Chinese Herbal Medicine ("3 Medicines and 3 Formulations") for COVID-19: Rapid Systematic Review and Meta-Analysis

Xiao-Yang Hu
University of Southampton  https://orcid.org/0000-0002-3143-7999

Yangzihan Wang
Newcastle University School of Population and Health Sciences: Newcastle University Institute for Health and Society

Junqiao Chen
ISCTE-University Institute of Lisbon: ISCTE-Instituto Universitario de Lisboa

Trisha Greenhalgh
University of Oxford Nuffield Department of Primary Care Health Sciences

Jon Wardle (✉ Jon.Wardle@scu.edu.au )
Southern Cross University  https://orcid.org/0000-0001-8813-3542

Research

Keywords: COVID-19, herbal medicine, systematic review, meta-analysis

Posted Date: February 17th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-209399/v2

License: ©  This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

Background: To evaluate the evidence behind claims that Chinese Herbal Medicine, specifically “three medicines and three formulations” (3M3F, comprising Jinhua Qinggan, Lianhua Qingwen, Xuebijing, Qingfei Paidu, Huashi Baidu and Xuanfei Baidu), is an effective treatment for COVID-19.

Methods: We searched PubMed, MEDLINE and CNKI databases, preprint servers, clinical trial registries and supplementary sources for Chinese- or English-language randomised trials or non-randomised studies with comparator groups, which tested the constituents of 3M3F in the treatment of COVID-19 up to September 2020. Primary outcome was change in disease severity. Secondary outcomes included various symptoms. Meta-analysis (using generic inverse variance random effects model) was performed when there were two or more studies reporting on the same symptom.

Results: Of 607 articles identified, thirteen primary studies (six RCTs and seven retrospective non-randomised comparative studies) with 1467 participants met our final inclusion criteria. Studies were small and had significant methodological limitations, most notably potential bias in assessment of outcomes. No study convincingly demonstrated a statistically significant impact on change in disease severity. Eight studies reported sufficiently similar secondary outcomes to be included in a meta-analysis. Some statistically significant impacts on symptoms, chest CT manifestations, laboratory variables and length of stay were demonstrated, but such findings were sparse and many remain unreplicated.

Conclusions: These findings neither support nor refute the claim that 3M3F alters the severity of COVID-19 or alleviates symptoms. More rigorous studies are required to properly ascertain the potential role of Chinese Herbal Medicine in COVID-19.

Systematic review registration: This review was registered on PROSPERO (CRD42020187502) prior to data collection and analysis.

Background

China was the first country to be seriously affected by COVID-19. The first version of the Novel Coronavirus Pneumonia Treatment Plan was published on 16th January 2020 (1), and the Plan was soon revised into the 7th edition (2). From the 4th revision the Treatment Plan included standardised Chinese herbal medicine (CHM) and claimed CHM to be effective to be recommended to patients with all stages of disease from observation period to critical phase (3, 4). Six CHM recipes were selected for use, specific combination Chinese herbal drugs known as the “3 Medicines and 3 Formulations” (3M3F, \[\text{Jinhua Qinggan granule-JHQG, Lianhua Qingwen-LHQW, Xuebijing-XBJ}\]).

The ‘3 Medicines’ (Jinhua Qinggan granule-JHQG, Lianhua Qingwen-LHQW capsule/granule, Xuebijing-XBJ) are repurposed existing medicines, used for symptomatic relief of respiratory illnesses including SARS, H1N1 influenza and pneumonia (5-7). The ‘3 Formulations’ (Lung Cleansing and Detoxifying Decoction, Qingfei Paidu-QFPD decoction, Huashi Baidu-HSBD formula and XuanFei Baidu-XFBD granule) are novel preparations, developed from existing CHM formulas for treatment of COVID-19.

The 3M3F were claimed to have significant efficacy after observation of population data, and the role of 3M3F in COVID-19 treatment was officially announced in a Chinese government press conference on 23rd March 2020, with 3M3F promoted as being able to relieve symptoms, and reduce the number of mild of moderate cases progressing to severe cases (8). Specific claims included that the compound significantly improves immunological indicators for both mild and severe COVID-19; that one of the Medicines (LHQW) and the 3 Formulations are effective in improving radiologically-assessed lung infiltrates; that one of the Formulations (XFBD) improves lymphocyte count by 17% and cure rate by 22%; and that another of the Formulations (HSBD) reduces the time for viral testing to turn negative and shortens hospital stay by 3 days. One Medicine (LHQW) was suggested to have antiviral and anti-inflammatory effects by inhibiting the SARS-COV-2 replication and reducing the pro-inflammatory cytokines production at the mRNA levels (9). These claims were widely reported in the Chinese press and endorsed by politicians (both in China and internationally, including the World Health Organisation).

Due to the concise nature of the official statement, the above findings were communicated in concise language with little detail of data supporting claims. However, despite the paucity of available data, 3M3F was readily and significantly implemented into COVID-19 treatment management. The South China Morning Post then reported that over 90% of Chinese patients with COVID-19 had been treated with CHM (10). Large quantities of 3M3F were shipped as part of the Chinese government’s aid package to other countries (11). Despite unclear evidence of efficacy and some negative press in the West (12), they have been distributed by local civic organisations such as the Red Cross and Chinese embassies (13). It is critically important to independently review the evidence base behind such claims, because these medicines have been formally promoted in China and internationally, and have been adopted into multiple national guidelines (14).

Whilst multiple reviews have reviewed the role of herbal medicine - and CHM specifically - for COVID-19, none of these reviews look at 3M3F specifically. Independent review is essential to shed light on the debate around the effectiveness of CHM in the global COVID-19 pandemic. As such, our review is the first systematic review to evaluate whether 3M3F improves outcome in COVID-19 and test the specific efficacy claims outlined above.

