Characterising long COVID: a living systematic review

Melina Michelen,1,2 Lakshmi Manoharan,2 Natalie Elkheir,3,4 Vincent Cheng,5,4 Andrew Dagens,2 Claire Hastie,5,6 Margaret O’Hara,5 Jake Suett,6,7 Dania Dahmash,2 Polina Bugaeva,7 Ishmeeal Rigby,2 Daniel Munblit,8,9,10 Eli Harriss,11 Amanda Burls,1 Carole Foote,12 Janet Scott,13 Gail Carson,1,2 Piero Olliaro,2 Louise Sigfrid,2 Charitini Stavropoulou1

ABSTRACT

Background While it is now apparent clinical sequelae (long COVID) may persist after acute COVID-19, their nature, frequency and aetiology are poorly characterised. This study aims to regularly synthesise evidence on long COVID characteristics, to help inform clinical management, rehabilitation strategies and interventional studies to improve long-term outcomes.

Methods A living systematic review. Medline, CINAHL (EBSCO), Global Health (Ovid), WHO Global Research on COVID-19 database, LitCovid and Google Scholar were searched till 17 March 2021. Studies including at least 100 people with confirmed or clinically suspected COVID-19 at 12 weeks or more post onset were included. Risk of bias was assessed using the tool produced by Hoy et al. Results were analysed using descriptive statistics and meta-analyses to estimate prevalence.

Results A total of 39 studies were included: 32 cohort, 6 cross-sectional and 1 case–control. Most showed high or moderate risk of bias. None were set in low-income countries and few included children. Studies reported on 10951 people (48% female) in 12 countries. Most included previously hospitalised people (78%, 8520/10 951). The longest mean follow-up time was 221.7 (SD: 10.9) days post COVID-19 onset. Over 60 physical and psychological signs and symptoms with wide prevalence were reported, most commonly weakness (41%; 95% CI 25% to 59%), general malaise (33%; 95% CI 24% to 39%), concentration impairment (26%; 95% CI 21% to 32%) and breathlessness (25%; 95% CI 18% to 34%) as the most common symptoms reported.

Conclusion Long COVID is a complex condition with prolonged heterogeneous symptoms. The nature of studies precludes a precise case definition or risk evaluation. There is an urgent need for prospective, robust, standardised, controlled studies into aetiology, risk factors and biomarkers to characterise long COVID in different at-risk populations and settings.

Key questions

What is already known?

► A significant number of people continue to describe ongoing symptoms long after the acute phase of COVID-19, often referred to as long COVID.

► Long COVID is a heterogeneous condition with an uncertain prevalence, for which there is currently no precise case definition.

What are the new findings?

► The breadth of reported symptoms suggests a complex, heterogeneous condition affecting both those who were hospitalised and those managed in the community.

► Our review identifies weakness (41%; 95% CI 25% to 59%), general malaise (33%; 95% CI 15% to 57%), fatigue (31%; 95% CI 24% to 39%), concentration impairment (26%; 95% CI 21% to 32%) and breathlessness (25%; 95% CI 18% to 34%) as the most common symptoms reported.

What do the new findings imply?

► The current evidence base of the clinical spectrum of long COVID is limited, based on heterogeneous data, and vulnerable to biases, hence caution should be used when interpreting or generalising the results.

► Our review identifies areas where further long COVID research is critically needed to help characterise long COVID in different populations and define its aetiology, risk factors and biomarkers, as well as the impact on variants of concern and vaccination on long-term outcomes.

INTRODUCTION

SARS-CoV-2 first emerged in December 2019 causing a widespread pandemic. Most people experience asymptomatic or mild-to-moderate acute COVID-19 symptoms, while around 15% of people are estimated to progress to more severe disease requiring hospitalisation and approximately 5% become critically ill.1

While the acute phase of the disease was characterised early, there are still limited data on long-term outcomes.2 Symptoms of long-lasting COVID-19 sequelae and complications, termed long COVID by people living with long COVID,3 have been
reported worldwide. Yet the underlying aetiology behind prolonged or fluctuating symptomatology is limited and there is no widely accepted uniformed case definition. Instead, long COVID has been defined pragmatically as ‘not recovering for several weeks or months following the start of symptoms’. Others have distinguished between postacute COVID-19, referring to symptoms beyond 3 weeks, and chronic COVID-19, referring to symptoms beyond 12 weeks, while the National Institute for Health and Care Excellence distinguishes between ongoing symptomatic COVID-19 lasting from 4 to 12 weeks and post COVID-19 syndrome continuing for over 12 weeks.

The number of people living with long COVID is unknown. Attempts to quantify the prevalence of long COVID use different methods, including national surveys and patient-led studies, making it difficult to compare across studies. The UK’s Office for National Statistics has estimated that on average 1 in 5 people have symptoms beyond 5 weeks, while 1 in 10 have symptoms persisting over 12 weeks. A patient-led survey found that in survival analysis, the chance of full recovery by day 50 was smaller than 20% and a COVID-19 symptom app study found that 13.3% (558/4182) patients had symptoms lasting 28 days or more, 4.5% (189/4182) patients had symptoms for 8 or more weeks and 2.3% (95/4182) patients had symptoms lasting over 12 weeks.

