Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Responses in COVID-19 patients are distinct across sexes. Yu et al. show that, compared with male patients, female COVID-19 patients exhibit evidence of a major recruitment wave of circulating mucosal-associated invariant T (MAIT) cells into the airways and demonstrate that these cells display an immunologically active phenotype.
SUMMARY
Background: Sexual dimorphisms in immune responses contribute to coronavirus disease 2019 (COVID-19) outcomes, but the mechanisms governing this disparity remain incompletely understood.
Methods: We carried out sex-balanced sampling of peripheral blood mononuclear cells from hospitalized and non-hospitalized individuals with confirmed COVID-19, uninfected close contacts, and healthy control individuals for 36-color flow cytometry and single-cell RNA sequencing.
Findings: Our results revealed a pronounced reduction of circulating mucosal-associated invariant T (MAIT) cells in infected females. Integration of published COVID-19 airway tissue datasets suggests that this reduction represented a major wave of MAIT cell extravasation during early infection in females. Moreover, MAIT cells from females possessed an immunologically active gene signature, whereas cells from males were pro-apoptotic.
Conclusions: Our findings uncover a female-specific protective MAIT cell profile, potentially shedding light on reduced COVID-19 susceptibility in females.
Funding: This work was supported by NIH/NIAID (U01AI066569 and UM1AI104681), the Defense Advanced Projects Agency (DARPA; N66001-09-C-2082 and HR0011-17-2-0069), the Veterans Affairs Health System, and Virology Quality Assurance (VQA; 75N93019C00015). The content is solely the responsibility of the authors and does not necessarily represent the official view of the National Institutes of Health.

INTRODUCTION
Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has led to a global pandemic of coronavirus disease 2019 (COVID-19) and a death toll of more than 2.3 million people and rising. According to reported sex-disaggregated data, males are affected disproportionately by SARS-CoV-2, with a higher incidence of cases, mortality, and morbidity. This follows a similar trend of higher case fatality rates for males in SARS-CoV and Middle East respiratory syndrome CoV (MERS-CoV) and experiments using SARS-CoV mouse models. Sex differences in the immune response are thought to be a key contributing factor to these CoV disease
outcomes, agreeing with the current body of knowledge indicating that innate and adaptive immune responses are altered substantially according to sex.\textsuperscript{7–10} Specific to SARS-CoV-2 infection, responses of lymphocytes and myeloid cells have been shown to be associated with COVID-19 outcomes.\textsuperscript{11–19} Correspondingly, a recent study of sex differences in COVID-19 immune responses uncovered an association between poor disease outcomes in males and weak T cell responses in CD4\textsuperscript{+} and CD8\textsuperscript{+} compartments, whereas poor outcomes in females were associated with high innate immune cytokines, tumor necrosis factor superfamily 10 (TNFSF-10), and interleukin-15 (IL-15).\textsuperscript{20} The sex differences elucidated in this seminal study further cement the need to better understand the mechanisms governing sex-specific susceptibility to SARS-CoV-2.

RESULTS

In this study, we carried out sex-balanced sampling of peripheral blood mononuclear cells (PBMCs) from individuals with COVID-19 and control subjects for 36-color flow cytometry and single-cell RNA sequencing (scRNA-seq) analyses. A total of 88 samples were analyzed from 45 individuals. Details regarding subject demographics and sample information are summarized in Figure 1A and Tables 1 and S1). Briefly, we analyzed samples from 28 individuals with COVID-19 as confirmed by positive SARS-CoV-2 PCR and/or immunoglobulin G (IgG) seroconversion. These included 9 hospitalized subjects (20%), 7 requiring intensive care, hereafter referred to as “hospitalized.” An additional 19 subjects were identified in non-hospital settings (42.2%), hereafter referred to as “infected.” Most of these confirmed COVID-19 cases were sampled longitudinally (range, 1–28 days), including pre- and post-anti-SARS-CoV-2 IgG seroconversion. The dates of symptom onset for all subjects with confirmed COVID-19 were recorded at enrollment, providing an illness range of 1–40 days. We also recorded symptom severity, obtained via investigator survey on 39 symptoms related to COVID-19 (STAR Methods). Additionally, we included 7 subjects (15.6%), hereafter referred to as “exposed,” who were also sampled at multiple time points. These subjects, despite being close contacts of infected individuals, remained with negligible symptom scores, were negative for SARS-CoV-2 by PCR, and did not demonstrate detectable anti-SARS-CoV-2 IgG for at least 2 months after enrollment. Last, we included a group of 10 “healthy” subjects (22.2%) who were enrolled prior to the pandemic in 2019 and did not show any symptoms associated with COVID-19 or other respiratory illness.\textsuperscript{21}

Immune profiling of PBMCs from individuals with COVID-19 reveals the association of CD8\textsuperscript{+}CD161\textsuperscript{hi} T cell frequencies with disease severity

With these flow cytometry data, we generated a map of immune cell populations and their subsets by down-sampling to 3,000 viable CD45\textsuperscript{+} singlets per sample and concatenated all data for uniform manifold approximation and projection (UMAP)\textsuperscript{22} and unbiased clustering via flow self-organizing maps (FlowSOM).\textsuperscript{23} Unique marker expression of respective populations facilitated our annotation of major PBMC populations (Figures 1B and 1C), including CD4\textsuperscript{+} and CD8\textsuperscript{+} (αβ) T cells, γδ T cells, B cells, plasmablasts, natural killer (NK) cells, monocytes (MOs), and dendritic cells (DC), which were also confirmed by manual analysis (Figures S1A–S1H). Subpopulations were also annotated in this manner, such as CD45RA\textsuperscript{-}CD27\textsuperscript{+}CCR7\textsuperscript{-} naive, CD45RA\textsuperscript{+}CCR7\textsuperscript{-} central memory (CM), CD45RA\textsuperscript{+}CCR7\textsuperscript{-} effector memory (EM), and CD45RA\textsuperscript{-}CD27\textsuperscript{-}CCR7\textsuperscript{-} terminally differentiated EM (terminally differentiated effector memory cells re-expressing CD45RA [EMRA]) CD8\textsuperscript{+} T cells and CD8\textsuperscript{+}CD161\textsuperscript{hi} T cells and other indicated subpopulations (Figures 1B, 1C, and S1A–S1H). We noted that a minor population of basophils (Basos) and
Please cite this article in press as: Yu et al., Mucosal-associated invariant T cell responses differ by sex in COVID-19, Med (2021), https://doi.org/10.1016/j.medj.2021.04.008
neutrophils (polymorphonuclear neutrophils [PMNs]), primarily from hospitalized individuals, were detected (Figure 1B) despite using a PBMC isolation protocol.