Methods

This rapid systematic review is reported following the PRISMA checklist. We largely followed Cochrane Interim Rapid Reviews Guidance produced specifically for the COVID-19 pandemic (15), except for tailoring our search to Chinese bibliographic database. Our team included bilingual authors experienced undertaking systematic review tasks in English and Chinese and familiar with both health systems.

Search strategy and selection criteria
In early May 2020, we searched PubMed, MEDLINE and CNKI (China National Knowledge Infrastructure) databases with date restrictions (2019-2020). We used keywords and MeSH terms in domains of COVID (e.g. "COVID-19", "Coronavirus"), Chinese and herbal medicine (e.g. "Herbal medicine", "Traditional Chinese Medicine"), official terms for the 6 Medicines, (e.g. "Lianhua Qingwen") and Chinese, English and botanical terms associated with the 3 Formulations (e.g. "Ma Huang"). Using the same or similar keywords, we searched pre-print servers (MedRxiv and BioRxiv), clinical trial registries (ChiCTR, Clinicaltrials.gov, WHO ICTRP, PROSPERO), as well as Cochrane Task Exchange, Public Health England and a hand-search of references from selected articles. A detailed search strategy and search term alternatives are available as supporting information; see Supplementary Material 1.

The search was repeated in September 2020. Web pages of Chinese Center for Disease Control and Prevention (16), National Health Commission of People's Republic of China (17) and State Administration of Traditional Chinese Medicine (18) were searched for reference to clinical studies. Studies identified from English [JW, XYH] and Chinese databases [YW, JC] were screened independently by two authors.

We included all Chinese- and English-language comparative studies of 3M3F, including randomised controlled trials (RCTs) or non-randomised studies of interventions. We included any of the 3M3F used separately or together, and alone or in conjunction with other medicines. To be included, a study of any of the 3 Formulations had to report reasonable details of the formulation which were consistent with guidelines from the State Administration of Traditional Chinese Medicine. We included any study on confirmed COVID-19 patients, including those suspected initially and diagnosed retrospectively. We placed no limitation on age, disease severity or ethnicity (in practice, most participants would have been Chinese).

Quality appraisal of studies

We used the version 2 of the Cochrane Risk of Bias for randomized trials (RoB 2) (19) and the Newcastle-Ottawa Scale for non-randomised studies (20). One reviewer extracted data and critically appraised the studies [YW, JW]. A second reviewer double checked [JC, XYH]. Disagreements were resolved by a third reviewer.

Data extraction

Data were extracted by YW, JC and XYH for Chinese-language sources, and by JW and XYH for English-language sources. We charted the following fields onto a data extraction sheet: geographic location of recruitment, care setting, inclusion criteria including participants' starting disease severity category, age, gender, proportion of immuno-depression, pre-existing conditions, and pregnancy status.

Outcome Measurements

We predefined a primary outcome domain ('change in disease severity category at the end of treatment'), since this was a major claim at the government press conference. We sought clearly-defined categories (preferably from guidelines) and used clinically in the study settings.

China standardised definitions of disease severity early in the COVID-19 outbreak. The Chinese national guideline categorises disease severity into mild, moderate, severe and critical; the Chinese Center for Disease Control and Prevention has mild, severe and critical categories (21). This definition is cited in the US CDC guideline (22). In other countries, 'triage category' is used in regional or local settings (23). Usually, these categorisations take many clinical characteristics into consideration, including vital signs, symptoms, laboratory and radiographic findings. We did not include 'disposition' (for example, home care or hospital admission) on its own as a marker of disease severity unless the triage criteria were clearly stated. We included categories 'dead' and 'cured' if the definition of 'cured' was clear, and we did not apply time limitations for disease progression or treatment. Only categorisations from studies using the same definition were eligible for meta-analysis.

We took an emergent approach to secondary outcomes, adjusting our data extraction sheet to reflect outcomes reported in primary studies. Although a disease severity category is already a composite measure, we analysed changes in symptoms separately as secondary outcomes, because of official claims that 3M3F could relieve symptoms. We extracted treatment outcomes of the symptoms reported in COVID-19 patients.

At the time of this review, there was no international consensus on the outcomes that should be reported when studying COVID-19, so we extracted non-symptom outcomes if they were reported in the primary studies; these included laboratory, radiology and healthcare utilisation measures. All these outcomes were mentioned in the press conference (8).

Data analysis

When there were two or more studies reporting on the same outcome measures, we conducted meta-analysis using RevMan [v5.4]. For continuous variables, because of variability in diagnostic and inclusion criteria, interventions, and length of treatments and follow-up, a generic inverse variance random effects model was utilised to pool the mean difference (MD) with 95% confidence interval (CI) to incorporate heterogeneity (24). When the units of the outcome measures used across studies were not consistent, the effects as standardised mean differences (SMD) were reported. For dichotomous variables, we compared groups using risk ratio (RR) with 95% CI. Heterogeneity was judged moderate when $I^2>30\%$, substantial when $I^2>50\%$, and considerable when $I^2>75\%$ (24). Potential sources were investigated in a sensitivity analysis if appropriate when interpreting the findings.
Results

Description of dataset

The study flowchart is shown in Figure 1. Thirteen studies – Six randomised controlled trials and seven retrospective non-randomised comparative studies covering a total of 1467 participants – met our final inclusion criteria. All the studies were conducted in China: seven in Wuhan, Hubei; one study (25) in Qiongdonngan, Guizhou; one (26) in Beijing; one (27) in Changsha, Hunan; one (28) in Shiyan and one (29) in Xiangyang, Hubei; another one (30) was a large scale of study recruiting patients from 23 hospitals of nine provinces of mainland China. They covered three Medicines (LHQW, JHQG and XBJ) and one Formulation (QFPD decoction). No relevant study was identified from China CDC, NHC and SATCM’s websites.

The key characteristics of the included studies are shown in Table 1. A table of excluded studies with reasons for exclusion is given in the Supplementary Material 2.

