Abstract: Respiratory syncytial virus is the most common cause of serious lower respiratory tract illness in infants and children and causes significant disease in the elderly and immunocompromised. Recently there has been an acceleration in the development of candidate RSV vaccines, monoclonal antibodies and therapeutics. However, the effects of RSV genomic variability on the implementation of vaccines and therapeutics remain poorly understood. To address this knowledge gap, the National Institute of Allergy and Infectious Diseases and the Fogarty International Center held a workshop to summarize what is known about the global burden and transmission of RSV disease, the phylogeographic dynamics and genomics of the virus, and the networks that exist to improve the understanding of RSV disease. Discussion at the workshop focused on the implications of viral evolution and genomic variability for vaccine and therapeutics development in the context of various
immunization strategies. This paper summarizes the meeting, highlights research gaps and future priorities, and outlines what has been achieved since the meeting took place. It concludes with an examination of what the RSV community can learn from our understanding of SARS-CoV-2 genomics and what insights over sixty years of RSV research can offer the rapidly evolving field of COVID-19 vaccines.

**Keywords:** respiratory syncytial virus; genomics; evolution; vaccines; therapeutics; epidemiology; transmission; global health

**Introduction**

Discovered in 1956 [1], respiratory syncytial virus (RSV) is now the leading cause of hospitalization in infants and children due to bronchiolitis and pneumonia globally [2]. In addition to infants and children, those with compromised immune systems, those with chronic heart or lung disease [3], and the elderly [4], are at significant risk of severe infection and death. Due to recent technological advances, there are multiple vaccine and therapeutic interventions currently in development, including live-attenuated, chimeric, vector-based, particle-based, subunit, and monoclonal antibody-based platforms [5]. However, our understanding of viral genomic variability, and other factors that may interact with this such as seasonality, phylogeographic dynamics, and viral genotype and host phenotype interactions remains limited [6]. These factors are likely to impact on timing of vaccinations [7], decisions on the use of different vaccination strategies, and the possibility of immune pressure impacting on viral fitness and the development of escape mutants [8]. With this in mind, in September 2019 NIAID and Fogarty International Center co-sponsored a workshop to 1) summarize the existing evidence base for our understanding of the phylogeographic dynamics of RSV infection, and how these relate to genomic variability, disease phenotype and burden of disease; 2) learn from related viruses the potential impacts of genomic variability on immunization and therapeutic effectiveness; 3) predict the potential impact of
RSV genomic variability on interventions aiming to reduce disease burden; 4) discuss the need for standardization of RSV case definitions and clinical metadata curation; and 5) identify research priorities that will inform the introduction of, and surveillance subsequent to, the introduction of RSV immunizations and therapeutics.

**RSV global burden of disease**

RSV almost universally affects children worldwide before the age of 2 [9]; and cases appear to occur as early as the first day of life [10]. In low and middle income countries (LMICs), RSV is now the most common cause of childhood lower respiratory tract infections [2], with an estimated 33.1 [21.6-50.3] million cases, 3.2 [2.7-3.8] million hospitalizations, and 118,000 [94,600-149,400] deaths in children under 5 years of age in 2015 [11]. In a recent study in South Asia, RSV was the most common single etiological agent identified among neonates presenting with signs of severe bacterial infection (including fast breathing and chest indrawing) who subsequently died [10]. An RSV mortality study has shown that most childhood RSV deaths occur in infancy, with a peak in deaths in LMICs at 3-4 months of age, compared to 7 months in high income countries [12]. Most child mortality data have been based on hospital studies and in settings which under-represent remote and vulnerable populations and so extrapolations from this to whole populations in order to capture community deaths in LMIC may have under-estimated RSV mortality. Furthermore, correct RSV attribution as a cause of death is complex. RSV is found widely in children during RSV seasons and may not be the cause of death in a sick child. Conversely, a prior RSV infection (undetectable at the time of hospital presentation) may predispose a child to a subsequent fatal pneumococcal infection or result in hospitalization of a child who then dies from a nosocomial infection, as shown by the extended risk period of death after RSV infection [13]. Investment has been made in a number of ongoing community mortality surveillance studies [14–16] which seek to measure RSV mortality directly and study causal attribution and may lead to improved mortality estimates.
Less is known about the global burden of RSV acute lower respiratory tract infections in the elderly or those with underlying health conditions. A review and meta-analysis looking at RSV disease in those >65 years of age found that in 2015 there were an estimated 1.5 million (95% confidence interval [CI], 0.3 million–6.9 million) episodes in high income countries with 214,000 (95% CI, 100,000–459,000) hospitalizations[4]. Globally, estimates in this age group were 336,000 [186,000–614,000] hospitalizations and 14,000 [5,000–50,000] in-hospital deaths [4] although this was based on very limited LMIC data. There is currently little information on RSV morbidity and mortality in non-hospitalized elderly persons, in both high- and low- income settings.

