ABSTRACT

Objectives Despite the publication of hundreds of trials on gout and hyperuricemia, management of these conditions remains suboptimal. We aimed to assess the quality and consistency of guidance documents for gout and hyperuricemia.

Design Systematic review and quality assessment using the appraisal of guidelines for research and evaluation (AGREE) II methodology.

Data sources PubMed and EMBASE (27 October 2016), two Chinese academic databases, eight guideline databases, and Google and Google scholar (July 2017).

Eligibility criteria We included the latest version of international and national/regional clinical practice guidelines and consensus statements for diagnosis and/or treatment of hyperuricemia and gout, published in English or Chinese.

Data extraction and synthesis Two reviewers independently screened searched items and extracted data. Four reviewers independently scored documents using AGREE II. Recommendations from all documents were tabulated and visualised in a coloured grid.

Results Twenty-four guidance documents (16 clinical practice guidelines and 8 consensus statements) published between 2003 and 2017 were included. Included documents performed well in the domains of scope and purpose (median 85.4%, range 66.7%–100.0%) and clarity of presentation (median 79.2%, range 48.6%–98.6%), but unsatisfactory in applicability (median 10.9%, range 0.0%–86.7%) and editorial independence (median 28.1%, range 0.0%–83.3%). The 2017 British Society of Rheumatology guideline received the highest scores. Recommendations were concordant on the target serum uric acid level for long-term control, on some indications for urate-lowering therapy (ULT), and on the first-line drugs for ULT and for acute attack. Substantially inconsistent recommendations were provided for many items, especially for the timing of initiation of ULT and for treatment for asymptomatic hyperuricemia.

Conclusions Methodological quality needs improvement in guidance documents on gout and hyperuricemia. Evidence for certain clinical questions is lacking, despite numerous trials in this field. Promoting standard guidance development methods and synthesising high-quality clinical evidence are potential approaches to reduce recommendation inconsistencies.

Strengths and limitations of this study

The first systematic review to assess the quality of clinical practice guidelines and consensus statements on the diagnosis and treatment for both hyperuricemia and gout.

The first systematic review to summarise recommendations for best practice in hyperuricemia and gout.

The appraisal of guidelines for research and evaluation II instrument is an international, structured, validated and rigorously developed tool.

Only guidance documents in English and Chinese were included.

Literature search was >1 year old at the time of publication.

BACKGROUND

Gout is an inflammatory arthritis occurring in response to monosodium urate (MSU) crystals formation, a common and necessary pathogenic factor of which is hyperuricemia. The prevalence of gout and hyperuricemia, as well as their disease burden, are rising globally. However, although >600 related clinical studies have been published to date, the quality of care for gout and hyperuricemia remains suboptimal. The goal of treatment is to reduce the body’s total uric acid pool and consequently to minimise the risk of acute flares, arthropathy, nephrolithiasis and other complications. A study in the USA found that only 22% of patients with gout received therapy adhering to all quality indicators. A nationwide population study in the UK reported that only 48% of prevalent patients received proper consultation and only 27%...
of incident patients were provided with urate-lowering therapy (ULT) within 1 year of diagnosis.6

High-quality guidance documents are important for improving the quality of care for gout and hyperuricemia at individual, community and national levels.13 Current guidance documents for gout and hyperuricemia have been developed by rheumatology, endocrinology and cardiology groups, at regional, national or international levels. Among these documents, the American College of Rheumatology (ACR) guidelines,14 15 updated in 2012, and the European League Against Rheumatism (EULAR) guidelines,16–18 updated in 2016, have the most substantial global influence. Besides, the most recent documents (released in 2017) are two national guidelines, from the American College of Physicians (ACP)19 20 and the British Society for Rheumatology (BSR),21 respectively, and one consensus statement, from the Chinese Multidisciplinary Expert Task Force on Hyperuricemia and Its Related Diseases.22

Despite the variety of documents, current guidelines and consensuses on gout and hyperuricemia provide inconsistent recommendations, even those released by highly respected professional organisations, such as the ACP and the ACR.25 Some distinct differences lie in key aspects for patient care, such as the pharmacological treatment for asymptomatic hyperuricemic patients, the timing of initiation of ULT in patients with gout flare,24 and indications for ULT.25 These discrepancies may result from ethnic and social differences, but can be a consequence of inconsistent guideline development.25 Low-quality guidance documents put individual patients and communities at risk, and impede the application of guideline recommendations in clinical practice.26 Hence, we conducted this study to systematically evaluate the quality of clinical practice guidelines and consensus statements on gout and hyperuricemia and to compare key recommendations on patient care from all included documents.

METHODS
Detailed methods of the study have been published previously27 and this study was registered with PROSPERO.

Literature search and selection criteria
We systematically searched PubMed and EMBASE from inception to 27 October 2016 using a comprehensive search strategy (online supplementary tables 1 and 2) to identify guidelines and consensus recommendations pertaining to the diagnosis and treatment of gout and hyperuricemia. We searched two academic databases for Chinese publications (the Chinese Biomedical Literature Database and the Wanfang Data) and eight guideline databases from inception to 24 July 2017 using search strategies tailored to different databases (online supplementary table 3). We also searched Google and Google scholar in July 2017 for potentially eligible guidelines and consensus recommendations that were not indexed in the aforementioned databases.