The symptoms of long COVID are equally ill-defined, with patients describing it as a fluctuating illness of disparate symptoms. Indeed, the National Institute for Health Research has suggested that postacute COVID-19 may consist of several distinct clinical syndromes including: a postintensive care syndrome, chronic fatigue syndrome, long-term COVID-19 syndrome and disease from SARS-CoV-2 inflicted organ damage. Additionally, even with an expanding knowledge of risk factors in the acute phase, little is currently known on predictive factors for developing long COVID. Despite suggested classifications, there is yet no clear consensus.

Our early understanding of long COVID has been accumulated from case reports and cross-sectional online survey studies as the pandemic global research focus has largely been on studies of hospitalised patients during the acute phase. As the pandemic progresses, emerging studies have followed up patients to present the fluctuating multiorgan sequelae of acute COVID-19, yet evidence is still scarce. There continues to be a call to further understand and acknowledge this condition by incorporating patient knowledge and experiences, together with standardised studies, exploring underlying aetiologies behind different syndromes.

Given the enormous number of people worldwide who have suffered from COVID-19, it is essential to establish a precise categorisation of long COVID. Such categorisation will not only help people better understand their symptoms but also direct research into prevention, treatment and support, ultimately allowing us to understand and prepare to respond to the long-term consequences inflicted by the COVID-19 pandemic. Our review seeks to synthesise and continually update the evidence on the character and prevalence of long COVID.

METHODS

Systematic reviews conducted early during the COVID-19 pandemic soon became redundant due to the rapidity with which new research was released. In recognition of this, many reviewers have moved towards the concept of a ‘living systematic review’ (LSR), which compared with traditional systematic reviews has in-built mechanisms for regular update and renewal. We conducted a ‘living’ systematic review to provide frequently updated evidence on the symptoms and complications of long COVID. This review was developed in collaboration with infectious disease clinicians, public health professionals, information specialists, review methodologists with experience in clinical epidemic research and members of the global Long COVID Support Group, which includes people living with long COVID. This is the first version of this LSR, which will be updated approximately every 6 months as new evidence emerges, using the established protocol and review platform. The updates will be led by the International Severe Acute Respiratory and emerging Infection Consortium (ISARIC) systematic review team in collaboration with members of Long COVID Support. Previous versions will be archived in online supplemental materials. The findings will be disseminated via BMJ Global Health and on a dedicated webpage with infographics and a brief summary for lay people and professionals.

Protocol registration

This report was structured according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement guidelines. The protocol was registered with PROSPERO and published in a peer-reviewed journal.

Search strategy

The following databases were searched: Medline and CINAHL (EBSCO), Global Health (Ovid), WHO Global Research Database on COVID-19 and LitCovid from 1 January 2020 to 17 March 2021. Additionally, we searched Google Scholar on 17 March 2021, screening the first 500 titles. A ‘backwards’ snowball search was conducted of the references of systematic reviews. Full search terms are included in online supplemental file 1. The search terms and inclusion criteria have, for this first version, been designed to cast a wide net and will be modified in line with new evidence, research priorities and clinical and policy needs.

Eligibility criteria

Peer-reviewed studies were considered eligible if they included at least 100 people with laboratory confirmed and/or clinically diagnosed COVID-19. Without a clear, internationally agreed case definition, we included studies that reported symptoms or outcomes assessed at 12 or more weeks post COVID-19 onset.
There were no language restrictions. Reviews and opinion pieces were excluded. Studies were excluded if they included fewer than 100 participants, to avoid small study effects,18 or the follow-up was unclear or less than 12 weeks post onset.

**Screening**

Screening was performed independently by two systematic reviewers. Any disagreements were resolved via consensus or a third reviewer. Non-English articles were translated using Google Translate and assessed by a systematic reviewer with good knowledge of the language. The data were managed using the review software Rayyan.19

**Data extraction**

Data extraction was performed using Microsoft Excel. A data extraction template informed by a previous review20 was reviewed, updated and piloted before being finalised. Data extracted included study design, population characteristics, outcomes, prevalence, duration of symptoms and risk factors. Data extraction was performed by one systematic reviewer and checked by a second reviewer. Disagreements were resolved through consensus. To avoid duplication of data in future updates and ensure robustness, data extraction was not performed for non-peer-reviewed preprints.

**Risk of bias assessment**

The included studies were assessed for risk of bias using the tool produced by Hoy et al21 (online supplemental file 2). This assessment checklist is a validated tool for assessing risk of bias in prevalence studies. The checklist has 10 domains for assessing risk of bias, used to calculate a cumulative overall risk of bias for the whole study.