Although uncertainty remains about the exact morbidity and mortality attributable to RSV at a global level, it is clear that a successful vaccine could lead to a number of direct and indirect health benefits, including a direct reduction in the number of hospitalizations and deaths in both children and adults. Reduction in RSV related cases would free up hospital bed spaces for treatment of other life threatening condition and contribute towards a reduction in the overuse of antibiotics[17,18] therefore positively impacting on antimicrobial resistance, a significant and growing problem globally [19,20]. Finally, a reduction in RSV disease would also have direct and indirect economic advantages. In 2017, the direct global costs of RSV treatment were estimated to be close to $6 billion dollars [13]. An indirect benefit of reduction in disease episodes and hospitalizations would be increased productivity for caregivers, for whom a single episode of hospitalized RSV infection in a child is estimated to lead to a loss of 7 days of economic productivity in high income countries [21], and 12 days in low and middle income settings [22].

Disease in infants has been a research focus, as vaccines are most likely to be introduced first for this age group. However there remain significant gaps in our understanding of burden of disease in elderly populations and pregnant women, particularly with respect to community
mortality data. There are gaps in our understanding of the economic burden of disease, including a paucity of cost effectiveness studies to determine the impact of interventions, and a lack of quantification of the long-term consequences of RSV disease, particularly recurrent wheezing and asthma. Detailed disease burden studies, in association with high resolution data on timing of RSV seasonality, epidemic seeding, persistence and community spread, are important factors for governments to consider in making decisions on whether and how to initiate new interventions for RSV.

**RSV seasonality**

In temperate climates, RSV disease occurs as regular yearly outbreaks usually between late fall and early spring. Timing of outbreaks may vary from year to year, and in the same year timing may vary between different geographic regions and even between nearby communities [19]. In tropical regions, disease timing is often more variable and prolonged than in temperate regions, sometimes occurring year round [23,24]. An improved understanding of RSV seasonality will be important in identifying the most cost-effective implementation of RSV preventive interventions (such as long-acting monoclonal antibodies or maternal immunization) and therapeutics. For many countries (particularly LMICs) there is limited information on RSV epidemic dynamics, and to address this gap a number of predictive models have been created to assess the impact of geography and climate on RSV seasonality. A recent analysis of global seasonal patterns of respiratory virus epidemics including RSV, influenza, parainfluenza virus and human metapneumovirus found an association between latitude and yearly epidemics, with seasonal patterns occurring in both the northern and southern hemispheres [25]. The study examined the relationship between temperature, relative humidity and virus activity, and found that lower temperature was associated with increased activity of RSV and influenza. However, the activity of the two viruses differed in their response to relative humidity and temperature. RSV incidence was
higher when relative humidity was higher and temperature lower than the seasonal average. By contrast, influenza activity was shown to increase at higher relative humidity and temperatures (although the role of climatic factors on influenza is still very much debated) [26].

