Sex Differences in Influenza: The Challenge Study Experience

Luca T. Giurgea,1,6 Adriana Cervantes-Medina,1 Kathie-Anne Walters,2 Kelsey Scherler,2 Alison Han,1,6, Lindsay M. Czajkowski,1 Holly Ann Baus,1 Sally Hunsberger,1 Sabra L. Klein,1 John C. Kash,3 Jeffery K. Taubenberger,2 and Matthew J. Memoli1

1LID Clinical Studies Unit, Laboratory of Infectious Diseases, Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA, 2Institute for Systems Biology, Seattle, Washington, USA, 3Biostatistics Research Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA, 3W. Harry Feinstone Department of Molecular Microbiology and Immunology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, USA, and 4Viral Pathogenesis and Evolution Section, Laboratory of Infectious Diseases, Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA

Background. Preclinical animal studies and retrospective human studies suggest that adult females have worse outcomes from influenza than males. Prospective studies in humans are missing.

Methods. Data from 164 healthy volunteers who underwent influenza A/California/04/2009/H1N1 challenge were compiled to compare differences between sexes. Baseline characteristics, including hormone levels, hemagglutination inhibition (HAI) titers, neuraminidase inhibition (NAI) titers, and outcomes after challenge were compared. Linear and logistic regression models were built to determine significant predictor variables with respect to outcomes of interest.

Results. HAI titers were similar between the sexes, but NAI titers were higher in males than females at 4 weeks and 8 weeks postchallenge. Females were more likely to have symptoms (mean, 0.96 vs 0.80; P = .003) and to have a higher number of symptoms (median, 3 vs 4; P = .011) than males. Linear and logistic regression models showed that prechallenge NAI titers, but not HAI titers or sex hormone levels, were predictive of all shedding and symptom outcomes of interest.

Conclusions. Females in our cohorts were more likely to be symptomatic and to have a higher number of symptoms than males. NAI titers predicted all outcomes of interest and may explain differential outcomes between the sexes.

Keywords. influenza; sex differences; neuraminidase; hemagglutinin; estrogen; testosterone.

Among humans, males and females reportedly differ in the prevalence, pathogenesis, prognosis, and treatment responses for many viral infectious diseases, including coronavirus disease 2019, hepatitis C, human immunodeficiency virus (HIV), and influenza [1, 2]. Biological sex (ie, differences caused by sex chromosome complement and sex steroid hormone concentrations) can impact susceptibility to viruses by affecting antiviral immune responses that restrict virus replication as well as responses that promote inflammation and contribute to tissue damage and prolonged disease [3]. Male–female differences in exposure to and outcomes of viral infections, including influenza, can also be caused by differences in the behaviors, occupations, and even societal norms that define our genders [4]. The complex psychological and cultural interactions between biological sex and gender can complicate interpretation of observational, epidemiological studies reporting male–female differences in infectious disease outcomes.

Preclinical animal models have been used to test hypotheses about the mechanisms mediating sex differences in influenza pathogenesis and responses to inactivated influenza vaccines. These studies collectively show that adult female mice develop greater inflammation and immunity following infection with either H1N1 or H3N2 viruses, which contribute to more severe outcomes in females compared with males [5–8]. Protection against severe outcomes from influenza A virus infections in males is mediated by androgens, which dampen inflammation, including recruitment of monocytes and eosinophils into the lungs of male mice [9]. In humans, epidemiological studies suggest that women of reproductive age have higher rates of influenza [10] and influenza-related hospitalizations [11] compared to men of the same age, though this trend is reversed prior to puberty and at older ages [12]. Pregnancy is a known female-specific risk factor for hospitalization among patients with influenza, but it does not solely explain all female-biased severe outcomes from influenza [13, 14].

