Original research

Parenchymal lung abnormalities following hospitalisation for COVID-19 and viral pneumonitis: a systematic review and meta-analysis

Laura Fabbri 1,2, Samuel Moss 1,2, Fasihul A Khan 1,2, Wenjie Chi 3, Jun Xia 3, Karen Robinson 4, Alan Robert Smyth 1,2, Gisli Jenkins 1,2, Iain Stewart 1,2

ABSTRACT

Introduction Persisting respiratory symptoms in COVID-19 survivors may be related to development of pulmonary fibrosis. We assessed the proportion of chest CT scans and pulmonary function tests consistent with parenchymal lung disease in the follow-up of people hospitalised with COVID-19 and viral pneumonitis.

Methods Systematic review and random effects meta-analysis of proportions using studies of adults hospitalised with SARS-CoV-2, SARS-CoV, MERS-CoV or influenza pneumonia and followed up within 12 months. Searches performed in MEDLINE and Embase. Primary outcomes were proportion of radiological sequelae on CT scans; restrictive impairment; impaired gas transfer. Heterogeneity was explored in meta-regression.

Results Ninety-five studies (98.9% observational) were included in qualitative synthesis, 70 were suitable for meta-analysis including 60 SARS-CoV-2 studies with a median follow-up of 3 months. In SARS-CoV-2, the overall estimated proportion of inflammatory sequelae was 50% during follow-up (0.50; 95% CI 0.41 to 0.58; I²=95%), fibrotic sequelae were estimated in 29% (0.29; 95% CI 0.22 to 0.37; I²=94.1%). Follow-up time was significantly associated with estimates of inflammatory sequelae (−0.036; 95% CI −0.068 to −0.004; p=0.029), associations with fibrotic sequelae did not reach significance (−0.021; 95% CI −0.051 to 0.009; p=0.176). Impaired gas transfer was estimated at 38% of lung function tests (0.38 95% CI 0.32 to 0.44; I²=92.1%), which was greater than restrictive impairment (0.17; 95% CI 0.13 to 0.23; I²=92.5%), neither were associated with follow-up time (p=0.207; p=0.864).

Discussion Sequelae consistent with parenchymal lung disease were observed following COVID-19 and other viral pneumonitis. Estimates should be interpreted with caution due to high heterogeneity, differences in study casemix and initial severity.

PROSPERO registration number CRD42020183139.

Key messages

What is the key question?
⇒ What proportion of chest CT scans and pulmonary function tests are consistent with parenchymal lung disease in the follow-up of people hospitalised with COVID-19?

What is the bottom line?
⇒ A substantial proportion of respiratory symptoms following hospitalisation with COVID-19 or other viral pneumonitis could be related to the development of lung fibrosis, but high heterogeneity in estimates should be interpreted with caution.

Why read on?
⇒ We include meta-analysis of 46 studies evaluating radiological changes, inflammatory or fibrotic, and 50 studies of lung function, impaired gas transfer or restrictive impairment, including sensitivity analysis, comparisons with findings during hospitalisation and meta-regression to explore heterogeneity.

INTRODUCTION

Since COVID-19, the disease caused by Severe Acute Respiratory Syndrome (SARS)-CoV-2, was declared a global pandemic, over 280 million individuals have been infected (December 2021). The clinical spectrum of COVID-19 is wide, and can range from asymptomatic or mild flu-like symptoms, to severe viral pneumonia, requiring hospital admission, oxygen administration and mechanical ventilation. Emerging data suggest that approximately half of COVID-19 survivors experience a long-term multisystemic syndrome characterised by chronic breathlessness and chronicity of symptoms, particularly following hospitalisation. The causes for the persistent respiratory symptoms have not been clearly elucidated, however, post-mortem studies on COVID-19 patients have highlighted diffuse parenchymal alterations, with alveolar damage, exudation and development of pulmonary fibrosis.

Pulmonary fibrosis is characterised by a dysregulated remodelling of the lung parenchyma. It can occur after a lung injury, although the cause cannot always be identified. Viral agents are considered important insults, with scientific rationale to implicate their role in fibrosis pathogenesis, although empirical evidence that suggests they can promote parenchymal lung disease is limited. Fibrotic lung sequelae have been highlighted in the follow-up of SARS-CoV and Middle Eastern Respiratory Syndrome (MERS)-CoV. Similarly, influenza viruses have also been proposed to promote the development of pulmonary fibrosis.
Given the exceptional rate of COVID-19 spread and the longer-term impact on quality of life, particularly breathlessness, it is possible that lung fibrosis may be a long-term consequence in survivors. We undertook a systematic review and meta-analysis to assess the prevalence of lung sequelae in people hospitalised with viral pneumonitis, focusing on CT scans and pulmonary function tests as non-invasive diagnostic exams routinely used.17–18

METHODS
Search strategy and selection criteria
The review has been reported following Preferred Reporting Items in Systematic review and Meta-Analysis (PRISMA) and population, intervention, comparison, and outcome (PICO) guidelines.19–20

All original research reporting outcomes in populations of hospitalised adult patients (aged ≥18) with presumed or confirmed viral infection by SARS-CoV-2, SARS-CoV, MERS-CoV or influenza viruses were considered eligible for inclusion. No intervention was assessed relative to a control group. Comparisons were made between radiological sequelae types and metrics of lung function impairment, and compared with findings during hospitalisation where available. The prespecified primary outcomes within 12 months of hospitalisation were: (1) presence of radiological sequelae at follow-up CT scans; (2) presence of restrictive lung function impairment and (3) presence of reduced diffusing capacity for carbon monoxide (DL\textsubscript{CO}). Inflammatory radiological findings were defined as ground glass opacification or consolidation. Radiological patterns suggestive of fibrosis were defined as either reticulation, lung architectural distortion, interlobular septal thickening, traction bronchiectasis or honeycombing. Restrictive lung impairment was defined as a total lung capacity (TLC) <80% predicted value or forced vital capacity (FVC) <80% predicted value with normal-to-high forced expiratory volume in 1 s/FVC ratio. Impaired gas transfer was defined as percent predicted DL\textsubscript{CO} <80%.