We next set out to screen our flow cytometry dataset for immune populations that exhibited major quantitative changes associated with COVID-19 severity. Regression analysis revealed a significant (p < 0.05) association of several major PBMC populations, including CD8+ T cells, NK cells, B cells, and MOs relative to disease severity rank (i.e., healthy, exposed, infected, and hospitalized) (Figure 1D). The highest degree of association with the most significance was found in CD8+ T cells (Figure 1D). Additional analysis of CD8+ T subpopulations also revealed significant changes in CD8+ EM T cells (Figure S1I) and in CD8+ CD161hi T cells with high significance (p = 0.0006) (Figure S1I). Other significant changes were also observed, such as in B cell subpopulations (naive, IgD+ non-class switch, and plasmablasts), NK cells (CD56lo populations), DCs (CD11c+ DCs and plasmacytoid DCs [pDCs]), MOs (classical, intermediate, and nonclassical), and CD4+ (EM) αβ T cells (Figure S1I). Our findings therefore suggest that the frequencies of certain lymphocyte and myeloid populations are affected in COVID-19, including a major effect on CD8+ CD161hi (αβ) T cells.

The robust statistical changes in total CD8+ T cells and the CD8+ CD161hi T subpopulation prompted us to look at these cells more closely. Regarding annotation of this CD161hi cluster, because the overwhelming majority of events were low to negative for CD56 and for T cell receptor γδ (TCRγδ) (Figure 1E), the phenotype was largely consistent with mucosal-associated invariant T (MAIT) cells but not NK T or γδ T cells. This designation is congruent with recent work with COVID-19 PBMCs.24–27 Next we performed a more focused analysis of this cluster in COVID-19 by assessing their frequencies by severity rank using samples taken within 3 days of enrollment, the time point most proximal to the initial symptom score recordings. Results were displayed via UMAP contour plots, revealing a reduction in these cells in the SARS-CoV-2 setting (Figure 1F). Manual gating (CD3+ TCRγδ− CD8+ CD161hi) from all flow cytometry events revealed a significant reduction when comparing healthy (p = 0.0036) or exposed (p = 0.0488) subjects versus hospitalized subjects and a negative correlation (p = 0.0002) with disease severity rank (Figure 1G). Moreover, the ages of the hospitalized group are significantly higher than those of other groups (Figure 1H), although the ages of sampled subjects were similar between sexes at each severity rank (Figure 1I). To address the potential confounding factor of age, we built a multiple regression model with age and severity rank as two independent variables (Figure 1J). Our results showed that age does not significantly (p = 0.1165) affect CD161hi cells but severity rank does (p = 0.0034). Therefore, we concluded that changes of these CD161hi cells are associated with COVID-19 disease severity.

Figure 1. Loss of peripheral CD8+CD161hi T cell frequencies correlates with increased severity of COVID-19

(A) Overview of groups in this study.
(B) UMAP visualization of PBMC subsets identified by FlowSOM clustering. Samples from all participants were pooled and down-sampled to 3,000 live CD45+ cells per sample. MO, monocye; NK, natural killer; DC, dendritic cell; PMN, polymorphonuclear neutrophil; Baso, basophil.
(C) Representative marker expression by CD4+ T cell, CD8+ T cell, MO, NK cell, and B cell subsets.
(D) Correlation analysis of immune cell subsets as shown in (C) with disease severity rank.
(E) Expression of CD161, TCR γδ and CD56 in CD8+ T cell subsets. NK cells were used as a positive control for CD56 expression.
(F) UMAP of samples grouped by disease severity rank. Samples collected within 3 days from initial symptom score recording were included.
(G) Frequencies of CD161hi T cells (mean ± standard error) in different severity groups (left) and their correlation with severity rank (right).
(H and I) Age comparisons among severity rank and between sexes.
(J) Multiple linear regression of age (x1) and severity rank (x2) with CD161hi cell frequencies. Regression models with p values are shown for age and severity rank.

Significance was determined by Kruskal-Wallis test with Dunn’s test (E) or ANOVA (H and I). *p < 0.05, **p < 0.01, ***p < 0.001. Correlation efficiency was calculated by Spearman’s rank correlation (D and G). See also Figure S1 and Tables 1 and S1.
Elucidation of sex differences in CD8\(^+\) lymphocytes during SARS-CoV-2 infection

Given the previously reported sex differences in immune responses in COVID-19\(^,\)\(^2\),\(^20\) we further grouped the data by sex and characterized the frequencies of CD8\(^+\) sub-sets stratified by disease severity rank (Figures 2A, 2B, S2A, and S2B), time after symptom onset (early, \(\leq 14\) days; middle, 15–21 days; late, >21 days) (Figures 2C–2E), and seroconversion status (Figure 2F). The severity rank results showed that healthy females had greater frequencies of CD161\(^hi\) cells relative to males and that females had a more dramatic loss of peripheral CD161\(^hi\) cells compared with males in COVID-19 (Figures 2A–2C). We also found that, within the CD8\(^+\) compartments of healthy subjects, males had greater frequencies of CD8\(^+\) memory T cells (combined EMRA, EM, and CM) (Figures 2D and 2E). Although the memory cell predominance in males was preserved at all time points, the greater abundance of CD161\(^hi\) cells in females was lost at early and middle time points but not in late disease (Figures 2D and 2E). The loss of this difference was due to a precipitous drop of CD161\(^hi\) cells in females at early and middle time points (Figure 2F). No obvious changes of naive CD8\(^+\) T cells were found (Figure S2C). Last, when analyzing data from individuals with confirmed COVID-19 by seroconversion status, we found that CD161\(^hi\) cells were higher in females relative to males prior to seroconversion, whereas CD8\(^+\) memory cells were higher in males in seroconverted subjects (Figure 2G). We identified a female-specific decline in circulating CD161\(^hi\) cell frequencies upon exposure to/infection with SARS-CoV-2. This sex-specific reduction may be due to extravasation into airway tissue, suggesting a sex-specific role of these CD161\(^hi\) cells in COVID-19.

scRNA-seq uncovers the role of CD161\(^hi\) T lymphocytes in SARS-CoV-2 immune responses

Next we analyzed scRNA-seq (10x Genomics) data of 48 different PBMC samples from 24 subjects across all groups (Tables 1 and S1) to further characterize these

