Mutations associated with severity of the pandemic influenza A(H1N1)pdm09 in humans: a systematic review and meta-analysis of epidemiological evidence

E. A. Goka · P. J. Vallely · K. J. Mutton · P. E. Klapper

Received: 25 April 2014 / Accepted: 12 July 2014 © Springer-Verlag Wien 2014

Abstract Mutations in the haemagglutinin (HA), non-structural protein 1 (NS1) and polymerase basic protein 2 (PB2) of influenza viruses have been associated with virulence. This study investigated the association between mutations in these genes in influenza A(H1N1)pdm09 virus and the risk of severe or fatal disease. Searches were conducted on the MEDLINE, EMBASE and Web of Science electronic databases and the reference lists of published studies. The PRISMA and STROBE guidelines were followed in assessing the quality of studies and writing-up. Eighteen (18) studies, from all continents, were included in the systematic review (recruiting patients 0 – 77 years old). The mutation D222G was associated with a significant increase in severe disease (pooled RD: 11 %, 95 % CI: 3.0 % - 18.0 %, p = 0.004) and the risk of fatality (RD: 23 %, 95 % CI: 14.0 %–31.0 %, p = < 0.0001). No association was observed between the mutations HA-D222N, D222E, PB2-E627K and NS1-T123V and severe/fatal disease. The results suggest that no virus quasispecies bearing virulence-conferring mutations in the HA, PB2 and NS1 predominated. However issues of sampling bias, and bias due to uncontrolled confounders such as comorbidities, and viral and bacterial coinfection, should be born in mind. Influenza A viruses should continue to be monitored for the occurrence of virulence-conferring mutations in HA, PB2 and NS1. There are suggestions that respiratory virus coinfections also affect virus virulence. Studies investigating the role of genetic mutations on disease outcome should make efforts to also investigate the role of respiratory virus coinfections.

Introduction

The virulence genes of influenza A viruses

The history of influenza viruses has shown that the viruses that caused pandemics had changes in the surface and internal proteins [1]. The 1957 pandemic H2N2 influenza virus had haemagglutinin (HA), neuraminidase (NA) and polymerase basic protein 1 (PB1) genes from an avian virus and the remainder of the genes derived from a previously circulating human virus. Similarly, the 1968 pandemic H3N2 virus had avian HA and PB1 segments in a background of human viral genes [1–7]. It has been suggested that changes in the HA, PB and NS1 genes affect disease severity.

Changes in the antigenic sites, or in the vicinity of the antigenic sites, of the HA gene result in antigenic drift and compromise vaccine and immune effectiveness, with some novel strains causing pandemics with high virulence [8–13]. Also, changes in the HA receptor-binding sites or in molecules in their vicinity affect viral virulence [14]. Similarly, glycosylation in the vicinity of the cleavage site promotes folding of the HA and can lead to masking or unmasking of proteolytic cleavage or antigenic sites and affect viral pathogenicity [8]. The polymerase basic protein
2 (PB2) gene has been shown to affect host specificity in influenza H5N1, H7N7 and in the 1918 H1N1 virus [15–17]. The NS1 protein has been implicated in the pathogenicity of the highly pathogenic H5N1 avian influenza virus and in the 1918 H1N1 virus by aiding in the circumvention of the innate immune response such as the production of proinflammatory cytokines [18, 19]. Thus mutations in the PB2, HA, and NS1 genes affect viral virulence by determining host specificity, enhancing the ability of hybrid viruses to efficiently replicate in humans, enhancing the efficiency of viral attachment, entry and release, and facilitating antagonism of the interferon pathways and evasion of the immune response [15–21].

Functional sites in HA related to virulence

HA is an envelope glycoprotein of 550 amino acids [8]: 329 in HA1, and 221 in HA2 [8]. The HA gene contains three major functional sites: the receptor-binding site (RBS), the antigenic site and the cleavage site. Amino acids 187 and 222 of HA1 determine the receptor-binding specificity of the HA: D187/D222 for α(2,6) receptors in humans, D187/G222 for α(2,6) and α(2,3) receptors in birds [12, 21–28]. In the 1918 virus, a single mutation (D222G) reduced the binding affinity for α(2,6) receptors [26, 27] and the infectivity of the virus [21], while a double variant, D187/E/D222G, rendered the HA non-binding to α(2,6) receptors and the virus non-infectious [26, 27, 29].

Out of the 329 HA1 residues, 131 are antigenic sites, 181 are non-antigenic sites, and the status of the others is unknown. The antigenic sites are further classified into five antigenic epitopes (Supplementary Table S1), designated Sa, Sb, Ca1, Ca2 and Cb in influenza H1N1 viruses and A, B, C, D and E in influenza H3N2 viruses [8, 30–33]. In general, most of the antigenic changes in influenza A(H1N1) and A(H3N2) occur through at least three substitutions in positively selected sites [11]. However, a deletion of K at position 130 (K134 in H3 numbering) was responsible for the antigenic difference between the A/Bayern/7/95 and A/Beijing/262/95 influenza A(H1N1) subtypes [9]. The positively selected antigenic sites on H1N1-HA1 have previously been summarized by Cox et al. [34], Brownlee and Fodor [33], Igarashi et al. [13], Liao et al. [11], Huang et al. [35], Stray and Pitman [36], and Shen et al. [12]. We have summarized the findings from these studies in supplementary Table S1. Further, supplementary Table S2 details the amino acid mutations in these HA antigenic sites that have occurred in the main influenza A(H1N1) virus strains that have circulated since 1918.

Cleavage of the HA gene is important for viral infectivity as it is a prerequisite for the fusion of the viral and cell membranes [37]. N-glycosylation is achieved by addition of glycans to asparagine (N) residues [38]. Influenza A (H1N1) viruses possess five glycosylation sites: N71, N104, N142 (variant – 144), N177 (variant – 172 and 179), and N286 (H1 numbering) [39, 40]. The A/SC/1918 and CA/07/2009 isolates had only one glycosylation site at N104 [13, 39]; however, after 1918, H1N1 viruses gained the N286 site, and N142 and N177 before 1957. The other site, N71, appeared in 1987 [39]. The seasonal influenza A(H1N1) viruses that circulated prior to the 2009 pandemic virus had four glycosylation sites, having lost the 286 site [39].