A comprehensive study conducted using data from the United States and Mexico examined how climatic factors explained [27] latitudinal differences in the seasonality of RSV in temperate and subtropical latitudes. In northern regions, RSV epidemics demonstrated biennial and annual RSV seasons with the peak occurring in winter months; coastal regions tended to show an annual pattern, whereas central regions demonstrated a biennial one. In tropical regions there was persistent incidence during the year, leading to annual epidemics with shallow troughs. Modeling state-specific datasets of weekly RSV and bronchiolitis hospitalizations with climate data showed that in northern temperate localities, the primary driver of RSV timing was specific humidity, with peak incidence occurring during dry winter months. In southern tropical regions, the primary driver of RSV transmission was increasing levels of precipitation. Using projections of global warming trends, scenarios for changing RSV seasonality in the year 2100, showed that changes in precipitation patterns could lead to increased year-round incidence in northern regions.

These studies demonstrate the variability of RSV seasonality, the relevance of which is likely to vary depending on the immunization approach implemented. A better understanding of the global seasonality of RSV incidence is likely to inform the timing of treatment of infants with monoclonal antibodies, or scheduling of maternal immunization. These studies also highlight that these patterns may be subject to change in the context of global warming. What has not yet been established is the exact mechanism by which climatic factors affect RSV transmission, and how a reduction in overall RSV case burden might impact these patterns. On one hand, decades of influenza vaccination have not changed the seasonality of virus
activity, likely due to the suboptimal effectiveness of existing vaccines and uncertainty regarding benefits of immunization on reduction of viral transmission [28]. In contrast, childhood rotavirus vaccination has changed the periodicity of winter outbreaks in the US, with later than usual epidemics in the first few years of the vaccination program giving way to bi-annual epidemic cycles [29,30]. Such changes in epidemic dynamics are well-understood feature of vaccination programs that reduce susceptible hosts; the exact dynamic changes however heavily depend on whether high-transmitter groups are vaccinated, and the nature, strength, and duration of vaccine-induced immunity. Another factor which has now been shown to dramatically disrupt the pattern of RSV annual epidemics are the social distancing measures implemented to slow the transmission of SARS-CoV-2 in 2020/21. In a number of countries (such as Australia) these have led to the suppression of the usual autumn/winter peaks in infection, and to the appearance of unusual, delayed spikes of infection in the spring and summer [31]. This suggests that climate is unlikely to be the only driver of RSV epidemics, and that the shape of future RSV epidemics is likely to depend on an interplay between climate, population behaviours, and in the future the implementation of vaccination campaigns.

**RSV genotypes and phylogeography**

An added complication for our understanding of RSV transmission dynamics is the presence of 2 RSV subtypes, A and B. These two subtypes are based on antigenic and sequence-based variations predominately associated with the RSV G gene [32] (Figure 1A). Each subtype has multiple genotypes that are in co-circulation with one subtype dominating in any given year [33–35]. In recent years, several unique genetic modifications in RSV G have been identified, including a 72-nucleotide duplication (referred to as the ON genotype) in RSV-A [36] and another with a 60-nucleotide duplication (referred to as the BA genotype) in RSV-B [37] (Figure 1B). With time, new genotypes for each subtype arise and spread globally,
replacing previous strains; the predominant strains seen at present are ON1 (RSV-A) and BA9 (RSV-B) (Figure 1C). Many questions remain unanswered in our understanding of RSV genetic evolution, in part due to the paucity of sequence data when compared to other RNA viruses like influenza. It is important to note that although the G gene has been widely used to classify RSV-A and RSV-B genotypes, and study their evolution, most vaccine development has focused on the F gene, which shows less genetic variability than the G gene [8,38].

The ecology and phylodynamics of RSV remain unclear on both local and global levels. At a local level, the number of genetic lineages within a typical epidemic remains poorly understood, as does the genetic diversity of those lineages, and the rapidity with which lineages disseminate and persist within a community. It is not clear whether lineages are consistently reimported every year into regions with marked RSV seasonality, in a model reminiscent of influenza ecology. If RSV lineages are imported each season, the degree to which local evolution, adaptation to host populations, and competition between different lineages shape viral diversity remains to be defined. At a global level, the relationships and migrations between viral populations in different countries, regions and hemispheres are yet to be fully delineated. More research is needed to determine how frequently, and why, new global lineages emerge, and why some become dominant in successive seasons (such as ON1 within RSV-A). A lack of clarity on RSV spread locally and globally affects the ability to model the likely effects of vaccine introduction on evolutionary dynamics within and between communities. If as has been shown RSV subtype can affect the effectiveness of long-lasting monoclonal antibodies[39], then it will be important to understand how frequently genotypes with point mutations that can affect antibody binding arise, and how they increase in frequency to become the predominant strains globally (more details in Implications of genomic variability for vaccine and therapeutics development).
Relationship between RSV genotype and clinical phenotype