**Data analysis**

We undertook individual descriptive analysis for each study. We presented symptom proportions by different settings, as presented in the individual studies: hospitalised, non-hospitalised or a mix of both populations if no subset data were available. Symptoms were broadly grouped into physiological clusters through discussion with clinicians. Proportion of symptoms and its 95% CIs were estimated using the exact method.22 If there were two or more studies for each symptom, a meta-analysis was performed using a random intercept logistic regression model with Hartung-Knapp modification due to the heterogeneity and skewed sample sizes.23 24 Heterogeneity between estimates was assessed using the I² statistic.25 Additional subgroup analysis was conducted to explore the modification of the following factors on proportion of symptoms: hospitalisation, settings, continents and follow-up timing. We also conducted meta-regression analysis on the percentage of females and intensive care unit (ICU) patients where there were more than 10 studies for the symptom. Sensitivity analyses were conducted to examine the impact of high risk of bias studies and statistical methods, Freeman-Tukey double arcsine transformation using inverse variance meta-analysis, on the estimates. Funnel plots were plotted using proportion of the symptom against the precision and sample sizes22 where there were more than 10 studies for the symptom to explore risk of publication bias. All analysis and data presentation were performed using metaprop26 and ggplot227 in R (V.4.0.5) via RStudio (V.1.3.1093).28 The data are presented using a combination of infographics, prepared by a design company (Design Science29) and scientific tables to facilitate interpretation by different stakeholders, including non-specialists.
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Country</th>
<th>Population size</th>
<th>Age (years)</th>
<th>Sex (% female)</th>
<th>COVID-19 confirmation method</th>
<th>Follow-up time (days)</th>
<th>Follow-up timepoint</th>
<th>Follow-up mode</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-hospitalised</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hopkins et al</td>
<td>Cross sectional</td>
<td>UK</td>
<td>434</td>
<td>Median (range): 40 (19–77)</td>
<td>75</td>
<td>PCR or serological assays (26.3%)</td>
<td>6 months</td>
<td>First survey</td>
<td>Electronic survey</td>
</tr>
<tr>
<td>Klein et al</td>
<td>Cohort (P)</td>
<td>Israel</td>
<td>103</td>
<td>Mean (SD): 35 (12)</td>
<td>38</td>
<td>PCR (RT-PCR)</td>
<td>6 months</td>
<td>Onset</td>
<td>Phone interview</td>
</tr>
<tr>
<td>Petersen et al</td>
<td>Cohort (P)</td>
<td>Faroe Islands</td>
<td>180</td>
<td>Mean (SD; range): 39.9 (19.4; 0–93)</td>
<td>54</td>
<td>PCR (RT-PCR)</td>
<td>Mean (SD) 125 (17)</td>
<td>Onset</td>
<td>Phone interview</td>
</tr>
<tr>
<td>Stavem et al</td>
<td>Crosssectional</td>
<td>Norway</td>
<td>451</td>
<td>Mean (SD): 49.8 (15.2)</td>
<td>56</td>
<td>PCR (RT-PCR)</td>
<td>Median (range): 117 (41–193)</td>
<td>Onset</td>
<td>Outpatient visit and survey</td>
</tr>
<tr>
<td><strong>Non-hospitalised and hospitalised</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parente-Arias et al</td>
<td>Cohort (P)</td>
<td>Spain</td>
<td>151</td>
<td>Mean (range): 55.2 (18–88)</td>
<td>65</td>
<td>PCR (RT-PCR)</td>
<td>Mean (SD): 100.5 (3.3) Admission</td>
<td>Phone interview</td>
<td></td>
</tr>
<tr>
<td>Venturelli et al</td>
<td>Cohort (P)</td>
<td>Italy</td>
<td>767</td>
<td>Mean (SD): 63 (13.6)</td>
<td>33</td>
<td>PCR (RT-PCR) (94%); serology (5%); Clinician diagnosis (1.2%)</td>
<td>Median (IQR): 105 (84–127)</td>