In addition to sex-specific differences in infectious disease pathogenesis, a number of preclinical and clinical studies illustrate that females tend to have greater immune responses, reporting of adverse reactions, and even efficacy of vaccines that protect against a number of viral diseases, including influenza.
For example, following vaccination with trivalent inactivated vaccine (TIV), women of reproductive age have worse injection site discomfort and greater hemagglutination inhibition (HAI) titers and neutralizing antibody titers than men [17], with these differences observed in response to either a full dose or a half dose of TIV [18]. Although HAI and neutralizing antibody titers have been considered a correlate of protection, there is a growing appreciation that neuraminidase inhibition (NAI) titers also serve as a correlate of protection because these antibodies reduce influenza viral burden through impairment of viral budding from infected cells [19–21]. Sex differences in NAI, however, have not been evaluated or reported.

Influenza challenge studies provide a unique opportunity for the careful study of a homogenous group of participants controlled for many of the confounders present in retrospective studies. Participants are typically all young, healthy volunteers, thereby controlling for confounding by comorbidities and age. Pregnancy is an exclusion factor for participation in challenge studies. All participants are subjected to equal doses of challenge virus, thereby eliminating exposure differences secondary to occupation and other gender-related factors, which may underlie some of the differences between sexes [22]. Finally, data collection and follow-up in challenge studies are standardized, attenuating differences in reporting and eliminating differences in access to care. We reviewed data from 4 different influenza challenge studies performed by our group, abbreviated as H1N1 pdMIST [23], HAI pdMIST [19], FLU-V [24], and CR6261 [25]. H1N1 pdMIST was the first challenge study using Good Manufacturing Practice–produced wild-type influenza, designed to determine in a dose-escalating manner the optimal dose of virus capable of causing mild to moderate disease in >60% of participants. HAI pdMIST recruited participants with high and low HAI titers to evaluate anti-hemagglutinin (HA) and anti-neuraminidase (NA) antibodies as correlates of protection. The FLU-V study evaluated the efficacy of a novel influenza vaccine against placebo in reducing shedding and symptoms after influenza challenge. Finally, the CR6261 study evaluated the efficacy of a monoclonal antibody infusion (compared to placebo) in reducing shedding and symptoms after influenza challenge. Using data from these 4 influenza challenge studies, we set out to characterize differences between males and females and explore the underlying factors contributing to these differences.

**MATERIALS AND METHODS**

**Study Design**

Data collected from 4 previous H1N1 influenza challenge studies was aggregated. Selection criteria for analysis included only participants who received a dose of challenge virus of \(10^7\) median tissue culture infectious dose \(50\) and excluded participants who received other experimental vaccine or therapeutic products as a part of the treatment arms in these studies (Table 1).

**Table 1. Datasets Included in Study With Study Descriptions and Selection Criteria Applied to Each Study**

<table>
<thead>
<tr>
<th>Study</th>
<th>Study Description</th>
<th>Original Study HAI Titer Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1N1 pdMIST</td>
<td>Influenza dose escalation challenge study designed to determine the optimal dose for influenza challenge</td>
<td>≤1:40</td>
</tr>
<tr>
<td>HAI pdMIST</td>
<td>Influenza challenge comparing participants with low and high HAI titers designed to investigate correlates of immunity</td>
<td>&lt;1:40 and ≥ 1:40</td>
</tr>
<tr>
<td>FLU-V</td>
<td>Phase 2 placebo-controlled study assessing efficacy of a novel conserved capsid peptide vaccine through vaccination and subsequent influenza challenge</td>
<td>&lt;1:10</td>
</tr>
<tr>
<td>CR6261</td>
<td>Phase 2 placebo-controlled study assessing efficacy of a monoclonal antibody against HA stem through early infusion after influenza challenge</td>
<td>≤1:10</td>
</tr>
</tbody>
</table>