### Table 1. Summary of demographics and sample information

<table>
<thead>
<tr>
<th>Group</th>
<th>Healthy</th>
<th>Exposed</th>
<th>Infected</th>
<th>Hospitalized</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (mean ± SD [range])</td>
<td>39.70 ± 13.68 (25–61)</td>
<td>42.57 ± 13.93 (17–60)</td>
<td>36.73 ± 13.88 (20–65)</td>
<td>59.44 ± 15.11 (31–76)</td>
<td>42.84 ± 16.06 (17–76)</td>
</tr>
<tr>
<td>Number of subjects (female:male [F:M ratio])</td>
<td>n = 10 (5:5)</td>
<td>n = 7 (3:4)</td>
<td>n = 19 (8:11)</td>
<td>n = 9 (4:5)</td>
<td>n = 45 (20:25)</td>
</tr>
<tr>
<td>Race (%</td>
<td>African American</td>
<td>3 (30.00%)</td>
<td>0</td>
<td>1 (5.26%)</td>
<td>5 (55.56%)</td>
</tr>
<tr>
<td></td>
<td>Asian</td>
<td>0</td>
<td>1 (14.29%)</td>
<td>2 (10.53%)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>White</td>
<td>7 (70.00%)</td>
<td>5 (71.42%)</td>
<td>16 (84.21%)</td>
<td>3 (33.33%)</td>
</tr>
<tr>
<td></td>
<td>Others/unknown</td>
<td>0</td>
<td>1 (14.29%)</td>
<td>0</td>
<td>1 (11.11%)</td>
</tr>
<tr>
<td>Days since onset when enrolled (mean ± SD [range])</td>
<td>N/A</td>
<td>15.17 ± 11.25 (5–33)</td>
<td>11.18 ± 4.30 (3–19)</td>
<td>8.25 ± 6.14 (1–18)</td>
<td>11.19 ± 6.73 (1–33)</td>
</tr>
<tr>
<td># samples of flow cytometry (F:M ratio)</td>
<td>n = 10 (5:5)</td>
<td>n = 20 (9:11)</td>
<td>n = 44 (21:23)</td>
<td>n = 9 (4.5)</td>
<td>n = 83 (39:44)</td>
</tr>
<tr>
<td># samples of scRNA-seq (F:M ratio)</td>
<td>n = 5 (3:2)</td>
<td>n = 8 (3.5)</td>
<td>n = 29 (12.17)</td>
<td>n = 6 (1:5)</td>
<td>n = 48 (19:29)</td>
</tr>
<tr>
<td>Days since onset when samples collected</td>
<td>N/A</td>
<td>26.94 ± 14.94 (5–61)</td>
<td>18.27 ± 8.44 (3–40)</td>
<td>8.25 ± 6.14 (1–18)</td>
<td>19.32 ± 11.44 (1–61)</td>
</tr>
</tbody>
</table>

See also Table S1.
Figure 2. Sex-specific responses of circulating CD8+ T cells in individuals with COVID-19

(A) UMAP of samples stratified by sex and severity rank as shown in Figure 1F.
(B) Frequencies of CD161hi cells between females and males at each rank group.
(C) Linear regression of CD161hi cells with severity rank between sexes. Dashed lines indicate 95% confidence intervals. Regression models with p values are shown for each sex.
(D) UMAP of samples stratified by sex and time after symptom onset (early, ≤14 days; middle, >15 days and ≤21 days; late, >22 days).
(E) Frequencies of CD161hi and memory CD8+ T cells between sexes and time points.
(F) Sex-specific changes of CD161hi cells frequencies shown in (E).
(G) Frequencies of CD161hi and memory CD8+ T cells in samples from subjects with confirmed COVID-19 pre- and post-seroconversion.

Data were plotted as mean ± standard error (B and E–G). Significance was determined by Mann-Whitney test (B, E, and G) and Kruskal-Wallis test with Dunn’s test (F). *p < 0.05. See also Figure S2.
CD161<sup>hi</sup> cells. Data were processed using the Seurat 3 package, and subsequent transcript-based annotation was carried out (Figures 3A, S3A, and S3B). Within T cell subsets (Figure 3B; Table S2), we were able to identify CD161<sup>hi</sup> cells in a single cluster containing high KLRB1 (i.e., CD161) expression and co-expression of CD3D, CD8A, and TRAV1-2 (Figures 3C and S3B), which encodes the Vα7.2 invariant TCR α chain on MAIT cells. Grouping these data by disease severity showed that hospitalized individuals had lower frequencies of T cells, including CD161<sup>hi</sup> cells (Figure 3D), agreeing with our flow cytometry findings and consistent with reported lymphopenia in individuals with severe COVID-19. Also showing the same trend as our flow cytometry data was the higher frequency of CD161<sup>hi</sup> cells in healthy adults (Figure 3E), although it did not reach statistical significance because of the

Figure 3. Characterization of CD8<sup>+</sup>CD161<sup>hi</sup> T cells among COVID-19 PBMCs using scRNA-seq

(A) UMAP and unsupervised cluster analysis of PBMCs. RBC, red blood cell; PB, plasmablast; PLT, platelet.

(B and C) Visualization of T cell subsets with high resolution in UMAP (B) and expression of their marker genes as indicated in violin plots (C). The T<sup>+</sup> cluster likely represents a dropout population with low unique molecular identifier (UMI) counts. N, naive; EM, effector memory; CM, central memory; DN, double negative; rep, replicating.

(D) Changes of T cell subsets with severity rank. N, number of individuals.

(E) Frequencies of CD161<sup>hi</sup> clusters relative to all T cell subsets. Females were plotted in red and males in blue. The red dashed box delineates healthy females.

(F) Top enriched pathways of CD161<sup>hi</sup> clusters in the Reactome Pathway Database, ranked by false discovery rate (FDR; −log<sub>10</sub> scale). Data were plotted as mean ± standard error. Significance was determined by Kruskal-Wallis test (E). See also Figure S3 and Table S2.
variations between females and males. Next, to address the functional role of this CD161 hi cluster in COVID-19, we performed gene enrichment analysis using differentially expressed genes (DEGs) calculated by comparing the CD161 hi cluster from all samples with all other T cell clusters. With several top-ranked hits consisting of immune pathways related to cytokine signaling (IL-4, IL-13, IL-10, and transforming growth factor β [TGF-β] signaling) and signal transduction and transcription (an NGF-TRKA signaling axis, nuclear events, and RAF-independent mitogen-activated protein kinase (MAPK) 1/3 activation) and an estrogen-dependent pathway (Figure 3F), our results inferred a possible sex-specific immune response of these CD161 hi cells in COVID-19.

Considering the critical roles of MOs in activating MAIT cells 33 and the association of dysregulated MOs with severe COVID-19 outcomes 12,13,20,34, we applied the CellphoneDB package 35 and analyzed ligand-receptor interactions between MO and T cell clusters in our data. We identified 5 MO clusters, including three conventional subsets (classical, non-classical, and intermediate) and two subsets of activated MOs (CD16 MOA and CD14 MO A) associated with COVID-19 (Figures 4A and 4B). These two activated MO subsets, which were also reported by other studies, 13,34 express high levels of interferon-related genes (Figure S3B). Although cell-cell communications may not take place in the blood in the same way as in solid tissue, our results inferred unique interactions between CD161 hi cells and MOs with the following gene pairs: KLRB1_CLEC2D, CCL5_CCR1, CXCR6_CXCL16, and IL18_IL-18R (Figures 4C and 4D). Moreover, the number of interaction counts of MOs was most abundant with the CD161 hi cluster relative to all major T cell populations (Figure 4E). A recent study reported IL-18-dependent MAIT cell activation by SARS-CoV-2-infected macrophages, which supports the existence of unique interactions of MAIT cells with MOs/macrophages in COVID-19.27 These transcriptome findings further suggest a role of circulating CD161 hi cells in the SARS-CoV-2 immune response and their possible interactions with MO-derived cells.