Functional sites in PB2 related to virulence

The PB1, PB2, and PA proteins form the polymerase, which is responsible for replication and transcription of the viral RNA in the nucleus of infected cells using a capsid-snatching mechanism [41, 42]. The location of the cap-binding site on PB2 is controversial, Honda et al. [43], Palese and Shaw [44], and Li et al. [42] have suggested the involvement of N-site residues 242–252 and C-site residues 533–577, while Fechter et al. [45] have suggested a more central site involving residues F363 and F404, and Poole et al. [44, 46] have suggested that the C-terminal residues 1–269 and 580–683 are involved. One of the most commonly identified virulence markers is mutation (PB2-E627K); the glutamic acid (E) is generally found in avian influenza viruses, while human viruses have a lysine (K). Therefore, this mutation determines host specificity [47]. Other important PB2 mutations include D701N, S714R, E627K, and D567G; the glutamic acid (E) is generally found in avian influenza viruses, while human viruses have a lysine (K). Therefore, this mutation determines host specificity [47]. Other important PB2 mutations include D701N, S714R, E627K, and D567G [39, 53–55]. The pandemic influenza A(H1N1)pdm09 virus has an avian PB2 gene [5] and hence needs the 627 K mutation for it to efficiently replicate in humans. However, some researchers [48, 52, 56–58] have indicated that E627K has little effect on the transmissibility of Flu Apdm09, while others found that it does have an effect [59], and others have suggested that this might be compensated by Q591R [52, 60].

Functional sites in NS1 related to virulence

Depending on the virus strain, NS1 consists of 124–237 amino acids, is expressed exclusively in infected cells, and contains an N-terminal RNA-binding domain (residues 1–73) and a C-terminal effector domain (residues 73–237) [61]. The effector domain contains several other domains, including the cleavage and CPSF-binding domain (residues 175–210), the poly(A)-binding protein (PABP) domain (residues 218–225), nuclear localization signal 1 (residues 34–38), nuclear localization signal 2 (residues 211–216), the nuclear export signal (residues 132–141), and the
Mutations associated with A(H1N1)pdm09 influenza severity

e1F4G1 domain (residues 81-113) [62]. Residue 186, 103 and 106 prevent transport of cellular mRNA to the cytoplasm by interaction with poly(A)-binding protein II (PABII) [63–65], whereas amino acids 215 to 237 have been identified as the binding site for PABII [63].

Wang et al. [66], Cheng et al. [67], and Yin et al. [68] indicated that the dsRNA-binding residues in the N-terminus include T5, P31, D34, R35, R38, K41, G45, R46 and T49. Further, a five-residue short peptide sequence (123-IMDKN-127) counteracts the PKR-mediated antiviral response [69, 70]. The effector domain residues T89/M93, P164/P167 and L141/E142 bind to p85b and induce the phosphatidylinositol 3-kinase (PI3K)/Akt signalling pathway [71, 72], whereas R38A/K41A and E96A/E97A are responsible for TRIM25 binding [73–76]. Studies by Donelan et al. [77] and Talon et al. [78] indicated that the NS1 R38A/K41A mutations abolished dsRNA binding and IFN antagonism, while Gack et al. [73], Mibayashi et al. [74], Bornholdt and Prasad [75], and Chien et al. [76] indicated that mutations R38A/K41A and E96A/E97A NS1, in contrast to wild-type NS1, led to abolishment of TRIM25 binding and dysfunctional RIG-I. The NS1 mutations implicated in the pandemic influenza A(H1N1)pdm virus include T123V [79, 80].

Rationale for conducting the review

An understanding of the positively selected important mutations in the influenza A virus HA, PB2 and NS1 genes that are associated with severe outcome could help future research for vaccine and drug development. It could also help with patient management and public-health surveillance and preparedness. Due to the ever-changing nature (antigenic drift and shift) of influenza viruses, the World Health Organization (WHO) Global Influenza Surveillance and Response System (GISRS) coordinates the global risk assessment of influenza viruses yearly. One hundred twenty-two (122) national influenza centres, designated by national authorities in 92 countries, send around 150,000 to 200,000 respiratory specimens to the five global collaborating centres (the United Kingdom and Australia National Institute for Medical Research, US Centers for Disease Control and Prevention, Chinese Centres for Disease Control and Prevention, and National Institute of Infectious Diseases in Japan) for antigenic characterization to identify novel variants twice a year [81, 82]. An additional approach is field vaccine trials, which are undertaken by multiple sites in North America, Europe and Australia. The case–control and cohort studies identify and characterise influenza viruses, document observed mutations and observed responsiveness of the circulating viruses to that season’s vaccine [83, 84]. Lastly, independent studies are conducted by hospitals, university institutions, government ministries of health and other bodies, of which some are published and others are not. Some of the virus sequences are deposited in GenBank, the Influenza Research Database, and the Research Collaboration for Structural Bioinformatics Protein Data Bank. A major problem with the above methods is that most of the studies, while using laboratory-confirmed influenza viruses, do not include virulence (i.e., severe disease or death) as an endpoint or give a clear link between specific mutations and disease outcome [85, 86]. Therefore, most of the evidence on the role of mutations on severity are based on animal and cell-line experiments [51, 53, 59, 87–94], and evidence from human infections is scanty. The December 2009 WHO report [95] and the November 2010 European Centres for Disease Control and Prevention report [96] provided a summary of mutations in the pandemic virus and their effect on disease severity, but both dwelled only on the HA D222G mutation, and no study has systematically reviewed the evidence for an association between mutations in the HA, PB2 and NS1 genes in the pandemic influenza A(H1N1)pdm09 virus and patient outcome; hence the importance of this review.