Abbreviation: HA, hemagglutinin; HAI, hemagglutination inhibition; TCID\(50\), median tissue culture infectious dose.
plates were read at 490 nm. Titers were assessed from a minimum in the dark, the reaction was quenched with sulfuric acid and the H6N1 and H6N2 viruses were mixed with serial dilutions of heat-inactivated participant sera and incubated overnight in 96-well plates coated with fetuin. Plates were washed and then peanut agglutinin conjugated to horseradish peroxidase was added. After a 2-hour incubation in the dark, plates were washed and o-phenylenediamine dihydrochloride was added. After 10 minutes in the dark, the reaction was quenched with sulfuric acid and the plates were read at 490 nm. Titers were assessed from a minimum dilution of 10 to a maximum of 640 for the FLU-V and CR6261 studies and 2560 for H1N1 pdMIST and HAI pdMIST. Titers below the minimum were coded as 1. All measurements were made in triplicate. Nasal washes were analyzed for viral shedding using 1-step real-time quantitative reverse-transcription polymerase chain reaction for the influenza A virus matrix 1 gene [28]. A standard curve with an external standard was used to calculate copy number.

Outcome Measures

Ages were presented as means. HAI and NAI titer raw data were available as titers and transformed (log2). Data were available at week 4 and week 8 postchallenge for H1N1 pdMIST and HAI pdMIST. Data were available for FLU-V at baseline, day 35 (approximated as week 4), and day 63 postchallenge (approximated as week 8). Data were available for CR6261 at baseline, day 29 (approximated as week 4), and day 66 (approximated as week 8). Presence of symptoms included proportions of participants who had any influenza-related symptoms at any time point. Presence of shedding included proportions of participants who had shedding of influenza at any time point. Mild to moderate influenza disease, defined as presence of shedding and symptoms, included proportions of participants who had both presence of symptoms and presence of shedding at any time point. Days of symptoms and days of shedding were presented as medians. Number of symptoms includes the number of different symptoms each participant had throughout the study period and was presented as a median.

Statistical Analysis

Data was collected in Excel and processed using RStudio version 1.1.463. Comparisons between sexes were performed using t tests for normally distributed variables (age, testosterone, prechallenge NAI titers, week 4 NAI titers, week 8 NAI titers), using Wilcoxon rank-sum tests for nonnormally distributed variables (estradiol, prechallenge HAI titers, week 4 HAI titers, week 8 HAI titers, days of shedding, days of symptoms, number of symptoms), and using tests of proportions or Fisher exact test for proportions. The analysis was repeated for the subgroup of participants with documented viral shedding. Differences and

![Figure 1](https://example.com/figure1.png)

**Figure 1.** Flow diagram of number of subjects from each dataset included for analysis after application of selection criteria, with breakdown by sex. Abbreviations: F, female; HAI, hemagglutination inhibition; M, male.
95% confidence intervals were provided for t tests. Interquartile ranges were provided for non–normally distributed variables. P values < .05 were considered statistically significant. No adjustment for multiple analyses was performed. Linear regression models were performed to examine outcomes (days of shedding, days of symptoms, number of symptoms) with age, sex, prechallenge estrogen level, prechallenge testosterone level, prechallenge HAI titer, and prechallenge NAI titer as predictor variables. Logistic regression models were performed to analyze outcomes (presence of shedding, presence of symptoms, mild to moderate influenza disease) with age, sex, prechallenge estrogen level, prechallenge testosterone level, prechallenge HAI titer, and prechallenge NAI titer as predictor variables. There was a small amount of missing data, particularly in antibody assays at day 28 (10 participants, all male), due to participant nonadherence with appointments.

RESULTS

Baseline Characteristics

Of the 164 participants included in the analysis, 91 were male and 73 were female. Age was similar between the sexes (Table 2). At baseline, testosterone levels were higher in males than females and estradiol levels were lower in males than females. Prechallenge geometric mean HAI titers were similar between males and females. Similarly, prechallenge geometric mean NAI titers were not significantly different between males and females.