Searches were performed in MEDLINE (1946 to latest), Embase (1974 to latest) and Google Scholar. Handsearches were conducted of the reference lists of eligible primary studies and relevant review articles. No language criteria were applied. Preprints, abstracts and non-original studies were excluded. Searches were last updated on 29 July 2021. Searches were carried out using patient-related, treatment-related and outcomes-related terms (online supplementary figure 1). Titles and abstracts were screened in duplicate, followed by full-text review. Disagreements between reviewers were resolved by consensus with a third reviewer.

Data analysis
Data from the selected articles were extracted independently using a proforma by reviewers and mutually confirmed. Extracted data included study design, viral agent, methods of diagnosis, participant demographics, severity of acute infection (ventilatory requirements), as well as CT and lung function outcomes. Baseline investigations were defined as those performed during hospitalisation, and follow-up as obtained after discharge; baseline data were only extracted where studies reported follow-up. If more than one follow-up visit was reported, the most complete sample size followed by the latest examination within 12 months from discharge was extracted in a hierarchical manner. Where data were not reported in the text, we contacted corresponding authors. Absolute values of the number of people meeting outcome criteria and number of people with exam results available were extracted as numerator and denominator, respectively.

Meta-analyses of proportions were performed where sufficient studies reported data, enabling an estimation of the prevalence of outcomes. Cohorts with fewer than ten cases (SARS-CoV, influenza) or 25 cases (SARS-CoV-2) were excluded from quantitative synthesis owing to risk of selection bias when estimating proportions. Separate analyses were performed in each viral subtype (SARS-CoV-2, SARS-CoV, influenza) and according to the type of radiological (suggestive inflammatory and fibrotic patterning) or physiological (restrictive impairment, impaired gas transfer) outcome. Quantitative synthesis and random effects meta-analysis were performed in Stata SE V.16 (StataCorp) using the metaprop command, which computes 95% CIs based on binomial distribution and applies the Freeman-Tukey double arcsine transformation to support inclusion of observations of 0% and 100%.21 Heterogeneity was assessed with I\textsuperscript{2}; we report all estimates regardless of heterogeneity.

Meta-regression was performed where there were sufficient studies of a viral strain (n≥10). For SARS-CoV-2 studies, meta-regression was performed to assess the associations with key study characteristics, timing of follow-up (months), severity of cohort (mild, moderate, severe), prospective design, evidence of selection bias (strict inclusion criteria based on indication for CT or where less than 60% of screened patients tested for outcomes), and approach to radiological classification (study author defined, or by review). Residual heterogeneity is assessed with I\textsuperscript{2}, R\textsuperscript{2} is used to describe the variance in estimate explained by adjusted models. Reliability of estimates was assessed through sensitivity analysis in a restricted timeframe of 3–6 months follow-up, and in subanalysis on studies that reported baseline quantifications for population-based summary estimates of change.

The risk of bias in individual studies and overall quality of evidence were assessed by two authors independently. Any disagreements were resolved by consensus with a third reviewer. The risk of bias assessment followed the appropriate tools available from the CLARITY Group at McMaster University, through criteria specific for study design. We assessed exposure, the outcomes of interest, prognostic factors, interventions, adequacy of follow-up and cointerventions. Randomised controlled trials were evaluated on random sequence generation, allocation concealment, blinding, adequacy of follow-up, selective reporting and other possible causes of risks of bias.

The quality of the evidence for each overall estimate of proportion was evaluated using the GRADE guidance.23 Observational studies were considered very low but could be upgraded. Analytical and publication risks of bias, inconsistency, indirectness, and imprecision in reporting were assessed. An overall judgement of ‘high’, ‘moderate’, ‘low’, or ‘very low’ was provided for the quality of the cumulative evidence for review outcomes.

RESULTS
A total of 8321 records were identified from databases and hand searches. After title and abstract screening, 131 unique full-text manuscripts were assessed for eligibility, and 95 were included for qualitative synthesis (89 in English, 6 in Chinese). A total of 70 studies were included in the quantitative synthesis (figure 1). Among the manuscripts included, 60 reported infections by SARS-CoV-214–23; 18 by SARS-CoV\textsuperscript{13} 14 84–100; 1 by MERS-CoV\textsuperscript{101}; 16 by Influenza (11 subtype H1N1, 1 subtype H5N1, 1 subtype H3N2, 2 subtype H7N9 and 1 study both H1N1 and H7N9), 102–107 All studies were observational in design, with the exception of a single randomised control trial.108
We focus reporting on changes subsequent to a SARS-CoV-2 infection, quantitative synthesis for SARS-CoV and influenza are provided in online supplemental material.