Aims and objectives of the review

This study aims to review available epidemiological evidence on mutations in the PB2, NS1 and HA genes of influenza A (H1N1)pdm09 virus and their association with disease outcome. Also, it has been suggested by some studies that coinfections with respiratory virus affect disease severity [97–102]. Currently, most of the studies on genetic mutations and virus virulence have not investigated the role of coinfections in those circumstances. Our other study [103] investigated the association between coinfection with influenza A viruses and other respiratory viruses that have circulated between 2007 and 2012, including the pandemic influenza A virus, and disease outcome. It is therefore imperative to investigate what other factors, apart from coinfections, could contribute to severe disease. This review seeks to determine whether mutations previously described to increase the severity of disease caused by influenza A viruses occurred in the pandemic influenza A(H1N1)pdm09 virus that circulated between 2009 and 2012.

The objectives of the review therefore were:

1. To investigate the occurrence of HA mutations D222E, D222G and D222N, PB2 mutations E627K and D701N, and NS1 mutations T123V, R38A, K41A, E96A and E97A in mild, severe and fatal cases of influenza caused by strain A(H1N1)pdm09.

2. To determine which mutations should be regarded as posing a major public-health concern.
3. To determine whether any other mutations in the HA, PB2 and NS1 genes are associated with severe disease outcome.

Methodology

Protocol for the review

The authors of this manuscript developed the protocol of this review. The protocol was not published or deposited to any online server.

Search strategy

The MEDLINE, EMBASE and Web of Science databases were searched for primary epidemiological studies on the role of the PB2, NS1 and HA genes of influenza A virus in determining disease severity. Specifically, the search aimed at identifying literature on the PB2 gene and its role in promoting virus replication at high temperature, NS1 and inhibition of interferon production, and the effect of HA mutations within glycosylation sites, antigenic sites and receptor-binding sites or in their vicinity and the observed outcomes in patients (mild, severe or fatal). Websites of health organisations, e.g., WHO, UK Health Protection Agency, the US Centers for Disease Control, and the World Influenza Network Centre, were visited to check influenza bibliography references listed therein or any published reports on influenza. Also, reference lists of good-quality studies, identified through the electronic sources, were examined to look for studies addressing the question under review.

For the electronic databases, the search technique involved combining a number of subject headings, keywords, and scoring for text words including: orthomyxovirus, influenza virus, influenza A virus type H1N1, 2009 pandemic influenza virus, influenza A(H1N1)pdm09, influenza virus haemagglutinin, HA gene, PB2 gene, non-structural protein 1, NS1, evolution, molecular evolution, mutation, genetic mutation, virus mutation, genetic evolution, antigenicity, prognosis, virulence, virulence, virus virulence, severity, severe disease, mild disease, pathogenicity, hospitalization, admission, ICU, death, fatality, and mortality. A copy of the electronic search conducted for this review on EMBASE is provided in Supplementary Table S4.

Study assessment tool and study selection criteria

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guideline [104] and the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) [105] tools were used as checklists for writing this review and for critical appraisal of identified epidemiological studies. The principal investigator of this review carried out the online and manual searches. Titles and abstracts were first reviewed for relevance, and those that were clearly not relevant (e.g., “Mutations on HA, PB2 and NS1 genes of pandemic Flu A(H1N1)pdm09 induces virulence in ferrets”, “Mouse adapted influenza A(H1N1) PB2, E627K mutation of H9N2 avian influenza virulence”) were removed. Papers that met the minimum inclusion criteria were downloaded and read in full. A range of epidemiological studies, i.e., cross-sectional, case–control, and cohort studies, with endpoints of laboratory-confirmed influenza virus infection and prognosis of the patients, i.e., mild, severe or fatal (according to the November 2009 WHO criteria for clinical classification of the disease due to the pandemic virus [106]) were included. Due to inadequacy of genetic studies of large sample sizes, studies with a sample size of \( \geq 5 \) were included. Studies that did not investigate the role of PB2, NS1 or HA genes in the virulence of influenza virus A viruses or did not report the outcome, which was being considered in this study, were not written in English, or were poorly designed and conducted (e.g., the analysis did not give the exact mutations that were identified or did not link the observed mutations to prognosis and disease outcome) were excluded.

Assessment of bias in the studies

Since the included studies are not randomized control trials, they would be prone to bias due to uncontrolled confounding factors such as the study population and period of virus identification, which could affect the degree to which the study samples were representative of the general population of viruses that were circulating, mainly due to differential surveillance (systematic differences in the way the virus was identified from various study groups, i.e., random sampling of severe or mild cases) or what laboratory protocols were used (PCR or other methods), and the possibility of contamination of samples. Bias could also be introduced by different data sources and methods of measurement of outcomes (mutations and clinical disease outcomes), e.g., classifying diseases as mild or severe (clinicians use different guidelines; what was classified as severe by one clinician might be classified as mild by another). Bias may also be due to coinfection with other respiratory viruses and bacteria, and some of the patients may have been infected with more than one respiratory virus, and this also might have influenced the outcome. Bias may occur due to
comorbidities, immune status, or other confounders such as the age of the patients, vaccination status, immune status, and the kind of clinical treatment they were given. All of the studies included in this review were assessed using the PRISMA guideline, and a summary of the scores for the 18 included studies is provided in Supplementary Table S3.

Data extraction and statistical analysis

A data extraction form was designed by the principal investigator of this study and confirmed by the coauthors. Data extracted from the studies included the study setting, i.e., author and year of publication, where and when it was conducted, study design, i.e., the viruses and genes that were analysed, the observed mutations, and the prognosis for the patients (mild, moderate, severe, disease or death(s) associated with specific mutations). Association between mutation type and disease outcome was assessed using risk differences. The differences between these statistics were assessed using the difference of two proportions (at significance level of $\alpha = 0.05$), and the differences and results were summarized using forest plots.

**Results**

Characteristics of the included studies

Two thousand three hundred ninety-three (2,393) search results were obtained, of which 307 were from MEDLINE, 703 were from EMBASE and 1,383 were from Web of Science. A manual search for other published papers yielded 25 papers. After removing the duplicates and irrelevant papers, 98 papers were reviewed. Eighty (80) more papers were found ineligible to be included and were excluded, and 18 papers were included in the systematic review (Fig. 1). The included studies were conducted in 2009 and 2010 and published between 2010 and 2013.