<td>Onset</td>
<td>Outpatient visit</td>
</tr>
<tr>
<td>Anastasio et al</td>
<td>Cohort (P)</td>
<td>Italy</td>
<td>379</td>
<td>Median IQR; range: 56 (49–63; 20–80)</td>
<td>54</td>
<td>PCR (RT-PCR)</td>
<td>Median IQR: 135 (102–175)</td>
<td>Onset</td>
<td>Outpatient visit</td>
</tr>
<tr>
<td>Einvik et al</td>
<td>Crosssectional</td>
<td>Norway</td>
<td>538</td>
<td>Mean (SD): 57.7 (14.2) (hospital); 49.6 (15.3)</td>
<td>42 (hospital) 56</td>
<td>PCR (RT-PCR)</td>
<td>Mean (SD): 112 (30 (hospital) 118 (27)</td>
<td>Onset</td>
<td>Outpatient visit and survey</td>
</tr>
<tr>
<td>Jacobson et al</td>
<td>Cohort (P)</td>
<td>USA</td>
<td>118</td>
<td>Mean (SD): 43.3 (14.4)</td>
<td>47</td>
<td>PCR (RT-PCR)</td>
<td>Mean (SD): 119.3 (33) Diagnosis</td>
<td>Outpatient visit</td>
<td></td>
</tr>
<tr>
<td>Logue et al</td>
<td>Cohort (P)</td>
<td>USA</td>
<td>177 (C)</td>
<td>Mean (SD): 48 (15.2)</td>
<td>57</td>
<td>Lab confirmed</td>
<td>Median (range): 169 (31–300)</td>
<td>Onset</td>
<td>Electronic survey</td>
</tr>
<tr>
<td>Mazza et al</td>
<td>Cohort (P)</td>
<td>Italy</td>
<td>226</td>
<td>Mean (SD; range): 58 (12.8; 26–87)</td>
<td>34</td>
<td>PCR (RT-PCR)</td>
<td>Mean (SD): 90 (13.4) Discharge</td>
<td>Onset</td>
<td>Phone interview</td>
</tr>
<tr>
<td>Rass et al</td>
<td>Cohort (P)</td>
<td>Austria</td>
<td>135</td>
<td>Median IQR; range: 56 (48–68; 19–87)</td>
<td>39</td>
<td>PCR (RT-PCR)</td>
<td>Median IQR: 102 (91–110)</td>
<td>Onset</td>
<td>Outpatient visit</td>
</tr>
<tr>
<td>Sonnweber et al</td>
<td>Cohort (P)</td>
<td>Austria</td>
<td>145</td>
<td>Mean (SD): 57 (14)</td>
<td>43</td>
<td>PCR (RT-PCR)</td>
<td>Mean (SD): 103 (21) Diagnosis</td>
<td>Outpatient visit</td>
<td></td>
</tr>
<tr>
<td><strong>Hospitalised</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alharthy et al</td>
<td>Cohort (P)</td>
<td>Saudi Arabia</td>
<td>127</td>
<td>Mean (SD): 47 (11.38)</td>
<td>21</td>
<td>PCR (RT-PCR)</td>
<td>4 months</td>
<td>Discharge</td>
<td>Outpatient visit</td>
</tr>
<tr>
<td>Arnold et al</td>
<td>Cohort (P)</td>
<td>UK</td>
<td>110</td>
<td>Median IQR: 60 (46–73)</td>
<td>38</td>
<td>PCR or radiological diagnosis</td>
<td>Median IQR: 90 (80–97)</td>
<td>Onset</td>
<td>Outpatient visit</td>
</tr>
<tr>
<td>Baricich et al</td>
<td>Crosssectional</td>
<td>Italy</td>
<td>204</td>
<td>Mean (SD): 57.9 (12.8)</td>
<td>40</td>
<td>NR</td>
<td>Mean (SD): 124.7 (17.5) Discharge</td>
<td>Outpatient visit</td>
<td></td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Country</th>
<th>Population size</th>
<th>Age (years)</th>
<th>Sex (% female)</th>
<th>COVID-19 confirmation method</th>
<th>Follow-up time (days)</th>
<th>Follow-up timepoint</th>
<th>Follow-up mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bellan et al&lt;sup&gt;12&lt;/sup&gt;</td>
<td>Cohort (P)</td>
<td>Italy</td>
<td>238</td>
<td>Median (IQR): 61 (50–71)</td>
<td>40</td>
<td>PCR (RT-PCR) (RT-PCR) (97.5%); bronchoalveolar lavage (0.4%); serology/radiological (2.1%)</td>
<td>3–4 months</td>
<td>Discharge</td>
<td>Onset</td>
</tr>
<tr>
<td>Blanc et al&lt;sup&gt;14&lt;/sup&gt;</td>
<td>Cohort (P)</td>
<td>Spain</td>
<td>100</td>
<td>Mean (SD) TLco&lt;90: 54.98 (10.72) TLco&gt;90: 54.75 (9.83)</td>
<td>36</td>
<td>PCR (RT-PCR) Median (IQR): 104 (89.25–126.75)</td>
<td>Onset</td>
<td>Onset</td>
<td>Outpatient visit</td>
</tr>
<tr>
<td>Doyle et al&lt;sup&gt;16&lt;/sup&gt;</td>
<td>Cohort (P)</td>
<td>UK</td>
<td>129</td>
<td>Mean: 62 (Cambridge) 56 (London)</td>
<td>31 (Cambridge) PCR (RT-PCR) Median (range): 113 (96–138)</td>
<td>Discharge</td>
<td>6 months</td>
<td>NR</td>
<td></td>
</tr>
<tr>