Individual SARS-CoV-2 study characteristics are presented in Table 1 and online supplemental table 1. Risk of bias assessment identified a number of limitations and possible causes of biases (online supplemental tables 2 and 3). Five studies did not specify whether any serological or molecular testing was performed and seventeen referred to local guidelines at the time the study was conducted. Inclusion and exclusion criteria differed among studies, indicating that the severity of patients enrolled, and care pathways followed may represent a possible selection bias. Few studies investigated the presence of previous respiratory diseases or considered it as an exclusion criteria, others were restricted to include only symptomatic patients or perform follow-up CT where there was a clinical indication, such as abnormalities on chest X-ray or reduced DLco. Details for all the studies are presented in online supplemental tables 1–3; online supplemental figure 2A,B.

A total of 70 studies described thoracic CT findings, 46 were included in meta-analysis of radiological sequelae of SARS-CoV-2. Causes of exclusion are listed in Figure 1. The median follow-up time was 3 months. Within 12 months following hospitalisation for SARS-CoV-2 infection, the overall estimated proportion of chest CT inflammatory changes was 0.50 (95% CI 0.41 to 0.58; I²=95.0%) on a total of 2670 CT scans, while radiological changes suggestive of fibrosis were estimated at a proportion of 0.29 (95% CI 0.22 to 0.37; I²=94.1%) on 2811 exams. Severe heterogeneity was observed in overall estimates (Figure 2).

Lung function sequelae were described in a total of 64 papers, with 50 reaching sample size criteria for inclusion in quantitative synthesis. A total of 3146 tests for restrictive impairment and 3419 for impaired DLco were included following SARS-CoV-2 infection. Follow-up lung function tests were performed at a median of 3 months after discharge. The estimated proportion of individual tests with impaired gas transfer during follow-up was 0.38 (95% CI 0.32 to 0.44; I²=92.1%), while the estimated proportion with restricted impairment was 0.17 (95% CI 0.13 to 0.23; I²=92.5%) (Figure 3). Estimates were similar when restricted to the 3–6 months subgroup, with lower heterogeneity (0.14; 95% CI 0.10 to 0.19; I²=86.6%).

In meta-regression, adjustment for timing of follow-up was significantly associated with the overall estimate of inflammatory changes (−0.036; 95% CI −0.068 to −0.004; p=0.029) and explained 14.7% of the variance in effect (Figure 4A, online supplemental table 4). Adjustment for timing of follow-up was not significantly associated with estimates of changes suggestive of fibrosis (−0.021; 95% CI −0.051 to 0.009; p=0.176), explaining 4.9% of the variance in effect (Figure 4B, online supplemental figure 2A,B).
Table 1  SARS-CoV-2 studies overview

