Journal of Infection xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Infection

journal homepage: www.elsevier.com/locate/jinf

Review

Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis

Zhaohai Zheng, M.D.[a,b,1], Fang Peng, B.D.[a,1], Buyun Xu, M.D.[a], Jingjing Zhao, M.D.[a,b], Huahua Liu, M.D.[c], Jiahao Peng, M.D.[d], Qingsong Li, B.D.[e], Chongfu Jiang, B.D.[e], Yan Zhou, M.D.[a], Shuqing Liu, M.D.[a,f], Chunji Ye, M.D.[a], Peng Zhang, M.D.[a], Yangbo Xing, M.D.[a], Hangyuan Guo, M.D., Ph.D.[a], Weiliang Tang, M.D.[a,*]

[a] *Department of Cardiology, Shaoxing People's Hospital (Shaoxing Hospital, Zhejiang University School of Medicine), Shaoxing City, Zhejiang Province, 312000, PR China*
[b] *Department of Cardiology, The Run Run Shaw Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou City, 310003, Zhejiang, PR China*
[c] *Cardiovascular Research Center, School of Basic Medical Science, Xi'an Jiaotong University Health Science Center, 710049, Xi'an, China*
[d] *Loma Linda University School of Public Health, 24951 Circle Dr, Loma Linda, CA 92354, USA*
[e] *Department of Cardiology, Awati County People's Hospital, No.1 North Jiefang Road, Awati County, Xinjiang Uygur Autonomous Region, 843000, PR China*
[f] *Shaoxing University School of Medicine, 312000, Shaoxing, Zhejiang, China*

ARTICLE INFO

Article history:
Accepted 16 April 2020
Available online xxx

Keywords:
COVID-19
Risk factor
Comorbidity
Clinical manifestation
Laboratory examination

SUMMARY

Background: An epidemic of Coronavirus Disease 2019 (COVID-19) began in December 2019 and triggered a Public Health Emergency of International Concern (PHEIC). We aimed to find risk factors for the progression of COVID-19 to help reducing the risk of critical illness and death for clinical help.

Methods: The data of COVID-19 patients until March 20, 2020 were retrieved from four databases. We statistically analyzed the risk factors of critical/mortal and non-critical COVID-19 patients with meta-analysis.

Results: Thirteen studies were included in Meta-analysis, including a total number of 3027 patients with SARS-CoV-2 infection. Male, older than 65, and smoking were risk factors for disease progression in patients with COVID-19 (male: OR = 1.76, 95% CI (1.41, 2.18), P < 0.00001; age over 65 years old: OR =6.06, 95% CI(3.98, 9.22), P < 0.00001; current smoking: OR =2.51, 95% CI(1.39, 3.32), P = 0.0006). The proportion of underlying diseases such as hypertension, diabetes, cardiovascular disease, and respiratory disease were statistically significant higher in critical/mortal patients compared to the non-critical patients (diabetes: OR=3.68, 95% CI (2.68, 5.03), P < 0.00001; hypertension: OR = 2.72, 95% CI (1.60,4.64), P = 0.0002; cardiovascular disease: OR = 5.19, 95% CI(3.25, 8.29), P < 0.00001; respiratory disease: OR = 5.15, 95% CI(2.51, 10.57), P < 0.00001). Clinical manifestations such as fever, shortness of breath or dyspnea were associated with the progression of disease [fever: 0R = 0.56, 95% CI (0.38, 0.82), P = 0.003;shortness of breath or dyspnea: 0R=4.16, 95% CI (3.13, 5.53), P < 0.00001]. Laboratory examination such as aspartate amino transferase(AST) > 40U/L, creatinine(Cr) ≥ 133mol/L, hypersensitive cardiac troponin I(hs-cTnI) > 28pg/mL, procalcitonin(PCT) > 0.5ng/mL, lactatede hydrogenase(LDH) > 245U/L, and D-dimer > 0.5mg/L predicted the deterioration of disease while white blood cells(WBC)$<4\times10^9$/L meant a better clinical status[AST > 40U/L:OR=4.00, 95% CI (2.46, 6.52), P < 0.00001; Cr ≥ 133μmol/L: OR = 5.30, 95% CI (2.19, 12.83), P = 0.0002; hs-cTnI > 28 pg/mL: OR = 43.24, 95% CI (9.92, 188.49), P < 0.00001; PCT > 0.5 ng/mL: OR = 43.24, 95% CI (9.92, 188.49), P < 0.00001;LDH > 245U/L: OR = 43.24, 95% CI (9.92, 188.49), P < 0.00001; D-dimer > 0.5mg/L: OR = 43.24, 95% CI (9.92, 188.49), P < 0.00001; WBC $< 4\times10^9$/L: OR = 0.30, 95% CI (0.17, 0.51), P < 0.00001].

Conclusion: Male, aged over 65, smoking patients might face a greater risk of developing into the critical or mortal condition and the comorbidities such as hypertension, diabetes, cardiovascular disease, and

* Corresponding author: Prof. Weiliang Tang, Address: NO. 568 North zhongxing Road, Yuecheng district, Shaoxing city, Zhejiang province, 312000, PR China.
E-mail address: twl-sxyz@163.com (W. Tang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jinf.2020.04.021

Please cite this article as: Z. Zheng, F. Peng and B. Xu et al., Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis, Journal of Infection, https://doi.org/10.1016/j.jinf.2020.04.021

respiratory diseases could also greatly affect the prognosis of the COVID-19. Clinical manifestation such as fever, shortness of breath or dyspnea and laboratory examination such as WBC, AST, Cr, PCT, LDH, hs-cTnI and D-dimer could imply the progression of COVID-19.

Introduction

Since December 2019, a cluster of pneumonia have attacked all human beings.[1] The pathogen was designated as SARS-CoV-2 by the International Committee on Taxonomy of Viruses, and this pneumonia was named as Coronavirus Disease2019 (COVID-19) by World Health Organization (WHO).[2] Nowadays, there were more than one million confirmed cases and over 100000 deaths occurred in 208 countries/territories according to the report of WHO until April 12th, 2020. The rapidly increasing of patients, especially the critical or mortal patients, brought a big challenge to the public health. Lai et al.[3] found that mortality was correlated with country health care resources. However, in many countries, the invasive ventilator and intensive care unit (ICU) were far from adequate for the treatment of critical patients. Clinical workers should pay attention to the risk factors of COVID-19 critical disease and death, identify critical patients early, allocate medical resources rationally and timely adjust the treatment plan to enhance the efficacy and reduce the risk of death. In this article, we analyzed the clinical characteristics of COVID-19 patients with critical/mortal illness and non-critical illness in 13 literatures with 3027 patients, to identify the risk factors for COVID-19 patients to develop critical disease or death, in order to effectively predict the progression of the disease, make early treatment response and allocate medical resources in a better way.

Data and Methods

Search strategy and selection criteria

This systematic review and meta-analysis is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)Statement.

We selected relevant studies published between Jan 1, 2020 and Mar 20, 2020, by searching Pubmed, Embase, Web of Science, and CNKI. We applied no language restrictions. The search terms and relative variants were as follows: severe acute respiratory syndrome coronavirus 2 OR Wuhan coronavirus OR Wuhan seafood market pneumonia virus OR COVID-19 OR COVID19 OR coronavirus disease 2019 virus OR SARS-CoV-2 OR SARS2 OR 2019-nCoV OR 2019 novel coronavirus AND Mortalities OR Mortality OR Fatality OR Death OR acute respiratory distress syndrome (ARDS) OR ICU. We also reviewed the references of included articles to guarantee the comprehensiveness and accuracy of our research.

The inclusion criterions for the 13 articles are as follows: (1) groups involving critical illness or death and non-critical illness; (2) patients should be confirmed to have been infected by 2019 novel coronavirus; (3) study designs included randomized controlled trials, nonrandomized controlled trials, case-control studies, cohort studies, cross-sectional studies, and also case reports; (4) at least one outcome reported among demographical characteristics, comorbidities, clinical manifestations or laboratory examinations; (5) study sample was larger than 20. We excluded duplicate reports, abstracts from conferences and commentary articles.

Data Extraction and Quality Assessment

Data extraction and the evaluation of literature quality were conducted independently by 2 investigators (Z.Z.H&T.W.L). Microsoft Excel database was used to record all available information, including baseline details, comorbidities, clinical manifestations and laboratory examinations. Any disagreement was resolved by another investigator (P.F). The MINORS[4] was used to assess bias risk.

Statistical Analysis of Data

All analyses were performed using Microsoft Excel, State software version 15.0 and RevMan software version 5.3. The results of the included studies were performed with fixed-effect models (Mantel–Haenszel method) or random-effect models in cases of significant heterogeneity between studies. We used the I^2 statistics to assess the magnitude of heterogeneity: 25%, 50%, and 75% represented low, moderate, and high degrees of heterogeneity, respectively. The chosen of the proper effect model was based on the analysis results: the fixed effect model was used if $I^2 \leq 50\%$ and the random effect model was used if $I^2 > 50\%$. If there was statistical heterogeneity among the results, a further sensitivity analysis was conducted to determine the source of heterogeneity. After the significant clinical heterogeneity was excluded, the randomized effects model was used for meta-analysis. $P < 0.05$ was considered as statistical significance.

Results

Research Selection and Quality Assessment

Based on the previous search strategy, 523 studies were searched from the online database. After deleting duplicate records, a total of 343 records were retained. Then, 311 articles were excluded by the titles and abstracts, and 19 of the remaining 32 articles were deleted for various reasons. The last 13 articles were included in the meta-analysis. Finally, a total of 13 studies with 3027 patients were included[5–17] (Fig. 1).All of the selected studies were published in 2020 with different sample patient sizes that ranged from 27 to 1099 patients. Clinical outcome was defined as ICU admission in 3 studies,[6,8,12] refractory in 1 study,[7] severity in 2 studies,[9,11] $Sp0_2 < 90\%$in 1 study,[13] onset of ARDS in 1 study[14] and death in the remaining investigation.[5,10,15–17] The risk of bias and applicability concerns included studies are showed in Table 1. Over all, none of the studies was considered to be seriously flawed according to the MINORS assessment. The 13 included studies scored between 18 and 21. All studies were considered to have a low risk of bias for selection.