Details of the studies included in this review are summarized in Table 1. Identified studies were from all over the world; One from North America, three from South America, six from Europe, one from the Middle East, five from Asia, and two from Africa. Out of the 18 studies, four (Potdar et al. [107], Tse et al. [108], Baldanti et al. [109], and the Promed Email [110]) were conducted during the first wave (April–September 2009), nine (Puzelli et al. [111], Kilander et al. [112], Mak et al. [113], Miller et al. [114], Graham et al. [115], Farooqui et al. [116], Chen...
Table 1  Characteristics of the studies included in the analysis of the effect of HA, NS1 and PB2 mutations on virulence of pandemic influenza A(H1N1)pdm09 viruses

<table>
<thead>
<tr>
<th>Author &amp; study design</th>
<th>Period</th>
<th>Sample size</th>
<th>Patients’ age (yrs)</th>
<th>Mutations Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Puzelli et al. (2010), cross-sectional survey - Italy</td>
<td>May/09 – Feb/10</td>
<td>169</td>
<td>1 – 77</td>
<td>HA- D222G &amp; D222E [111]</td>
</tr>
<tr>
<td>2. Kilander et al. (2010), cross-sectional study - Norway</td>
<td>May/09 – Jan/10</td>
<td>266</td>
<td>All ages</td>
<td>HA- D222G, D222E, &amp; D222N [112]</td>
</tr>
<tr>
<td>3. Mak et al. (2010), cross-sectional survey – Hong Kong</td>
<td>May/09 – Jan/10</td>
<td>458</td>
<td>All ages</td>
<td>HA-D222G, D222E &amp; D222N [113]</td>
</tr>
<tr>
<td>8. Tse et al. (2011), cohort study of patients admitted with influenza – China</td>
<td>Apr – May/09</td>
<td>20</td>
<td>18 – 85</td>
<td>HA-D222G/N [108]</td>
</tr>
<tr>
<td>10. Wedde et al. (2013), cross-sectional survey of fatal, hospitalized and community pdm09 positive cases – Germany</td>
<td>Oct – Apr/09</td>
<td>357</td>
<td>0 – 73</td>
<td>HA-D222G/E/N [121]</td>
</tr>
<tr>
<td>11. Baldanti et al. (2010), cross-sectional study of hospitalized and outpatients with ILI – Pavia &amp; Milan, Italy</td>
<td>Apr – Nov/09</td>
<td>131</td>
<td>All ages</td>
<td>HA-D222G/E/N [109]</td>
</tr>
<tr>
<td>12. Moussi et al. (2013), outpatient and hospitalized individuals with pdm09 – Tunisia</td>
<td>09 – 10 and 10 – 11 specific months not given</td>
<td>50</td>
<td>15 – 47</td>
<td>HA-D222G/E, K374E [122]</td>
</tr>
<tr>
<td>13. Ferreira et al. (2011), cross-sectional study of fatal and nonfatal patients with ILI – Brazil</td>
<td>Apr/09 – Jun/10</td>
<td>90</td>
<td>All ages</td>
<td>HA-D222G/N/S, S203T, Y230H, E235K, M257V, Q293H, &amp; E391K [123]</td>
</tr>
<tr>
<td>15. Venter et al. (2012) cross-sectional study of outpatients and hospitalized individuals - South-Africa</td>
<td>Feb/09 – Dec/10</td>
<td>30</td>
<td>All ages</td>
<td>HA-D222G, S185T, Q293H, PB2-E627K, [118]</td>
</tr>
<tr>
<td>16. Promed Email (2009) outbreak investigation &amp; hospital based surveillance - Netherlands</td>
<td>Jul – Aug/09</td>
<td>33</td>
<td>All ages</td>
<td>PB2-E627K [110]</td>
</tr>
</tbody>
</table>
Mutations associated with A(H1N1)pdm09 influenza severity

et al. [117], Venter et al. [118], and Akcay Ciblak et al. [119] covered the first and second wave, one (Vazquez-Perez et al. [120]) covered only the second wave, two (Wedde et al. [121] and Moussi et al. [122]) covered the first and third waves, and two (Ferriera et al. [123] and Barrero et al. [124]) covered all three waves of the 2009 influenza A(H1N1) pandemic (Table 1). The number of viruses sequenced ranged from 13 to 357. In 11 studies, only the HA gene was sequenced; in three, the HA, PB2 and NS1 genes were sequenced; in two, the HA and PB2/NS1 genes were sequenced; and in one, only the PB2 gene was sequenced (Supplementary Table S3).

Possible sources of bias in the included studies

The PRISMA guideline was used to assess bias in the included studies, and the results are presented in Supplementary Table S3. The first issue concerns the possibility of selection bias. None of the included studies reported sequences of all of the Flu Apdm09 viruses identified during the study periods. This is understandable, because such an exercise would be very expensive. In those cases, samples were selected from a pool of Flu Apdm09 viruses, with each study adopting its own criteria to achieve representativeness. Severe and fatal cases occurred mainly in people aged >5-30 years and in very few children ≤5 years old. This is not surprising, as epidemiological studies have documented that the pandemic virus peaked in 5- to 19-year-olds [125]. Some studies did not control for comorbidities and other factors known to affect disease severity. Only one study each investigated co-infections with bacteria and other respiratory viruses. Therefore, these shortfalls should be born in mind when interpreting the results of this review.

Association between mutations in influenza A(H1N1)pdm09 virus and severe disease

HA-D222G mutation and severity

Overall, the evidence suggests that the mutation HA-D222G was associated with severe disease and mortality. Most of the risk differences regarding disease severity and death as well as the pooled risk differences (RD) were statistically significant (pooled RD: 11 %, 95 % CI: 3.0 %–18.0 %, p = 0.004 for severity and RD: 23 %, 95 % CI: 14.0 %–31.0 %, p = < 0.0001 for fatality) (Fig. 2). It should, however, be noted that the I² statistic for severity indicates that the heterogeneity was statistically significant (I²: 86.5, p = < 0.0001), whereas that for mortality was not. Observational studies are subject to bias due to uncontrolled confounders. Of the eight studies included in the severity analysis, three [117, 120, 121] controlled for comorbidities such as pregnancy, hypertension, congenital heart disease, myocarditis, asthma, diabetes and obesity, whereas the other five studies [111–114, 122] did not (Supplementary Table S3). Further, only Vazquez-Peres investigated and reported bacterial coinfection (Supplementary Table S3), and together, these factors might have caused the slight differences. However, despite this, we think that the evidence supports the hypothesis that D222G increased the risk of severity and mortality.