| Author(s)       | Year | Study design | Sample size | Age reporting (years) | FU | Severity | Selection bias | Quantitative synthesis |
|-----------------|------|--------------|-------------|-----------------------|----|----------|-----------------|------------------------|
| Anastasio et al | 2021 | P Cohort     | 222         | Median +IQR 58(53–67) | 4  | 1        | 0               | d,r                    |
| Arnold et al    | 2020 | P Cohort     | 110         | Median +IQR 60 (46–73) | 3  | 1        | 0               | r                      |
| Barisione et al | 2021 | P Cohort     | 94          | Mean+SD 61 (12.1)     | 1  | 1        | 0               | l,t,d                  |
| Bellan et al    | 2021 | P Cohort     | 238         | Median +IQR 61 (50–71) | 4  | 1        | 1               | d,r                    |
| Boari et al     | 2021 | P Cohort     | 94          | Mean+SD 66 (11)       | 4  | 1        | 1               | f,d                    |
| Bonnesen et al  | 2021 | P Cohort     | 12          | Median +IQR 62 (57–67) | 3  | 2        | 1               |                        |
| Cao et al       | 2021 | P Cohort     | 81          | Mean+SD 45 (15)       | 3  | 1        | 0               | l,t,r                  |
| Crisafulli et al| 2021 | P Cohort     | 81          | Mean+SD 66.5 (11.2)   | 4  | 1        | 0               | d,r                    |
| Daher et al     | 2020 | P Cohort     | 33          | Mean+SD 64 (3)        | 1.5| 0        | 0               | d,r                    |
| de Graaf et al  | 2021 | P Cohort     | 113         | Mean+SD 59(27–82)     | 4  | 2        | 1               | d,r                    |
| Ekbom et al     | 2021 | P Cohort     | 12          | Median +IQR 62.5 (51–71) | 3  | 1        | 0               | l,t,d,r                |
| Finney et al    | 2021 | P Cohort     | 134         | Median +IQR 60 (48–65) | 3  | 2        | 1               | l,t,d,r                |
| Gianella et al  | 2021 | P Cohort     | 114         | Mean+SD 54 (12)       | 6  | 2        | 1               | l,t,d                  |
| Huang et al     | 2021 | P Cohort     | 1733        | Median +IQR 54.5 (44–59)| 1.5| 2        | 1               | –                      |
| Huang et al     | 2020 | P Cohort     | 137         | Median +IQR 59 (50–68)| 3  | 1        | 1               | l,t,d,r                |
| Labarca et al   | 2021 | P Cross-Sectional | 103    | Mean +sd 64.3 (13.78)| 1  | 1        | 1               | l,t,d                  |
| Lago et al      | 2021 | P Cross-Sectional | 4       | Median +sd 43.6 (17.4)| 6  | 1        | 1               | l,t,d,r                |
| Li et al        | 2020 | R Cohort     | 53          | Mean+SD 50.2 (15.2)   | 8  | 1        | 0               | –                      |
| Li et al        | 2021 | P Cohort     | 289         | Mean+SD 43.6 (17.4)   | 6  | 1        | 1               | l,t,d,r                |
| Li et al        | 2020 | P Cohort     | 76          | Mean+SR 41.3 (13.38)  | 3  | 1        | 1               | d,r                    |
| Liu et al       | 2020 | R Cohort     | 51          | Mean+SR 46.6 (13.9)   | 2  | NA       | 0               | f                      |
| Liu et al       | 2021 | P Cohort     | 41          | Mean+SD 50(14)        | 7  | 1        | 0               | f                      |
| Liu et al       | 2020 | P Cohort     | 149         | Mean+IQR 43 (36–56)   | 1  | 1        | 0               | f                      |
| Liu et al       | 2020 | R Cohort     | 99          | Mean+SD 56.13 (20.7)  | 2  | 1        | 1               | –                      |
| Lombardi et al  | 2021 | P Cohort     | 86          | Mean+SD 58 (13)       | 1  | 1        | 1               | d,r                    |
| Lv et al        | 2021 | P Cohort     | 137         | Mean+SD 47 (13)       | 0.5| 1        | 0               | –                      |
| McGruder et al  | 2021 | P Cohort     | 76          | Mean+SD 54 (13.7)     | 4  | 1        | 1               | f                      |
| Miwa et al      | 2021 | R Case series | 17       | Median +IQR 63 (59–67)| 3  | 2        | 1               | –                      |
| Morin et al     | 2021 | P Cohort     | 177         | Mean+SD 56.9 (13.2)   | 4  | 1        | 1               | l,t,d                  |
| Myall et al     | 2021 | P Cohort     | 325         | Mean+SD 60.5 (10.7)   | 1.5| 1        | 1               | –                      |
| Noel-Savina et al| 2021 | P Cohort   | 72          | Mean+SD 60.5 (12.8)   | 4  | 1        | 0               | l,t,d,r                |
| Núñez-fernández et al | 2021 | P Cohort | 225         | Median +IQR 62 (50–71)| 3  | 1        | 0               | d,r                    |
| Polese et al    | 2021 | P Cohort     | 41          | Mean+SD 51(14)        | 1  | 2        | 1               | –                      |
| Qin et al       | 2021 | P Cohort     | 81          | Mean+SD 59 (14)       | 3  | 1        | 1               | l,t,d,r                |
| Raman et al     | 2021 | P Cohort     | 58          | Mean+SD 55.4 (13.2)   | 3  | 1        | 0               | r                      |
| Raman et al     | 2021 | P Case series | 28        | Mean+SD 55.5 (11.9)   | 1.5| 2        | 0               | d,r                    |
| Santus et al    | 2021 | P Cohort     | 20          | Mean+SD 58.3 (15.5)   | 1.5| 1        | 0               | –                      |
| Schandl et al   | 2021 | P Cohort     | 113         | Mean+SD 58 (12.8)     | 6  | 2        | 1               | d,r                    |
| Shah et al      | 2020 | P Cohort     | 60          | Median +IQR 67 (54–74)| 3  | 1        | 0               | l,t,d,r                |
| Sibila et al    | 2021 | P Cohort     | 172         | Mean+SD 56.1 (19.8)   | 3  | 1        | 0               | d,r                    |
| Smet et al      | 2021 | P Cross-Sectional | 220    | Mean+SD 53 (13)       | 1.5| 1        | 0               | l,d,r                  |
supplemental table 5). No other characteristics were observed to be significantly associated with proportion of CT changes, including severity of cohort, prospective design, risk of selection bias or method of radiological classification.

Within a sensitivity analysis restricted to between 3 and 6 months follow-up, we observed similar estimated effects and associations in meta-regression (online supplemental figure 3, online supplemental tables 4 and 5). The estimated proportion of chest CT inflammatory changes restricted to this subgroup was 0.49 (95% CI 0.39 to 0.59, I²=93.6%), while timing, prospective design and the severity of the cohort contributed to variance in the estimated effect: R² 9.3%, 11.7% and 2.6%, respectively (online supplemental table 4). The estimated proportion of radiological change suggestive of fibrosis in this subgroup online supplemental tables 4 and 5). The estimated proportion of chest CT inflammatory changes restricted to this subgroup was 0.49 (95% CI 0.39 to 0.59, I²=93.6%), while timing, prospective design and the severity of the cohort contributed to variance in the estimated effect: R² 9.3%, 11.7% and 2.6%, respectively (online supplemental table 4). The estimated proportion of radiological change suggestive of fibrosis in this subgroup.

**Table 1** Continued

| Study | Year | Study design | Sample size | Age reporting (years) | FU | Severity | Selection bias | Quantitative synthesis |
|-------|------|--------------|-------------|-----------------------|----|----------|----------------|------------------------|
| Fabbri et al | 2023 | | | | | | | |

**Figure 2** Radiological findings at follow-up in SARS-CoV-2 studies. Estimates are reported as proportion of CT scans showing the outcome of interest (n) on the total number of exams performed (N) and 95% CI. Inflammatory radiological findings were defined as ground glass opacification or consolidation. Fibrotic radiological findings were defined as either reticulation, lung architectural distortion, interlobular septal thickening, traction bronchiectasis or honeycombing.