The association between RSV genotype and disease severity has been examined for several decades but remains elusive, with contradictory results emerging from disparate observational clinical studies. While the majority of studies [40–42] appear to indicate that RSV-A causes more severe disease (e.g. increased need for ventilation, higher risk of hospitalization, higher disease severity score) than RSV-B in infants, others have shown either that RSV B infection [43] leads to longer hospital stays or increased severity indices than RSV-A [44], or alternatively that there is no significant difference in disease severity between RSV-A and RSV-B infections [45,46]. Moreover, studies looking at differences in symptoms between children infected with specific RSV genotypes have also shown discrepant results. For example, a study in Italy [43] showed that RSV A NA1 infection in children led to greater hospitalization and a higher propensity for lower respiratory tract infection than ON1 infection did, while another study in Japan found that children infected with the ON1 genotype were admitted more rapidly to hospital, had increased wheezing, and a greater risk of chest X-ray abnormalities than those with NA1 infection did [47].

These inconsistent findings can be attributed to many factors [48,49] including study design and laboratory approaches such as testing methodology, sample size and enrollment criteria. Patient characteristics, including age of infection, previous immune history, family medical history (e.g. asthma) and environmental factors such as household smoking or air pollution can also be confounders. Moreover, there is no consensus within the scientific community in how to score disease severity with changing criteria over time and with difficulty in assessing symptoms in the infant population under study. In addition to these factors, diversity in RSV circulating strains over time, changing methods for defining genotypes, and a lack of consensus among the scientific community about how to name RSV strains has
caused some confusion in interpretation of results, and as outlined below are important priorities for standardization.

**Importance of standardization for leveraging information from RSV networks**

A number of international studies examining RSV epidemiology and genomics are ongoing [50–53], and information from these is most likely to be useful if collected and presented in a standard format. Participants at the workshop were unanimous in agreeing on the importance of standardization of a number of aspects of RSV research, and ways in which this would improve our understanding of RSV disease [6]. Standardization could be applied across all areas of RSV virologic and clinical research, from a better understanding of viral genetics to the determination of the host response. A standardized nomenclature for virus sequence upload would facilitate curation and evolutionary analyses on databases such as GenBank (https://www.ncbi.nlm.nih.gov/genbank/) or GISAID (www.gisaid.org). A consistent approach to genotyping would facilitate the study of genotype/phenotype relationships, and also simplify the study of the emergence, growth, replacement and extinction of different virus strains in a global context. Two recent publications have moved towards a more consistent approach to viral genotyping [54,55], but work remains to establish a system that is accepted and used by the RSV research community as a whole.

The use of standard protocols for clinical metadata and sample collection, and for clinical trial study design, will also be important. The recent establishment of international standards for antiserum to RSV-A [56] and RSV-B [57] suggest that it should also be possible to establish similar standards for establishing viral load and correlates of protection. A consistent approach to the study of RSV and its effects on human hosts will be key in predicting and tracking whether genomic variability is likely to impact on the effectiveness of vaccines and therapeutics.
Potential approaches to RSV immunization