Demographical characteristics

The demographical characteristics of the included studies are shown in Table 2. The results from the 13 included studies (with a total amount of 3027 patients) showed that the proportion of male was significant higher in critical/death group compared to the non-critical group [male: OR = 1.77, 95% CI (1.43, 2.19), $P < 0.00001$] (Fig. 2).

The median ages ranged from 49 to 70.5 years old in the critical/mortal group across the enrolled studies. The median ages ranged from 37 to 62 years old in the non-critical group (Table1).

Please cite this article as: Z. Zheng, F. Peng and B. Xu et al., Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis, Journal of Infection, https://doi.org/10.1016/j.jinf.2020.04.021

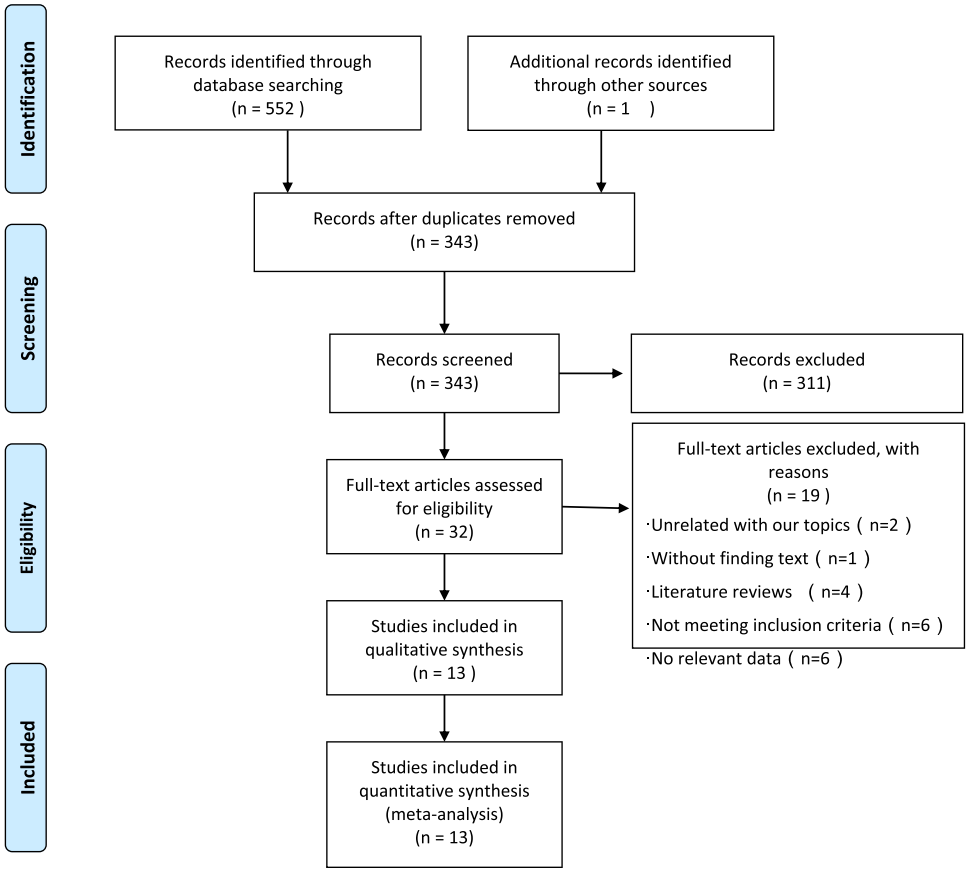


Fig. 1. Flow diagram of the study selection process.

Table 1
MINORS rating scale: ①A clearly stated aim;②Inclusion of consecutive patients; ③Prospective collection of data;④ Endpoints appropriate to the aim of the study;⑤Unbiased assessment of the study endpoint;⑥Follow-up period appropriate to the aim of the study;⑦Loss to follow up less than 5%;⑧ Prospective calculation of the study size.⑨Appropriate selection of control group;⑩Synchronization of control group; ⑪Baseline comparable between groups ⑫Appropriately statistical analysis. The global ideal score being 24 for comparative studies.

Study	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	⑪	⑫	Score
Guan WJ	2	2	2	2	2	0	0	0	2	2	2	2	18
Huang C	2	2	2	2	2	1	2	0	2	2	2	2	21
Mo P	2	2	2	2	2	0	2	0	2	2	2	2	20
Peng YD	2	2	2	2	2	0	1	0	2	2	2	2	19
Shi Y	2	2	2	2	2	1	2	0	2	2	2	2	21
Tang N	2	2	2	2	2	1	0	0	2	2	2	2	19
Tian S	2	2	2	2	2	0	1	0	2	2	2	2	19
Wang D	2	2	2	2	2	1	2	0	2	2	2	2	21
Wang Z	2	2	2	2	2	1	2	0	2	2	2	2	21
Wu C	2	2	2	2	2	1	0	0	2	2	2	2	19
Yang X	2	2	2	2	2	0	1	0	2	2	2	2	19
Yuan ML	2	2	2	2	2	1	2	0	2	2	2	2	21
Zhou F	2	2	2	2	2	1	2	0	2	2	2	2	18

The median ages were generally higher in critical/death group compared to the non-critical group. Furthermore, age over 65 years was analyzed as a subgroup by Guan[8] and Tian[11] (with a total amount of 1273 patients). Meta-analysis showed that the proportion of patients older than 65 years was higher in critical/death group compared to the non-critical group. [age over 65 years old: OR = 6.01, 95% CI (3.95, 9.16), P < 0.00001] (Fig. 2).

Five studies showed that the proportion of current smoker was statistically significant higher in critical/mortal group compared to the non-critical group [current smoking: OR = 2.04, 95% CI (1.32, 3.15), P = 0.0006] (Fig. 2).

There was no heterogeneity in the estimates of male, age over 65 years old and current smoking among the identified studies with $I^2 = 0$.

Comorbidities

The comorbidities of patients of the included studies are shown in Table 3. We then compared the difference of the prevalence of the comorbidities between critical/mortal patients and non-critical patients. For diabetes, cardiovascular disease and respiratory disease, the heterogeneity test results were calculated as $I^2 = 45\%$, 37%

Table 2
Demographics of the included studies.

Study	Year	Research type	Country	Number of patients n		Age median, y		Male n (%)		Current Smoking n (%)	
				Critical/Mortal	Non-critical	Critical/Mortal	Non-critical	Critical/Mortal	Non-critical	Critical/Mortal	Non-critical
Guan WJ et al. [5]	2020	Retrospective study	China	67	1032	63	46	45(67.2%)	592(57.5%)	17(25.4%)	120(11.7%)
Huang C et al.[6]	2020	Retrospective study	China	13	28	49	49	11(84.6%)	19(67.9%)	0	3(10.7%)
Mo P et al.[7]	2020	Retrospective study	China	85	70	61	46	55(64.7%)	31(44.3%)	4(4.7%)	2(2.9%)
Peng YD et al.[8]	2020	Retrospective study	China	16	96	57.5	62	9(56.3%)	44(45.8%)	–	–
Shi Y et al.[9]	2020	Retrospective study	China	49	438	56	45	36(73.5%)	223(50.9%)	6(12.2%)	34(7.8%)
Tang N et al.[10]	2020	Retrospective study	China	21	162	64	52.4	16(76.2%)	82(50.6%)	–	–
Tian S et al.[11]	2020	Retrospective study	China	46	216	61.4	44.5	26(56.5%)	101(46.8%)	–	–
Wang D et al.[12]	2020	Retrospective study	China	36	102	66	51	22(61.1%)	53(52.0%)	–	–
Wang Z et al.[13]	2020	Retrospective study	China	14	55	70.5	37.0	7(50.0%)	25(45.5%)	–	–
Wu C et al.[14]	2020	Retrospective study	China	84	117	58.5	48	60(71.4%)	68(58.1%)	–	–
Yang X et al.[15]	2020	Retrospective study	China	32	20	64.6	51.9	21(65.6%)	14(70.0%)	–	–
Yuan ML et al.[16]	2020	Retrospective study	China	10	17	68	55	4(40.0%)	8(47.1%)	–	–
Zhou F et al.[17]	2020	Retrospective study	China	54	147	69	52	38(70.4%)	81(55.1%)	5(9.3%)	6(4.1%)

Table 3
Comorbidities of patients of the included studies.