Eleven studies reference China’s national guideline (4th to 7th revisions) to select study participants. The diagnosis criteria evolved in these revisions. The 5th revision published in early February allowed a clinical diagnosis for patients from high-risk areas (Hubei Province) without laboratory confirmation, if chest imaging was typical. This was later cancelled in the 6th revision. The 7th revision published in early March added antibody test as an option of laboratory tests. Two studies (30, 31) followed the 4th guideline to select patients, one of which (31) only involved suspected cases. These suspected cases would be considered “clinically diagnosed” if the 5th guideline criteria were applied. Eight studies followed 5th or 6th guideline (25, 28, 29, 32-36) with confirmation of laboratory testing, and one of them (32) included a special inclusion requirement of being hospitalized for more than 6 days. One study (35) recruited both suspected and diagnosed cases according to the 7th treatment guideline, and used epidemiological history, clinical symptoms, CT images and etiological evidence as criteria. Two studies (26, 27) did not mention guideline-based diagnosis. Two studies captured post-acute COVID data (26, 29), while none followed long enough to observe potential chronic COVID symptoms. Eight studies (26, 31-33, 36, 37) provided a breakdown of participants’ underlying conditions, most commonly hypertension (ranging from 12.2% to 33.3%), coronary heart disease (2.1% to 16.2%), stroke (5.9% to 15.9%), diabetes (7.8% to 25.6%). Two studies (35, 37) reported a small number of patients with COPD (1.1% to 4.9%). One study (37) included a small number of patients with pre-existing respiratory disease (chronic obstructive pulmonary disease and tuberculosis, about 3%). Other small proportion of underlying condition reported including chronic kidney and liver disease, cirrhosis, bronchial asthma, hyperlipidaemia and diseases not specified.

In all studies except one arm in (35), CHM medicines or formulations were used in conjunction with usual care (as recommended in the current version of the Chinese national guideline), and compared with usual care alone. ‘Usual care’ in all the studies included three main approaches: nutrition and supportive treatment, symptomatic treatment and antiviral and antibacterial treatment.

Quality appraisal of included studies

The results of quality appraisal of the included studies are shown in Figure 2.

Results of quality assessment of the RCTs are shown in Figure 2a. All six trials had some form concerns or were considered having high risk of bias. When evaluating the randomization process, three trials (30, 35, 37) produced random sequences through SPSS or SAS software, whilst three trials (25, 27, 36) used random number table. Two trials (30, 37) concealed the allocation until the completion of enrolment. The studies (25, 27, 36) did not report allocation concealment. One study (35) was designed as non-blind and patients were grouped using a block random method, and this trial was assessed of high risk in the randomization process. This trial (35) was judged to be of high risk of bias in measurement of the outcome, since assessors’ and patients’ knowledge of highly promoted interventions could influence assessment on outcomes such as symptom improvement. The other four studies (30, 35-37) were appraised as high risk in the same aspect. The other two RCT were open labelled as well. However, because their main outcomes are laboratory tests, they were judged as being low risk of bias. Three studies (25, 27, 36) did not report information about whether patients were aware of their allocation. Also four studies (25, 27, 35, 37) reported no trial registration information on the manuscripts, and we cannot match the studies to protocols we retrieved from Chinese Clinical Trial Registry, we judge they of some concerns with the domain of ‘selection of reported result’. The study (30) was the only one registered with the Number: CTR-TRC-2000029434, but this study did not include intention-to-treat analysis which was considered as inappropriate to estimate the effect of assignment to intervention.

Of the non-randomised studies (all of which were retrospective cohort studies) (31, 32, 34) three were found to be of fair quality, the other four (26, 28, 29, 33) were of outstanding quality (Figure 2b). All studies had extensive exclusion criteria for major diseases (including renal disease, cancer and immunodeficiency), and all but one study (37) excluded comorbid respiratory diseases, though as the presence of these comorbidities is low for Chinese COVID-19 patients the population is likely to be representative of patients with COVID-19 (38). The exposed and non-exposed cohort were from the same community. Two studies (31, 34) failed to have comparability on the basis of study design, the other studies normally controlled for age or disease severity of patients. All the studies were completed, but only two were considered of enough follow-up length: one study (26) conducted lasted for 25 days, and clearly-recorded data of nucleic acid test and pneumonia recovery situation were collected till the 15th day of hospitalization. Another one (29) lasted for 22 days. The others were finished within 7 to 10 days. All studies used medical records to ascertain exposure and did not stipulate the outcome of interest at the beginning of the studies, suggesting a potentially significant source of bias.

Effects of interventions on outcome measures
Primary outcome

Our primary outcome measure (change in disease severity category according to clinical guidelines) was adequately reported in only one (non-randomised) study. One study (32) claimed that there was a significantly lower proportion of patients becoming severe in the treatment group compared to the comparator group, as judged by a p value less than 0.05 (see Table 2 for numbers). However, our own calculation reveals that p=0.0503 (see Supplementary Material 3 for details).

One randomised controlled trial (36) reported changes in disease severity but we choose not to include these findings because the definition of category used as treatment outcome was not clear. There was also inconsistency in the numbers presented in this study (see Supplementary Material 4 for details).

Moreover, the study included both mild and moderate patients, but only presented data on progression to severe or dead, missing progression from mild to moderate and progression to critical. We wrote to the corresponding author for clarification, but received no response.

One retrospective analysis (33) of QFPD decoction showed no significant difference in the numbers of patients being cured (as defined by the Chinese national guideline).

Secondary outcomes

Improvement in symptoms

Primary studies measured symptom resolution differently. Fever resolution, for example, was measured in three ways: time taken for fever to resolve, whether fever was resolved after at the end of treatment, and change in symptom score. Assigning a score to a symptom is a common practice in CHM studies, although it has been criticised for systematic errors, non-standardized use in each study and statistical inappropriateness (39). As a result, we will not report on the Traditional Chinese Medicine (TCM) scoring of symptoms, but have included additional information in Supplementary Material 5.