Two recently published trials, one on maternal vaccination with an F-protein based vaccine [58] and the second on infant passive immunization with a long-lasting monoclonal antibody [59], show the potential of different approaches for the prevention of RSV disease in young infants. The rationale for maternal vaccination is to stimulate the production of IgG to the RSV target (most commonly the F protein), which can then be transferred through the placenta to infants and protect them against disease in the first few months of life, when they are most vulnerable to severe disease. A mathematical modeling study reported at the workshop assumed that maternal vaccination would extend the duration of maternal antibody protection against infection from 1 month to 4 months, and demonstrated that the elevated level of maternal antibody passed to the infants could reduce the severity of RSV disease and consequently the risk of infant hospitalization by a factor of 30%. Additionally, the model predicted that maternal vaccination could reduce the risk of RSV infection of the mother during the epidemic season by 50%, thus indirectly protecting the infant (Kaiyuan Sun, unpublished results). The first published Phase III trial for maternal immunization [58], results of which were also presented at the workshop, did not meet the primary endpoint of reducing medically significant RSV lower respiratory tract infection in infants, but as predicted by the modelling study, did show a significant reduction (by 44%) in hospitalization from RSV disease. A number of other maternal RSV vaccines are currently progressing through Phase I and II trials [60]. It is possible that newer maternal vaccine products, which utilize the more antigenic pre-fusion F protein [61], may lead to higher neutralizing titres in the mother and extended protection for the infant.

The second trial administered the long-lasting monoclonal antibody nirsevimab/MEDI8897 (which targets the F protein) or placebo as a single dose to infants who had been born preterm at the start of the RSV season, and found a 70.1% reduction in medically attended RSV
infection in the treated group. Further preclinical, Phase II and III trials are ongoing to identify whether the same benefits will also be seen in infants born at term [60]. In addition to passive maternal antibody transfer and direct protection via prophylaxis, a third possible approach is to reduce transmission from older infants or children, who in addition to being at risk of disease themselves are known to be a risk factor for the introduction of RSV into households [62,63]. The most common active immunization approaches under development for RSV-naïve infants and young children in this age-group are live-attenuated RSV or live chimeric paramyxovirus-based vector vaccine candidates for intranasal administration, all of which are currently in Phase I or I/II trials.

**Implications of genomic variability for vaccine and therapeutics development**

An effective RSV vaccine would likely impact the epidemic cycling and population dynamics of the virus. These effects could be predicted based on an understanding of the natural history of RSV infection and immunity, studies of which have been published from a LMIC setting[62,64,65], combined with experience from other vaccination programs such as measles and rotavirus, where detailed birth cohort studies contributed to our understanding of protection from reinfection and severe disease. Conversely, whether RSV phylogeographic dynamics, and the consequent distribution of mutations in different strains, are likely to impact on vaccine effectiveness, is likely to depend on the target product profile and age-specific coverage of the chosen vaccine.

Typically, resistance to vaccines is less likely to emerge than resistance to drugs. While drugs target a specific biological mechanism that can be abrogated by a point mutation, vaccine responses are generally polyclonal and host dependent [66]. It has been shown that the immune response to the F protein in natural infection and after vaccine challenge is
polyclonal, with antibodies being produced to all protein epitopes [67,68](Figure 1D). In addition to this polyclonal response, viral populations are expected to remain low in a vaccinated host where infection is prevented, limiting the potential for viral escape [69]. If a gene-based or protein-based RSV vaccine targeting the F glycoprotein is introduced, genomic variability is unlikely to affect the effectiveness of immunization, as there are few situations in which viral vaccines that present a combination of epitopes have been associated with escape mutants that have rendered them less effective [69]. As a warning note however, recent laboratory work (summarized in [70]) and clinical trials have demonstrated that mutations in the SARS-CoV-2 Spike protein, similar in structure and function to the RSV F protein, can impact on the effectiveness of antibody neutralization. Mutations such as the point substitution E484K mean that previous COVID-19 infection, and vaccines, are less effective in protecting against infection by novel strains such as B.1.351[71].