Fabbri L, et al. Thorax 2023;78:191–201. doi:10.1136/thoraxjnl-2021-218275
was 0.34 (95% CI 0.25 to 0.43; I²=93.3%), timing explained 21.0% of variance in this subgroup, while risk of selection bias and approach to radiological classification also contributed to variance in the estimate: R² 21.1% and 2.9%, respectively (online supplemental table 5). The lowest unadjusted heterogeneity in estimate was observed at the 4-month follow-up, where inflammatory changes were estimated at a proportion of 0.53 (95% CI 0.41 to 0.64, I²=81.4%) while radiological changes

Figure 3  Pulmonary function testing at follow-up in SARS-CoV-2 studies. Estimates are reported as proportion of tests showing the outcome of interest (n) on the total number of exams performed (N) and 95% CI. Restrictive lung impairment was defined as a total lung capacity <80% predicted value or forced vital capacity (FVC) <80% predicted value with normal-to-high FEV1/FVC ratio. Impaired gas transfer was defined as percent predicted DLCO<80%. DLCO, diffusing capacity for carbon monoxide; FEV1, forced expiratory volume in 1 s.

Figure 4  Bubble plots of the association between follow-up time and proportion in meta-regression. follow-up time reported in months. (A) Meta-regression bubble plot of estimated proportion of inflammatory changes on thoracic CT, −0.036 (95% CI −0.068 to −0.004, p=0.029). (B) Meta-regression bubble plot of estimated proportion of changes suggestive of fibrosis on thoracic CT, −0.021 (95% CI −0.051 to 0.009, p=0.176). (C) Meta-regression bubble plot of estimated proportion of impaired gas transfer in lung function (LF) tests, −0.018 (95% CI −0.046 to 0.010, p=0.207). (D) Meta-regression bubble plot of estimated proportion of restrictive impairment in LF tests, −0.002 (95% CI −0.031 to 0.026, p=0.864).
suggestive of fibrosis were estimated at 0.32 (95% CI 0.22 to 0.43; $I^2=84.9\%$) (online supplemental figures 4 and 5).

In subanalysis of studies that reported baseline CT outcomes, estimates of inflammatory changes were 0.92 (95% CI 0.87 to 0.96; $I^2=89.4\%$) at baseline, 0.44 (95% CI 0.35 to 0.53; $I^2=89.3\%$) at follow-up, resulting in an estimated difference in proportion of $-0.47 (95\% \text{ CI } -0.56 \text{ to } -0.37; I^2=87.8\%)$ over time (figure 5A, online supplemental figure 4). Estimates of changes suggestive of fibrosis at baseline were 0.32 (95% CI 0.15 to 0.52; $I^2=98.0\%$) and 0.26 (95% CI 0.17 to 0.36; $I^2=92.9\%$) at follow-up, with an estimated difference in proportion of $-0.09 (95\% \text{ CI } -0.25 \text{ to } 0.07; I^2=96.4\%)$ over time (figure 5B, online supplemental figure 5). Timing of follow-up was not significantly associated with estimates in sub analysis of matched cohorts, the residual heterogeneity of estimates of difference in inflammatory changes was 57%, while residual heterogeneity in estimates of differences in fibrotic changes was 58.3% (online supplemental tables 4 and 5; online supplemental figures 4 and 5). Prospective design contributed 5.3% of variance in estimates of inflammatory changes, while selection bias explained 8.3% of variance in estimates of fibrotic changes.

In meta-regression of lung function estimates, adjustment for timing of follow-up was not significantly associated with impaired gas transfer ($-0.018; 95\% \text{ CI } -0.046 \text{ to } 0.010; p=0.207$) or restrictive impairment ($-0.002; 95\% \text{ CI } -0.031 \text{ to } 0.026; p=0.864$) (figure 4C,D, online supplemental tables 6 and 7). No significant associations were observed in meta-regression of lung function estimates, although differences in severity of the cohorts explained 35.3% of the variance in estimated effect in 3–6 months follow-up sensitivity analysis.

In separate viral agent strata, the estimated proportion of patients with inflammatory changes during follow-up CT scans was 0.81 (95% CI 0.58 to 0.97; $I^2=91.8\%$), and 0.61 (95% CI 0.27 to 0.90; $I^2=93.3\%$) following SARS-CoV and Influenza infections, respectively. The overall estimate of radiological change suggestive of fibrosis during follow-up was 0.66 (95% CI 0.43 to 0.86; $I^2=92.8\%$) and 0.27 (95% CI 0.15 to 0.40; $I^2=57.1\%$) following SARS-CoV and Influenza infections, respectively (online supplemental figure 6). Estimates of the proportion of restrictive impairment on lung function tests were low across other viral pneumonias, 0.10 (95% CI 0.05 to 0.17; $I^2=80.2\%$) for SARS-CoV and in 6/73 participants with MERS-CoV (online supplemental figure 7). Estimates of the proportion of impaired gas transfer on tests were similar in SARS-CoV compared with SARS-CoV-2 (0.36; 95% CI 0.27 to 0.46; $I^2=84.4\%$), while estimates were higher following influenza (0.54; 95% CI 0.43 to 0.63), and a single study of MERS-CoV identified gas transfer impairments in 25/73 participants.

Based on the GRADE framework, we have low confidence in estimates for all outcomes. All studies included in the quantitative synthesis had an observational design. Risk of bias was low to moderate as possible confounding factors were not extensively assessed and could not be modelled in estimates of proportion. Inconsistency between studies was considered serious due to the substantial heterogeneity that could be only partially reduced by adjustment for timing. No causes of indirectness were detected since all study subjects had confirmed viral pneumonia, although severity and eligibility criteria were inconsistent. We judged the risk of imprecision as moderate, due to the possible influence of sample size on proportion. Risk of publication bias evaluation identified symmetry and very low risk of bias in funnel plots (online supplemental table 8; online supplemental figures 8 and 9).