Mutation HA-D222E and D222N and disease outcome

Figures 3 and 4 present the results of the analysis of the association between D222E and D222N and disease severity. For the former, there was no evidence for or against a role of this mutation in severe disease (pooled RD: -2.0 %, 95 % CI: -5.0 % - 1.0 %, p = 0.2), and for mortality (RD: 1.0 %, 95 % CI: -3.0 % - 6 %, p = 0.54) as none of the associations were statistically significant. More research could, in due course, shed more light on the role of D222E.

As for the mutation D222N, all of the risk differences (for individual studies and the pooled RDs) suggested a positive association between this mutation and more-severe disease and death (pooled RD: 2.0 %, 95 % CI: 1.0 %– 5.0 %, p = 0.24 for severity analysis and RD: 5.0 %, 95 % CI: 1.0 % - 9.0 %, p = 0.09 for the fatality analysis). However, apart from Tse et al. [108], none of the RDs for the individual studies were statistically significant (Fig. 4). This could also be due to bias by uncontrolled confounders, as the I² values in the mortality analysis were statistically significant (I²: 73.5, p = 0.001).

Association between the PB2-E627K mutation and severe disease

Five (5) of the included studies reported on the impact of the E627K mutation on severe or fatal disease (Fig. 5). Four of the five studies did not identify E627K either in the mild or severe cases. Similarly, no association was noted in a suspected outbreak of the E627K mutant in the Netherlands, as reported in a 2009 Promed-Email. An investigation of a suspected E627K mutant was conducted among individuals who had camped at the West Frisian Islands in the Netherlands, following isolation of this virus in a diabetic index case. The mutation was later also identified in two of the 10 contact patients from across the country who had also camped at the same site and time with the index case, and they all had mild disease. The email further reports failure to identify the mutation in a further 22 samples of patients hospitalized with pandemic influenza A(H1N1)pdm09 between July and August, 2009.
The absence of the PB2-E627K mutation in patients supports the proposition that the pandemic Flu A pdm09 virus did not require the E627K mutation to cause serious disease [48, 52, 56–58]. However, the fact that these patients had severe disease suggests that other factors might have caused the serious illness. In two of the included studies, HA-D222G was identified in 2/8 (25 %) and in 2/6 (33.3 %) of patients who had died [107, 119]. However in the former, the PB2 genes in the two dead patients were not sequenced. Either the D222G or other factors could explain the outcomes. None of the studies included in this analysis controlled for comorbidities or bacterial and viral coinfections, and this should be borne in mind when interpreting our analysis.

### NS1-T123V and other mutations and disease severity

Mutation NS1-T123V was not associated with increased risk of severe or fatal disease, and this mutation was not found in either of the studies included in this analysis in either mild or severe/fatal cases (Fig. 6).

A number of other mutations in the pandemic influenza A(H1N1)pdm09 virus have been reported, and some have been indicated to be associated with disease outcome [126, 127]. We identified three studies that investigated the association between the mutation HA-Q293H (Q310H) and disease outcome. The evidence was inconclusive. Two of the three studies reporting on Q293H found that it increased severity, whereas one found that it did not.

![Fig. 2](image-url)  
**Fig. 2** Mutation HA-D222G and risk of severe disease and mortality. RD (risk difference) = risk in severe or fatal cases minus risk in mild cases, calculated using the random model. The squares represent the estimated risk difference, the diamond represents their average. Horizontal lines indicate 95 % confidence intervals, and the sizes of the squares represent the weight of the study. HA, haemagglutinin gene of influenza A(H1N1)pdm09 virus. All studies recruited patients of all age groups presenting with ILI at hospital, as outpatients or hospitalized patients, or patients who had died.

### Table: Risk difference and 95% CI

<table>
<thead>
<tr>
<th>Study name</th>
<th>D222G &amp; Severe disease</th>
<th>D222G &amp; Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RD</td>
<td>95% CI</td>
</tr>
<tr>
<td></td>
<td>Severe</td>
<td>Mild</td>
</tr>
<tr>
<td>Puzelli et al., (2010)</td>
<td>0.04 (-0.03 - 0.10)</td>
<td>0.253</td>
</tr>
<tr>
<td>Kilander et al., (2010)</td>
<td>0.09 (-0.02 - 0.19)</td>
<td>0.083</td>
</tr>
<tr>
<td>Mak et al., (2010)</td>
<td>0.002 (-0.01 - 0.01)</td>
<td>0.706</td>
</tr>
<tr>
<td>Miller et al., (2010)</td>
<td>0.00 (-0.14 - 0.14)</td>
<td>1.000</td>
</tr>
<tr>
<td>Chen et al., (2010)</td>
<td>0.12 (0.03 - 0.21)</td>
<td>0.007</td>
</tr>
<tr>
<td>Wedde et al., (2013)</td>
<td>0.42 (-0.28 - 0.56)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Moussi et al., (2013)</td>
<td>-0.05 (-0.30 - 0.19)</td>
<td>0.662</td>
</tr>
<tr>
<td>Vazquez-Perez et al., (2013)</td>
<td>0.21 (0.04 - 0.38)</td>
<td>0.016</td>
</tr>
<tr>
<td>Overall</td>
<td>0.10 (-0.02 - 0.18)</td>
<td>0.014</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study name</th>
<th>95% CI</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Severe</td>
<td></td>
</tr>
<tr>
<td>Puzelli et al., (2010)</td>
<td>0.10 (-0.10 - 0.31)</td>
<td>0.329</td>
</tr>
<tr>
<td>Kilander et al., (2010)</td>
<td>0.30 (0.13 - 0.47)</td>
<td>0.001</td>
</tr>
<tr>
<td>Miller et al., (2010)</td>
<td>0.09 (-0.05 - 0.22)</td>
<td>0.198</td>
</tr>
<tr>
<td>Potdar et al., (2010)</td>
<td>0.25 (-0.12 - 0.62)</td>
<td>0.182</td>
</tr>
<tr>
<td>Tse et al., (2011)</td>
<td>0.20 (-0.22 - 0.62)</td>
<td>0.351</td>
</tr>
<tr>
<td>Chen et al., (2010)</td>
<td>0.75 (0.35 - 1.15)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Wedde et al., (2013)</td>
<td>0.33 (0.18 - 0.48)</td>
<td>0.019</td>
</tr>
<tr>
<td>Ferreira et al., (2011)</td>
<td>0.25 (0.09 - 0.41)</td>
<td>0.002</td>
</tr>
<tr>
<td>Moussi et al., (2013)</td>
<td>0.13 (-0.36 - 0.61)</td>
<td>0.611</td>
</tr>
<tr>
<td>Vazquez-Perez et al., (2013)</td>
<td>0.15 (0.005 - 0.30)</td>
<td>0.043</td>
</tr>
<tr>
<td>Overall</td>
<td>0.23 (0.14 - 0.31)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>
although only one of the three statistics was significant (Fig. 6).