Therefore, if the F protein used for a vaccine is RSV-A or RSV-B specific, there is the possibility that increased effectiveness against one of the subtypes might favor spread of the other subtype, if the vaccine is widely enough used. In a community setting, it is not clear whether previous infection with one subtype protects against subsequent infection with the same subtype [72]. The fact that A and B epidemics tend to alternate on an annual or biannual basis suggests a level of subtype- or strain-specific immunity at a population level [73]. Most current vaccine proteins are based on the A2 F protein, a laboratory RSV-A strain [61]. Studies have shown that vaccinations targeting RSV-A or B F glycoprotein can lead to differences in elicited subtype neutralizing activity, supporting the theoretical possibility of vaccine subtype-specific host immunity [74]. If deployed to a majority of the population, a maternal or childhood vaccination which leads to preferential protection against a particular subtype could impact the long-term dynamics and cycling of RSV strains (echoing patterns of interaction and replacement in the rotavirus and pneumococcus vaccine programs) [75]. A better understanding of RSV subtype-specific and cross-immunity would help to build
models that could predict these effects. It will be important to better understand vaccine responses in different target populations, which will yield insight into the likely effectiveness and dynamical consequences of different vaccination approaches.

Compared to the uncertainties outlined above, if a long-lasting monoclonal antibody were to be widely introduced, the importance of genomic variability in determining immunization success becomes much clearer. When grown in cell culture in the presence of monoclonal antibodies, other viruses such as measles and polio eventually develop mutations that make them resistant to the antibody [76,77]; the same is true for RSV (summarized in [8]). Small changes of one or two residues can have a dramatic impact on binding affinity. This became evident when the monoclonal antibody suptavumab failed to meet its primary endpoint, due to changes in the binding site as a result of mutations in site V of the RSV F protein in circulating RSV B viruses. Despite presenting a protective trend against medically attended infections caused by RSV subgroup A, two mutations found in circulating RSV B viruses affected its overall performance. The reduced efficacy against RSV B was due to a two-amino acid change at positions 172 and 173 in the antigenic site V region of the F protein, the epitope of suptavumab, which reduced susceptibility to antibody neutralization in vitro [39] (Figure 1E). It is therefore important that clinical studies involving anti-RSV F monoclonal antibodies monitor for amino acid substitutions in antigenic binding regions of RSV isolates from subjects experiencing virologic failure, and to assess the impact of these changes on phenotypic susceptibility and viral fitness. In vitro studies have already found a number of mutations in RSV that affect the binding of nirsevimab (MEDI8897) without having a clear impact on viral fitness [78]. This clear risk of monoclonal antibody failure, due either to drift (as with suptavumab) or potentially to selection, if monoclonal antibody administration is sufficiently widespread, highlights the importance of global surveillance of RSV for genomic variability. It is possible that cocktails of monoclonals, as have been used for Ebola [79,80] might ameliorate the risk of point mutations leading to prophylaxis failure. A potentially
reassuring point is that the F protein in both RSV-A and B shows a limited range of variability, with amino acid changes that tend to toggle back and forth, rather than the drift seen with other RNA viruses [8]. This suggests that there may be a natural limit to the variation that can be tolerated by the protein and still maintain function. However, the broader the polyclonal response elicited by a vaccine, the more resistant it should be to natural sequence variation, although the trade-off may well be reduced protection, as seen in the results from the maternal vaccine trial [58].

**Strengths and weaknesses of different immunization approaches in the context of RSV genomic variability**

In the absence of an effective vaccine, such as that for Hepatitis B, which can be given soon after birth and induce long-lasting immunity [81] there are a number of forms that RSV immunization might take, each with its advantages and disadvantages (Figure 2). For maternal immunization, genomic variability is less likely to threaten the effectiveness of the vaccine. However, this approach will only protect infants for a short period of time and is unlikely to interrupt transmission of RSV across communities, or lead to protection for older children or the elderly.