Sex-specific heterogeneity of circulating MAIT cells in COVID-19

To analyze our scRNA-seq dataset for potential sex differences in circulating CD161 hi cells, we first sought to examine phenotypic heterogeneity in this population. To do this, we performed a focused subcluster analysis, which generated 3 distinct clusters (Figure 5A). However, this added resolution revealed a cluster that expressed TRDC, encoding the constant region of the δ chain expressed by γδ T cells (Figure 5B), which was excluded from subsequent analyses. In contrast, the other two clusters had higher KLRB1 and TRAV1-2 expression (Figure 5B), and they are therefore referred to here as MAITα and MAITβ clusters. Of note, these 2 clusters make up approximately 80% of CD161 hi PBMCs, which is consistent with a previously reported frequencies of circulating MAIT cells.24 Our results showed that the MAITα cluster possessed upregulated genes associated with cytotoxic T cells (GNLY, CD8A, and CD8B), migration/adhesion (CXCR4 and ITGB2), and cytokine signaling (IRF1, B2M, NFkB1, JUNB, and FOS) (Figure 5C; Table S3). Although the MAITα cluster showed expression of the hemoglobin B gene (HBB), this cluster did not express other hemoglobin genes (Figures S4A and S4B), and its HBB expression level was much lower than in red blood cells (RBCs), suggesting that this cluster is not contaminated with RBCs. The MAITβ cluster was enriched for genes of ribosomal proteins, apoptosis (BAX and STUB1), and the linker histone H1 associated with apoptosis (HIST1H1C, HIST1H1D, and HIST1H1E) (Figure 5C; Table S3). Gene enrichment analysis further supported a functional dichotomy for α and β clusters. Although MAITα was enriched with several immune process pathways (e.g., interferon γ [IFN-γ], IL-4, and IL-13 signaling as well as antigen processing and
Figure 4. Receptor-ligand interaction inferences uncover unique interactions between CD8+CD161hi T cells and MOs

(A and B) Visualization (A) and percentage distribution (B) of different MO subsets identified in Figure S3A. Three MO subsets represent resting classical (cl) CD14, non-classical (nc) CD16, and intermediate (int) MO as seen in healthy subjects. Two MO subsets involved in IFN signaling are associated with individuals with COVID-19. Based on differential expression of CD14 and CD16, they are referred to as activated CD14 and activated CD16 MO (CD14 MOA and CD16 MOA, respectively).

(C) Overview of selected ligand-receptor interactions inferred using CellPhoneDB in the COVID-19 PBMC single-cell dataset. The red box delineates specific interactions of CD161 hi T cells with MOs. p values and scales are indicated by circle size and color, respectively.
MAITβ was enriched in cellular responses to external stimuli, metabolism of RNA, viral infection, and programmed cell death but not immune processes (Figures 5D and 5E; Table S4). Hence our results suggest MAIT cell heterogeneity, with the MAITα signature representing an immunologically poised/active phenotype and the MAITβ signature representing a stressed/apoptotic phenotype.

Last for this series of experiments, we sought to determine the dynamics of the two phenotypically distinct clusters by sex over the COVID-19 disease course. By first grouping our data by severity and time after symptom onset, we found that MAITα was the major phenotype in healthy individuals, whereas MAITβ predominated in exposed and infected groups (Figures 5F and 5G) and at early, middle, and late time points (Figure 5H). There was a noted exception for hospitalized individuals (Figure 5F), who had very few cells, as seen in our flow cytometry data, consistent with lymphopenia, which occurs in severe COVID-19.36 When further stratified by sex, we found that MAIT cell frequencies were higher in healthy females (Figures 5I and 5J), corroborating our flow cytometry results. In healthy females, these cells were skewed toward the MAITα cluster, whereas the few cells present in healthy males consisted mostly of the MAITβ cluster (Figure 5I). However, this difference was lost in the exposed/infected setting, where the cells of individuals of both sexes were comprised mostly of the MAITβ cluster (Figure 5J). Nonetheless, MAITβ cluster percentages were statistically greater in females in late disease (Figure 5J), which reflects the increased MAIT cells during late infection in females, as shown by our flow cytometry findings. Regarding expression of CD69, a T cell activation marker, we did not observe major differences across cluster or sex but did observe elevated expression in the hospitalized group (Figures S4C–S4F). This possibly suggests an altered MAIT cell response in hospitalized individuals.24–27 These results suggest sex-specific MAIT cell differences at the quantitative and phenotypic levels in health and COVID-19.

**Respiratory tract MAIT cell responses differ by sex in COVID-19**

To test our hypothesis that the sex-specific change of circulating MAIT cells is due to their extravasation into airway tissues, we utilized published scRNA-seq datasets of bronchoalveolar lavage fluid (BALF)37 and nasopharyngeal swabs (NPSs)38 to assess MAIT cells in airway tissues of individuals with COVID-19. Beginning our analysis with the BALF dataset, we identified the MAIT cell cluster by expression of TRAV1-2, CD3D, KLRB1, and SLC4A10 (Figures 6A–6C). With this annotation, we found a significant increase (p = 0.0188) of MAIT cells in individuals with COVID-19 (including mild and severe cases) relative to normal control individuals and a higher MAIT cell frequency (p = 0.0332) in females relative to males among subjects with COVID-19 (Figure 6D). This detection of increased MAIT cells in BALF from females, along with the drop in these cells we observed in peripheral blood from females, suggest female-dominant extravasation of MAIT cells in COVID-19.

To perform the same analysis with NPS samples, we integrated the T cell data from NPSs with those from BALF to identify MAIT cells in NPSs (Figure 6E and 6F) because TCR genes were not aligned in the NPS dataset.38 With this annotation, we identified 232 and 631 MAIT cells in BALF and NPS, respectively, and quantified MAIT cell frequencies in the NPS dataset, again observing a significant increase (p = 0.0139) in individuals with COVID-19 (Figure 6G). We also analyzed the data across severity,

(D) Expression of representative ligand and receptor pairs between MAIT cells and MOs as indicated. Red and blue circles indicate CD161hi clusters and MOs, respectively.
(E) Heatmap of interaction counts between major T cell and MO subsets.
observing a significant increase ($p = 0.0038$) in subjects with moderate disease relative to normal subjects and a decrease ($p = 0.0366$) relative to individuals with critical COVID-19 (Figures 6H and 6I). We could not perform the same analysis by sex.
Figure 6. Sex differences of MAIT cells in airway tissue samples from individuals with COVID-19

(A and B) Clustering analysis of scRNA-seq data from the COVID-19 BALF dataset with subtracted T and NK cells.37

(C) MAIT cell cluster indicated by marker genes.

