CoV, pandemic A/H1N1pdm09, avian influenza’s A/H5N1 and A/H7N9, and MERS-CoV), it is also clear that the surveillance of both ‘mild versus severe’ and ‘community versus hospitalized’ acute respiratory infections is becoming more crucial for the detection of such novel viruses.

The recent and rapid development of quicker and cheaper (high-throughput) deep sequencing platforms now readily allows the profiling of pathogen genetic sequence diversity directly from clinical samples. This could be applied very productively to characterize the
global population of NIRVs, but will require substantial funding support and collaboration between clinical, public health and research experts.

So how could this large-scale sequencing of NIRVs be funded? In principle, there are two options, but they are not necessarily mutually exclusive. An initial large-scale whole genome sequencing project funded by either government or private research institutions to demonstrate feasibility and utility would be the most likely initial source of funding. As the clinical and public health utility of this approach becomes well-established, governmental funding is likely to follow, perhaps with support from commercial companies, such as those developing antiviral drugs and/or vaccines against these NIRVs.

This is the case for Pandemic Influenza Preparedness Framework (PIPF), which includes an annual contribution by vaccine and diagnostic pharmaceutical companies towards this partnership with the WHO. The most direct benefit and savings to the healthcare system from a successful vaccine will be fewer admissions to hospital and visits to GPs, as fewer people will become infected and develop disease from these NIRVs. Failing this, effective antivirals will reduce morbidity and mortality for those people who do need hospital admission due to NIRV infections, or may prevent the need for admission if such drugs are available and prescribed in a timely manner by their GPs, therefore reducing absenteeism from school or work. Both of these benefits will strengthen the case for the cost-effectiveness of this investment and support the argument for its longer-term maintenance and sustainability.

So for example, given an average reagent cost per sample of £200 (~US$246) for deep (high throughput) sequencing, with an average of 24 samples per year (i.e. 2 samples per month) from 10 global sampling sites, for each of these 5 groups of NIRVs (RSV, PIV, HMPV, CoV, RV - leaving out their individual species/subtypes for this illustration), and
excluding additional staff, overheads and dry-ice sample shipments to the nearest laboratory with deep sequencing capability, over a 5-year surveillance period, the cost of this venture would be in the region of £1.2 million (~US$1.48 million), yet would generate 6,000 complete genomes of NIRVs and significantly expand their current sequence dataset by three times. Including the individual viral subtypes and species (PIV type 1-4 and CoV OC43, 229E, NL63, HKU1), this becomes £2.64 million (~US$3.25 million). Increasing the number of global sampling sites from 10 to 20 doubles this to £5.28 million (~US$6.49 million), etc.

At sites where samples are already collected for routine influenza surveillance, the residual volume from these samples can be used relatively easily for the surveillance of these NIRVs, where the presence of these viruses has been detected through routine diagnostic testing, as for influenza. Further savings can be obtained by batching these samples for periodic, large-scale, sequencing runs to ensure that all lanes are filled to capacity.

The greater number of viruses and the more sites sampled will enhance the resolution and our understanding of this viral diversity. If this surveillance can be maintained (as with influenza) on an annual, global basis, this may well capture the most important viral variation that has the potential to impact on the clinical effectiveness of any antiviral drug and/or vaccine developed for combatting these viruses. In fact, an analysis starting with a greater number of sites may allow a subsequent reduction in the number of sites and samples if similar patterns of genetic diversity are seen within samples obtained from neighbouring regions or populations. A minimum number of samples could then be obtained from fewer key sentinel sites (which may not necessarily all match the same sites currently used for influenza surveillance), which would make for a much more affordable and efficient annual surveillance program – similar to the present situation for global influenza surveillance.
The actionable outcomes of such a project would include (but are not limited to): i) a greater understanding of the pattern and spectrum of genetic diversity for each of these viruses in different populations in different parts of the world, which will also identify conserved genomic regions that could then be targets for the development of antiviral drugs and vaccines with long-lasting effectiveness; ii) the identification of viral strains with specific mutations that may link to possible increase of clinical virulence, leading to the development of routine diagnostic assays to detect such strains in patients to allow clinicians and public health teams to prepare for any potential system impacts (e.g. increased emergency department visits, and poorer clinical outcomes); iii) the identification of any unusual patterns of increased mutation rates for any of these viruses in specific populations or regions (i.e. mutant hotspot regions for any of these viruses) that may warrant particular attention and perhaps even customized antiviral and vaccine development (e.g. combination therapy or multi-epitope vaccine targets) to control the emergence of mutant viruses that could then spread worldwide.

Thus, considering a catch-all approach to the surveillance of the NIRVs may be more cost effective than following the single pathogen path, particularly where residual samples from existing influenza surveillance testing can be used. There is no doubt that the NIRVs impact on all age groups, either directly or indirectly via the exacerbation of pre-existing comorbidities. The development of specific antiviral drugs and vaccines for these viruses will have a substantial beneficial impact on the health and well-being of our children and our elderly. For this reason, we should strive together, as an increasingly interconnected global community, to make this happen.