Postchallenge Antibody Titers and Outcomes

Postchallenge geometric mean HAI titers rose at week 4 and 8 but were similar between males and females (Table 3, Supplementary Figure 1). The geometric mean NAI titers, however, were significantly greater in males compared to females at weeks 4 and 8 (Table 3, Supplementary Figure 2). Males were 16% less likely to have symptoms of influenza than females, though the likelihood of viral shedding was similar. Males had a median 4 days of symptoms, whereas females had a median 5 days of symptoms, though this did not reach statistical significance. Males had a significantly fewer median number of symptoms than females. Days of shedding did not significantly differ between males and females. Similar differences between sexes were observed in a subgroup analysis of participants with viral shedding (Table 4), though the only outcome to retain statistical significance was the number of symptoms.

Predictive Modeling

Linear and logistic models (Table 5, Supplementary Tables 1–6) demonstrated that prechallenge NAI titers significantly predicted all outcomes (Supplementary Figures 3–8), which were presence of symptoms (P = .042), presence of shedding (P < .001), days of symptoms (P = .001), days of shedding (P < .001), number of symptoms (P < .001), and mild to moderate influenza disease (P < .001). Age, sex, baseline testosterone level, baseline estradiol level, and prechallenge NAI titers were not statistically significant predictors of outcomes. Tests of interactions between sex and all other covariates were not significant, with exception of an interaction between sex and prechallenge HAI titer for presence of shedding as outcome.

DISCUSSION

These findings add to a growing body of literature exploring sex differences in disease from influenza. Through analysis of data from healthy volunteer influenza challenge studies, we were able to overcome many of the limitations and confounding factors found in prior studies that have demonstrated sex differences with respect to clinical outcomes and antibody responses. We have demonstrated that clinical influenza disease experienced by females in our cohorts was worse compared to males. Specifically, a higher proportion of women experienced symptoms after challenge and women had more symptoms during their illness than men. Even among a subgroup of participants with evident infection demonstrated by active viral shedding, females had a higher number of symptoms. The rest of the observations lost their statistical significance, likely due to low sample size, though the trends of higher NAI titers in males and more severe symptoms in females persisted. These findings are in agreement with observations from retrospective studies showing that women have worse outcomes with influenza, even when accounting for comorbidities and exposure risk [10, 11].

Prior work has hypothesized that hormonal differences leading to transcriptional changes in immune cells may account for some of the variation in responses between sexes to influenza [5, 16, 29, 30]. The linear and logistic regression models

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Male (n = 91)</th>
<th>Female (n = 73)</th>
<th>Difference (95% CI)</th>
<th>PValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y, mean</td>
<td>29.3</td>
<td>29.8</td>
<td>-0.47 (–2.59 to 1.65)</td>
<td>.664</td>
</tr>
<tr>
<td>Testosterone, ng/mL, mean</td>
<td>5.84</td>
<td>0.93</td>
<td>4.91 (4.48–5.33)</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Estradiol, pg/mL, median (IQR)</td>
<td>16.22 (12.55–24.21)</td>
<td>31.58 (17.38–64.8)</td>
<td>…</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Prechallenge log₂ HAI titer, median (IQR)</td>
<td>0 (0–3.82)</td>
<td>0 (0–4.32)</td>
<td>…</td>
<td>.962</td>
</tr>
<tr>
<td>Prechallenge log₂ NAI titer, mean</td>
<td>6.55</td>
<td>5.94</td>
<td>0.60 (~21 to 1.42)</td>
<td>.147</td>
</tr>
</tbody>
</table>

Means were compared by t test and medians by Wilcoxon-rank sum test.