(D) Frequencies of MAIT cells in BALF between normal subjects and individuals with COVID-19 (left) and across sexes within subjects with COVID-19 (right).
because of insufficient numbers of samples from females (Figure 6I). Nonetheless, these results match the reduced circulating MAIT cell frequencies seen in our hospitalized individuals, suggesting that a lymphopenic state that occurs in severe COVID-19 affects MAIT cells, consistent with other reports.24–26,39

In a final experiment, we sought to characterize MAIT cell transcriptomes by sex in the BALF and NPS datasets and determine whether these cells resembled α and β phenotypes we identified in circulating MAIT cells. Cluster analysis was not warranted here, given the low cell numbers in these datasets. Instead, we leveraged gene modules derived from our respective α and β clusters of circulating MAIT cells. We found that MAIT cells in BALF and NPS data were largely skewed toward the β module, with minimal sex differences (Figure S5A). However, when we directly examined DEGs (DEGs between sexes), we were able to detect sex differences associated with α and β phenotypes. Specifically, in BALF, we found increased IL7R expression in females (Figure 6J) and other IL-7 signaling-associated genes (CISH and SOCS1) (Figure 6K). Given the key role of this signaling in T cell survival, we explored additional pathway genes, finding that female MAIT cells had upregulated anti-apoptotic genes (BCL2 and FOXP1) and downregulated pro-apoptotic genes (BAX and CASP3) (Figures S5B and S5C). Also observed in cells from females were upregulated anti-proliferative genes (CDKN1B and BTG2) (Figure S5D). These patterns matched MAITα gene changes in our PBMC dataset. We were able to find other sex differences, including increased expression of several transcription factors (KLF2, MYC, and CEBPD) (Figure 6L). Conversely, males had higher expression of CCL2 (Figure 6M), which has been linked to COVID-19 immunopathology.40 In short, our results infer sex differences at the qualitative level in COVID-19, with MAIT cells from females possessing a pro-survival and immunologically active phenotype.

DISCUSSION

Despite the knowledge of sex differences in the immune response as an underlying factor in COVID-19 disease outcomes, the sexually dimorphic responses of MAIT cells, an unconventional T cell population deemed important in this disease, has remained unknown. Although MAIT cells may be known primarily for their role in host defense against bacteria, responding via their invariant TCRs to microbial vitamin B metabolites presented on major histocompatibility complex class I-related molecules,33,41 they are also capable of responding in a TCR-independent manner to cytokines such as IL-12 and IL-18 and can be critical players in immune responses against certain viruses.42–44 We now demonstrate that MAIT cells in females are quantitatively and qualitatively more robust in the SARS-CoV-2 setting, potentially helping us understand the immunological reasons for reduced COVID-19 susceptibility of females.

Our finding that MAIT cell recruitment to airway tissues may be more robust in females with COVID-19 was aided first by our observation of higher frequencies of circulating MAIT cells in females in the healthy setting. This difference can be explained by the
rate of physiological aging-related attrition of MAIT cells, which is substantially less pronounced in the blood of females.45–47 The resultant higher frequencies in circulation enabled us to readily uncover the precipitous percentage drop we saw with MAIT cells relative to exposed/infected females. When trying to elucidate the potential cause of this drop, we considered two possible scenarios, (1) lymphopenia and (2) extravasation, which are not necessarily mutually exclusive. It is accepted that lymphopenia is associated with severe COVID-19 infection,32,36,48,49 which agrees with our observations in our hospitalized group (comprised of 77.8% of individuals in intensive care). Similarly, lymphopenia could partially explain the reduction in MAIT cells described by Jouan et al.,26 who studied male-dominated samples from critically ill individuals with COVID-19, although extravasation also likely occurred. In our study, however, we demonstrated that the frequencies of circulating MAIT cells dropped in our group of infected individuals. Because these subjects were not critically ill, our findings point to extravasation as a major reason for the sex-specific drop in the frequencies of circulating MAIT cells. The same pattern may also exist in several other studies.24,25,27 For example, although it has been demonstrated that circulating MAIT cells are reduced in individuals with moderate COVID-19 relative to healthy subjects in aggregated data,24 it is possible that the frequencies in healthy females contributed to reaching a statistical difference. Further supporting our conclusion, we were able to show, with publicly available scRNA-seq data from COVID-19 BALF samples,37 that females in that study had an increased MAIT cell percentage relative to males, allowing us to conclude that MAIT cell extravasation during COVID-19 may be quantitatively more robust in females.

Our results also suggest that MAIT cells may be qualitatively superior in females with respect to anti-viral immune activity in COVID-19. Leading us to this conclusion, our scRNA-seq analysis of PBMCs from affected individuals revealed two distinct clusters of MAIT cells, MAITα and MAITβ. The α cluster was enriched for various immune pathways, such as IFN-γ signaling, inferring a capacity for anti-viral immune function. In contrast, the β cluster was enriched for cell stress and apoptosis pathways, inferring a frail phenotype roughly similar to a previously described population of double-negative MAIT cells.50,51 We showed in the healthy setting that MAIT cells in females were skewed toward the α cluster, whereas those in males were comprised of the β cluster. Although, from these results, it could be inferred that the α cluster should be overrepresented in airways of females with COVID-19, this was not the case with BALF. However, we reasoned that such a finding would be very difficult to prove for two main reasons. First, extravasated MAIT cells with an α phenotype would be restricted to the early wave of recruitment because circulating cells are almost completely skewed to the β module in exposed/infected individuals. Second, a certain level of transcriptional reprogramming would occur upon immune cell extravasation into the tissue and potentially again upon accessing the alveolar space. Still, we were able to show in BALF that certain gene patterns remained consistent with the α signature in females versus males. In addition, our finding that BALF samples from females had quantitatively more MAIT cells gives further credence that differences revealed in the blood would likewise extend to tissue.

Limitations of study
The following limitations of this study should be considered when interpreting results. It is possible that decreased COVID-19 severity and elevated levels of MAIT cells are associated independently with the female sex. We showed the immunologically active signature of MAIT cells from females at the transcriptomic level, but functional analysis of MAIT cells in SARS-CoV-2 infection would be needed to demonstrate whether MAIT cells directly protect females with COVID-19. The same argument is relevant with respect to MAITα and MAITβ subsets in lung
tissues of individuals with COVID-19. Last, although our analyses included over 80 sex-balanced samples, the current study may not be powered to distinguish other immune cell types or specific sequelae that are associated with sex differences in COVID-19, and larger cohorts with more comprehensive longitudinal sampling would be needed in the future.

We conclude that MAIT cells in females are quantitatively and qualitatively distinct from those of males, and we surmise that this distinction provides a protective advantage in COVID-19. Indeed, females in general tend to have elevated frequencies of circulating MAIT cells, as also gleaned from large independent studies with European, South Korean, and Chinese populations. Further supporting this argument, it has now been recorded that fatality rates of adults with COVID-19 tend to be lower in females at all ages in 37 of 38 different countries or regions where MAIT cell frequencies tend to be higher in females, including North America, Europe, and Asia. These points also argue against the possibility that an immunologically more robust MAIT cell response has a net negative effect; for example, by immunological misfiring or cytokine storm-related immunopathology. However, one open question our findings raise is whether the males in our study, who had more circulating CD8+ memory T cells, would have an advantage in the reinfection setting or following vaccination. Future studies are needed to explore this question and to better understand sex differences in MAIT cells in general and in COVID-19.