Figures 3a to 3o show the results of meta-analysis of studies which tested the effectiveness of 3M3F on 13 reported COVID-19 symptoms. Limited findings suggested that 3M3F may reduce time took to fever recovery by SMD -0.98 days, 95% CI -1.78 to -0.17; participants = 163; studies = 3; I^2= 83%. There were larger proportion of COVID-19 patients benefited from 3M3F in recovery of fever, cough, fatigue/tiredness, phlegm, short of breath and muscle pain, but not in the other seven symptoms reported (Table 2).

One RCT comparing LHQW granule as an add on to antiviral and antimicrobial treatment in line with 7th edition of national guidelines failed to show a reduction in the proportion of patients with improved fever RR 1.00 [0.91, 1.10], cough RR 0.86 [0.69, 1.06], fatigue RR 1.05 [0.84, 1.33], diarrhoea RR 1.00 [0.80, 1.25], nausea/vomiting RR 0.98 [0.75, 1.26], or loss in appetite RR 1.00 [0.80, 1.25], comparing LHQW granule to usual care (35).

Data from three retrospective cohort studies (32-34) showed a statistically significant effect in favour of 3M3F in reducing time to fever resolution at 0.98days, 95% CI -1.78 to -0.17; participants = 163; studies = 3; I^2= 83% (Figure 3a). Two retrospective cohort studies (31, 32, 34) and a single RCT (37) tend to suggest larger proportion of patients with fever resolved by taking LHQW (granule) and JHQG together with usual care RR 1.38, 95% CI 1.19 to 1.61; participants = 318; I^2 = 0% (Figure 3b).

There was large heterogeneity among studies reporting the proportion of patients with cough resolved and they showed conflict findings. Three retrospective cohort studies (31, 32, 34) favoured LHQW group RR 1.90, 95% CI 1.24 to 2.90; participants = 199; I^2 = 18%, while a RCT failed to prove the favourable effects of JHQG plus usual care versus usual care RR 1.54, 95% CI 0.97 to 2.45 (37) (Figure 3c).

Similar positive findings from RCTs or retrospective cohort studies were observed in the proportion of patients with symptom resolution in fatigue/tiredness (RR 1.48, 95% CI 1.18 to 1.86; participants = 219; studies = 3; I^2 = 0%, Figure 3d), phlegm (RR 1.97, 95% CI 1.08 to 3.61; participants = 176; studies = 4; I^2 = 52%, Figure 3e), shortness of breath (RR 3.93, 95% CI 1.89 to 8.17; participants = 83; studies = 3; I^2 = 0%, Figure 3f), and muscle pain (RR 1.83, 95% CI 1.02 to 3.27; participants = 49; studies = 3; I^2 = 2%, Figure 3g). On the contrary, studies with small samples failed to show a favourable effect over 3M3F in the resolution of chest tightness (RR 2.00, 95% CI 0.81 to 4.96; participants = 89; studies = 3; I^2 = 64%), diarrhoea (RR 1.09, 95% CI 0.65 to 1.82; participants = 35; studies = 3; I^2 = 0%), nausea/vomiting (RR 1.25, 95% CI 0.82 to 1.90; participants = 43; studies = 3; I^2 = 0%), loss in appetite (RR 0.63, 95% CI 0.14 to 2.84; participants = 33; studies = 3; I^2 = 55%), sore throat (RR 1.35, 95% CI 0.68 to 2.70; participants = 26; studies = 3; I^2 = 0%), headache (RR 1.21, 95% CI 0.83 to 1.77; participants = 47; studies = 3; I^2 = 0%), or block/running nose (RR 1.00, 95% CI 0.64 to 1.57; participants = 23; studies = 3; I^2 = 0%).

Table 3 shows the impact on symptom resolution in studies which were not amenable to meta-analysis. Statistically significant differences were shown for LHQW capsule (time to resolution of fever, cough, and fatigue), LHQW granule (time to resolution of cough, shortness of breath, symptom scores for fever, dry and sore throat), and QFPD decoction (time to resolution of cough).

Recovery or improvement of chest CT manifestations

Significant changes were shown in two retrospective cohort studies in time to reduction in lung lesion on CT scan, in QFPD (decoction) -4.80 days, 95% CI -5.82, -3.77, and JHQG (decoction) - 0.53 days, 95% CI -0.98, -0.08 at day 15, as adds on to usual care. In addition, there was a larger proportion of patients...
experiencing recovery/improvement of chest CT manifestations (RR 1.16, 95% CI 1.03 to 1.30; participants = 521; 3 retrospective cohort studies; I² = 0%, Figure 3o).

**Other secondary outcome measure**

Inconclusive findings on blood test results, length of hospital stay, viral conversion, and medication used are reported narratively (Table 4). One non-randomised study found statistically significant differences in favour of LHQW in four laboratory tests (white cell count, lymphocyte count, C-reactive protein and procalcitonin). The clinical significance of these results is not clear and the authors do not discuss them. Inconclusive findings were observed in reduction in length of stay: one small, non-randomised study (33) showed a statistically significant reduction in length of stay in those received QFPD decoction, while one (29) failed to show the same.

**Adverse events**

No study reported any serious adverse events (AE). Four studies did not discuss AE in their results (29, 32, 34, 35). Among those that discussed AEs, three suggested no AE was observed either in the 3M3F or the comparator groups (25, 27, 31) and one reported no serious side effects (36). One RCT (30) reported 45.8% (65/142) cases of AEs including abnormal liver function, renal dysfunction, headache, nausea, vomiting, diarrhoea and loss of appetite in the add-on LHQW capsule, while the control group reported 54.2% (77/142) cases with adverse events, including abnormal liver function, renal dysfunction, headache, nausea, vomiting, diarrhoea and loss of appetite. However, such comparison of this study (30) was found with no statistical significance at 0.84, 95% CI 0.67 to 1.07. The RCT of (37) using JHQG reported diarrhoea in 27 out of 82 (33%) participants in TG versus 0 in CG, and this result has statistically significant difference.