Study	Diabetes n (%)		Hypertensionn (%)		Cardiovascular disease n (%)		Respiratory disease n (%)		Malignancy n (%)	
	Critical/Mortal	Non-critical	Critical/Mortal	Non-critical	Critical/Mortal	Non-critical	Critical/Mortal	Non-critical	Critical	Non-critical
Guan WJ	67 (26.9%)	63 (6.1%)	24 (35.8%)	141 (13.7%)	6 (9.0%)	21 (2.0%)	7 (10.4%)	5 (0.5%)	1 (1.5%)	9 (0.9%)
Huang C	13 (7.7%)	7 (25.0%)	2 (15.4%)	4 (14.3%)	3 (23.1%)	3 (10.7%)	1 (7.7%)	0	0	1 (3.6%)
Mo P	85 (14.1%)	3 (4.3%)	22 (25.9%)	15 (21.4%)	14 (16.5%)	0	4 (4.7%)	0	5 (5.9%)	2 (2.9%)
Peng YD	16 (25.0%)	19 (19.8%)	10 (62.5%)	82 (85.4%)	–	–	–	–	–	–
Shi Y	49 (14.3%)	22 (5.0%)	26 (53.1%)	73 (16.7%)	4 (8.2%)	7 (1.6%)	–	–	2 (4.1%)	3 (0.7%)
Wang D	36 (22.2%)	6 (5.9%)	21 (58.3%)	22 (21.6%)	9 (25.0%)	11 (10.8%)	3 (8.3%)	1 (1.0%)	4 (11.1%)	6 (5.9%)
Wang Z	14 (42.9%)	1 (1.8%)	5 (35.7%)	4 (7.3%)	5 (35.7%)	3 (5.5%)	2 (14.3%)	4 (7.3%)	1 (7.1%)	3 (5.5%)
Wu C	84 (19.0%)	6 (5.1%)	23 (27.4%)	16 (13.7%)	5 (6.0%)	3 (2.6%)	–	–	–	–
Yang X	32 (21.9%)	2 (10.0%)	–	–	3 (9.4%)	2 (10.0%)	2 (6.3%)	2 (10.0%)	1 (3.1%)	1 (5.0%)
Yuan ML	10 (60.0%)	0	5 (50.0%)	0	3 (30.0%)	0	–	–	0	1 (5.9%)
Zhou F	54 (31.5%)	19 (12.9%)	26 (48.1%)	32 (21.8%)	13 (24.1%)	2 (1.4%)	4 (7.4%)	2 (1.4%)	0	2 (1.4%)

and 50%. Thus, the fixed-effect model was used for further analyses. The proportion of diabetes, cardiovascular disease and respiratory disease was statistically significant higher in critical/mortal group compared to the non-critical group [diabetes:0R = 3.68, 95% CI (2.68, 5.03), P < 0.00001;cardiovascular disease: OR = 5.19, 95% CI (3.25, 8.29), P < 0.00001; respiratory disease: OR = 5.15, 95% CI (2.51, 10.57), P < 0.00001](Fig. 3). For hypertension, the heterogeneity test showed that $I^2 = 72\%$. Given that the severity of illness and severity of epidemic might contribute to the heterogeneity, we classified the studies into two subgroups according to whether the study site was located in Wuhan. However, heterogeneity still exists. So the random effect model was used. The result indicated a higher proportion of hypertension in critical/mortal group[OR = 2.72, 95% CI (1.60,4.64), P = 0.0002] (Fig. 3).For malignancy, the fixed-effect model($I^2 = 0$) meta-analysis showed that the proportion of malignancy was higher in critical/death group yet without statistical significance[OR = 1.60, 95% CI (0.81, 3.18), P = 0.18] (Fig. 3).

Clinical Manifestation

The study of clinical manifestation included 13 studies, a total of 3025 cases. The clinical features are showed in Table 4 and the results of meta-analysis are showed in Table 5. For fever (temperature ≥ 37.3°C), the proportion of fever was statistically lower in critical/mortal group [0R = 0.56, 95% CI (0.38, 0.82), P = 0.003]. The proportion of headache and myalgia/arthralgia were lower in critical/mortal group compared to the non-critical group [headache: OR = 0.82, 95% CI (0.50, 1.36), P = 0.45; myalgia/arthralgia: OR = 0.77, 95% CI (0.58, 1.04), P = 0.09] but there was no statistical significance. For cough, sputum production, fatigue, diarrhea and nausea/ vomiting, the proportion of them was higher in critical/mortal patients[cough: OR = 1.08, 95% CI (0.85,1.38), P = 0.52; sputum production: OR = 1.14, 95% CI (0.84, 1.54), P = 0.39; fatigue: OR = 1.13, 95% CI (0.88, 1.44), P = 0.34;diarrhea: OR = 1.41, 95% CI (0.82, 2.43), P = 0.22;nausea/ vomiting: OR = 1.32, 95% CI (0.72, 2.42), P = 0.37], however, without statistical significance. For shortness of breath/ dyspnea, the proportion of this clinical manifestation was statistically significant higher in critical/mortal group[0R = 4.16, 95% CI (3.13, 5.53), P < 0.00001].

Laboratory examination

The laboratory examination of the included studies (4 studies, a total of 1286 cases) are shown in Table 6. For "WBC $< 4 \times 10^9$per L", the analysis results of the fixed effect-model ($I^2 = 0$)showed that the proportion of "WBC $< 4 \times 10^9$per L" was statistically lower in critical/mortal group[0R = 0.30, 95% CI (0.17, 0.51), P < 0.00001] (Fig. 4). For "AST > 40U/L", "Cr ≥ 133μmol/L" and