**DISCUSSION**

We systematically investigated the prevalence of radiological and functional sequelae post-hospitalisation for viral pneumonitis, particularly for that caused by SARS-CoV-2. Within 12 months of hospitalisation, radiological patterns of inflammation were estimated in 50% of scans during follow-up, while changes suggestive of fibrosis were estimated in 29% of scans. In studies with matched baseline scans during hospitalisation we estimated inflammatory changes in over 90% of CT scans, which reduced to 44% at a median follow-up of 3 months, with timing of follow-up strongly associated with estimates across all studies. In contrast, radiological changes suggestive of fibrosis were estimated in a smaller percentage of CT scans of matched follow-up (26%), though proportions remained similar to hospitalisation and follow-up timing was not significantly associated with estimates, suggesting a more persistent change. In analyses of lung function across all follow-up, impaired gas transfer
was estimated in 38% of tests and showed a similar association with follow-up time as radiological change. Restrictive impairment was estimated in 17% of tests and was not associated with follow-up timing. Heterogeneity in overall estimates were frequently substantial and therefore results should be interpreted with caution. We demonstrate that parenchymal lung damage by viral insult may be common and has the potential to explain COVID-19-related respiratory symptoms in the months following hospitalisation.

A high proportion of people with inflammatory findings such as ground glass opacities and consolidation were observed at baseline following SARS-CoV-2, consistent with the radiological signs commonly described for viral pneumonitis. The difference in inflammatory changes reduced over the course of matched follow-up, as we would expect with the resolution of the acute inflammation. However, radiological changes suggestive of fibrosis were observed in a similar proportion of people during hospitalisation and at follow-up, suggesting a potential lack of resolution, also demonstrated by a single study comparing CT scans at 6 and 12 months in which fibrotic abnormalities and traction bronchiectasis did not improve.

Meta-regression indicated that estimates of radiological sequelae reduced over time, particularly for inflammatory changes and more slowly for fibrotic changes, supporting the hypothesis that parenchymal abnormalities observed after infection may lead to long-term sequelae. Radiological and functional sequelae were estimated in approximately 20% of cases at 12 months in meta-regression of time and outcomes, and have been described up to 5 years after Influenza infections, and up to 15 years after SARS-CoV.

In individuals with SARS-CoV-2, restrictive and gas transfer impairment were associated with infection severity, with similar findings reported in SARS-CoV, although not always statistically significant. We observe that the estimated prevalence of impaired gas transfer is greater than the prevalence of restrictive impairment following SARS-CoV-2 infection, with similar findings following other viral pneumonias. Meta-regression suggested that estimates of impaired gas transfer reduced over time, while the lower estimates of restrictive impairment did not change. Unresolved radiological changes and impaired lung function are important signs suggestive of fibrotic interstitial lung disease, and prospective studies should accurately define the prevalence of post-COVID pulmonary fibrosis.

Other systematic reviews have been published addressing radiological changes on CT and impairment to lung function in response to COVID-19, often limited to smaller numbers of studies, shorter follow-up, qualitative review alone or lack of a preregistered protocol. We included over 40 studies in quantitative synthesis of each radiological and physiological sequelae based on a preregistered protocol, including up to 12 months of follow-up, representing the largest systematic review and meta-analysis. High levels of heterogeneity are routinely reported in meta-analysis of proportions, so we perform sensitivity analysis, subanalysis and meta-regression to provide further reliable insights. We additionally model potential sources of heterogeneity in meta-regression, identifying timing of follow-up as an important characteristic to interpret estimates. A high risk of selection bias commonly contributed to variance in fibrotic estimates, while prospective design more commonly contributed to variance in inflammatory estimates, both of which highlight the impact of study inclusion criteria on generalisability of systematic review findings. Unique to our protocol, we separately report estimates from Influenza and SARS-CoV studies, which suggest similar changes in response to non-COVID-19 viral pneumonitis.

There are limitations to this systematic review and meta-analysis. As our search strategy focused on follow-up tests, the number of included articles that reported baseline findings were limited, and no studies included CT findings prior to hospitalisation. Similarly, we cannot exclude that all functional impairments were caused by the infections rather than underlying respiratory conditions, however, study criterion often excluded patients with known history of pulmonary disease. Estimates of proportion are based on the number of tests performed, not patients infected, which would be affected by selection bias toward symptomatic patients as well as lost to-follow-up. We assess for selection bias according to large discrepancies between screened and included numbers of participants, and also demonstrate minimal lost to-follow-up in subanalyses of studies with matched time points. Interpreting estimates requires caution as heterogeneity was frequently substantial and not completely attributable to the study-level features evaluated, consistent reasons for outlying study estimates were not identified. We observed that overall estimates of radiological patterns suggestive of fibrosis were consistent in sensitivity analyses restricted to 3–6 months of follow-up, with lowest unadjusted heterogeneity observed at 4 months, suggesting similar timeframes may be suitable for radiological follow-up. It is likely that variability in casemix demographic and severity of acute infection contributed to the heterogeneity between studies, which may be further addressed by individual patient data approaches. All estimates represent individuals hospitalised with infection, which may not reflect prevalence in non-hospitalised cases. We defined radiological sequelae attributable to inflammatory and fibrotic changes, however, these were not always reported specifically or exclusively and there are limitations to classifying radiological patterns. Ground glass opacities are not exclusive to inflammation, and could reflect retractile fibrosis during follow-up, but are frequently consistent with inflammation or atelectasis in the acute period. Approach to radiological classification only explained minor variance in fibrotic estimates, specific patterning likely contributes to residual heterogeneity. Internationally standardised approaches to reporting of post-COVID radiological change would support patient management and epidemiological study. Similarly, we acknowledge the limits of diagnosing restrictive impairment without TLC measures, where some results may represent pseudorestriction or mixed pattern.