**Discussion**

**Summary of key findings**

Despite strong official endorsement of 3M3F as an effective treatment for COVID-19, the evidence base for this intervention rests on thirteen studies covering a total of 1467 participants. While the limited studies appear to suggest that 3M3F, when used on top of usual care, may offer some relief for some symptoms and changes in lung lesion on CT scan experienced by mostly mild or moderate COVID-19 patients, the results do not support the high-level claims that 3M3F could prevent disease from progressing to a more severe type. There were methodological concerns in all studies, with especially high risk of bias in outcomes assessment in the four RCTs. Missing and wrong protocol registration information intensifies our concern over the overall integrity of these studies.

Of the six remedies that make up 3M3F, four had been tested in any experimental study that met our inclusion criteria. Our primary outcome measure (reduction in severity of disease) did not achieve convincing statistical significance in any of the primary studies. In relation to the secondary outcomes, the positive effects of LHQW, JHQG and QFPD on various symptoms could be explained by bias in assessment of outcome (and in particular, the widespread use of the 'symptom score' in TCM), and would need to be replicated before being viewed as definitive. Similarly, the positive impacts of different 3M3F remedies on radiological outcome (two studies), laboratory tests of biomarkers (one study) and length of stay (two studies) need to be replicated before being viewed as definitive.

With the exception of diarrhoea with JHQG, the thirteen studies did not report any adverse events linked to 3M3F use. Adverse events have, however, been reported in the past when LHQW was used to treat influenza (40, 41). Previous studies have also reported some digestive system side effects from using JHQG to treat influenza, though not significantly more than the control group (6). Duan and colleagues attributed the high incidence of diarrhoea in their treatment group to the high dose of JHQG they used to treat COVID-19, and also invoked classical TCM theories to suggest that diarrhoea may have a curative role in this condition.

Although we did not limit the publication language or geography, unsurprisingly all studies were conducted in China, thus the findings may not be generalizable to other countries. There is also no placebo-controlled study, making it impossible to assess the effect of 3M3F when used alone. Most of the articles are of low quality and sample size, potentially limiting their use in informing practice. We also observe some concerning practices in these studies, for example, the number of trials registered in Clinical Trial Registry is small, and in one case we cannot even find the registered protocol using the protocol number given by the authors, and informed consent was collected only verbally in some studies. However, it should be recognized that these studies were often performed quickly and opportunistically in the early acute phase of a sudden pandemic without proper planning, and some limitations in study design and execution are understandable. Moreover, these issues are not unique to studies of 3M3F. Analysis of outcome reporting from protocols of clinical trials found similar levels of inconsistencies in outcome reporting in both Western medicine and traditional Chinese medicine clinical trial protocols (42). However, given the limitations of the studies published so far, the results of the studies do not appear to support the definitive public claim that 3M3F could prevent the progression of COVID-19.

**Strengths and limitations of this review**

To our knowledge, this is the first systematic review and meta-analysis of a group of CHM specifically promoted for COVID-19. Whilst some systematic reviews have examined the impact of integrating any CHM with conventional treatment (43-45), our review has expanded these findings by concentrating on more specific aspects to avoid overgeneralization. Firstly, comparing with (43), we had examined the impact of 3M3F which was explicitly promoted for use in COVID-19, and till January 2021, our review has identified all the published clinical studies using 3M3F as interventions. Secondly, comparing with (44), we included both RCTs and non-RCTs to provide more comprehensive information to examine the work of 3M3F, because large-scale of RCTs are insufficient in this field of research and data from other types of studies also works as evidence. Thirdly, comparing with (44), we had provided more accurate and detailed
information in quality appraisals of included studies and, comparing with (45), independent analysis of outcomes of each intervention. We followed Cochrane interim guidance for rapid reviews during this pandemic (15), and undertook independent statistical analysis of key findings from primary studies.

One limitation is the small number of primary studies identified. The relative success of China in managing the initial and second waves of COVID-19 may have limited the ability to conduct trials once detailed protocols based on early clinical experience had been developed, given that no studies outside China were identified. It is also possible that the Chinese government had access to additional unpublished data before developing its official statement on 3M3F. At least 39 clinical trials for CHM interventions were registered in the Chinese Clinical Trials Registry by January 2021 before this review was initiated, though it is unclear how many of these relate to the interventions assessed in this study (42). If such data exist, we recommend that they are placed in the public domain to ensure clinicians, researchers and policy-makers are appropriately informed. Another limitation is that other traditional medicines used for treatment of COVID-19 were not included in our review. We prioritised 3M3F as it has been officially sanctioned and promoted by the Chinese government for use, and has since been implemented into multiple national guidelines.

Suggestions for further research

Larger, multi-centre trials of CHM, and especially 3M3F, are urgently needed, with consistent inclusion criteria and objective outcome measures designed to contribute to meta-analyses. Better reporting of adverse events is needed to confirm the safety profile of 3M3F. It was beyond the scope of this review to explore the pharmaceutical properties and alleged antiviral mechanisms of the various ingredients; there is much scope for further studies in this area, perhaps with a view to developing new chemical entities for mainstream medicine. Many of these studies were performed before much was known about the disease, or which outcomes were most appropriate for inclusion. Only one study attempted to measure or report viral load of COVID-19 patients or whether this was reduced with the intervention; such variables should be included in further research. Additionally, as our examination focused primarily on the use of CHM in acute COVID-19 treatment, future research examination of CHM for longer-term symptomatic relief may be warranted given that many outcomes measured in the studies are also often reported as significant in post-acute COVID-19 (46).