**STAR METHODS**

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Ethics statement
  - Participants in this study
- METHOD DETAILS
  - Collection of peripheral blood mononuclear cells (PBMCs)
  - Sample processing for flow cytometry and single cell RNA-sequencing (scRNA-seq)
  - Panel and Staining for Flow Cytometry
  - 36-color Full Spectrum Flow Cytometry
  - High-dimensional data analysis of flow cytometry data
  - ScRNA-seq using 10x Genomics platform
  - Processing and quality control of scRNA-seq
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Dimensionality reduction and clustering analysis
  - Differential gene expression analysis
  - Pathway enrichment analysis
  - Inference of ligand-receptor interactions between T cells and monocytes
  - Integration of BALF and NPS dataset
  - Calculations of the feature scores in MAIT cells
  - Statistical analysis

**SUPPLEMENTAL INFORMATION**

Supplemental information can be found online at https://doi.org/10.1016/j.medj.2021.04.008.
ACKNOWLEDGMENTS

This work was supported by NIH/NIAID (U01AI066569, UM1AI104681), the U.S. Defense Advanced Projects Agency (DARPA, N66001-09-C-2082 and HR0011-17-2-0069), the Veterans Affairs Health System, and Virology Quality Assurance (VQA) 75N93019C00015. COVID-19 samples were processed under Biosafety level (BSL)-2 with aerosol management enhancement or BSL-3 in the Duke Regional Biocontainment Laboratory which received partial support for construction from NIH/NIAID (UC6AI058607). The flow cytometry was performed at Duke Eye Center with the support from NIH/NEI (P30EY005722, R01EY02179) and Research to Prevent Blindness (Unrestricted, Duke Eye Center). We would like to thank Monica DeLay and Patrick Duncker (Cytek Biosciences) for help with spectral flow cytometry and Chris Ciccolella and Geoff Kraker (Omiq, Inc.). We would also like to thank Maria Miggs, Deborah Murray, Tyffany Locklear, Robert Rolfe, Jack Anderson, Allison Fullenkamp, Raul Louzuo, Thad Gurley, and Julie Steinbrink for their work and the Durham Veterans Affairs Health Care System and Duke Regional Hospital for support.

AUTHOR CONTRIBUTIONS


DECLARATION OF INTERESTS

M.T.M. reports grants for biomarker diagnostics from the Defense Advanced Research Projects Agency (DARPA), National Institutes of Health (NIH), Sanofi, and the Department of Veterans Affairs. T.W.B. reports grants from DARPA and is a consultant for Predigen. M.T.M., T.W.B., E.L.T., and C.W.W. report pending patents on molecular methods to diagnose and treat respiratory infections. E.L.T. reports grants on biomarker diagnostics from DARPA and the NIH/Antibacterial Resistance Leadership Group (ARLG) and an ownership stake in Predigen. G.S.G. reports an ownership stake in Predigen. C.W.W. reports grants for biomarker diagnostics from DARPA, NIH/ARLG, Predigen, and Sanofi and has received consultancy fees from bioMerieux, Roche, Biofire, Giner, and Biomeme.

Received: December 22, 2020
Revised: February 9, 2021
Accepted: April 7, 2021
Published: April 13, 2021

REFERENCES


### KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antibodies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BV421 anti-human CCR7 (clone G043H7)</td>
<td>BioLegend</td>
<td>Cat# 353208</td>
</tr>
<tr>
<td>BV711 anti-human CCR6 (clone G034E3)</td>
<td>BioLegend</td>
<td>Cat# 353436</td>
</tr>
<tr>
<td>BV750 anti-huamn CXCR5 (clone RF8B2)</td>
<td>BD Biosciences</td>
<td>Cat# 747111</td>
</tr>
<tr>
<td>BV563 anti-huamn CCR5 (clone 2D7)</td>
<td>BD Biosciences</td>
<td>Cat# 741401</td>
</tr>
<tr>
<td>PE-Cyanine7 anti-human CXCR3 (clone CEW33D)</td>
<td>ThermoFisher</td>
<td>Cat# 25-1839-42</td>
</tr>
<tr>
<td>PerCP-eFluor710 anti-human TCR γδ (clone B1.1)</td>
<td>ThermoFisher</td>
<td>Cat# 46-9959-42</td>
</tr>
<tr>
<td>APC-Fire810 anti-human CD38 (clone HIT2)</td>
<td>BioLegend</td>
<td>Cat# 303549</td>
</tr>
<tr>
<td>FITC anti-human CD57 (clone HNK-1)</td>
<td>BioLegend</td>
<td>Cat# 359604</td>
</tr>
<tr>
<td>PE-Cyanine5 anti-huamn CD95 (clone DX2)</td>
<td>ThermoFisher</td>
<td>Cat# 15-0959-42</td>
</tr>
<tr>
<td>Spark-NIR685 anti-huamn CD19 (clone H1B19)</td>
<td>BioLegend</td>
<td>Cat# 302270</td>
</tr>
<tr>
<td>Spark-NIR550 anti-human CD14 (clone 63D3)</td>
<td>BioLegend</td>
<td>Cat# 367148</td>
</tr>
<tr>
<td>PerCP anti-human CD45 (clone H130)</td>
<td>ThermoFisher</td>
<td>Cat# MHCD4531</td>
</tr>
<tr>
<td>PE anti-huamn CD25 (clone BC96)</td>
<td>ThermoFisher</td>
<td>Cat# 12-0259-42</td>
</tr>
<tr>
<td>APC anti-human CD27 (clone O323)</td>
<td>ThermoFisher</td>
<td>Cat# 17-0279-42</td>
</tr>
<tr>
<td>APC-eFluor780 anti-human HLA-DR (clone L243)</td>
<td>ThermoFisher</td>
<td>Cat# 47-99952-42</td>
</tr>
<tr>
<td>PerCP-Cyanine5.5 anti-human CD11b (clone ICRF44)</td>
<td>BioLegend</td>
<td>Cat# 301328</td>
</tr>
<tr>
<td>PE-eFluor610 anti-human CD24 (clone eBioSN3 (SN3 A5-2H10))</td>
<td>ThermoFisher</td>
<td>Cat# 61-0247-41</td>
</tr>
<tr>
<td>Alexa Fluor®647 anti-human CD1c (clone L161)</td>
<td>BioLegend</td>
<td>Cat# 331510</td>
</tr>
<tr>
<td>APC-R700 anti-human CD127 (clone HIL-7R-M21)</td>
<td>BD Biosciences</td>
<td>Cat# 565185</td>
</tr>
<tr>
<td>BUV496 anti-human CD16 (clone 3G8)</td>
<td>BD Biosciences</td>
<td>Cat# 612944</td>
</tr>
<tr>
<td>BV480 anti-human IgD (clone IA6-2)</td>
<td>BD Biosciences</td>
<td>Cat# 566138</td>
</tr>
<tr>
<td>BUV395 anti-human CD45RA (clone SH9)</td>
<td>BD Biosciences</td>
<td>Cat# 740315</td>
</tr>
<tr>
<td>BUV737 anti-human CD56 (clone NCAM16.2)</td>
<td>BD Biosciences</td>
<td>Cat# 612766</td>
</tr>
<tr>
<td>BUV805 anti-human CD8 (clone SK1)</td>
<td>BD Biosciences</td>
<td>Cat# 612889</td>
</tr>
<tr>
<td>BUV661 anti-human CD11c (clone B-5y6)</td>
<td>BD Biosciences</td>
<td>Cat# 612967</td>
</tr>
<tr>
<td>SuperBright436 anti-human CD123 (clone 6H6)</td>
<td>ThermoFisher</td>
<td>Cat# 62-1239-42</td>
</tr>
<tr>
<td>BVS70 anti-human IgM (clone MHM-88)</td>
<td>BioLegend</td>
<td>Cat# 314517</td>
</tr>
<tr>
<td>BV650 anti-human CD28 (clone CD28.2)</td>
<td>BioLegend</td>
<td>Cat# 302946</td>
</tr>
<tr>
<td>BB515 anti-human CD141 (clone 1A4)</td>
<td>BD Biosciences</td>
<td>Cat# 566017</td>
</tr>
<tr>
<td>eFluor®450 anti-human CD161 (clone HP-3G10)</td>
<td>ThermoFisher</td>
<td>Cat# 48-1619-42</td>
</tr>
<tr>
<td>BV510 anti-human CD3 (clone OKT3)</td>
<td>BioLegend</td>
<td>Cat# 317332</td>
</tr>
<tr>
<td>Pacific Orange anti-human CD20 (clone HIL47)</td>
<td>ThermoFisher</td>
<td>Cat# MHCD2030</td>
</tr>
<tr>
<td>BVS605 anti-human IgG (clone G18-145)</td>
<td>BD Biosciences</td>
<td>Cat# 563246</td>
</tr>
<tr>
<td>BV785 anti-human PD-1 (clone EH12.2H7)</td>
<td>BioLegend</td>
<td>Cat# 329929</td>
</tr>
<tr>
<td>BUV615 anti-human CD4 (clone SK3)</td>
<td>BD Biosciences</td>
<td>Cat# 612987</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Biological samples</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Peripheral blood from healthy and COVID-19 related subjects</td>
<td>Duke University Health System</td>
<td>NA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chemicals, peptides, and recombinant proteins</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Live/Dead UV Blue</td>
<td>ThermoFisher</td>
<td>Cat# L34962</td>
</tr>
<tr>
<td>Brilliant Stain Buffer Plus</td>
<td>BD Biosciences</td>
<td>Cat# 566385</td>
</tr>
</tbody>
</table>