Similarly, one of the three studies included in the S203T analysis found that the mutation significantly reduced risk (RD: -0.23, 95% CI: -0.7%–33.0%, p = 0.0001), while of the other two, one found no difference and the other found a positive association (Fig. 6). The analysis of additional mutations in HA therefore had significantly heterogeneous outcomes ($I^2 = 84.1$, p = 0.002 and $I^2 = 79.8$, p = 0.01), and this could again be due to bias caused by differences in the patients’ underlying conditions, vaccination status, age and other factors. None of the studies included in these two analyses investigated bacterial and viral coinfections.

The NS1 mutations T123V and E55S lie in the region responsible for induction of the PKR signalling pathway and in the RNA-binding domain, respectively [71], whereas HA-293 (Q310) and 203 (S220) are antigenic sites [13, 35]. In general, our findings on the role of these mutations should be taken cautiously, as we did not find many studies that clearly reported the number of patients with these mutations who had mild or severe disease. The importance of these mutations has not yet been elucidated.

### Discussion and conclusion

In this review, we found that mutation HA-D222G significantly increased the risk of severe disease and fatality. Mutation HA-D222N was also found to be positively associated, but this was not statistically significant. However, there was no evidence linking mutation HA-D222E with severe disease. The D222G and D222N single and mixed variants have been found in pandemic viruses from approximately 20 countries, including Norway, Mexico, Ukraine and the USA [95, 126, 127]. Influenza viruses constantly change their genetic material due to antigenic drift [8, 9, 12], and a historical understanding of positively
<table>
<thead>
<tr>
<th>Study name</th>
<th>RD</th>
<th>95% CI</th>
<th>p value</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>D222N &amp; Severe disease</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kilander et al., (2010)</td>
<td>0.05</td>
<td>(-0.03 - 0.13)</td>
<td>0.185</td>
<td>[112]</td>
</tr>
<tr>
<td>Mak et al., (2010)</td>
<td>0.01</td>
<td>(-0.01 - 0.03)</td>
<td>0.273</td>
<td>[113]</td>
</tr>
<tr>
<td>Miller et al., (2010)</td>
<td>0.00</td>
<td>(-0.24 - 0.24)</td>
<td>1.000</td>
<td>[114]</td>
</tr>
<tr>
<td>Chen et al., (2010)</td>
<td>0.05</td>
<td>(-0.11 - 0.01)</td>
<td>0.120</td>
<td>[115]</td>
</tr>
<tr>
<td>Baldanti et al., (2010)</td>
<td>0.07</td>
<td>(-0.03 - 0.18)</td>
<td>0.175</td>
<td>[116]</td>
</tr>
<tr>
<td>Wedde et al., (2013)</td>
<td>0.03</td>
<td>(-0.02 - 0.09)</td>
<td>0.254</td>
<td>[117]</td>
</tr>
<tr>
<td>Vazquez-Perez et al., (2013)</td>
<td>0.13</td>
<td>(-0.02 - 0.27)</td>
<td>0.099</td>
<td>[118]</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td>0.02</td>
<td>(-0.01 - 0.05)</td>
<td>0.236</td>
<td></td>
</tr>
</tbody>
</table>

**I² = 33.4, p = 0.17**

<table>
<thead>
<tr>
<th>Study name</th>
<th>RD</th>
<th>95% CI</th>
<th>p value</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>D222N &amp; Mortality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kilander et al., (2010)</td>
<td>0.03</td>
<td>(-0.02 - 0.14)</td>
<td>0.104</td>
<td>[119]</td>
</tr>
<tr>
<td>Miller et al., (2010)</td>
<td>0.07</td>
<td>(-0.21 - 0.08)</td>
<td>0.079</td>
<td>[120]</td>
</tr>
<tr>
<td>Tse et al., (2011)</td>
<td>0.30</td>
<td>(-0.60 - 0.002)</td>
<td>0.002</td>
<td>[121]</td>
</tr>
<tr>
<td>Baldanti et al., (2010)</td>
<td>0.44</td>
<td>(0.24 - 0.63)</td>
<td>0.634</td>
<td>[122]</td>
</tr>
<tr>
<td>Wedde et al., (2013)</td>
<td>0.04</td>
<td>(-0.03 - 0.11)</td>
<td>0.113</td>
<td>[123]</td>
</tr>
<tr>
<td>Ferreira et al., (2011)</td>
<td>0.08</td>
<td>(-0.05 - 0.20)</td>
<td>0.198</td>
<td>[124]</td>
</tr>
<tr>
<td>Vazquez-Perez et al., (2013)</td>
<td>0.08</td>
<td>(-0.04 - 0.20)</td>
<td>0.197</td>
<td>[125]</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td>0.05</td>
<td>(0.01 - 0.09)</td>
<td>0.089</td>
<td></td>
</tr>
</tbody>
</table>

**I² = 73.5 p = 0.001**

Fig. 4  Mutation HA-D222N and risk of severe disease and mortality. RD (risk difference) = risk in severe or fatal cases minus risk in mild cases, calculated using the random model. The squares represent the estimated risk difference, the diamond represents their average. Horizontal lines indicate 95% confidence intervals, and the sizes of the squares represent the weight of the study. HA, haemagglutinin gene of influenza A(H1N1)pdm09 virus. All studies recruited patients of all age groups presenting with ILI at hospital, as outpatients or hospitalized patients, or patients who had died.