Conclusions

The findings from this rapid systematic review neither support nor refute the claim by the Chinese government that CHM (specifically 3M3F) alters the severity of COVID-19 or provides alleviation of symptoms. While the limited studies appear to suggest that 3M3F, when used on top of usual care, may offer some relief for some symptoms experienced by mostly mild or moderate COVID-19 patients, the results do not support the high-level claims that 3M3F could prevent disease from progressing to a more severe type. Studies were few in number, small in size, and had significant methodological limitations (most notably, potential bias in assessment of outcomes), though the positive nature of some individual findings do suggest further examination may be warranted. More rigorous studies are required to properly ascertain the potential role of CHM in treatment of COVID-19.

List Of Abbreviations
Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

This review contains secondary analyses of published data. The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

All authors have read and agreed to the published version of the manuscript.

Funding

The authors did not receive specific funding for this review. TG's research group is supported by the UK National Institute of Health Research (NIHR) and Wellcome Trust. XYH is supported by the NIHR School for Primary Care Research. The views expressed are those of the authors. Funders had no role in study design, data collection, data analysis, data interpretation or writing of the report. All authors had full access to all data, though not all authors speak Chinese.
Authors' contributions

Conceptualisation: All authors
Investigation: All authors
Methodology: All authors
Acquisition of data: YW, JC, XYH, JW
Formal analysis: XYH
Writing – original draft: XYH, YW, JW
Writing – review & editing: All authors

Acknowledgements

Not applicable

References

1. CHINA NHCO/T/PSRO. The Diagnosis and Treatment Plan for Pneumonia Caused by Novel Coronavirus Beijing, China, ADD. 2020.


Tables

Table 1: Studies reporting treatment of COVID-19 patients with 3M3F
<table>
<thead>
<tr>
<th>Component tested</th>
<th>Author / year</th>
<th>Study type</th>
<th>Source</th>
<th>Intervention</th>
<th>Quality control reported? (Y/N)</th>
<th>Chemical analysis reported? (Y/N)</th>
<th>Care Setting</th>
<th>Disease severity</th>
<th>Sample size</th>
<th>Mean age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cheng 2020</td>
<td>Retrospective cohort</td>
<td></td>
<td></td>
<td>N⁹</td>
<td>N⁹</td>
<td>Inpatient</td>
<td>Moderate</td>
<td>102</td>
<td>CG 56</td>
</tr>
<tr>
<td></td>
<td>Yao 2020</td>
<td>Retrospective cohort</td>
<td></td>
<td></td>
<td>N⁹</td>
<td>N⁹</td>
<td>Inpatient</td>
<td>Moderate</td>
<td>42</td>
<td>CG 62</td>
</tr>
<tr>
<td></td>
<td>Lv 2020</td>
<td>Retrospective cohort</td>
<td></td>
<td></td>
<td>N⁹</td>
<td>N⁹</td>
<td>Inpatient</td>
<td>Moderate</td>
<td>101</td>
<td>CG 59</td>
</tr>
<tr>
<td></td>
<td>Liu 2020b</td>
<td>Retrospective cohort</td>
<td></td>
<td></td>
<td>N</td>
<td>N</td>
<td>Inpatient</td>
<td>Moderate and Severe</td>
<td>80</td>
<td>TG 51</td>
</tr>
<tr>
<td>QFPD decoction</td>
<td>Li 2020</td>
<td>Retrospective cohort</td>
<td>N/A</td>
<td><em>Ephedra sinica</em> 9g, Baked <em>Glycyrrhiza uralensis</em> Fisch. 6g, <em>Semen Armeniacae Amarum</em> 9g, Raw <em>Gypsum Fibrosum</em> 15-30g, <em>Cinnamomum cassia</em> Presl. 9g, <em>Alisma</em></td>
<td>N</td>
<td>N</td>
<td>Inpatient</td>
<td>92% mild or moderate</td>
<td>60</td>
<td>TG 54</td>
</tr>
<tr>
<td></td>
<td>Xin 2020</td>
<td>Retrospective cohort</td>
<td></td>
<td></td>
<td>Y⁹</td>
<td>N</td>
<td>Inpatient</td>
<td>Mild and moderate</td>
<td>63</td>
<td>TG 46</td>
</tr>
</tbody>
</table>

⁹ National Medicine Permission No. Z20100040
XBJ injection

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Wen RCT 2020</td>
<td>N N Inpatient Severe 60 50ml TG 49 100ml TG 47 CG 48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhang Retrospective cohort 2020</td>
<td>N N Inpatient Moderate 44 TG 49 CH 46</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*TG: Treatment Group. CG: Comparator Group. RCT: Randomised controlled trial

1. The study conducted by Yao et al. (2020) did not report the source of the medication, however Yiling Pharmaceutical is the only company produces this medication.
2. Not reported in article, but the information was found in Pharmacopoeia of the People's Republic of China 2020 (First Part)
3. Ultra-performance liquid chromatography fingerprint identified 9 of the 32 common peaks were compared with chemical standards, they are Neochlorogenic acid, chlorogenic acid, cryptochlorogenic acid, isoforsythoside A, forsythoside A, quercitrin, isochlorogenic acid C, forsythin, and glycyrrhizic acid
5. Not reported in article, but the information was found in Pharmacopoeia of the People's Republic of China 2020 (First Part) that an active ingredient of Lianhua Qingwen Granule is Forsythin (C_{27}H_{34}O_{11}) for at least 0.69mg in each bag of the medication.
6. The article reported as “Jianhua Qinggan Granule is based on Maxing Shigan Tang and Yinqiao San. The active ingredients of Yinqiao San including Chlorogenic acid (C_{16}H_{18}O_{9}), Forsythin (C_{27}H_{34}O_{11}), Arctigenin (C_{21}H_{24}O_{6}), Buddlenoid (C_{20}H_{20}O_{13}), Acacetin (C_{16}H_{12}O_{5}), Liquiritigenin (C_{15}H_{12}O_{4}) etc. are effective for anti-virus.”
7. The study conducted by Xin et al. (2020) reported batch number for each herb
8. The study conducted by Liu et al. (2020a) and Wen et al. (2020) did not report the source of the medication, however Hongri Pharmaceutical is the only company produces this medication.
Table 2: Impact on symptoms: findings from meta-analysis