<td>Garrigues et al&lt;sup&gt;15&lt;/sup&gt;</td>
<td>Cohort (P)</td>
<td>France</td>
<td>120</td>
<td>Mean (SD): 63.2 (15.7)</td>
<td>38</td>
<td>PCR (RT-PCR) Mean (SD): 110.9 (11.1)</td>
<td>Admission</td>
<td>Phone interview</td>
<td>Phone interview</td>
</tr>
<tr>
<td>Gherlone et al&lt;sup&gt;17&lt;/sup&gt;</td>
<td>Cohort (P and R)</td>
<td>Italy</td>
<td>122</td>
<td>Median (IQR): 62.5 (53.9–74.1)</td>
<td>25</td>
<td>PCR (RT-PCR) Median (IQR): 104 (95–132)</td>
<td>Discharge</td>
<td>Outpatient visit</td>
<td></td>
</tr>
<tr>
<td>Han et al&lt;sup&gt;18&lt;/sup&gt;</td>
<td>Cohort (P)</td>
<td>China</td>
<td>114</td>
<td>Mean (SD; range): 54 (12); 24–82</td>
<td>30</td>
<td>PCR (RT-PCR) Mean (SD): 175 (20)</td>
<td>Onset</td>
<td>Outpatient visit</td>
<td></td>
</tr>
<tr>
<td>Huang et al&lt;sup&gt;19&lt;/sup&gt;</td>
<td>Cohort (P and R)</td>
<td>China</td>
<td>1733</td>
<td>Median (IQR): 57 (47–65)</td>
<td>48</td>
<td>Lab confirmed Median (IQR): 186 (175–199)</td>
<td>Onset</td>
<td>Onset</td>
<td>Onset</td>
</tr>
<tr>
<td>Zhang et al&lt;sup&gt;21&lt;/sup&gt;</td>
<td>Cohort (R/S)</td>
<td>China</td>
<td>527</td>
<td>Median (IQR; range): 42.5 (32–54; 0–91)</td>
<td>44</td>
<td>NR Mean (SD): 221.7 (10.9)</td>
<td>Discharge</td>
<td>Outpatient visit</td>
<td></td>
</tr>
<tr>
<td>Lerum et al&lt;sup&gt;22&lt;/sup&gt;</td>
<td>Cohort (P)</td>
<td>Norway</td>
<td>103</td>
<td>Median (25th–75th percentile): 59 (49–72)</td>
<td>48</td>
<td>Nasopharyngeal swab 3 months</td>
<td>Discharge</td>
<td>Onset</td>
<td>Outpatient visit</td>
</tr>
<tr>
<td>Méndez et al&lt;sup&gt;23&lt;/sup&gt;</td>
<td>Cohort (R/S)</td>
<td>Spain</td>
<td>215</td>
<td>Median (IQR): 55 (47–68)</td>
<td>40</td>
<td>Lab confirmed Median (IQR): 87 (62–109)</td>
<td>Discharge</td>
<td>Outpatient visit</td>
<td></td>
</tr>
<tr>
<td>Nguyen et al&lt;sup&gt;24&lt;/sup&gt;</td>
<td>Cohort (P)</td>
<td>France</td>
<td>125</td>
<td>Median (IQR; range): 36 (27–48; 16–85)</td>
<td>55</td>
<td>PCR (RT-PCR) Mean (SD): 221.7 (10.9)</td>
<td>Onset</td>
<td>Phone interview</td>
<td>Phone interview</td>
</tr>
<tr>
<td>Nugent et al&lt;sup&gt;25&lt;/sup&gt;</td>
<td>Cohort (R/S)</td>
<td>USA</td>
<td>182</td>
<td>Median (IQR): 67.4 (58.3–80.1)</td>
<td>47</td>
<td>PCR (RT-PCR) Median (IQR): 92.9 (52.6–127.7)</td>
<td>Discharge</td>
<td>Outpatient visit</td>
<td></td>
</tr>
<tr>
<td>Qin et al&lt;sup&gt;26&lt;/sup&gt;</td>
<td>Cohort (P)</td>
<td>China</td>
<td>647</td>
<td>Mean (SD): 58 (15)</td>
<td>56</td>
<td>PCR (RT-PCR) Mean (SD): 221.7 (10.9)</td>
<td>Discharge</td>
<td>Outpatient visit</td>
<td></td>
</tr>
<tr>
<td>Qu et al&lt;sup&gt;27&lt;/sup&gt;</td>
<td>Cohort (P)</td>
<td>China</td>
<td>540</td>
<td>Median (IQR): 47.50 (37–57)</td>
<td>50</td>
<td>PCR (RT-PCR) Mean (SD): 221.7 (10.9)</td>
<td>Discharge</td>
<td>Electronic survey</td>
<td>Electronic survey</td>
</tr>
<tr>
<td>Sibila et al&lt;sup&gt;28&lt;/sup&gt;</td>
<td>Cohort (P)</td>
<td>Spain</td>
<td>172</td>
<td>Mean (SD): 56.1 (19.8)</td>
<td>43</td>
<td>NR Mean (SD): 101.5 (19.9)</td>
<td>Discharge</td>
<td>Onset</td>
<td>Outpatient visit</td>
</tr>
<tr>
<td>Simani et al&lt;sup&gt;29&lt;/sup&gt;</td>
<td>Cohort (P)</td>
<td>Iran</td>
<td>120</td>
<td>Mean (SD): 54.62 (16.94)</td>
<td>33</td>
<td>PCR or radiological diagnosis Mean (SD): 101.5 (19.9)</td>
<td>Discharge</td>
<td>6 months</td>
<td>NR</td>
</tr>
<tr>
<td>Suárez-Robles et al&lt;sup&gt;30&lt;/sup&gt;</td>
<td>Crosssectional</td>
<td>Spain</td>
<td>134</td>
<td>Mean (SD): 58.53 (18.53)</td>
<td>54</td>
<td>PCR (RT-PCR) Mean (SD): 221.7 (10.9)</td>
<td>Discharge</td>
<td>Onset</td>
<td>Onset</td>
</tr>
</tbody>
</table>
RESULTS