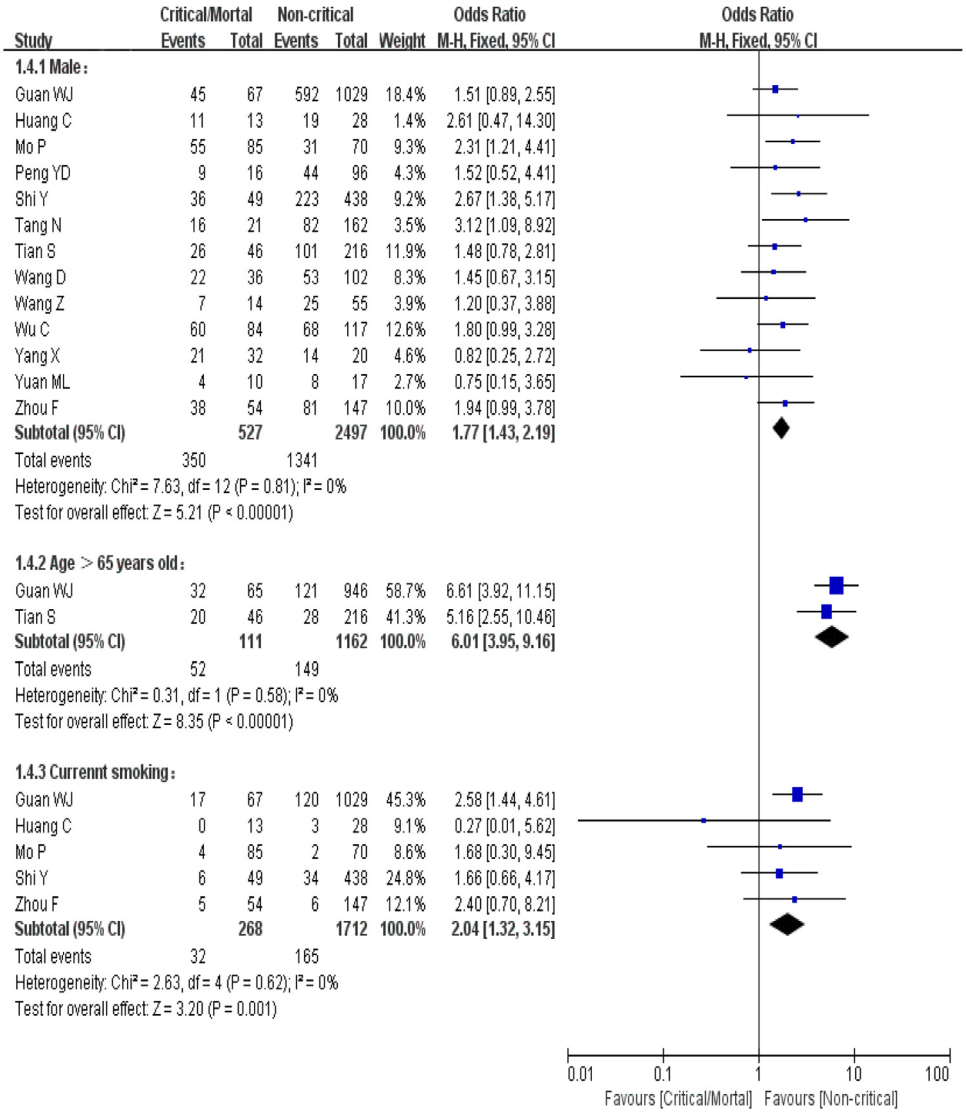


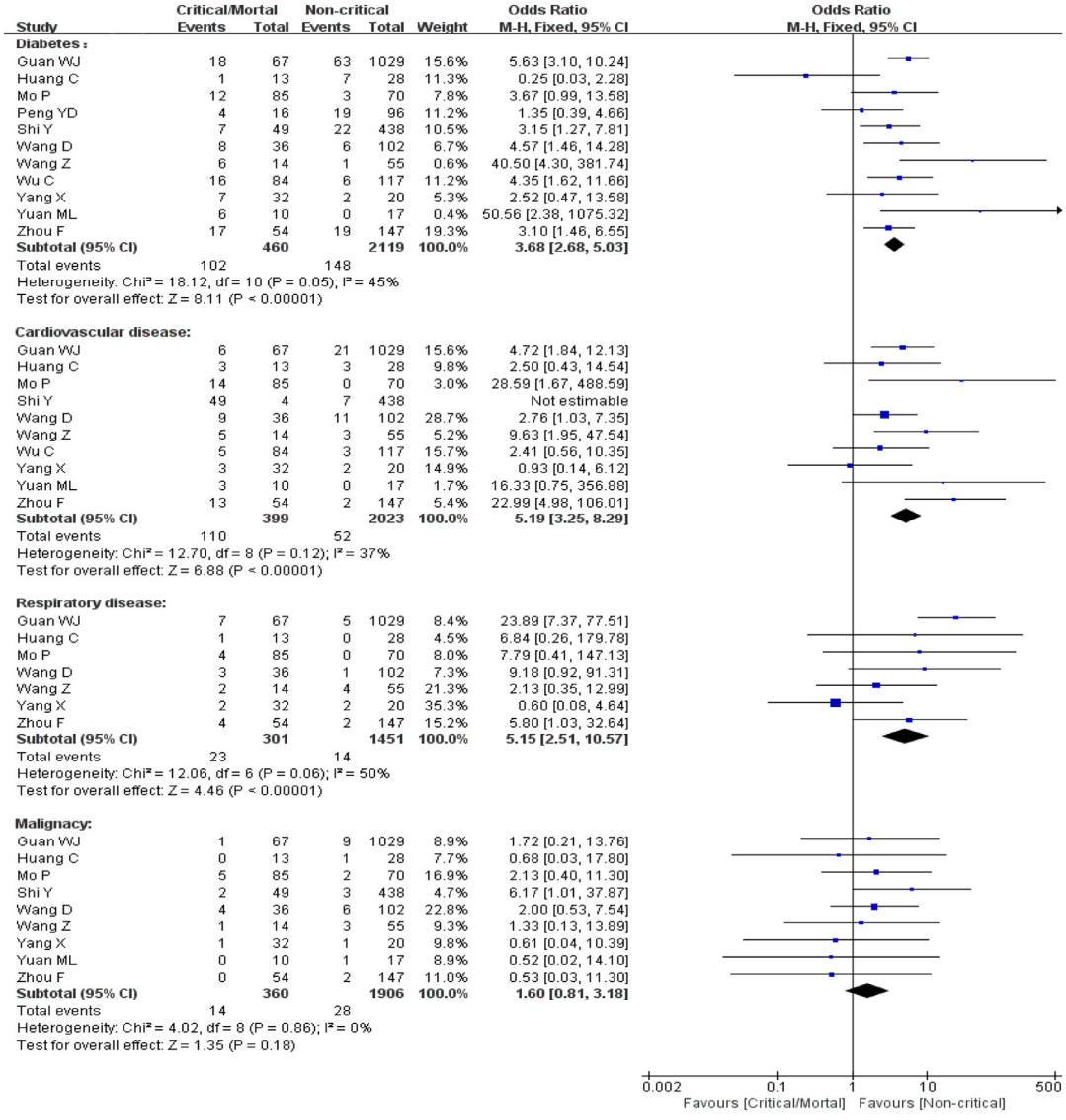
Fig. 2. Meta-analysis for male, age>65 years old and current smoking in COVID-19 cases. Heterogeneity analysis was carried out using Q test, the among studies variation (I^2 index). Forest plots depict the comparison of the incidences of male, age>65 years old and current smoking in critical/mortal and non-critical patients.

"hs-cTnI > 28 pg/mL", the heterogeneity test results were calculated as $I^2 = 0\%$, 0% and 34%. We used the fixed-effect model for further analyses. The proportion of "AST > 40U/L", "Cr ≥ 133μmol/L" and "hs-cTnI > 28 pg/mL" was statistically significant higher in critical/death group compared to the non-critical group ["AST > 40U/L":0R = 4.00, 95% CI (2.46, 6.52), P < 0.00001;"Cr ≥ 133μmol/L":0R = 5.30, 95% CI (2.19, 12.83), P = 0.0002; "hs-cTnI > 28 pg/mL":0R = 43.24, 95% CI (9.92, 188.49), P < 0.00001] (Figure 5). The I^2 value of "PCT > 0.5 ng/mL", "LDH > 245U/L" and "D-dimer > 0.5mg/L" was, respectively, 52%, 59% and 72%. So the random effect model was used. The results indicated a higher proportion of "PCT > 0.5 ng/mL", "LDH > 245U/L" and "D-dimer > 0.5mg/L" in critical/mortal patients with statistical significance ["PCT > 0.5 ng/mL": OR = 43.24, 95% CI (9.92, 188.49), P< 0.00001; "LDH > 245U/L":OR=43.24, 95% CI (9.92, 188.49), P < 0.00001;"D-dimer0.5mg/L": OR = 43.24, 95% CI (9.92, 188.49), P < 0.00001] (Figure 5).

Discussion

The results of meta-analysis showed that male, aged over 65 and smoking patients might face a greater risk of developing into the critical or mortal condition and the comorbidities such as hypertension, diabetes, cardiovascular disease or respiratory diseases could also greatly affect the prognosis of the COVID-19. We found that patients with shortness of breath/dyspnea were more likely to develop into critical illness or even die, but patients with fever progressed better than those without fever. Laboratory examinations such as WBC, AST, Cr, hs-cTnI, PCT, LDH and D-dimer could imply the progression of COVID-19.

Coronavirus is an enveloped, non-segmented, single-stranded RNA virus.[18] At present, six human coronaviruses have been identified. And the SARS-CoV-2, which isolated from the lower respiratory tract of pneumonia patients with unknown causes in Wuhan, is identified as the seventh human coronaviruses.[19] SARS-CoV-2 attacks the alveolar epithelial cells via angiotensin-converting enzyme 2 (ACE2). ACE2 is the ACE of isozyme, mainly distributed in cardiovascular, kidneys, testes, lung and colon, and other organizations.[20] The main role of ACE2 is to incise Ang II to generate Ang 1-7, which mediates the protective effects of vasodilation, anti-inflammatory and anti-proliferation, to antagonize Ang II-induced vascular smooth muscle contraction, cell proliferation, fibrosis promotion and vascular inflammation.[21–23] When SARS-CoV-2 binds to ACE2 receptor on the surface of alveolar epithelial cells, The



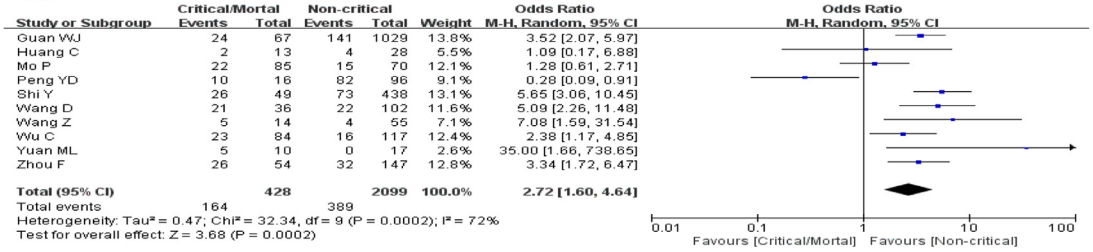


Fig. 3. Meta-analysis for comorbidities in COVID-19 cases. Fix-effect model for diabetes, cardiovascular disease, respiratory disease and malignancy. Random-effect model for hypertension. Heterogeneity analysis was carried out using Q test, the among studies variation (I^2 index). Forest plots depict the comparison of the incidences of the 5 diseases in critical/mortal and non-critical patients.

expression of ACE2 in alveolar epithelial cells is down-regulated by mechanisms such as internalization, shedding and viral replication.[24] Then the increased concentration of Ang II leads to inflammatory response, and exudation of neutrophils, macrophages and fibrinous, resulting in loss of pulmonary ventilation function and difficulty in maintaining oxygenation.[25] At the same time, viral infection will cause the imbalance of T helper-1 and T helper-2 responses, and induce an inflammatory storm by increasing the levels of inflammatory factors such as interleukin-4, interleukin-10 and interleukin-6.[26] Inflammatory storm in critical patients releases cytokines, causing systemic immune injury, which may be an important cause of multiple organ failure and even death.[27]

Studies have found that women are less susceptible to viral infection than men, possibly because of the protection of X chromosome and sex hormones, which play an important role in innate and adaptive immunity.[28] At the same time, men tend to be as-