We have demonstrated the presence of substantial radiological and functional sequelae following viral pneumonias that may be consistent with postviral interstitial lung disease. These parenchymal sequelae of viral infection could have a considerable impact given the large numbers of people discharged from hospital with COVID-19. While the certainty of the presented estimates is low, they justify vigilant radiological and functional follow-up of individuals hospitalised with viral pneumonia.

Twitter Laura Fabbri @Istamina

Contributors All authors contributed to either the conception, design, data acquisition, data analysis, or interpretation of the work; all authors contributed to the drafting and revising of the manuscript and all authors gave final approval for publication. IS accepts full responsibility for the work and conduct of the study, had access to the data, and controlled the decision to publish.

Funding National Institute for Health Research (NIHR) Professorship (RP-2017-08-ST2-014); Rayne Foundation Fellowship.

Competing interests GJ reports NIHR BRC salaries, studentships, professorship (RP-2017-08-ST2-014).

Patient consent for publication Not applicable.
Intestinal lung disease

Ethics approval This study does not involve human participants. Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data sharing not applicable as no datasets generated and/or analysed for this study. Data were obtained from published studies. This article is made freely available for personal use in accordance with BMJ’s website terms and conditions for the duration of the covid-19 pandemic or until otherwise determined by BMJ. You may download and print the article for any lawful, non-commercial purpose (including text and data mining) provided that all copyright notices and trade marks are retained.

ORCID iDs
Laura Fabbri http://orcid.org/0000-0002-8250-6464.
Fasihul A Khan http://orcid.org/0000-0002-0796-5724.
Alan Robert Smyth http://orcid.org/0000-0001-5494-5438.
Gisli Jenkins http://orcid.org/0000-0002-1340-2688.

REFERENCES
1. WHO. General’s opening remarks at the media briefing on COVID-19, 2020.
2. WHO. Coronavirus disease (COVID-19). Available: https://covid19.who.int/.
3. Wielopolski L, Colombo P, Fabbri L, et al. Thorax 2023;78:191–201. doi:10.1136/thoraxjnl-2021-218275.
4. Granholm A, Alhazzani W, Miller MH. Use of the grade approach in systematic reviews and guidelines. Br J Anaesth 2013;110:554–9.
Interstitial lung disease

53 Liu X, Zhou H, Zhou Y et al. Temporal radiographic changes in COVID-19 patients: relationship to disease severity and viral clearance. Sci Rep 2020;10:10263.

54 Lombardi F, Calabrese A, Iovene B et al. Residual respiratory impairment after COVID-19 pneumonia. BMC Pulm Med 2021;21:241.

55 Lv D, Chen X, Wang X et al. Pulmonary function of patients with 2019 novel coronavirus induced pneumonia: a retrospective cohort study. Ann Palliat Med 2020;9:3447–52.

56 McGroder CF, Zhang D, Choudhury MA et al. Pulmonary fibrosis 4 months after COVID-19 is associated with severity of illness and blood leucocyte telomere length. Thorax 2021;76:1242–5.

57 Miwa M, Nakajima M, Kaszynski RH et al. Abnormal pulmonary function and imaging studies in critical COVID-19 survivors at 100 days after the onset of symptoms. Respir Investig 2021;59:614–621.

58 Writing Committee for the COMEBAC Study Group, Morin L, Savale L et al. Four Month clinical status of a cohort of patients after hospitalization for COVID-19. JAMA 2021;325:1525–34.

59 Myall KI, Mukherjee B, Castanheira AM et al. Persistent Post-COVID-19 interstitial lung disease: An observational study of corticosteroid treatment. Ann Am Thorac Soc 2021;18:789–806.

60 Noel-Savina E, Viatte T, Faviez G et al. Severe SARS-CoV-2 pneumonia: Clinical, functional and imaging outcomes at 4 months. Respir Med Res 2021;80:100822.

61 Núñez-Fernández M, Ramos-Hernández C, García-Río F, Qin W, Chen S, Zhang Y et al. Functional and imaging outcomes at 4 months. Eur Respir J 2021;60:2000366. [Epub ahead of print: 22 07 2021].

62 Zhao Y-M, Shang Y-M, Song W-B, et al. Follow-Up study of the pulmonary function and related physiological characteristics of COVID-19 survivors three months after recovery. EclinicalMedicine 2020;25:100463.

63 Zhong L, Zhang S, Wang J, et al. Analysis of chest CT results of coronavirus disease 2019 (COVID-19) patients at first follow-up. Can Respir J 2020;2020:5328267.

64 Antonio GE, Wong KT, Hui DSC et al. Thin-section computed tomography findings in non-severe discharged COVID-19 pneumonia patients after discharge and follow-up chest CT findings. Acta Anaesthesiol Scand 2023;76:thoraxjnl-2023-002865.

65 Ramani C, Davis EM, Kim JS et al. Temporal radiographic changes in COVID-19 patients: a prospective, observational cohort study. J Thorac Med 2021;10:1–17.

66 Schandl A, Hedman A, Lyngå P et al. Long-term consequences in critically ill COVID-19 patients: a prospective cohort study. Acta Anaesthesiol Scand 2021;65:1285–92.