Most studies were set in Europe (62%, 24/39), followed by North America (8%, 3/39) and the Middle East (8%, 3/39) (figure 1). There was no study set in a low-middle income country.30 Most were cohort studies (82%, 32/39), followed by cross sectional studies (15%, 6/39) and a case–control study (3%, 1/39). These studies present data on 10,951 (range: 100–1733) people in 12 countries, aged from 9 months to 93 years old and 48% (5206/10,951) were females.

The map shows the global distribution of the studies identified and the shading shows the combined studies population size by country. Most studies included adults, while 10% (4/39) also included children.31–34 Only 15% (6/39) of studies reported ethnicity of the participants,35–40 but without stratification. Table 1 presents the included study characteristics.

Most studies (67%, 26/39) were cohorts of hospitalised patients post discharge, 10% (4/39) followed up people who were not hospitalised, while 23% (9/39) included both (hospitalised and non-hospitalised populations). Of the inclusions in this review, 78% (8520/10,951) were previously hospitalised during the acute COVID-19 phase.31 33–35 37 38 40–55 Twenty-two studies included people requiring ICU admission during the acute phase.31 33–35 37 38 40–55 The longest follow-up period in any study was a mean of 221.7 (SD: 10.9) days post onset. Only 56% (22/39) of studies specified COVID-19 severity,31 33–35 37 38 40–55 31% (12/39) treatment received during the acute phase36 40 41 45 46 50 53 56–60 and 62% (24/39) ventilation support requirements. Pre-existing comorbidities were reported in the majority of studies (85%, 33/39), with hypertension and diabetes most commonly documented.33 35–37 39–63

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Country</th>
<th>Population size</th>
<th>Age (years)</th>
<th>Sex (% female)</th>
<th>COVID-19 confirmation method</th>
<th>Follow-up time (days)</th>
<th>Follow-up timepoint</th>
<th>Follow-up mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sykes et al39</td>
<td>Cohort (P)</td>
<td>UK</td>
<td>134</td>
<td>Median (range): 58 (25–89)</td>
<td>34</td>
<td>PCR (RT-PCR)</td>
<td>Median (range): 113 (46–167)</td>
<td>Discharge</td>
<td>Outpatient visit</td>
</tr>
<tr>
<td>Taboada et al62</td>
<td>Cross sectional</td>
<td>Spain</td>
<td>183</td>
<td>Mean (SD): 65.9 (14.1)</td>
<td>40</td>
<td>PCR (RT-PCR)</td>
<td>6 months</td>
<td>Discharge</td>
<td>Unstructured interview</td>
</tr>
<tr>
<td>Weng et al62</td>
<td>Cohort (P)</td>
<td>China</td>
<td>117</td>
<td>45.3% ≥ 60 years</td>
<td>44</td>
<td>PCR (RT-PCR)</td>
<td>Median (IQR: range): 97.0 (95.0–102.0; 91–116)</td>
<td>Discharge</td>
<td>Phone interview</td>
</tr>
<tr>
<td>Xiong et al44</td>
<td>Cohort (P)</td>
<td>China</td>
<td>538 184 (C)</td>
<td>Median (IQR: range): 52 (41–62; 22–79)</td>
<td>55</td>
<td>PCR (RT-PCR)</td>
<td>Median (IQR: range): 97.0 (95.0–102.0; 91–116)</td>
<td>Discharge</td>
<td>Phone interview</td>
</tr>
<tr>
<td>Xu et al43</td>
<td>Case–control</td>
<td>China</td>
<td>103 27 (C)</td>
<td>Median (IQR: range): 97.0 (95.0–102.0; 91–116)</td>
<td>50</td>
<td>PCR (RT-PCR)</td>
<td>Median (IQR: range): 97.0 (95.0–102.0; 91–116)</td>
<td>Discharge</td>
<td>Outpatient visit</td>
</tr>
<tr>
<td>Zhang et al52</td>
<td>Cohort (P)</td>
<td>China</td>
<td>310</td>
<td>Median (IQR): 51 (31.8–61)</td>
<td>50</td>
<td>PCR (RT-PCR)</td>
<td>Median (IQR): 92.0 (90–100)</td>
<td>Discharge</td>
<td>Outpatient visit</td>
</tr>
</tbody>
</table>

C, control group; M/M, mild/moderate; NR, not reported; P, prospective; PCR, polymerase chain reaction; R, retrospective; RT, Reverse transcription; S/C, severe/critical; TLco, carbon monoxide transfer factor.
Risk of bias
Overall, 12 studies were assessed as high risk of bias, 22 as moderate risk of bias and 5 as low risk of bias. Most studies had a high risk of bias with regard to the generalisability of their results to the wider population with COVID-19. High risk of bias ratings were most common for external validity, with item 1 (representation of target population) and item 3 (random selection) having the most high risk of bias ratings (online supplemental file 2). Further, the recruitment process and response rates were often not well described and several studies applied different data collection methods. Although many studies applied validated measurement methods to assess participants, most were not designed to detect symptoms arising from COVID-19. Only four studies included a comparative control group.35 36 43 44

Symptoms and signs
Patients suffering from long COVID report a wide range of new or persistent symptoms, in both the hospitalised and non-hospitalised populations. Symptoms were broadly organised into physiological ‘clusters’ for the purpose of presentation and interpretation of this review (figure 2).

The focus of each study included in our analysis varied. Some authors focused solely on a specialty, such as dentistry, or a specific symptom, such as cognition, making comparative analysis difficult. Even among those studies which took a broad approach, the prevalence of symptoms was diverse. Similarly, the prevalence of the more commonly reported symptoms varied markedly.

Within these limitations, we performed a meta-analysis of the most commonly reported symptoms and signs of long COVID. The most commonly described symptoms (with prevalence of 25% or greater) were weakness (41%, 95% CI 25.43 to 59.01), general malaise (33%, 95% CI 14.91 to 57.36), fatigue (31%, 95% CI 23.91 to 39.03), concentration impairment (26%, 95% CI 20.96 to 31.73) and breathlessness (25%, 95% CI 17.86 to 33.97). Across studies, 37% (95% CI 18.43 to 59.93) of patients reported reduced quality of life. Although high I² values (>80%) were observed, they resulted from narrow dispersions in the estimates and well-separated estimates and CIs between studies (online supplemental file 4). The differences between these symptoms and the heterogeneity within them are likely to be, to some extent, due to other factors (eg, study settings, populations and different measurement tools used).