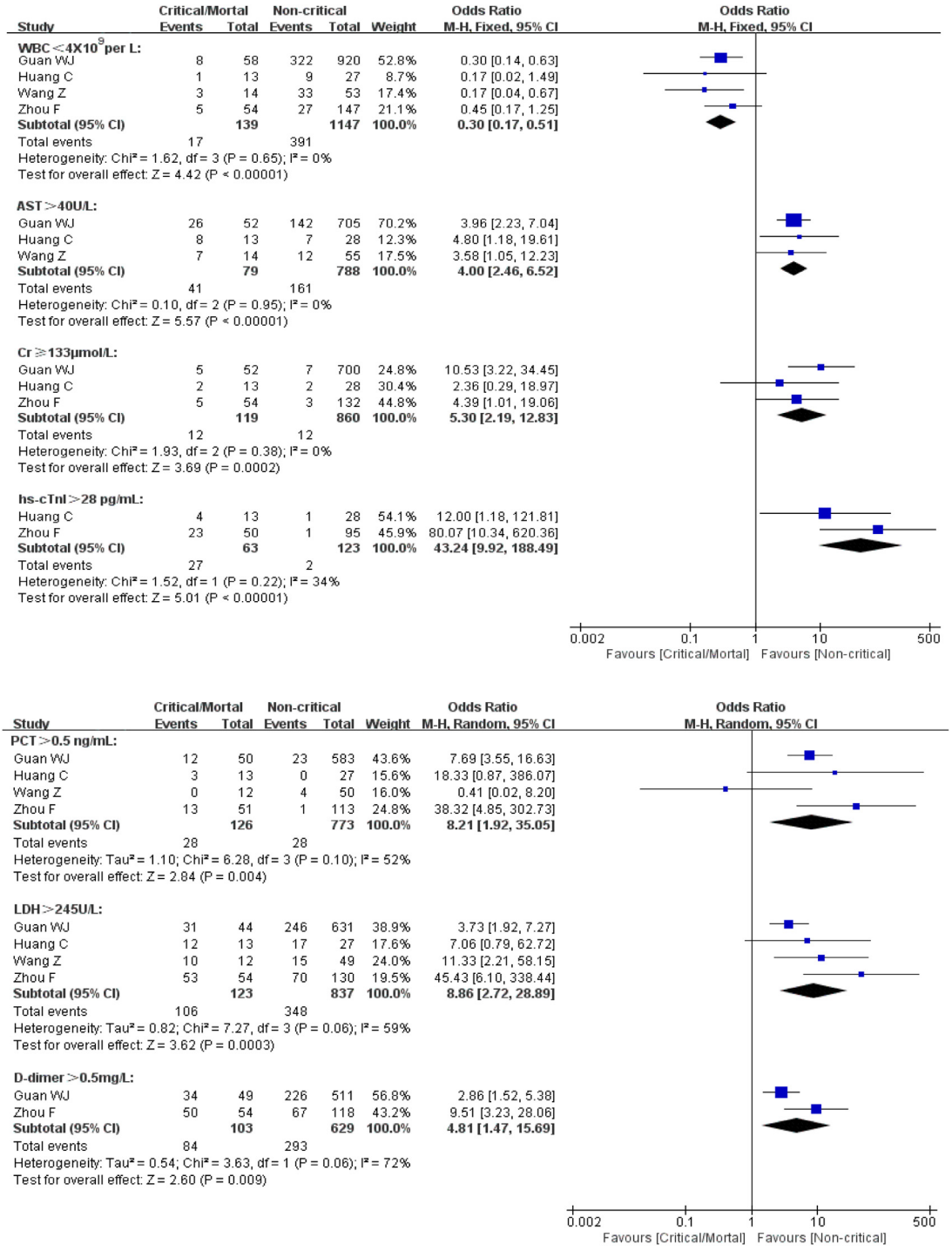


Fig. 4. Meta-analysis for laboratory examination in COVID-19 cases. Fix-effect model for "WBC $< 4 \times 10^9$per L" "AST > 40U/L" "Cr $\geq 133\mu$mol/L" and "hs-cTnI > 28 pg/mL". Random-effect model for "PCT > 0.5 ng/mL" "LDH > 245U/L" and "D-dimer > 0.5mg/L". Heterogeneity analysis was carried out using Q test, the among studies variation (I^2 index). Forest plots depict the comparison of the incidences of the laboratory examination in critical/mortal and non-critical patients.

sociated with bad lifestyle habits such as smoking and underlying diseases. As a result, the majority of critical or mortal patients are male. As the body's immunity declines with age, elderly patients are more likely to develop critical illness or even die. Therefore, when the patient is male, over 65 years old and smoking, the patient has a higher risk of developing critical illness or death.

When patients are combined with basic diseases such as diabetes and hypertension, the body is in a state of stress for a long time and the immunity tends to be low. Moreover, the long-term history of diabetes and hypertension will damage the vascular structure, and it is more likely to develop into critical disease in infection. Patients with chronic heart disease are more likely to be infected due to their weakened heart function and low immunity. When infected with SARS-CoV-2, they are more likely to have acute cardiovascular events and develop into severe diseases. When the patient has previous respiratory diseases such as chronic obstructive pulmonary disease, the patient's lung function is damaged. They have lower resistance to the virus and are prone to

Please cite this article as: Z. Zheng, F. Peng and B. Xu et al., Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis, Journal of Infection, https://doi.org/10.1016/j.jinf.2020.04.021

Table 4
Clinical manifestation of patients of the included studies.

Study	Fever (%)		Cough (%)		Sputum production (%)		Dyspnea (%)		Headache (%)		Nausea or vomiting (%)		Fatigue (%)		Myalgia or arthralgia (%)		Diarrhea (%)	
	Critical/Mortal	Non-critical	Critical/Mortal	Non-critical	Critical/Mortal	Non-critical	Critical/Mortal	Non-critical	Critical/Mortal	Non-critical	Critical/Mortal	Non-critical	Critical/Mortal	Non-critical	Critical/Mortal	Non-critical	Critical	Non-critical
Guan WJ	88.1%	89.0%	68.7%	67.9%	29.9%	34.0%	53.7%	16.4%	11.9%	13.8%	4.5%	5.1%	32.8%	38.6%	9.0%	15.4%	6.0%	3.7%
Huang C	100.0%	96.4%	84.6%	71.4%	38.5%	23.1%	92.3%	37.0%	0.0%	12.0%	–	–	53.8%	39.3%	53.8%	39.3%	0.0%	3.6%
Mo P	74.1%	90.0%	63.5%	61.4%	0.0%	0.0%	41.2%	21.4%	5.9%	4.3%	2.4%	2.9%	38.8%	38.6%	25.9%	40.0%	5.9%	2.9%
Peng YD	100.0%	88.5%	75.0%	66.7%	0.0%	0.0%	18.8%	10.4%	–	–	–	–	56.3%	64.6%	56.3%	64.6%	12.5%	13.5%
Tian S	80.4%	82.4%	54.3%	44.0%	–	–	32.6%	1.4%	6.5%	6.5%	–	–	32.6%	25.0%	–	–	–	–
Wang D	100.0%	98.0%	58.3%	59.8%	22.2%	28.4%	63.9%	19.6%	8.3%	5.9%	19.4%	11.8%	80.6%	65.7%	33.3%	35.3%	16.7%	7.8%
Wang Z	–	–	57.1%	54.5%	28.6%	29.1%	50.0%	23.6%	0.0%	18.2%	7.1%	3.6%	50.0%	40.0%	14.3%	34.5%	14.3%	14.5%
Wu C	92.9%	94.0%	81.0%	81.2%	48.8%	35.9%	59.5%	25.6%	–	–	–	–	32.1%	32.5%	32.1%	32.5%	–	–
Yang X	96.9%	100.0%	78.1%	75.0%	–	–	65.6%	60.0%	6.3%	5.0%	3.1%	5.0%	–	–	12.5%	10.0%	–	–
Yuan ML	60.0%	88.2%	50.0%	64.7%	–	–	100.0%	5.9%	–	–	–	–	–	–	10.0%	11.8%	–	–
Zhou F	94.4%	87.8%	72.2%	76.2%	25.9%	20.4%	63.0%	15.0%	–	–	5.6%	2.7%	27.8%	19.7%	14.8%	14.3%	3.7%	4.8%

Table 5
Results of meta-analysis of the clinical manifestation.

Clinical manifestation	OR	95%CI	P value
Fever	0.56	0.38-0.82	P=0.003
Headache	0.82	0.50-1.36	0.45
Myalgia or arthralgia	0.77	0.58-1.04	0.09
Cough	1.08	0.85-1.38	0.52
Sputum production	1.14	0.84-1.54	0.39
Fatigue	1.13	0.88-1.44	0.34
Diarrhea	1.41	0.82-2.43	0.22
Nausea or vomiting	1.32	0.72-2.42	0.37
Shortness of breath/Dyspnea	4.16	3.1- 5.53	<0.00001

developing ARDS. Thus, underlying diseases such as diabetes, hypertension, cardiovascular disease or respiratory disease are risk factors for disease progression. This is consistent with the analytical results in this paper.

The common clinical manifestations of COVID-19 patients are fever, cough and sputum.[6] When the patient's immune response is low, it may manifest as normal body temperature. Shortness of breath or dyspnea suggests poor lung function and lacking of oxygen. Therefore, when the patient is found to have difficulty in breathing or no fever, it is necessary to be alert for further deterioration of the patient's condition.

Viral infection will cause inflammation in human body. Various inflammatory factors produced by the inflammatory storm can cause systemic immune damage and even cause multi-organ failure. When the patient's laboratory indicator shows PCT $>$ 0.5 ng/mL, there is a higher risk of progression to critical illness. Study showed that the total number of WBC in the early stage of the disease is normal or reduced.[6] However, WBC $< 4 \times 10^9$/L means a better clinical outcome in this paper. Therefore, when the total number of WBC is found to be higher than the previous one, the patient may be associated with other infections that aggravated the disease. At this moment, the clinicians might pay more attentions to these patients and improve treatment. When AST $>$ 40U/L, LDH $>$ 245U/L and Cr $\geq$ 133 mol/L, it indicates that the liver and kidney dysfunctions have been involved, and corresponding treatments should be taken in time to prevent further deterioration of the disease. Current studies have shown that up to 20% of covid-19 patients have abnormal coagulation function.[6] Monocytes and tissue cells are activated after injury, causing the release of cytokines and the expression of tissue factors, and finally causing the hypercoagulability of blood. It will increase the risk of thrombosis and more likely to cause ischemia and hypoxia due to the embolization of the viscera, which leads to the progression of the disease to critical disease or death. When the D-dimer $>$ 0.5mg/L, it indicates the hypercoagulability of blood and suggests the deterioration of patients. At the same time, SARS-CoV-2 can cause myocardial injury by direct and/or indirect action. The direct injury is to infect cardiomyocytes by identifying ACE2 receptor, while the indirect injury may be caused by inflammatory storm inducing by immune response and/or oxygen supply imbalance inducing by acute respiratory distress syndrome. When hs-cTnI $>$ 28pg/mL, it strongly suggests the possibility of further deterioration of the patient's condition.

The quality of the literature included in this study is high, the analysis is rigorous, and the conclusions drawn by the study are highly credible. However, this meta-analysis also has some limitations: (a) most of the studies included in this meta-analysis were cross-sectional studies with insufficient demonstration ability. (b) most of the patients in our meta-analysis were Chinese, and our aim was to use the findings of this study to predict the overall profile of patients, including other countries and races; (c) more detailed patient information, such as iconograph and oxygen at ion index, was not available in most studies at the time of analysis.

Please cite this article as: Z. Zheng, F. Peng and B. Xu et al., Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis, Journal of Infection, https://doi.org/10.1016/j.jinf.2020.04.021

Table 6
Laboratory examination of patients of the included studies.