(Continued on next page)
RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Daniel R. Saban (daniel.saban@duke.edu).

Materials availability
This study did not generate new unique reagents.

Data and code availability
Publicly available scRNA-seq data of BALF\(^{37}\) were downloaded from GEO with the accession number GSE145926, and the count data of NPS\(^{38}\) were downloaded from https://figshare.com/articles/COVID-19_severity_correlates_with_airway_epithelium-immune_cell_interactions_identified_by_single-cell_analysis/12436517. All clinical metadata of participants and samples in this study are included in Table S1. The fcs files of flow cytometry are available to download from https://flowrepository.org/id/FR-FCM-Z3WR. The scRNA-seq of PBMC data generated in this study are available in Gene Expression Omnibus (GEO) with the accession number GSE171555. Additional Supplemental Items are available from Mendeley Data at https://dx.doi.org/10.17632/csc9p34d5t.1. All analytic scripts will be also made available per request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Ethics statement
This study and relevant protocols were approved by the Institutional Review Boards of Duke University Health System (DUHS). All procedures were performed in accordance with the Declaration of Helsinki, applicable regulations, and local policies.

Participants in this study
Inpatients (hospitalized) and outpatients (infected) with confirmed infection of SARS-CoV-2 were identified through the DUHS and enrolled into the Molecular and Epidemiological Study of Suspected Infection (MESSI, Pro00100241). The PCR testing for
SARS-CoV-2 was performed at either the North Carolina State Laboratory of Public Health or at clinical laboratories of the DUHS. The exposed group, who closely contacted with COVID-19 patients, presented negative PCR test and negative serology test during longitudinally sampling from the first visit to at least 2 months after, typically 0, 7, 14, and 28 days relative to enrollment. Initial severity scores of individuals were recorded through a self-reporting survey on 38 defined symptoms related to COVID-19 plus “other” when enrolled. The exposed group (average score = 9.71) showed a lower severity scores compared with infected (outpatients) group (average score = 18.16). The hospitalized patients presented severe disease symptoms with breath difficulty, cough, fever or chest pain when enrolled, and 77.8% of them for this study required intensive care unit (ICU) care. All COVID-19 patients were also longitudinally sampled with serology test from enrollment to convalescent phase. Healthy donors were enrolled in 2019 (Duke IRB Protocol Pro00009459) with no diagnosis or symptoms consistent with COVID-19 or other respiratory illness. Written informed consent was obtained from all subjects or legally authorized representatives. Patient Demographics are summarized in Tables 1 and S1.

In keeping with inclusion criteria above, all participants with the minimum age of 17 were maintained in the analysis. Investigators were blinded during experiments in terms of sex, time post symptom onset and anti-SARS-COV-2 IgG serology condition.

**METHOD DETAILS**

**Collection of peripheral blood mononuclear cells (PBMCs)**

Peripheral whole blood was collected in EDTA vacutainer tubes and processed within 8 hours via Ficoll-Hypaque density gradient method to obtain cryopreserved PBMCs. Blood was diluted 1:2 in PBS, layered onto the the Ficoll-Hypaque in 50 mL conicals, and centrifuged at 420 g for 25 min. Buffy coat was used to collect PBMCs and cell pellets were washed twice with D-PBS following centrifugation at 400 g for 10 min. Counts and cell viability were obtained using Vi-Cell automated cell counter (Beckman-Coulter) before PBMCs were adjusted to 10x10^6 cells/ml in cryopreservation media (90% FBS, 10% DMSO) and aliquoted into cryopreservation vials on ice. Finally, cells were subjected to controlled freezing at −80°C using CoolCell LX (BioCision) for 12-24 hours, and transferred to liquid nitrogen vapor phase.

**Sample processing for flow cytometry and single cell RNA-sequencing (scRNA-seq)**

Counts and cell viability of thawed PBMCs were measured by Countess II after a wash with DMEM 10% FBS. The cell viability of hospitalized patients ranged from 70%–80% whereas all other samples exceeded 80% viability. An additional dead cell removal step (Miltenyi Biotec) was conducted on hospitalized PBMC samples prior to aliquot for scRNA-seq. To perform scRNA-seq, 200,000 cells per sample were aliquoted, spun down, resuspended in 30 µL PBS supplemented with 0.04% BSA and 0.2U/µl RNase inhibitor and counted using Countess II.