Patients also reported a diverse array of less prevalent symptoms and signs, including sweating, chest pain, sore throat, anxiety and headaches, among others. The prevalence of these symptoms was lower, usually less than 20%. Figure 3 presents the range of documented patient symptoms and signs, including all the studies. Figure 4 displays these data by population, including the studies that specified hospitalised and non-hospitalised cohorts. We also performed subgroup analysis based on setting (hospitalised vs non-hospitalised) and follow-up time. In several symptoms and signs, the heterogeneity of the results was found to be associated with level of hospitalisation, hospital settings, location of the studies and follow-up timing using subgroup analysis (online supplemental files 5-8). Using meta-regression, the proportion of female patients in the studies was positively associated with headache and smell and taste disturbance (online supplemental file 9), while the proportion of ICU patients in the studies was positively associated with muscle pain (online supplemental file 10). No major difference was found in the sensitivity analyses (online supplemental files 11 and 12). Asymmetries found in the funnel plots suggest reporting biases and poor methodological quality in the included studies (online supplemental file 13).

Imaging and diagnostics
Multiple studies assessed lung sequelae and respiratory performance through outpatient visits follow-up (49%, 19/39).31 37–43 46 48 49 51–54 56 60 61 66 Imaging results were reported in 33% (13/39)31 37–39 43 46 48 52–54 56 61 66 of the
cohort studies, with one including controls \(^43\) and one with a population including children.\(^39\) Authors used heterogenous measurement techniques with an observed tendency towards novel imaging, including artificial intelligence and point-of-care ultrasound.\(^43\) \(^54\) Studies found abnormal CT results, including consolidation, reticulation, residual ground glass opacity, interstitial thickening and fibrotic changes. Some of these studies presented comparisons between initial CT findings and those at follow-up, showing improvements in pulmonary clinical measures and radiologic resolutions at follow-up visits.\(^37\) \(^39\) \(^46\) \(^48\) \(^54\) One study assessing thrombotic complications in COVID-19 with a minimum of 90-day follow-up from critical care admission found low rates of hospital-associated venous thromboembolism post discharge.\(^66\) 

Pulmonary function tests were reported in 26\% (10/39) of studies,\(^37\) \(^38\) \(^41\)–\(^43\) \(^48\) \(^49\) \(^51\) \(^53\) \(^61\) including spirometry, diffusion capacity, lung volume and exercise tests. These studies found evidence of altered pulmonary function, most frequently significant reduction of carbon monoxide transfer factor. 

One study assessed kidney function in people with COVID-19-associated acute kidney injury (AKI) compared...
with people with non-COVID-19-associated AKI, found that COVID-19-related AKI was associated with decreased kidney recovery during outpatient follow-up.36

**Risk factors**

Exploring the literature, we sought to produce a meta-analysis of risk factors for long COVID. We found a considerable diversity of reported risk factors, including age, sex, comorbidities, ethnicity and severity of the acute phase.

Several cohorts (64%, 25/39) assessed whether there was an association between the severity of initial COVID-19, including symptom load, level of hospital care, need for mechanical ventilation and the risk of persisting sequelae. An association between female gender and long COVID risk has also been noted in longitudinal studies (20.5%, 8/39), as has the association between presence of comorbidity, increasing age and minority ethnicity, with long COVID risk.

The limitations of the existing evidence base and inconsistency of reported findings preclude confident conclusions at this time. Instead, we have summarised the reported significant associations to date (online supplemental file 14) and suggest that these associations be explored in prospective controlled trials.

**Figure 4** Sign and symptoms in hospitalised and non-hospitalised cohorts. Note: The data on sign and symptoms from studies with data on hospitalised or non-hospitalised cohorts, it does not include studies that included mixed cohorts without subcategorisation. PTSD, post-traumatic stress disorder.
DISCUSSION

Our work represents the most comprehensive review of evidence regarding long COVID yet produced. Accurate to 17 March 2021, this LSR captures the breadth of persistent symptoms reported in 39 studies, including over 10,000 people. These data suggest long COVID is a syndrome affecting both previously hospitalised and non-hospitalised people, characterised by marked fatigue, weakness, general malaise, breathlessness and concentration impairment lasting for a prolonged period of time. Besides these common symptoms, there is a diverse array of secondary symptoms. The findings in this review show symptoms and prevalence aligned to current knowledge on long COVID. The Office for National Statistics (ONS) Cohort Study, including control participants, reports the most common symptoms persisting for 12 or more weeks included fatigue (8.3%), headache (7.2%), cough (7%) and myalgia (5.6%).

A deeper understanding of long COVID is currently prevented by the limitations of the published literature. The studies included in our review were highly heterogeneous due to differences in their study designs, settings, populations, follow-up time and symptom ascertainment methods. In addition, studies used inconsistent terminology describing symptoms and limited details and stratification on pre-existing comorbidities, the severity of COVID-19 and treatment methods. This inconsistency and limited reporting partly explain the high degree of variability observed. The lack of case–control studies prevent a direct attribution of symptoms solely to COVID-19; larger prospective studies with matched control groups are needed. We note that there are large, robust prospective cohort studies of hospitalised patients and non-hospitalised people. Simultaneously, qualitative studies are ongoing to better explore the long COVID patient experience.