References


17. Chen C, Zhao B, Yang X, Li Y. Role of two-way airflow owing to temperature difference in severe acute respiratory syndrome transmission: revisiting the largest


INSPIRE (International Network for the Sequencing of resPIRratory vIrusEs: http://www.global-inspire.org/) are (in alphabetical order): 

Authors/ Affiliations

1. Asha Mary Abraham, Christian Medical College, Vellore, India.
2. Amal Baraket, Centre National de référence de la grippe, Rabat, Morocco.
3. Seweryn Bialasiewicz, The University of Queensland, Brisbane, Australia.
4. Miguela Caniza, St Jude Children's Research Hospital, Memphis, TN, USA.
5. Paul KS Chan, The Chinese University of Hong Kong SAR, Hong Kong, China.
6. Cheryl Cohen, National Institute for Communicable Diseases (NICD); University of the Witwatersrand, Johannesburg, South Africa.
7. André Corriveau, Government of the Northwest Territories, Yellowknife, NT, Canada.
8. Benjamin J Cowling, The University of Hong Kong, Hong Kong SAR, China.
9. Steven J Drews, Alberta Provincial Laboratory for Public Health, University of Alberta, Edmonton, AB, Canada.
10. Marcela Echavarria, Centro de Educacion Medica e Investigaciones Clinicas, Buenos Aires, Argentina.
11. Ron Fouchier, Erasmus Medical Centre, Rotterdam, The Netherlands.
13. Todd Hachette, Dalhousie University, Halifax, NS, Canada.
15. Hamid Jalal, Addenbrookes Hospital, Cambridge, UK.
16. Lance Jennings, University of Otago, Christchurch, New Zealand.
17. Alice Kabanda, National Reference Laboratory, Kigali, Rwanda.
18. Herve A Kadjo, Pasteur Institute, Abidjan, Côte d'Ivoire.
19. Mohammed Rafiq Khanani, Dow University of Health Sciences, Karachi, Pakistan.
20. Evelyn SC Koay, National University of Singapore, Singapore.
22. Mel Krajden, British Columbia Centre for Disease Prevention and Control, Vancouver, BC, Canada.
23. Tommy T Lam, The University of Hong Kong, Hong Kong SAR, China.
24. Hong Kai Lee, National University Hospital, Singapore.
25. W. Ian Lipkin, Columbia University, New York, NY, USA.
27. David Marchant, University of Alberta, Edmonton, AB, Canada.
28. Hidekazu Nishimura, Sendai Medical Center, Sendai, Japan.
29. Pagbajabyn Nymadawa, Ulaanbaatar, Mongolia.
30. Benjamin A Pinsky, Stanford University School of Medicine, Stanford, CA, USA.
31. Sanjiv Rughooputh, Central Health Laboratory, Candos, Mauritius.
32. Joseph Rukelibuga, National Reference Laboratory, Kigali, Rwanda.
33. Taslimarif Saiyed, Centre for Cellular and Molecular Platforms, Bangalore, India.
34. Anita Shet, St. John’s Medical College Hospital, Bangalore, India.
35. Theo Sloots, University of Queensland, Brisbane, Australia.
36. JJ Muyembe Tamfum, Institut National de Recherche Biomédicale (INRB), Kinshasa Gombe, Democratic Republic of Congo.
37. Julian W Tang, University Hospitals Leicester, University of Leicester, Leicester, UK.
38. Stefano Tempia, National Institute for Communicable Diseases, Sandringham, South Africa.
39. Sarah Tozer, Centre for Children’s Health Research, Brisbane, Australia.

40. Florette Treurnicht, National Institute for Communicable Diseases, Sandringham, South Africa.

41. Matti Waris, University of Turku, Turku, Finland.

42. Aripuana Watanabe, Federal University of Sao Paulo, Sao Paulo, Brazil.

43. Emile Okitolonda Wemakoy, Université de Kinshasa, Kinshasa, Democratic Republic of Congo.

44. Hassan Zaraket, American University of Beirut, Beirut, Lebanon.
Figure legend

**Figure 1.** Geographic sources of genome sequence data from different respiratory RNA viruses. Only field strains with complete genome sequences (or nearly complete, >80% of full-length) and geographic information in GenBank ([www.ncbi.nlm.nih.gov/genbank](http://www.ncbi.nlm.nih.gov/genbank)) were counted (accessed Nov 2015). **Panel A** shows the counts of human influenza A and B viruses (FluA, FluB). **Panel B** shows the counts of four human coronavirus (CoV) species (NL63, OC43, HKU1 & 229E), human parainfluenza virus (hPIV), human respiratory syncytial virus (hRSV), rhinovirus (RV) and human metapneumovirus (HMPV), which are represented by different colours shown in the legend box. Radius of pie chart is the $\log_2$ of the total number of sequences from the country/region.
Table 1. Numbers of ongoing clinical trials related to vaccine (V) and antiviral agent (A) development for the different respiratory viruses.

<table>
<thead>
<tr>
<th>Virus</th>
<th>Influenza</th>
<th>RSV</th>
<th>PIV</th>
<th>HMPV</th>
<th>Coronavirus</th>
<th>Rhinovirus</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>A</td>
<td>V</td>
<td>A</td>
<td>V</td>
<td>A</td>
<td>V</td>
</tr>
<tr>
<td>USA</td>
<td>1561</td>
<td>190</td>
<td>49</td>
<td>13</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>


Figure 1. Geographic sources of genome sequence data from different respiratory RNA viruses. Only field strains with complete genome sequences (or nearly complete, >80% of full-length) and geographic information in GenBank (www.ncbi.nlm.nih.gov/genbank) were counted (accessed Nov 2015). Panel A shows the counts of human influenza A and B viruses (FluA, FluB). Panel B shows the counts of four human coronavirus (CoV) species (NL63, OC43, HKU1 &
229E), human parainfluenza virus (hPIV), human respiratory syncytial virus (hRSV), rhinovirus (RV) and human metapneumovirus (HMPV), which are represented by different colours shown in the legend box. Radius of pie chart is the log2 of the total number of sequences from the country/region.