Study	White blood cells $<4\times10^9$ per L n (%)		Aspartate aminotransferase>40U/L n (%)		Creatinine $\geq133\mu$mol/L n (%)		Hypersensitive troponin I>28 pg/mL n (%)		Procalcitonin >0.5 ng/mL n (%)		Lactate dehydrogenase >245U/L n (%)		D-dimer>0.5mg/L n (%)	
	Critical/Mortal	Non-critical	Critical/Mortal	Non-Critical	Critical/Mortal	Non-critical	Critical/Mortal	Non-critical	Critical/Mortal	Non-critical	Critical/Mortal	Non-critical	Critical/Mortal	Non-critical
Guan WJ	8 (13.8%)	322 (35.0%)	26 (50.0%)	26 (50.0%)	5 (9.6%)	7 (1.0%)	–	–	12 (24.0%)	23 (3.9%)	31 (70.5%)	246 (39.0%)	34 (69.4%)	226 (44.2%)
Huang C	1 (7.7%)	9 (33.3%)	8 (61.5%)	8 (61.5%)	2 (15.4%)	2 (7.1%)	4 (30.8%)	1 (3.6%)	3 (23.1%)	0	1 (92.3%)	17 (63.0%)	–	–
Wang Z	3 (21.4%)	33 (62.3%)	7 (50.0%)	7 (50.0%)	–	–	–	–	0	4 (8.0%)	2 (83.3%)	15 (30.6%)	–	–
Zhou F	5 (9.3%)	27 (18.4%)	–	–	5 (9.3%)	3 (2.3%)	23 (46.0%)	1 (1.1%)	13 (25.5%)	1 (0.9%)	1 (98.1%)	70 (53.8%)	50 (92.6%)	67 (56.8%)

The conclusions of this meta-analysis still need to be verified by more relevant studies with more careful design, more rigorous execution, and larger sample size.

This study analyzed the risk factors for progression to critical illness or death in COVID-19 patients to help assessing patient status and identify critical patients early. Pay close attention to these risk factors, and when relevant laboratory risk value appears, timely and personalized treatment regimens are needed to enhance the efficacy and reduce the risk of death.

References

1. Chan J F, Yuan S, Kok K H, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster.[J]. *Lancet* 2020;**395**(10223):514–23. doi:10.1016/s0140-6736(20)30154-9.
2. Zhou P, Yang X L, Wang X G, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin.[J]. *Nature* 2020;**579**(7798):270–3. doi:10.1038/s41586-020-2012-7.
3. Lai C C, Wang C Y, Wang Y H, et al. Global epidemiology of coronavirus disease 2019: disease incidence, daily cumulative index, mortality, and their association with country healthcare resources and economic status.[J]. *Int J Antimicrob Agents* 2020:105946. doi:10.1016/j.ijantimicag.2020.105946.
4. Slim K, Nini E, Forestier D, et al. Methodological index for non-randomized studies (minors): development and validation of a new instrument.[J]. *ANZ J Surg* 2003;**73**(9):712–16. doi:10.1046/j.1445-2197.2003.02748.x.
5. Guan W J, Ni Z Y, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China.[J]. *N Engl J Med* 2020. doi:10.1056/NEJMoa2002032.
6. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China.[J]. *Lancet* 2020;**395**(10223):497–506. doi:10.1016/s0140-6736(20)30183-5.
7. Mo P, Xing Y, Xiao Y, et al. Clinical characteristics of refractory COVID-19 pneumonia in Wuhan, China.[J]. *Clin Infect Dis* 2020. doi:10.1093/cid/ciaa270.
8. Peng Y D, Meng K, Guan H Q, et al. [Clinical characteristics and outcomes of 112 cardiovascular disease patients infected by 2019-nCoV].[J]. *Zhonghua Xin Xue Guan Bing Za Zhi* 2020;**48**(0):E004. doi:10.3760/cma.j.cn112148-20200220-00105.
9. Shi Y, Yu X, Zhao H, et al. Host susceptibility to severe COVID-19 and establishment of a host risk score: findings of 487 cases outside Wuhan.[J]. *Crit Care* 2020;**24**(1):108. doi:10.1186/s13054-020-2833-7.
10. Tang N, Li D, Wang X, et al. Abnormal coagulation parameters are associated with poor prognosis in patients with novel coronavirus pneumonia.[J]. *J Thromb Haemost* 2020. doi:10.1111/jth.14768.
11. Tian S, Hu N, Lou J, et al. Characteristics of COVID-19 infection in Beijing.[J]. *J Infect* 2020;**80**(4):401–6. doi:10.1016/j.jinf.2020.02.018.
12. Wang D, Hu B, Hu C, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China.[J]. *Jama* 2020. doi:10.1001/jama.2020.1585.
13. Wang Z, Yang B, Li Q, et al. Clinical Features of 69 Cases with Coronavirus Disease 2019 in Wuhan, China.[J]. *Clin Infect Dis* 2020. doi:10.1093/cid/ciaa272.
14. Wu C, Chen X, Cai Y, et al. Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China.[J]. *JAMA Intern Med* 2020. doi:10.1001/jamainternmed.2020.0994.
15. Yang X, Yu Y, Xu J, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study.[J]. *Lancet Respir Med* 2020. doi:10.1016/s2213-2600(20)30079-5.
16. Yuan M, Yin W, Tao Z, et al. Association of radiologic findings with mortality of patients infected with 2019 novel coronavirus in Wuhan, China.[J]. *PLoS One* 2020;**15**(3):e0230548. doi:10.1371/journal.pone.0230548.
17. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study.[J]. *Lancet* 2020 doi:10.1016/s0140-6736(20)30566-3.
18. Fehr A R, Perlman S. Coronaviruses: an overview of their replication and pathogenesis.[J]. *Methods Mol Biol* 2015;**1282**:1–23. doi:10.1007/978-1-4939-2438-7_1.
19. Ding Q, Lu P, Fan Y, et al. The clinical characteristics of pneumonia patients coinfected with 2019 novel coronavirus and influenza virus in Wuhan, China.[J]. *J Med Virol* 2020:10 doi:10.1002/jmv.25781.
20. Tipnis S R, Hooper N M, Hyde R, et al. A human homolog of angiotensin-converting enzyme. Cloning and functional expression as a captopril-insensitive carboxypeptidase.[J]. *J Biol Chem* 2000;**275**(43):33238–43. doi:10.1074/jbc.M002615200.
21. Griendling K K, Sorescu D, Lassegue B, et al. Modulation of protein kinase activity and gene expression by reactive oxygen species and their role in vascular physiology and pathophysiology.[J]. *Arterioscler Thromb Vasc Biol* 2000;**20**(10):2175–83. doi:10.1161/01.atv.20.10.2175.
22. Vickers C, Hales P, Kaushik V, et al. Hydrolysis of biological peptides by human angiotensin-converting enzyme-related carboxypeptidase.[J]. *J Biol Chem* 2002;**277**(17):14838–43. doi:10.1074/jbc.M200581200.

23. Santos R A, Simoes e Silva AC, Maric C, et al. Angiotensin-(1-7) is an endogenous ligand for the G protein-coupled receptor Mas.[J]. *Proc Natl Acad Sci U S A* 2003;**100**(14):8258–63. doi:10.1073/pnas.1432869100.
24. Glowacka I, Bertram S, Herzog P, et al. Differential downregulation of ACE2 by the spike proteins of severe acute respiratory syndrome coronavirus and human coronavirus NL63.[J]. *J Virol* 2010;**84**(2):1198–205. doi:10.1128/jvi.01248-09.
25. Xu Z, Shi L, Wang Y, et al. Pathological findings of COVID-19 associated with acute respiratory distress syndrome.[J]. *Lancet Respir Med* 2020;**8**(4):420–2. doi:10.1016/s2213-2600(20)30076-x.
26. Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study.[J]. *Lancet* 2020;**395**(10223):507–13. doi:10.1016/s0140-6736(20)30211-716/s2213-2600(20)30076-x.
27. Castrucci M R. Factors affecting immune responses to the influenza vaccine.[J]. *Hum Vaccin Immunother* 2018;**14**(3):637–46. doi:10.1080/21645515.2017.1338547.
28. Gal-Oz S T, Maier B, Yoshida H, et al. ImmGen report: sexual dimorphism in the immune system transcriptome.[J]. *Nat Commun* 2019;**10**(1):4295. doi:10.1038/s41467-019-12348-6.

Please cite this article as: Z. Zheng, F. Peng and B. Xu et al., Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis, Journal of Infection, https://doi.org/10.1016/j.jinf.2020.04.021

ANZ J. Surg. 2003; **73**: 712–716

ORIGINAL ARTICLE

METHODOLOGICAL INDEX FOR NON-RANDOMIZED STUDIES (*MINORS*): DEVELOPMENT AND VALIDATION OF A NEW INSTRUMENT

KAREM SLIM,* EMILE NINI,* DAMIEN FORESTIER,* FABRICE KWIATKOWSKI,† YVES PANIS‡ AND JACQUES CHIPPONI*

**Department of General and Digestive Surgery, Hôtel-Dieu, Clermont-Ferrand, †Department of Statistics, Centre Jean-Perrin Clermont-Ferrand and ‡Department of Digestive Surgery, Hôpital Lariboisière, Paris, France*

Background: Because of specific methodological difficulties in conducting randomized trials, surgical research remains dependent predominantly on observational or non-randomized studies. Few validated instruments are available to determine the methodological quality of such studies either from the reader's perspective or for the purpose of meta-analysis. The aim of the present study was to develop and validate such an instrument.

Methods: After an initial conceptualization phase of a methodological index for non-randomized studies (MINORS), a list of 12 potential items was sent to 100 experts from different surgical specialities for evaluation and was also assessed by 10 clinical methodologists. Subsequent testing involved the assessment of inter-reviewer agreement, test-retest reliability at 2 months, internal consistency reliability and external validity.