Abbreviations: CI, confidence interval; HAI, hemagglutination inhibition; IQR, interquartile range; NAI, neuraminidase inhibition.
constructed in this study did not find that baseline testosterone or estradiol levels were predictive of outcomes. Despite prior studies showing that postvaccination HAI titers were higher in women than men [17, 18], this study found no differences at 4 or 8 weeks after inoculation with a live wild-type challenge virus. However, 3 of 4 data sets in this study restricted participation to participants with lower HAI titers, which potentially selected for participants with poor immunologic responses to HA compared to the general population, potentially obscuring differences between the sexes. Serum NAI titers were higher in males than females after challenge, reaching statistical significance at 4 weeks and 8 weeks postchallenge but not at the prechallenge timepoint. Interestingly, all models identified prechallenge NAI titers as the highly significant variable predictive of all disease outcomes of interest. Sex itself was not a significant predictor variable, suggesting that differences in prechallenge NAI titers, once adjusted for other variables, underlie some of differences in outcomes observed between males and females. This finding is consistent with previous work demonstrating that anti-NA antibodies are pivotal in decreasing symptoms and attenuating disease severity [19, 20, 31, 32]. Two of the challenge studies [19, 25] used for this analysis already determined the importance of anti-NA antibodies in protection against influenza challenge, but importantly, the addition of 2 additional datasets only reinforced this finding. This further supports recent efforts drawing attention to the importance of immunity against NA to develop novel therapeutics and vaccines with improved efficacy and breadth of protection [21, 33, 34]. The protection afforded by anti-NA antibodies is generally considered infection permissive and is consistent with the similar rates of infection in the male and female subgroups despite differences in NAI titers [35]. However, it is important to note that it is unknown whether the artificial manner of inoculation with a single spray of a large inoculum of virus manifests any clinical differences from natural infection, which typically arises from longer exposure times.

The overall higher NAI titers seen in male participants provide a potential mechanism to explain some of the differences in symptom-related outcomes between sexes observed in this study, while acknowledging that the difference in NAI titers between sexes did not quite reach statistical significance at baseline. However, the difference became significant at both postchallenge timepoints, suggesting the possibility that males have a better NAI memory response during the early days of the infection, from which we do not have data. Additional features of anti-influenza immunity that were not measured could also potentially play important roles in sex-based differences in the pathogenesis of influenza. Cell-mediated immunity was not

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Male (n = 91)</th>
<th>Female (n = 73)</th>
<th>Difference (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 4 log HAI titer, median (IQR)</td>
<td>5.32 (3.32–6.32)</td>
<td>4.82 (2.49–6.32)</td>
<td>...</td>
<td>.784</td>
</tr>
<tr>
<td>Week 8 log HAI titer, median (IQR)</td>
<td>4.32 (0–6.32)</td>
<td>5.32 (0–6.32)</td>
<td>...</td>
<td>.733</td>
</tr>
<tr>
<td>Week 4 log NAI titer, mean</td>
<td>7.95</td>
<td>7.33</td>
<td>0.62 (0.1–1.23)</td>
<td>.046</td>
</tr>
<tr>
<td>Week 8 log NAI titer, mean</td>
<td>8.03</td>
<td>7.39</td>
<td>0.64 (0.03–1.25)</td>
<td>.039</td>
</tr>
<tr>
<td>Presence of symptoms</td>
<td>0.80</td>
<td>0.96</td>
<td>–0.16 (0.6–25)</td>
<td>.003</td>
</tr>
<tr>
<td>Presence of shedding</td>
<td>0.62</td>
<td>0.60</td>
<td>–0.01 (–0.16 to 0.14)</td>
<td>.869</td>
</tr>
<tr>
<td>Mild to moderate influenza disease</td>
<td>0.56</td>
<td>0.59</td>
<td>...</td>
<td>.7129</td>
</tr>
<tr>
<td>Days of symptoms, median (IQR)</td>
<td>4 (1–7)</td>
<td>5 (3–7)</td>
<td>...</td>
<td>.064</td>
</tr>
<tr>
<td>No. of symptoms, median (IQR)</td>
<td>3 (1–5)</td>
<td>4 (2–7)</td>
<td>...</td>
<td>.011</td>
</tr>
<tr>
<td>Days of shedding, median (IQR)</td>
<td>1 (0–4.5)</td>
<td>2 (0–6)</td>
<td>...</td>
<td>.486</td>
</tr>
</tbody>
</table>

Medians were compared by Wilcoxon-rank sum test, means by t test, and proportions using test of proportions.