The findings have identified several research gaps and priorities. The majority of long COVID cohorts were conducted in Western Europe on patients recently discharged from hospital. There is a paucity of evidence on the long-term effects of COVID-19 in low-middle income countries and in people who were not hospitalised. Similarly, there were no studies identified focusing on children, despite evidence showing that children and young people are also affected by long COVID. Additionally, no study stratified by ethnicity, an important risk factor for the acute phase.

Our review also highlights a need for standardised and validated COVID-19 research tools to harmonise data collection, improve quality and reduce reporting variability. For instance, fatigue is one of the most commonly reported symptoms of long COVID. However, the symptom alone is not clearly defined and it is open to different interpretations, hence it requires a validated tool such as the Visual Analogue Scale, graded fatigue scale for robust, objective and comparative analysis. ISARIC has developed open access research tools available to sites globally to facilitate standardisation of data collection, analysis and interpretation for adults and children of an age. We support the broader use of this tool as well as initiatives to standardise outcome measures for long COVID.

Similarly, our study highlights the need for further research to refine the many circulating interim case definitions and precisely characterise long COVID, including the potential impacts of variants of concern and vaccination on long COVID.

As this is an LSR, emerging themes from this first version will inform future updates. The LSR will be updated periodically, as new research is published internationally, in order to provide relevant up to date information for clinicians, patients, researchers, policy-makers and health-service commissioners. Version changes will be identified and previous reports will be archived.

CONCLUSION

This LSR summarises published evidence on the spectrum of long-term COVID-19-associated symptoms and sequelae (as of 17 March 2021). It is clear that long COVID affects different populations, with a wide range of symptomatology. Our findings suggest this multiorgan syndrome is characterised by fatigue, weakness, malaise, breathlessness and concentration impairment, among other less frequent symptoms. Currently, the strength of the available evidence is limited and prone to bias. The long-term effects of COVID-19, in both hospitalised and non-hospitalised individuals, including children and at-risk populations, should be a priority for future research using standardised and controlled study designs. Robust research is needed to characterise and define long COVID and identify risk factors and underlying aetiology, in order to inform prevention, rehabilitation, clinical and public health management to improve recovery and long-term COVID-19 outcomes. This LSR will be updated approximately every 6 months as new evidence emerges for up to 2 years.

Author affiliations

1 School of Health Sciences, City University of London, London, UK
2 ISARIC Global Support Centre, Centre for Tropical Medicine and Global Health, University of Oxford, Oxford, UK
3 Department of Clinical Research, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, UK
4 Bristol Medical School, University of Bristol, Bristol, UK
5 Long Covid Support, Birmingham, UK
6 Anaesthetic Department, Queen Elizabeth Hospital, Kings Lynn, UK
7 Julius-Maximilians-Universität Würzburg, Würzburg, Bayern, Germany
8 Department of Paediatrics and Paediatric Infectious Diseases, Institute of Child’s Health, Sechenov First Moscow State Medical University (Sechenov University), Moscow, Russia
9 Inflammation, Repair and Development Section, National Heart and Lung Institute, Faculty of Medicine, Imperial College London, London, UK
10 Research and Clinical Center for Neuropsychiatry, Moscow, Russia
11 Bodleian Health Care Libraries, University of Oxford, Oxford, UK
12 Freelance, Soquel, California, USA
13 MRC-University of Glasgow Centre for Virus Research, University of Glasgow, Glasgow, UK

Acknowledgements The authors would like to thank the members of the Long Covid Support Group and the International Severe Acute Respiratory and emerging Infection Consortium Global Support Centre.
Contributors MM, CS, VC and LS developed the concept of the study and led on the development of the protocol in collaboration with EH and members of the Long Covid Support Group (CH, MO, JSue tt), MM, CS and LS led on the drafting of the manuscript with contributions from all coauthors. EH, VC and MM performed the online searches, MM, CS, NE and LM screened the articles for inclusion. MM, CS and PB extracted data from the included articles. DD, IR and CS critically appraised the studies. VC led on the meta-analysis and presentation of the figures, in collaboration with MM, DD, VC, NE, LM, CH, MO, JScott, DD, PB, IR, DM, AJeB, CF, GC, PO, CS and LS, who helped inform the analysis, interpretation of the results and formulation of recommendations. All coauthors reviewed and approved the manuscript.

Funding This work was supported by the UK Foreign, Commonwealth and Development Office and Welcome (215091/2/18/Z), the Bill & Melinda Gates Foundation (OPP1209135) and the EU FP7 project PREPARE (602525).

Competing interests J.Suet dr e a k s he is an individual living with long-term symptoms of probably COVID-19. All other authors declare no other relationships or activities that could appear to have influenced the submitted work.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: https://creativecommons.org/licenses/by/4.0/.

ORCID iDs
Melina Michelen http://orcid.org/0000-0003-3659-7788
Natalie Elkeir http://orcid.org/0000-0002-1786-8530
Vincent Cheng http://orcid.org/0000-0002-6162-4146
Claire Haside http://orcid.org/0000-0003-2075-2130
Jake Suet dr http://orcid.org/0000-0001-8568-9107
Gail Carson http://orcid.org/0000-0001-8439-9933
Charliniti Stavropoulou http://orcid.org/0000-0003-4307-1848

REFERENCES
4 Nabavi N. Long covid: how to define it and how to manage it. BMJ 2020;370:m3489.