Results: The final version of MINORS contained 12 items, the first eight being specifically for non-comparative studies. Reliability was established on the basis of good inter-reviewer agreement, high test-retest reliability by the κ-coefficient and good internal consistency by a high Cronbach's α-coefficient. External validity was established in terms of the ability of MINORS to identify excellent trials.

Conclusions: MINORS is a valid instrument designed to assess the methodological quality of non-randomized surgical studies, whether comparative or non-comparative. The next step will be to determine its external validity when used in a large number of studies and to compare it with other existing instruments.

Key words: comparative study, methodology index, non-randomized study.

Abbreviation: MINORS, methodological index for non-randomized studies.

INTRODUCTION

Although surgeons are now conducting an increasing number of randomized trials,[1] most of the available evidence in surgery comes from non-randomized studies, both comparative and non-comparative. Indeed surgical research remains an example of a situation where randomization is not always possible or feasible.[2] Beyond large randomized trials, systematic reviews are an important way to answer questions in surgery. However, the systematic review or meta-analysis of studies other than randomized trials may be difficult because combining the results of observational studies of heterogeneous quality could be highly biased.

Observational studies include comparative studies such as case-control and cohort designs, and patient series which may or may not involve comparisons between two or more groups.

Several papers have discussed the methodology of meta-analyses of observational studies[3,4] and checklists have been proposed but not formally validated.[5] Downs and Black used clinimetric criteria to develop a checklist which was applicable to both randomized and non-randomized studies without distinction.[6] The aim of the present study was to develop and validate a methodological index for non-randomized studies (MINORS) which could be used by readers, manuscript reviewers or journal editors to assess the quality of such studies.

K. Slim MD, FACS; **E. Nini** MD; **D. Forestier** MD; **F. Kwiatkowski** PhD; **Y. Panis** MD, PhD; **J. Chipponi** MD, PhD.

Correspondence: K. Slim, Department of General and Digestive Surgery, Hôtel-Dieu Boulevard Leon Malfreyt, F-63058 Clermont-Ferrand, France. Email: kslim@chu-clermontferrand.fr

Accepted for publication 13 May 2003.

METHODS

Conceptualization phase

After reviewing the literature on quality assessment of randomized trials and discussing the particular features of non-randomized studies, a panel of eight practising surgeons selected 12 items to be considered for inclusion in MINORS. These items were chosen because of their ability to characterize the methodological and scientific value of published articles. Seven items were selected for assessment of non-comparative studies and five for use with comparative studies. The list of 12 items was then sent to 100 surgeons throughout France who had clinical research expertise in different specialities, including digestive, cardiovascular and thoracic surgery, gynaecology, otorhinolaryngology, orthopaedics, urology, neurosurgery, and ophthalmology. They were asked to score the ability of each item to assess the quality of a given study using a 7-point-scale, according to the method proposed by Oxman and Guyatt.[7] The mean score for each item was then compared with that of every other item to see whether there were any significant differences. Subsequently each item was scored from 0 to 2; 0 indicating that it was not reported in the article evaluated, 1 indicating that it was reported but inadequately, and 2 indicating that

it was reported adequately. The form also included a section allowing the surgeons to suggest additional items.

Assessment of face validity and content validity

To determine whether MINORS items appeared appropriate and whether they covered all important considerations relevant to the methodology of non-randomized studies, a revised list of items was sent to 10 French clinical methodologists for assessment on 13 credibility criteria according to the method proposed by Feinstein.[8]

Clinimetric testing of MINORS

Inter-reviewer agreement
To test the consistency of MINORS between reviewers, a random sample of published non-randomized studies, both comparative and non-comparative, was selected from among several specialties. For this purpose a Medline search was undertaken using the MeSH 'surgery' and limits by publication type (clinical trial *not* randomized controlled trial) for the year 2001. A numerical list of original articles was then established and 80 articles were selected randomly. The title, authors' names, institutional affiliation and journal identity were removed. These articles were then assessed by two independent reviewers with different methodological expertise (one junior and one senior surgeon) using the revised version of MINORS.

Test-retest reliability
Two months after the first assessment, a randomly selected sample of 30 articles was scored again by the junior surgeon without reference to his first assessment.

Internal consistency
This evaluation indicated whether the items were related to one another and worked together in a similar manner in assessing the quality of articles.

Validity
The power of MINORS to differentiate between excellent, fair or poor studies was examined by selecting a random sample of 15 excellent randomized controlled trials. These articles were chosen as the gold standard against which to assess the external validity of MINORS on the basis that they had all been published in three major journals which had adopted the CONSORT Statement,[9] (namely *British Medical Journal*, *British Journal of Surgery and The Lancet*). These articles were then scored according to MINORS and the results compared with a selected group of the 15 best-scored comparative studies from the sample of 80 described previously. The reviewer was blinded as to the source of the 15 randomized trials.

Statistical analyses

Agreement between reviewers was measured by the κ-coefficient (unweighted model) with a value greater than 0.4 being accepted as satisfactory.[10] Global scores were obtained by summing all the item scores**.** Results were expressed as means (standard deviations). The matched pairs *t*-test was used to compare mean global scores between reviewers. Internal consistency was assessed by the calculation of Cronbach's α-coefficient.[11] A value of $P < 0.05$ was considered statistically significant.

RESULTS

Content and face validity

Expert phase
Ninety of the 100 experts returned a completed form. Table 1 summarizes the scores of the 12 items included in the first version of MINORS. No item was scored less than five. Furthermore there was no difference between the different specialities. The experts suggested no additional methodological items apart from a modification of item 11. As a result a supplementary sentence was incorporated in that item in the revised version of MINORS relating to the size of non-comparative studies. Item 11 thus became relevant to both comparative and non-comparative studies. The revised version of MINORS included 12 items: the first subscale of eight items related to non-comparative studies whereas all 12 items were relevant to comparative studies (Table 2).

Revision phase
Because there was no statistical difference between the mean item scores as evaluated by the experts, the items were not weighted and the scoring was simplified to a 3-point scale from 0 to 2. If one considers that MINORS involves eight items for non-comparative studies and 12 items for comparative studies and that the maximum item score is 2, the ideal global score would be 16 for the non-comparative studies and 24 for the comparative studies.

Methodologist phase
All 10 methodologists completed their assessment and scored the final version favourably, all item mean scores being above 4.5 on a 7-point scale (Table 3).

Table 1. Assessment of items in the first version of MINORS by 90 experts in several surgical specialities using a scale from 0 to 7

Item	Median	Mean (SD)
1. A stated aim of the study	7	6.6 (0.7)
2. Inclusion of consecutive patients	6	5.8 (1.1)
3. Prospective collection of data	6	5.5 (1.2)
4. Endpoint appropriate to the study aim	6	6.3 (0.8)
5. Unbiased evaluation of endpoints	5	5.4 (1.2)
6. Follow-up period appropriate to the major endpoint	6	6.2 (0.8)
7. Loss to follow up not exceeding 5%	6	5.5 (1.2)
And in the case of comparative studies		
8. A control group having the gold standard intervention	6	6.0 (1.1)
9. Contemporary groups	6	5.5 (1.4)
10. Baseline equivalence of groups	6	6.1 (0.9)
11. Prospective calculation of the sample size	6	5.5 (1.3)
12. Statistical analyses adapted to the study design	6	6.3 (0.8)

SD, standard deviation.

Inter-reviewer agreement, test-retest reliability and internal consistency reliability

There were 26 comparative and 54 non-comparative studies in the random sample assessed by the two reviewers. Table 4 summarizes the correlation between the scores of the reviewers. Agreement between the reviewers was considered satisfactory for all items. The mean global scores on a scale from 0 to 24 were, respectively, 13.93 (0.35) for the junior surgeon and 12.98 (0.54) for the senior surgeon. This difference was statistically significant ($P < 10^{-7}$) but corresponds to only 0.95 of a global score point. The mean global scores, which ranged between 0 and 20 in this agreement assessment, did not differ significantly between the comparative and non-comparative studies ($P = 0.11$).

The assessment of test-retest reliability showed a satisfactory correlation between the original and repeated scoring after a 2-month interval. The mean global score decreased significantly from 13.91 (3.3) at the first test to 12.28 (3.6) at the second ($P < 0.0001$). The internal consistency reliability of MINORS was high with a global α-value of 0.73. This demonstrated that all items worked in a complementary and coherent manner.

Validity

The 15 gold-standard randomized trials had a mean global score of 23.1. The comparison between the score of these randomized trials and that of the 15 best comparative non-randomized studies (19.8) showed a significant difference ($P = 0.00001$) in favour of the randomized trials.

DISCUSSION

This index for the assessment of non-randomized studies was developed by a group of surgeons because of the problems faced by clinicians as to the lack of randomized surgical trials and the large number of observational studies in surgery. To apply the principles of evidence-based medicine to clinical practice requires a method for assessing the quality of published data.

Table 3. Credibility criteria assessed by 10 clinical methodologists on a 7-point scale

Criterion	Mean	SD (range)
1. Wide applicability	5.5	0.5 (5–6)
2. Use by various groups	4.8	0.8 (4–6)
3. Clarity and simplicity	5.1	0.8 (4–6)
4. Adequate instructions	5.0	1.0 (4–7)
5. Information available	4.9	1.2 (3–7)
6. Need for subjective decision	4.7	0.9 (4–7)
7. Likelihood of bias	4.8	1.2 (3–7)
8. Single domain	5.1	0.9 (4–7)
9. Redundant items	5.6	0.8 (4–7)
10. Comprehensiveness	5.1	0.7 (4–6)
11. Item weights	5.4	1.0 (3–7)
12. Number of response options	5.4	1.1 (4–7)
13. Discrimination power	5.1	0.5 (4–6)

SD, standard deviation.