**Panel and Staining for Flow Cytometry**

Approximately 0.5-2x10^6 cells per cryopreserved sample were stained for flow cytometry analysis. Antibody titrations used in this study were previously established by Cytek Biosciences with slight modifications. All staining procedures were performed at room temperature. PBMCs were stained with live/dead Blue (Thermo Fisher) for 15 min, washed with FACS-EDTA buffer and spun down at 1500 rpm for 5 min. Samples were resuspended with Brilliant Stain Buffer Plus (BD Biosciences) and sequentially stained with anti-CCR7 for 10 min, all other chemokine receptor
mix of CCR6, CXCR5, CCR5 and CXCR3 for 5 min, anti-TCR gamma/delta for 10 min and the rest surface receptor mix for 30 min. After incubation, PBMCs were washed with FACS-EDTA buffer and spun down at 1500 rpm for 5 min. Samples were fixed with 1% PFA for 20 min, spun down and resuspended in FACS-EDTA buffer.

36-color Full Spectrum Flow Cytometry
Samples were acquired using a four-laser Cytek Aurora Spectral Flow Cytometry System. Single color controls for spectral unmixing were acquired with PBMCs from healthy control blood and UltraComp eBeads (ThermoFisher). Raw data were unmixed and further analyzed using either FlowJo for manual gating or Omiq (https://www.omiq.ai) for clustering visualization and analysis.

High-dimensional data analysis of flow cytometry data
Uniform Manifold Approximation and Projection (UMAP) and FlowSOM clustering analyses were performed on Omiq (https://www.omiq.ai), using equal random sampling of 3000 live CD45+ singlets, from each FCS file. The UMAP plot was generated with the parameters of 15 neighbors and 0.4 minimum distance. All markers in flow panel were used for analysis except live/dead and CD45.

ScRNA-seq using 10x Genomics platform
10x Genomics Single Cell 5’ v1 chemistry was used to generate Gel Bead-In Emulsions (GEM), and perform post GEM-RT cleanup, cDNA amplification, as well as library construction. An agilent DNA ScreenTape assay was used for quality control. Libraries were pooled and sequenced to saturation or 20,000 unique reads per cell on average using an Illumina NovaSeq6000 with 150-bp paired-end reads.

Processing and quality control of scRNA-seq
Raw sequencing data were initially processed with 10x Genomics Cell Ranger pipelines (V3.1.0). Briefly, BCL files were demultiplexed to generate FASTQ files. FASTQ files were aligned with STAR aligner to the human genome reference GRCh38 from Ensemble database. Feature barcode processing and UMI counting were then performed according to the standard workflow. (QC summary after sequencing). The following criteria were applied as quality control of single cells from all individual samples. Cells that had fewer than 1000 UMI counts or 500 genes, as well as cell that had greater than 10% of mitochondrial genes were removed from further analysis. Genes that were expressed by fewer than 10 cells were also excluded. After filtering, a total of 424,080 cells with 18,765 gene features were kept for the downstream analysis.

QUANTIFICATION AND STATISTICAL ANALYSIS
Dimensionality reduction and clustering analysis
The filtered gene-barcode matrix was analyzed using Seurat 3.28 All the procedures were conducted with the default parameters unless otherwise specified. Briefly, data were first normalized using log transformation and adjusted with a scale factor of 10,000. The top 2,000 variable genes were identified, and percentages of mitochondrial genes were regressed out when scaling data. Principle component analysis (PCA) was performed using these top variable genes, and top 25 principle components (PCs) were selected for graph-based clustering with Shared Nearest Neighbor (SNN) and visualization in UMAP. The resolution was set to 0.35 to identify major immune cell subsets in PBMCs. Sub-clustering of CD161hi T cells (21,610 cells) was also performed using the analytic pipeline mentioned above with two modifications: top 10 PCs were used, and the resolution was set to 0.1 to identify MAIT cell clusters.
Differential gene expression analysis
Differentially expressed genes (DEGs) were identified using Seurat 3 (FindAllMarkers or FindMarkers Functions) with either ‘wilcox’ for all cluster markers or ‘DESeq2’. Randomly downsampled data with 100,000 cells were used to find all markers of PBMC clusters. A gene was considered significant with adjusted p value or false discovery rate (FDR) < 0.05. DEGs results of all PBMCs and MAIT cells are listed in Tables S2 and S3.

Pathway enrichment analysis
Top 100 DEGs of MAIT clusters were used for pathway enrichment analysis using Reactome Pathway Database (https://reactome.org). A pathway was considered significantly over-presented with FDR < 0.05. The full pathway enrichment results are summarized in Table S4.

Inference of ligand-receptor interactions between T cells and monocytes
Ligand-receptor interactions between T cells and monocytes were inferred using CellPhoneDB. PBMC scRNA-seq data were randomly downsampled to 50,000 cells and T and monocyte clusters were extracted based on the expression of their lineage markers. CellPhoneDB was with default parameters (https://github.com/Teichlab/cellphonedb). The inferred interactions are considered significant when p value < 0.05.

Integration of BALF and NPS dataset
Publicly available scRNA-seq data of BALF and of NPS were downloaded and processed using Seurat 3 as previously described. All T cell clusters, were extracted from both datasets and integrated via the Single Cell Transform (SCT) method in Seurat 3. Top 3,000 variable features were selected for the integration. Dimensionality reduction was conducted using PCA and UMAP embedding of the top 100 PCs. Clusters were visualized at a resolution of 0.8 after constructing a SNN graph using the first 50 PCs.

Calculations of the feature scores in MAIT cells
The DEGs between MAIT1 and MAIT2 were used to generate their feature scores as previously described. The feature scores were calculated using AddModuleScore function in Seurat 3.

Statistical analysis
Data normality and homogeneity of variance were assessed using Kolmogorov-Smirnov test and Bartlett’s test, respectively. Due to the distribution and variance of human data, non-parametric statistical tests were favorably used throughout this study unless otherwise specified. Mann Whitney U test was used for two-group comparisons, and Kruskal-Wallis with post hoc Dunn’s test was used for comparisons of three groups and more. Spearman’s correlation efficiency was used to quantify the correlation of the ranked disease severity (from healthy as 1, to hospitalized as 4). Simple linear regression and multiple linear regression were calculated using GraphPad Prism 8. Data were transformed to meet the assumptions of linear regression when necessary. To adjust p values for multiple hypothesis testing, FDR correction was performed using the Benjamini-Hochberg procedure when appropriate. Two-tailed tests were used unless otherwise specified. A p value or FDR < 0.05 is consider statistically significant. Graphical data of quantifications presented throughout are expressed as the means ± SEMs and were plotted using Graphpad Prism 8. Other graphs in this study were generated using either the corresponding analytic packages or R package ggplot2.