<table>
<thead>
<tr>
<th>Study name</th>
<th>RD</th>
<th>95% CI</th>
<th>p value</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PB2-E627K &amp; Severe/Fatal disease</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Akcay Ciblak et al., (2013)</td>
<td>0.00</td>
<td>(-0.3 - 0.30)</td>
<td>1.000</td>
<td>[126]</td>
</tr>
<tr>
<td>Venter et al., (2012)</td>
<td>0.00</td>
<td>(-0.14 - 0.14)</td>
<td>1.000</td>
<td>[127]</td>
</tr>
<tr>
<td>Potdar et al., (2010)</td>
<td>0.00</td>
<td>(-0.27 - 0.27)</td>
<td>1.000</td>
<td>[128]</td>
</tr>
<tr>
<td>Graham et al., (2011)</td>
<td>0.00</td>
<td>(-0.15 - 0.15)</td>
<td>1.000</td>
<td>[129]</td>
</tr>
<tr>
<td>Promed-Email</td>
<td>-0.27</td>
<td>(-0.54 - 0.11)</td>
<td>0.043</td>
<td>[130]</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td>-0.03</td>
<td>(-0.10 - 0.45)</td>
<td>0.510</td>
<td></td>
</tr>
</tbody>
</table>

**I² = 88.0, p < 0.0001**

Fig. 5  Mutation PB2-E627K and risk of severe disease and mortality. RD (risk difference) = risk in severe or fatal cases minus risk in mild cases, calculated using the random model. The squares represent the estimated risk difference, the diamond represents their average. Horizontal lines indicate 95% confidence intervals, and the sizes of the squares represent the weight of the study. PB2, polymerase basic protein 2 of influenza A(H1N1)pdm09 virus. All studies recruited patients of all age groups presenting with ILI at hospital, as outpatients or hospitalized patients, or patients who had died.
selected sites may help to understand the relevance of the observed mutations. Position 222 resides in the receptor-binding site of the HA protein and may possibly influence binding specificity. The HA from the 1918 H1N1 pandemic switched from avian to human receptor specificity through mutation at two positions, G187D and D222G [27]. The A/New York/1/18 strain of the 1918 pandemic possessed a glycine (G) at position 222, and this markedly affected receptor binding, reducing α2-6 preference and increasing α2-3. The A/Memphis/42/1983 strain had an asparagine (N), whereas the 2009 pandemic influenza virus A/C/04/2009 had an aspartic acid (D) at this position (Supplementary Table S2). The results of our study could therefore be useful in public health applications. Other HA mutations identified in studies summarized in this review include D293G (Q310H), S203T (S220T), E374 K, N156D, and N370H in the UK substitution [128]. The importance of these mutations has not yet been elucidated.

No association was observed between mutations PB2-E627K and NS1-T123V and severe and fatal disease. The mutation PB2-E627K has previously been described in animal models to be associated with severe disease. Glutamic acid (E) is generally found in avian influenza viruses, while lysine (K) is found in human viruses; i.e., this mutation determines host specificity [47]. In experiments with single-gene reassortants of influenza viruses, it was shown that changes in the NP, basic polymerase-2 (PB2), and M1 proteins were involved in host restriction in monkeys [129], while attenuation of human viruses was achieved in human volunteers by changes in the NP, non-structural-1 (NS1), PB1, and PB2 proteins [130].

![Fig. 6](image)

The table below shows the risk difference and 95% CI for mutations associated with A(H1N1)pdm09 influenza severity:

<table>
<thead>
<tr>
<th>Study name</th>
<th>RD</th>
<th>95% CI</th>
<th>p value</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS1-T123V &amp; Severe/Fatal disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Potdar et al., (2010)</td>
<td>0.00</td>
<td>(-0.27 - 0.27)</td>
<td>1.00</td>
<td>[107]</td>
</tr>
<tr>
<td>Barrero et al., (2011)</td>
<td>0.00</td>
<td>(-0.17 - 0.17)</td>
<td>1.00</td>
<td>[124]</td>
</tr>
<tr>
<td>Overall</td>
<td>0.00</td>
<td>(-0.14 - 0.14)</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

\[ i^2 = 0, p = 1 \]

<table>
<thead>
<tr>
<th>HA-Q293H (Q310H) Severe/Fatal disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graham et al., (2011)</td>
</tr>
<tr>
<td>-0.13 (-0.07 - 0.33)</td>
</tr>
<tr>
<td>Ferreira et al., (2011)</td>
</tr>
<tr>
<td>-0.08 (-0.27 - 0.11)</td>
</tr>
<tr>
<td>Venter et al., (2012)</td>
</tr>
<tr>
<td>0.50 (0.24 - 0.76)</td>
</tr>
<tr>
<td>Overall</td>
</tr>
<tr>
<td>0.17 (-0.14 - 0.48)</td>
</tr>
</tbody>
</table>

\[ i^2 = 84.1, p = 0.002 \]