Abbreviations: CI, confidence interval; HAI, hemagglutination inhibition; IQR, interquartile range; NAI, neuraminidase inhibition.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Male (n = 56)</th>
<th>Female (n = 44)</th>
<th>Difference (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 4 log HAI titer, median (IQR)</td>
<td>4.82 (3.32–6.07)</td>
<td>4.32 (0–6.32)</td>
<td>...</td>
<td>.839</td>
</tr>
<tr>
<td>Week 8 log HAI titer, median (IQR)</td>
<td>4.32 (0–5.32)</td>
<td>4.32 (0–6.32)</td>
<td>...</td>
<td>.797</td>
</tr>
<tr>
<td>Week 4 log NAI titer, mean</td>
<td>7.58</td>
<td>7.17</td>
<td>0.41 (–0.41 to 1.23)</td>
<td>.328</td>
</tr>
<tr>
<td>Week 8 log NAI titer, mean</td>
<td>7.87</td>
<td>7.22</td>
<td>0.65 (–1.15 to 1.45)</td>
<td>.112</td>
</tr>
<tr>
<td>Presence of symptoms</td>
<td>0.91</td>
<td>0.98</td>
<td>...</td>
<td>.225</td>
</tr>
<tr>
<td>Days of symptoms, median (IQR)</td>
<td>6 (3–8)</td>
<td>6 (4–9)</td>
<td>...</td>
<td>.368</td>
</tr>
<tr>
<td>No. of symptoms, median (IQR)</td>
<td>3.5 (2–6)</td>
<td>5 (2–10)</td>
<td>...</td>
<td>.047</td>
</tr>
<tr>
<td>Days of shedding, median (IQR)</td>
<td>4 (1–6)</td>
<td>5 (2–6)</td>
<td>...</td>
<td>.084</td>
</tr>
</tbody>
</table>

Data are presented as medians, means, or proportions. Medians were compared by Wilcoxon-rank sum test, means by t test, and proportions using test of proportions or Fisher exact test.

Abbreviations: CI, confidence interval; HAI, hemagglutination inhibition; IQR, interquartile range; NAI, neuraminidase inhibition.
measured in this study but is known to be a significant factor in cross-protective immunity against influenza [36]. Furthermore, mucosal immunity is likely to play an underappreciated role in protection, especially against mild infections of the upper respiratory tract [37, 38]. Future models integrating these measures, including mucosal levels of anti-NA antibodies, may be better equipped to explain the relationships between sex and outcomes with influenza.

Despite the aforementioned benefits of influenza challenge studies, there are some limitations inherent to challenge studies in general, and to the specific characteristics of the cohorts used in this study. Volunteers with baseline HAI titers <1:40 were selected to participate in 2 of the 4 studies (and HAI titers <1:10 in a third study), leading to a specially curated population with decreased external validity.

Confounding may also have been introduced by differences in inclusion/exclusion criteria between studies and unequal proportions of male and female participants in each cohort. Unlike antibody levels, hormone levels are known to oscillate over short periods of time and our single measurement may be insufficient to completely rule out a relationship to clinical outcomes. Furthermore, even though data collection of symptoms was standardized, there may still be differences in reporting among sexes. Use of more objective outcome measures, such as markers of inflammation, may be helpful in supporting the conclusions of this study. Finally, the analysis between sexes and the regression models was not adjusted for multiple comparisons, which may weaken the meaningfulness of some observations but would be unlikely to impact the highly statistically significant predictive capacity of NAI titers in the regression models.

These data agree with published observations that females of reproductive age have worse outcomes during influenza infection. In this study, NAI titers were a predictor of clinical outcomes after influenza challenge, suggesting the possibility that anti-NA antibodies may underlie differences between sexes. The majority of participants in these studies, however, were specifically selected to have low HAI titers. Future studies in more general populations, ideally including additional variables like cell-mediated and mucosal immunity, are needed to confirm the differences in NAI titers between sexes observed here and to validate and refine the predictive models.

### Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

### Notes

**Financial support.** This research was supported by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases, National Institutes of Health.

**Potential conflicts of interest.** All authors: No potential conflicts of interest.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

### References