Infant passive immunization would consist of long-lasting monoclonal antibodies given shortly after birth, as in the recently published trial described above [59]. However, the relatively high cost of monoclonal antibody therapies, in both high- and low-income settings, is likely to mean that they are only given at points in the year where it is most likely to protect infants (i.e. in the run up to the RSV season), and thus a clear understanding of seasonality in different settings, from large, well run surveillance projects will be very important. Global monitoring will also be required to identify and track the spread of variants with mutations that could affect the efficacy of these monoclonal antibodies [39].
Mucosally-delivered live vaccines given in childhood offer the possibility of interrupting transmission dynamics and of generating a degree of herd immunity which would protect the most vulnerable: newborn infants and the elderly. However, a successful live vaccine could have unforeseen effects on the transmission dynamics of RSV more generally. Live vaccines against measles and rotavirus have shown that viral transmission dynamics can be dramatically altered by an effective vaccine. In the case of measles in Europe, a shift in the timing of seasonal epidemics from that seen in the pre-vaccination era was observed after the measles vaccine was introduced [82–85]. A similar result occurred in the US after the introduction of a rotavirus vaccine in the US, where there were changes in the epidemic cycling giving way to bi-annual epidemics and changes in the demographic features of those infected [29,30]. Such changes in epidemic dynamics seems to be a feature of vaccination programs that reduce susceptible hosts; the exact changes heavily depend on whether high-transmitter groups are vaccinated, and the nature, strength, and duration of vaccine-induced immunity [86]. With RSV, herd immunity may be difficult to achieve as natural immunity appears to be short-lived [87] and imperfect [88]; individuals are infected (and can transmit) throughout life [89]. If herd immunity could be achieved, genomic variability would be unlikely to threaten the effectiveness of any vaccine, as unlike the genomic reassortment seen in influenza, RSV does not demonstrate regular recombination events that lead to significant changes in the circulating virus that might render live vaccines eliciting a wide-ranging polyclonal response to a number of viral proteins ineffective [69]. In this scenario of an effective live vaccine, the key to predicting how viral transmission dynamics might be affected by vaccine implementation will be a comprehensive RSV evidence base facilitated by a standardized approach to viral nomenclature, genotyping and clinical case definitions.

**Conclusion**
Writing in the middle of a global pandemic caused by another RNA virus, SARS-CoV-2, it is instructive to set out what the RSV field has offered to research into COVID-19, and conversely what the RSV field can learn from the huge scientific effort and resources invested in better understanding this new virus. In terms of protein structure, work conducted in RSV has contributed enormously to progress in COVID-19 vaccine development: work on structure-based design of a prefusion-stabilized SARS-CoV-2 Spike protein for vaccine development [90] was based on the extensive work studying the prefusion form of the RSV F protein [61,91]. Many years of research have demonstrated the challenge of inducing protection to a virus like RSV[5], where natural infection only provides short-lived immunity, and shows that the development of a long-lasting, effective COVID-19 vaccine is also likely to prove challenging [92]; conversely the vulnerability of Spike based vaccines to mutations that reduce efficacy in reducing infection (although probably not severe disease), is a warning to the RSV field. It remains unclear whether SARS-CoV-2 will become an endemic virus such as RSV, and whether if, as seems possible from work in other endemic coronaviruses [92], natural or vaccine immunity will prove imperfect and result in seasonal epidemics [93]. If it does cause seasonal epidemics, the modelling work conducted for influenza and RSV highlighted earlier in this report will be highly relevant for our understanding of COVID-19 disease dynamics.

Equally, the RSV field has much to learn from the already considerable amount of research conducted into SARS-CoV-2. In little over 1 year almost 800,000 SARS-CoV-2 whole genome sequences from over 100 countries have been generated and made publicly available through GenBank (https://www.ncbi.nlm.nih.gov/genbank/) and GISAID (www.gisaid.org). A simple system for viral nomenclature was introduced in the first few weeks of the pandemic and has already been widely adopted. A standardized system to assign viral sequences to global lineages has been published and widely adopted [94,95]; this system allows for a better
understanding of the dynamics of viral importation [96] and spread [97,98] and investigation of nosocomial and other outbreaks [99]. The rapid development of effective mRNA, adenovirus vector and Spike protein vaccines for SARS-CoV-2 [70] will hopefully serve to inspire researchers to adapt these new technologies to improve upon existing RSV vaccine options.

For both COVID-19 and RSV, large-scale surveillance and collection of genetic data over long time-series within discrete locations are needed in order to improve our understanding of disease, and to allow for the modelling of whether vaccine implementation is likely to prove effective. Larger studies can characterize population dynamics and allow the estimation of competitive immunological dynamics between recently emerged genotypes/lineages and previously dominant ones. Birth cohort studies will allow an understanding the development of immunity against infection and prevention of severe disease [100], and of the degree of cross-immunity between strains. A combination of viral sequence data and data on the host immune response will allow for a better understanding of the strength and breadth of immune responses to different viral genotypes/lineages.