<table>
<thead>
<tr>
<th>HA-S203T (S220T) Severe/Fatal disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graham et al., (2011)</td>
</tr>
<tr>
<td>-0.23 (-0.34 - 0.12)</td>
</tr>
<tr>
<td>Farooqui et al., (2010)</td>
</tr>
<tr>
<td>0.00 (-0.32 - 0.32)</td>
</tr>
<tr>
<td>Ferreira et al., (2011)</td>
</tr>
<tr>
<td>0.11 (-0.08 - 0.30)</td>
</tr>
<tr>
<td>Overall</td>
</tr>
<tr>
<td>-0.06 (-0.30 - 0.19)</td>
</tr>
</tbody>
</table>

\[ i^2 = 79.8 p = 0.01 \]
affected by E627K in some studies [59], but not in others [48, 52, 56–58]. The finding in this review supports the latter. The NS1 mutations T123V and E55S lie in the region responsible for induction of the PKR signalling pathway and the RNA-binding domain, respectively [71], and therefore aid in virus replication. In this review, viruses that circulated had the mild PB2 and NS1 phenotypes [107, 115, 124, 127, 131, 132]. Other mutations in the NS1 and PB2 genes found in studies included in this review include PB2-K340N and NS1-G28S, E55G, G154R and T215P [95, 127]. It should be noted that our results are prone to sampling, diagnostic/information bias. None of the included studies reported sequences of all of the viruses that were identified. The authors of each study used their own sampling method to achieve representativeness of the strains circulating during the study period. There is a possibility that the samples included were not representative enough of the virulence mutations in circulation, and that could affect the outcome in this study. Also, not all of the studies included sequences the PB2 and NS1 genes; in most of the studies, only the HA gene was sequenced. Perhaps there could be a different outcome if, in addition to the HA gene, the PB2 and NS1 genes had been sequenced in all of the 18 included studies.

A number of factors, e.g., underlying chronic conditions, age, vaccination and patient’s immune status, and coinfections with other respiratory viruses or bacteria may affect the severity of influenza virus infections. However, some of the studies reviewed here did not control for these factors (Supplementary Table S3), and this might bias the reported outcome. We might not have been able to identify and include other studies. In addition, our inclusion of only studies that were written in the English language might have introduced selection or reporting bias. However, three electronic data bases were used (MEDLINE, EMBASE and Web of Science). In addition, official websites of different organisations were visited to identify recommended references. It is believed that such an extensive effort significantly reduced, if not eliminated, reporting, selection, and publication bias. Systematic error may have occurred in the design and execution of studies included in this review, e.g., preferential admission and sample collection from only severe cases, laboratory contamination or poor laboratory techniques would cause this study to inherit selection bias or information bias. However, a standardised method for scrutinising the quality of studies included in this study (PRISMA and STROBE) was adopted. Where studies had shortfalls, these shortfalls have been mentioned as part of the reporting process (Supplementary Table S3).

Studies have well documented the role of bacterial coinfection in causing ARIs and pneumonia. A 2012 review by Punpanich and Chotpitayasunondh [133] reported that 43 % of the paediatric deaths associated with pandemic influenza A(H1N1)pdm09 virus involved bacterial coinfection. Similar observations have been made by Ruuskanen et al. [134]. On the other hand, respiratory virus co-infections have been associated with severe disease outcome [97, 102, 135–138]. However, in this review, only Vazquez-Perez [120] and Barrero et al. [124] reported on patterns of bacterial coinfections and respiratory virus co-infections, respectively. Respiratory virus infection might lead to destruction of epithelial cells, making bacterial infection more likely [139, 140], or it could be the other way round. Sequential studies investigating initial virus vs. bacterial infection and the epidemiology of subsequent co-infections and disease outcome might help explain the causal relationships.

It would have been good if the role of mutations in seasonal influenza A (H1N1), influenza A (H3N2), influenza B virus (Flu B), and other respiratory viruses, including respiratory syncytial virus (RSV), rhinovirus (RV), adenovirus (AdV), human metapneumovirus (hMPV), and human parainfluenza virus types 1 to 4 (PIV1-4), on disease outcome had also been explored. However, mutations in the other respiratory viruses are not routinely investigated and reported. Li et al. [141] compared the positively selected sites of pandemic influenza vs. seasonal human, avian, and swine influenza viruses in 2009 and 2010. They identified a number of sites in the HA gene that underwent differential selection (HA-86, 94, 153, 160, 202, 234, 250, 303, 374, 399, 473, and 573), but none of these were observed in the studies included in our review. Mutations in the antigenic sites of seasonal influenza A(H1N1) virus have been summarized by Shih et al. [142]. In addition, the positively selected sites for influenza A(H3N2)-HA have been described by Wiley et al. [143], Bush et al. [144], Shih et al. [142], Suzuki and Gojobori [145], and Liao et al. [11], whereas the nine glycosylation sites [N63, N81, N122, N126, N133, N144, N165, N246 and N276] in influenza H3N2 viruses, which were acquired since the first appearance of this virus in 1968, have been documented by Abe et al. [146], Blackburne et al. [147], and Seidel et al. [148]. As influenza viruses also occur as coinfections, the impact of mutations in other respiratory viruses on disease severity should be borne in mind when interpreting our results.

In conclusion, this review has found an association between the mutation HA-D222G and severe and fatal disease. It has also established that during the two years the pandemic influenza A(H1N1)pdm09 virus circulated, no virus quasispecies bearing virulence-conferring mutations in all of the major virulence-conferring genes (HA, PB2 and NS1) predominated in humans. This result reaffirms previous reports suggesting the importance of PB2, NS1 and HA mutations working together to cause
serious disease. The circulating influenza A viruses bearing a glycine at position 222, the receptor-binding site of influenza A viruses, should continue to be monitored for the occurrence of other virulence-conferring mutations in HA, PB2 and NS1. Coinfection with respiratory viruses has been reported to increase disease severity [98], and future studies on the role of genetic mutation on the severity of disease caused by influenza virus should make efforts to control for other respiratory virus coinfections.

Acknowledgments The authors would like to acknowledge the University of Manchester, the Manchester Academic Health Science Centre.

Conflict of interest All authors declare that they have no conflict of interest.

Funding This work received no funding from any funding organization.

References


42. Li ML, Rao P, Krug RM (2001) The active sites of the influenza cap-dependent endonuclease are on different polymerase subunits. EMBO J 20:2078–2086
Mutations associated with A(H1N1)pdm09 influenza severity


82. WHO (2007) A description of the process of seasonal and H5N1 influenza vaccine selection and development


Mutations associated with A(H1N1)pdm09 influenza severity