Table 2. The revised and validated version of MINORS

Methodological items for non-randomized studies	Score†
1. **A clearly stated aim:** the question addressed should be precise and relevant in the light of available literature	
2. **Inclusion of consecutive patients**: all patients potentially fit for inclusion (satisfying the criteria for inclusion) have been included in the study during the study period (no exclusion or details about the reasons for exclusion)	
3. **Prospective collection of data**: data were collected according to a protocol established before the beginning of the study	
4. **Endpoints appropriate to the aim of the study**: unambiguous explanation of the criteria used to evaluate the main outcome which should be in accordance with the question addressed by the study. Also, the endpoints should be assessed on an intention-to-treat basis.	
5. **Unbiased assessment of the study endpoint**: blind evaluation of objective endpoints and double-blind evaluation of subjective endpoints. Otherwise the reasons for not blinding should be stated	
6. **Follow-up period appropriate to the aim of the study**: the follow-up should be sufficiently long to allow the assessment of the main endpoint and possible adverse events	
7. **Loss to follow up less than 5%**: all patients should be included in the follow up. Otherwise, the proportion lost to follow up should not exceed the proportion experiencing the major endpoint	
8. **Prospective calculation of the study size**: information of the size of detectable difference of interest with a calculation of 95% confidence interval, according to the expected incidence of the outcome event, and information about the level for statistical significance and estimates of power when comparing the outcomes	
Additional criteria in the case of comparative study	
9. **An adequate control group**: having a gold standard diagnostic test or therapeutic intervention recognized as the optimal intervention according to the available published data	
10. **Contemporary groups**: control and studied group should be managed during the same time period (no historical comparison)	
11. **Baseline equivalence of groups**: the groups should be similar regarding the criteria other than the studied endpoints. Absence of confounding factors that could bias the interpretation of the results	
12. **Adequate statistical analyses**: whether the statistics were in accordance with the type of study with calculation of confidence intervals or relative risk	

†The items are scored 0 (not reported), 1 (reported but inadequate) or 2 (reported and adequate). The global ideal score being 16 for non-comparative studies and 24 for comparative studies.

Table 4. List of 12 items of the definitive MINORS. Inter-reviewer correlation on a random sample of 80 articles and test-retest reliability on a random sample of 30 articles.

Methodological item for non-randomized studies	κ-coefficient for inter-reviewer agreement (SD)†	κ-coefficient for test-re-test reliability (SD)
1. A clearly stated aim	0.87 (0.07)	0.89 (0.11
2. Inclusion of consecutive patients	0.78 (0.06)	0.83 (0.09)
3. Prospective collection of data	0.79 (0.06)	0.82 (0.09)
4. Endpoints appropriate to the aim of the study	0.56 (0.09)	0.76 (0.12)
5. Unbiased assessment of the study endpoint	0.61 (0.08)	0.61 (0.13)
6. Follow-up period appropriate to the aim of the study	0.61 (0.08)	0.59 (0.11)
7. Loss to follow up less than 5%	0.69 (0.08)	0.74 (0.12)
8. Prospective calculation of the study size	1.00	1.00
Additional criteria in the case of comparative studies		
9. An adequate control group	0.86 (0.09)	1.00
10. Contemporary groups	0.79 (0.14)	0.61 (0.31)
11. Baseline equivalence of groups	0.87 (0.09)	1.00
12. Adequate statistical analyses	0.66 (0.14)	0.75 (0.22)

†A κ-coefficient of >0.4 was considered satisfactory.

This is an important consideration for the 'consumers' of clinical research. Our initial aim was to develop and validate an index which would be simple to use both by readers of published articles and reviewers of manuscripts submitted for publication, and be of sufficient sensitivity for use in meta-analysis of non-randomized studies. To achieve this we followed the recognized principles of scale construction[12] using a rigorous methodology. The results of the present study show clearly that the instrument we have developed has good reliability, internal consistency and validity. The high response rate from experts and the limited number of items used, suggest that MINORS is easy to apply. Its simplicity and objectivity is also demonstrated by its acceptability to surgeons having sound methodological expertise. Although the difference between the scores of the senior and the junior reviewers was statistically significant, its actual relevance was low as the difference did not exceed 1 point.

Similarly, the assessment of test-retest reliability showed a good correlation over an interval of 2 months. The reviewer scored perhaps more severely on the second occasion, which suggests greater expertise with further experience, but the difference was too small (1.6) to be important. Nevertheless this feature may need to be investigated further.

Instead of weighting, we chose to score the items from 0 to 2 according to whether they were reported or not and adequate or not. Weighting of items requires further investigation as we have no gold standard method to evaluate the relative importance of a given methodological item. In the light of the available literature, the most appropriate method of weighting would be based on consensus development among experienced epidemiologists before designing a large study to validate their conclusions. Few attempts have been made to estimate the respective values of some methodological items.[13] Furthermore the most significant findings regarding the weighting of items have been specifically related to randomized double blind studies. One could assume that the rationale for weighting in randomized trials can be extrapolated to non-randomized studies. However this needs to be confirmed by further investigations, especially in the field of surgery. Furthermore, the item weights could differ according to the type of study. For example unbiased evaluation of endpoints is important for functional disorders whereas the length of follow up and loss to follow up are important for hernia or cancer surgery. Currently, however, there is no sound evidence for the differential weighting of items in methodological indices or checklists for non-randomized studies.

Downs and Black[6] reported a checklist applicable to both randomized and non-randomized trials. It involved 27 items concerning external validity, bias, confounding factors, statistical power and reporting; however, the number of items and differences in scoring systems between items increased complexity and user burden. Several items were related to reporting and thus were not directly concerned with the methodological quality of a study. Also in their study, the period between the test and re-test was only 2 weeks and the reviewers were similar to one-another in their level of methodological skill. Furthermore their instrument was a checklist and was not developed as an index for scoring studies.

An important aspect of MINORS is its external validity; that is, its ability to identify high quality studies, which was established by comparison with the current standard for randomized trials, namely the CONSORT Statement. Since MINORS does not differentiate between randomized and non-randomized studies and includes several items derived from indices focusing on the quality of randomized trials, the fact that a given study has a randomized design is not sufficient to achieve a high score. MINORS was not developed specifically to assess the quality of randomized trials; however, we considered the randomized trial to be the best example of comparative studies and assumed that MINORS should be able to distinguish between different comparative studies. MINORS satisfied that expectation and clearly confirmed that a good randomized trial scores higher than a good non-randomized comparative study. The ability of MINORS to recognize the poor or fair quality of non-comparative studies is suggested in our study, but this needs to be further evaluated by comparison with the Downs and Black checklist.[6] This comparison will be the subject of a future study to develop a reliable standardized instrument for assessing the quality of non-randomized studies, especially for the purposes of meta-analysis. Nevertheless, as with randomized trials[14] for which there is no gold standard, it is possible that any newly proposed instrument might have internal flaws. An ideal index should be highly sensitive (by increasing the number of items) and applicable in daily practice (by minimizing user burden). This remains the challenge for epidemiologists and research in this field is in its infancy.

MINORS in our opinion has two important attributes. First, its simplicity in comprising only 12 items that are readily usable by both readers and researchers and second, its reliability, as demonstrated by clinimetric testing. Our aim now is to use MINORS in several more studies designed to evaluate the methodology of non-randomized studies. Only the repeated use of such an instrument can confirm the present preliminary clinimetric validation.

ACKNOWLEDGEMENTS

We thank the following participants for their help in the phases of conceptualization and validation: L. Audigé (from the Non-randomized Study Method Group of the Cochrane Collaboration), E. Albuisson, J. E. Bazin, F. Blanchard, G. Borges Da Silva, S. Bouee, D. Chopin, B. Dousset, C. Dziri, F. Fagnani, P-L Fagniez, L. Gerbaud, A. Kramar, F. Lacaine, B. Millat, E. Monnet, P. Perez, A. Perillat, T. Perniceni, L. R. Salmi and the surgeons who returned the completed form during the conceptualization phase of MINORS.

REFERENCES

1. Slim K, Haugh M, Fagniez P-L, Pezet D, Chipponi J. Ten-year audit of randomised trials in digestive surgery from Europe. *Br. J. Surg.* 2000; **87**: 1585–6.
2. McCulloch PM, Taylor I, Sasako M, Lovett B, Griffin D. Randomised trials in surgery: problems and possible solutions. *BMJ* 2002; **324**: 1448–51.
3. Stroup DF, Berlin JA, Morton SC *et al.* Meta-analysis of observational studies in epidemiology. A proposal for reporting. *JAMA* 2000; **283**: 2008–12.
4. Bhandari M, Morrow F, Kulkarni AV, Tornetta P. Meta-analyses in orthopaedic surgery. A systematic review of their methodologies. *J. Bone Joint Surg. Am.* 2001; **83**: 15–24.
5. Khan KS, ter Riet G, Popay J, Nixon J, Kleijnen J. Study quality assessment. In: *Undertaking systematic reviews of research on effectiveness. CRD's guidance for those carrying out or commissioning reviews* CRD Report 4, 2nd edn. York: York Publishing Services; March 2001.
6. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J. Epidemiol. Community Health* 1998; **52**: 377–84.
7. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J. Clin. Epidemiol.* 1991; **44**: 1271–8.
8. Feinstein AR. *Clinimetrics.* New Haven: Yale University Press, 1987.
9. Moher D, Schulz KF, Altman DG for the CONSORT Group. CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001; **357**: 1191–4.
10. King TS, Chinchilli VM. A generalized concordance correlation coefficient for continuous and categorical data. *Stat. Med.* 2001; **20**: 2131–47.
11. Bland JM, Altman DG. Cronbach's alpha. *BMJ* 1997; **314**: 572.
12. Bland JM, Altman DG. Validating scales and indexes. *BMJ* 2002; **324**: 606–7.
13. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; **273**: 408–12.
14. Jüni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 1999; **282**: 1054–60.