67 Shah AS, Wong AW, Hague CJ et al. Follow-up analysis of pulmonary function and radiological abnormalities in COVID-19 patients: a prospective, observational cohort study. J Clin Med 2021;10:1–17.

68 Wei J, Yang H, Lei P et al. Interstitial lung disease in discharged patients with severe COVID-19. Lancet Respir Med 2021;9:747–754.

69 Yasin R, Gomaa AAK, Ghazy T et al. Predicting lung fibrosis in post-COVID-19 patients after discharge with follow-up chest CT findings. Egyptian Journal of Radiology and Nuclear Medicine 2021;62:103280.

70 Yu M, Liu Y, Xu D et al. Comparison of patients with avian influenza A(H7N9) and influenza A(H1N1) complicated by acute respiratory distress syndrome. Medicine 2018;97:e0194.
Interstitial lung disease

108 Liu W, Peng L, Liu H, et al. Pulmonary function and clinical manifestations of patients infected with mild influenza A virus subtype H1N1: a one-year follow-up. PLoS One 2015;10:e0133698.

109 Lu P-xuan, Wang Y-xiang, Zhou B-ping, et al. Radiological features of lung changes caused by avian influenza subtype A H5N1 virus: report of two severe adult cases with regular follow-up. Chin Med J 2010;123:100–4.

110 Lust C-E, Combes A, Becquemin M-H, et al. Long-term outcomes of pandemic 2009 influenza A(H1N1)-associated severe ARDS. Chest 2012;142:583–92.

111 Mineo G, Ciccarese F, Modolon C, et al. Post-ARDS pulmonary fibrosis in patients with H1N1 pneumonia: role of follow-up CT. Radiol Med 2012;117:185–200.

112 Quispe-Laime AM, Fiore C, González-Ros MN, et al. Lung diffusion capacity and quality of life 6 months after discharge from the ICU among survivors of acute respiratory distress syndrome due to influenza A(H1N1). Med Intensiva 2012;36:15–23.

113 Saha A, Vaidya PJ, Chavhan VB, et al. Combined pirfenidone, azithromycin and prednisolone in post-H1N1 ARDS pulmonary fibrosis. Sarcoidosis Vasc Diffuse Lung Dis 2018;35:85–90.

114 Singh V, Sharma BB, Patel V. Pulmonary sequelae in a patient recovered from swine flu. Lung India 2012;29:277–9.

115 Costa ELV, Hirota AS, et al. Follow-Up after acute respiratory distress syndrome caused by influenza A (H1N1) virus infection. Clinics 2011;66:933–7.

116 Winterbauer RH, Ludwig WR, Hammar SP. Clinical course, management, and long-term sequelae of respiratory failure due to influenza viral pneumonia. Johns Hopkins Med J 1977;141:148–55.

117 Zarogoulidis P, Kouliatsis G, Papanas N, et al. Long-Term respiratory follow-up of H1N1 infection. Virol J 2011;8:319.

118 Franquet T. Imaging of pulmonary viral pneumonia. Radiology 2011;260:18–39.

119 Koo HI, Choi S-H, Sung H, et al. Radiographics Update: Radiographic and CT Features of Viral Pneumonia. Radiographics 2020;40:E8–15.

120 Han X, Fan Y, Alwaidi Q, et al. Fibrotic interstitial lung abnormalities at 1-year follow-up CT after severe COVID-19. Radiology 2021;301:E438–40.

121 Wang Q, Jiang H, Xie Y, et al. Long-term clinical prognosis of human infections with avian influenza A(H7N9) viruses in China after hospitalization. EClinicalMedicine 2020;20:100282.

122 Guo L, Han Y, Li J. Long-Term outcomes in patients with severe acute respiratory syndrome treated with oseltamivir: a 12-year longitudinal study. International Journal of Clinical and Experimental Medicine 2019;12:12464–71.

123 Zhang P, Li J, Liu H, Han N, et al. Long-Term bone and lung consequences associated with hospital-acquired severe acute respiratory syndrome: a 15-year follow-up from a prospective cohort study. Bone Res 2020;8:511.

124 Mo X, Jian W, Su Z, et al. Abnormal pulmonary function in COVID-19 patients at time of hospital discharge. Eur Respir J 2020;55:2001217.

125 Faverio P, Luppi F, Rebora P, et al. Six-Month pulmonary impairment after severe COVID-19: a prospective, multicentre follow-up study. Respiration 2021;100:1078–87.

126 Wild JM, Porter JC, Molyneaux PL, et al. Understanding the burden of interstitial lung disease post-COVID-19: the UK interstitial lung Disease-Long COVID study (UKILD-Long COVID). BMJ Open Respir Res 2021;8:e001049.

127 Ogihara K, Mani A, Pandey NN, et al. Clinical outcomes in coronavirus disease 2019 (COVID-19): a systematic review of chest CT findings in 4410 adult patients. Eur Radiol 2020;30:6129–38.

128 Polak SB, Van Gool IC, Cohen D, et al. A systematic review of pathological findings in COVID-19: a pathophysiological timeline and possible mechanisms of disease progression. Mod Pathol 2020;33:2128–38.

129 So M, Kabata H, Fukunaga K, et al. Radiological and functional lung sequelae of COVID-19: a systematic review and meta-analysis. BMC Pulm Med 2021;21:97.

130 Torres-Castro R, Vasconcello-Castillo L, Alisina-Restoy X, et al. Respiratory function in patients post-infection by COVID-19: a systematic review and meta-analysis. Pulmonology 2021;27:328–37.