In summary, we are in a period of incredible scientific progress in genomic surveillance, computational modeling and countermeasure development that have already lead to development of effective vaccination options for SARS-CoV-2 and may soon lead to the implementation of widespread RSV immunization. Lessons learnt about the evolution and global transmission dynamics of these RNA viruses will need to be factored into the implementation of these interventions to ensure their long-term efficacy.

Author Contributions: S.K., T.C.W., C.V., H.C., and D.J.S. contributed equally to conceptualization, original draft preparation, review and editing. J.C. contributed
substantially to visualization. All authors have read and agreed to the published version of
the manuscript.

**Funding:** Dr. Campbell reports grants from EU IMI, personal fees from WHO, grants and
personal fees from Sanofi, grants and personal fees from Bill and Melinda Gates Foundation,
and grants from UK NIHR, during the drafting of this manuscript. Dr. Williams is the
recipient of a Wellcome Trust Award [204802/Z/16/Z].

**Disclaimer:** The findings and conclusions in this report are those of the authors and do not
necessarily represent the official position of the U.S. National Institutes of Health or
Department of Health and Human Services.

**Acknowledgments:** The authors would like to thank all the presenters and participants at the
workshop (listed in Supplementary File 2). We would also like to thank Barney Graham for
extensive conversations, Justin Bahl and Vicente Mas for contributing towards Figure 1, and
Sabrina Russo for the design of Figure 2. We thank Ms. Susan Spring for help in preparation
of the meeting report.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


Causes of severe pneumonia requiring hospital admission in children without HIV
infection from Africa and Asia: the PERCH multi-country case-control study. Lancet


[87] Agoti CN, Mwihuri AG, Sande CJ, Onyango CO, Medley GF, Cane PA, et al.


Figure Captions

Figure 1. RSV genotypes and structure of the F protein. (A) Structure of the RSV genome. RSV has a negative stranded RNA genome which is approximately 15 kb long with 10 gene transcripts encoding 11 proteins. Two of these are the major surface glycoproteins, the attachment or G glycoprotein (green) and the fusion (F, orange) glycoprotein. (B) Differences within RSV-A and RSV-B. Genotypes are conventionally classified on the basis of genetic variation in the 2nd hypervariable region of G gene. In recent years, several unique genetic modifications in RSV G hypervariable domain have been identified. These include a 72-nucleotide duplication, referred to as the ON1 genotype, which in RSV-A which distinguish this group from the ancestral NA1 genotype. In RSV-B, a 60-nucleotide duplication and a short upstream deletion distinguish what is referred to as the BA clade from the ancestral GB1 genotype. (C) Phylogenetic tree of representative circulating RSV-A and B genotypes. The conventional classifications are shown in the legend. The pattern is of genotype dominance for several years followed by replaced by a different genotype. The discordance between the phylogenetic trees (constructed using all coding sequences for each sample) and the commonly used genotypes highlights the need for a standardized approach to RSV genotype classification. Details on methods used to construct tree in Supplementary File 1. (D) Antigenic structure of the RSV F protein. Surface representation of the 3D structures of hRSV F trimer folded in its prefusion [61] conformation, showing six main antigenic sites as defined by Gilman et al. [67]. (E) Amino acid differences between circulating RSV-B strains and the reference RSV-A2 strain at antigenic Site V [8]. The antigenic site is outlined in orange, and point substitutions compared to the reference in light (single variant) or dark blue (two variants). These variants include the L172Q and S173L substitutions that led to failure of binding of the monoclonal antibody suptavumab to circulating RSV-B strains in a recent Phase III trial [39]. Panels A, B, and C courtesy of Justin Bahl and Jiani Chen; Panels D and E courtesy of Vicente Mas.
Figure 2. Benefits and disadvantages of different RSV immunization approaches in the context of genomic variability. Figure courtesy of Sabrina Russo.