<table>
<thead>
<tr>
<th>Covid-19 symptoms reported</th>
<th>Time reduction (days)</th>
<th>Proportion of patients with symptom resolved – overall 3M3F</th>
<th>Proportion of patients with symptom resolved – LHQW</th>
<th>Proportion of patients with symptom resolved – JHQG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fever</td>
<td>-0.98 days [1.78, -0.17]</td>
<td>1.38 [1.19, 1.61]</td>
<td>1.35 [1.14, 1.60]</td>
<td>1.51 [1.07, 2.14]</td>
</tr>
<tr>
<td>Cough</td>
<td>-</td>
<td>1.74 [1.31, 2.30]</td>
<td>1.90 [1.24, 2.90]</td>
<td>1.54 [0.97, 2.45]</td>
</tr>
<tr>
<td>Fatigue/tiredness</td>
<td>-</td>
<td>1.48 [1.18, 1.86]</td>
<td>1.51 [1.13, 2.00]</td>
<td>1.44 [0.98, 2.11]</td>
</tr>
<tr>
<td>Phlegm</td>
<td>-</td>
<td>1.97 [1.08, 3.61]</td>
<td>2.46 [0.81, 7.51]</td>
<td>1.85 [1.01, 3.38]</td>
</tr>
<tr>
<td>Short of breath</td>
<td>-</td>
<td>3.93 [1.89, 8.17]</td>
<td>3.93 [1.89, 8.17]</td>
<td>-</td>
</tr>
<tr>
<td>Chest tightness</td>
<td>-</td>
<td>2.00 [0.81, 4.96]</td>
<td>2.00 [0.81, 4.96]</td>
<td>-</td>
</tr>
<tr>
<td>Diarrhoea</td>
<td>-</td>
<td>1.09 [0.65, 1.82]</td>
<td>1.04 [0.42, 2.58]</td>
<td>1.11 [0.60, 2.07]</td>
</tr>
<tr>
<td>Nausea/vomiting</td>
<td>-</td>
<td>1.25 [0.82, 1.90]</td>
<td>1.34 [0.59, 3.06]</td>
<td>1.17 [0.69, 1.99]</td>
</tr>
<tr>
<td>Loss in appetite</td>
<td>-</td>
<td>0.63 [0.14, 2.84]</td>
<td>1.04 [0.42, 2.58]</td>
<td>0.06 [0.00, 1.03]</td>
</tr>
<tr>
<td>Sore throat</td>
<td>-</td>
<td>1.35 [0.68, 2.70]</td>
<td>1.53 [0.38, 6.23]</td>
<td>1.30 [0.58, 2.87]</td>
</tr>
<tr>
<td>Headache</td>
<td>-</td>
<td>1.21 [0.83, 1.77]</td>
<td>1.29 [0.67, 2.46]</td>
<td>1.17 [0.73, 1.87]</td>
</tr>
<tr>
<td>Muscle pain</td>
<td>-</td>
<td>1.83 [1.02, 3.27]</td>
<td>1.83 [1.02, 3.27]</td>
<td>-</td>
</tr>
<tr>
<td>Block/running nose</td>
<td>-</td>
<td>1.00 [0.64, 1.57]</td>
<td>0.90 [0.53, 1.53]</td>
<td>1.31 [0.57, 3.05]</td>
</tr>
</tbody>
</table>

Table 3: Impact on symptoms: findings from analyses not amenable to meta-analysis

All studies' comparator group was usual care; treatment group was usual care plus the component of Chinese herbal medicine. See also Figure 3 for results of meta-analysis.
<table>
<thead>
<tr>
<th>Study ID (sample size)</th>
<th>Hu 2020 (n=284)</th>
<th>Cheng 2020 (n = 102)</th>
<th>Yu 2020 (n = 295)</th>
<th>Lv 2020 (n = 101)</th>
<th>Duan 2020 (n = 123)</th>
<th>Li 2020 (n = 60)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Proportion of patients becoming severe</strong></td>
<td>TG vs CG 2.1% vs. 4.2%, mean difference: -2.1%, 95%CI: -7.0% - 2.4%, p = 0.498</td>
<td>TG: 4/51 (7.8%)</td>
<td>CG: 11/51 (21.6%)</td>
<td>Inconsistent definitions and numbers reported</td>
<td>TG: 9/51 (11.0%)</td>
<td>CG: 10/41 (24.4%) P&gt;0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Proportion of patients becoming cured</strong></td>
<td>TG: 91.5%, CG: 82.4%, mean difference: 9.2%, 95%CI: 1.3% - 17.1%</td>
<td>TG: 4/51 (7.8%)</td>
<td>CG: 11/51 (21.6%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Time to resolution of fever (days)</strong></td>
<td>TG vs CG: 2 days vs. 3 days, HR: 1.39, 95%CI: 1.00 - 1.94, p = 0.017</td>
<td>TG: median 6d</td>
<td>CG: median 7d</td>
<td>P&lt;0.05</td>
<td>TG: median 6d</td>
<td>CG: median 7d</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Time to resolution of cough</strong></td>
<td>TG vs CG: 7 days vs. 10 days, HR: 1.71, 95%CI: 1.30 - 2.23</td>
<td>TG: 3.9±2.0</td>
<td>CG: 5.2±1.8</td>
<td>P&lt;0.05</td>
<td>TG: 4.9±0.7 days</td>
<td>CG: 6.6±0.4 days</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Time to resolution of fatigue/tiredness (days)</strong></td>
<td>Median (IQR)? TG: 3.0 (3.0 - 5.0), CG: 6.0 (4.0 - 8.0) HR95%CI 1.8 (1.3 - 2.5)</td>
<td>TG: 3.5±1.5d (n=51); CG: 4.8±1.53 (n=51) p=0.028</td>
<td>-1.30 [-1.89, -0.71]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

& Apart from one study (Hu et al., 2020) evaluated LHQW capsule, all the rest investigated the granule preparation of LHQW.

HR: hazard ratio; IQR: interquartile range

Table 4: Impact on other secondary outcome measures: findings from analyses not amenable to meta-analysis
### Study ID (sample size)

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Sample Size</th>
<th>Study ID</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hu 2020</td>
<td>n=284</td>
<td>Xiao 2020</td>
<td>n=188</td>
</tr>
<tr>
<td>Yu 2020</td>
<td>n=295</td>
<td>Li 2020</td>
<td>n=60</td>
</tr>
<tr>
<td>Xin 2020</td>
<td>n=63</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Proportion of patients whose chest CT improved after 7 days

- **TG vs CG:** 83.8% vs. 64.1%, mean difference: 19.7%, 95%CI: 9.6%-29.4%

- **TG:** 1 (1-2), **CG:** 1 (1-2) (P=0.482)

### White cell count (10^9/L): pre-treatment to post-treatment

- **TG:** 5.1 ± 0.4 to 5.9 ± 0.4, **CG:** 5.2 ± 0.4 to 5.5 ± 0.4, P < 0.05

### Lymphocytes (10^9/L): pre-treatment to post-treatment

- **TG:** 1.5 ± 0.1 to 1.7 ± 0.2, **CG:** 1.5 ± 0.1 to 1.6 ± 0.2, P < 0.05

### C-reactive protein (mg/L): pre-treatment to post-treatment

- **TG:** 26 ± 6 to 22 ± 4, **CG:** 27 ± 6 to 24 ± 4, P < 0.05

### Procalcitonin (ng/L): pre-treatment to post-treatment

- **TG:** 0.089 ± 0.025 to 0.058 ± 0.008, **CG:** 0.094 ± 0.022 to 0.094 ± 0.022, P < 0.05

### Length of stay

- **TG:** 13.6 ± 0.4 days, **CG:** 16.4 ± 0.3 days, P < 0.05

### Conversion rate of SAR-CoV-2 viral assay

- **TG vs CG:** 76.8% vs. 71.1%, mean difference: 5.6%, 95%CI: -4.6%-15.7%, P = 0.279

### Viral assay conversion time (median)

- **TG vs CG:** 11.0 vs. 12.0 days, **HR:** 1.21, 95%CI: 0.92-1.59

### Medications used

- **Antiviral medication use:** **TG:** 58 (100%); **CG:** 63 (100%)
- **Antibiotic use:** **TG:** 30 (51.7%); **CG:** 62 (98.4%)
- **Antibiotic use:** **P=0.269**; **Corticosteroid use:** **TG:** 7 (18.9%), **CG:** 5 (19.2%) (P=0.390); **Antiviral drugs:** **Interferon:** **TG:** 34 (91.9%), **CG:** 26 (100%), **P=0.140**; **Arbidol:** **TG:** 24 (64.9%), **CG:** 16 (61.5%), **P=0.997**; **Lopanivir:** **TG:** 29 (78.4%), **CG:** 25 (96.2%), **P=0.049**

& Apart from one study (Hu et al., 2020) evaluated **LHQW** capsule, all the rest investigated the granule preparation of **LHQW**.

**HR:** hazard ratio

**Figures**
Figure 1
Study selection PRISMA flow chart
### 2a: Applying Cochrane Risk of Bias tool to randomised trials

<table>
<thead>
<tr>
<th>Study</th>
<th>Randomisation process</th>
<th>Deviations from intended interventions</th>
<th>Missing outcome data</th>
<th>Measurement of the outcome</th>
<th>Selection of the reported result</th>
<th>Overall bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duan, 2020 (JHQG)</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>High risk</td>
<td>Some concerns</td>
<td>High risk</td>
</tr>
<tr>
<td>Yu, 2020 (LHQW)</td>
<td>Some concerns</td>
<td>Low risk</td>
<td>Low risk</td>
<td>High risk</td>
<td>Some concerns</td>
<td>High risk</td>
</tr>
<tr>
<td>Liu, 2020 (XBJ)</td>
<td>Some concerns</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Some concerns</td>
<td>Some concerns</td>
</tr>
<tr>
<td>Wen, 2020 (XBJ)</td>
<td>Some concerns</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Some concerns</td>
<td>Some concerns</td>
</tr>
<tr>
<td>Hu, 2020 (LHQW)</td>
<td>Low risk</td>
<td>Some concerns</td>
<td>Low risk</td>
<td>High risk</td>
<td>Low risk</td>
<td>High risk</td>
</tr>
<tr>
<td>Xiao, 2020 (LHQW)</td>
<td>High risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>High risk</td>
<td>Some concerns</td>
<td>High risk</td>
</tr>
</tbody>
</table>

### 2b: Applying Newcastle-Ottawa Score to non-randomised studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Representative</th>
<th>Selection</th>
<th>Comparability</th>
<th>Outcome</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cheng, 2020 (LHQW)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7/9</td>
</tr>
<tr>
<td>Yao, 2020 (LHQW)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6/9</td>
</tr>
<tr>
<td>Li, 2020 (LHQW)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6/9</td>
</tr>
<tr>
<td>Li, 2020 (QFPD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8/9</td>
</tr>
<tr>
<td>Zhang, 2020 (XBJ)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8/9</td>
</tr>
<tr>
<td>Liu, 2020 (JHQC)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8/9</td>
</tr>
<tr>
<td>Xin, 2020 (QFPD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8/9</td>
</tr>
</tbody>
</table>

Figure 2

Results of quality appraisal of primary studies

Figure 3

Forest plots of intervention studies where meta-analysis was possible (findings were interpreted separately)

### Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 4.RRCOVID3M3FSupplementv1.0.docx