We included the latest versions of all international and national/regional clinical practice guidelines and consensus statements for the diagnosis and/or treatment of gout and hyperuricemia, published in English or Chinese. Two reviewers (QL, XL) independently screened all searched documents. Reasons for exclusion were provided for documents excluded during the full-text review (online supplementary table 4). Disagreements were resolved through discussion with a third reviewer (SL).

Data extraction
We extracted the following data from each included document: document characteristics (eg, year of publication, funding body and evidence base), recommendations for diagnosis and monitoring of gout and hyperuricemia, and recommendations for management. Data were extracted by one investigator (QL) and checked by another (XL).

Appraisal of guidance documents
All included documents were assessed by four reviewers (QL, XL, JW and HL) independently using the appraisal of guidelines for research and evaluation (AGREE) II instrument.28 AGREE II is an internationally developed and validated tool to evaluate the quality of clinical practice guidelines29–31 and consensus statements.32 33

All reviewers completed an online training tutorial34 before the commencement of appraisal to ensure standardisation. We adapted detailed instructions for scoring from the AGREE II User’s Manual28 and provided objective scoring criteria for each item (online supplementary file 1). We selected four guidance documents for pilot scoring, during which we discussed and clarified our objective scoring criteria. When scoring for all included documents was completed, a meeting was held among reviewers and every item with scores differed more than one point was discussed. After the meeting, reviewers were given the opportunity to revise their scores or to keep the original scores. We recorded all original scores, revised scores and reasons for modifying scores for quality control purpose, and used the intraclass correlation coefficient (ICC) to test inter-rater reliability. The ICC was calculated via IBM SPSS Statistics 22.0 and an ICC ≥0.7 was considered acceptable.35

Recommendation synthesis
We manually extracted recommendations on key clinical questions from all included guidance documents and summarised them into four tables: (1) the diagnosis of gout and hyperuricemia, (2) the treatment of hyperuricemia, (3) the treatment of acute gout and (4) the treatment of tophi. Recommendations were extracted by one investigator (QL) and checked by another (XL). We further visualised these recommendations in a five-colour grid to illustrate inconsistencies. The most frequently recommended content was used as a reference. We used green to colour documents providing consistent recommendations, red to colour those providing contrary

Figure 1 Flow diagram for literature search. AGREE, appraisal of guidelines for research and evaluation; CBM, Chinese Biomedical Literature Database; GAIN, Guidelines and Audit Implementation Network; GIN, Guidelines International Network; NGC, National Guideline Clearinghouse; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; SIGN, Scottish Intercollegiate Guidelines Network; TRIP, Turning Research into Practice Database

recommendations and blue to colour those providing partially consistent recommendations. A partially consistent recommendation was defined as a recommendation that included but not the same as the reference content. Where recommendations were not given or were not applicable, the cell was coloured in yellow and in grey, respectively.

**Patient and public involvement**

Patients or the public were not involved in the conceptualisation or carrying out of this research.

**RESULTS**

**Search results**

Overall, we identified 5811 items across academic databases, guideline databases, Google and Google Scholar. After applying the inclusion and exclusion criteria, 24 guidance documents from 26 papers were included in the final appraisal and recommendation synthesis (figure 1). Studies excluded after full-text review and reasons for exclusion were provided as online supplementary table 4.

**Characteristics of included guidelines and consensus statements**

Table 1 summarised characteristics of included guidance documents, among which 16 were clinical practice guidelines and 8 were consensus statements. 16 national or regional organisations and three international groups (ie, the 3e (evidence, expertise, exchange) Initiative, the EULAR and the development group for the treat-to-target (T2T) recommendations) published these documents between 2003 and 2017. Sixteen documents were issued by rheumatology organisations and seven were developed by multinational development groups. Seventeen documents provided information on their guideline development group.
Table 1  Characteristics of included guidelines and consensus statements

<table>
<thead>
<tr>
<th>Document</th>
<th>Issuing organisation</th>
<th>Year of publication</th>
<th>Country</th>
<th>Funding body</th>
<th>Target population</th>
<th>Target audience</th>
<th>Guideline development group</th>
<th>Guideline review</th>
<th>Guideline update</th>
<th>Evidence base</th>
<th>LOE</th>
<th>SOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMA_2003</td>
<td>South African Medical Association</td>
<td>2003</td>
<td>South Africa</td>
<td>Pharmaceutical company</td>
<td>Gout</td>
<td>Phy</td>
<td>Multi</td>
<td>ER</td>
<td>Intermittent</td>
<td>NG</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>EULAR_2006</td>
<td>EULAR</td>
<td>2006</td>
<td>Europe</td>
<td>EULAR</td>
<td>Gout</td>
<td>NG</td>
<td>Rheu</td>
<td>NG</td>
<td>NG</td>
<td>SLR</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>MOH_MSR_AMM_2008</td>
<td>MOH, MSR, AMM</td>
<td>2008</td>
<td>Malaysia</td>
<td>Pharmaceutical company</td>
<td>Adults (&gt;16 years) with gout</td>
<td>Phy</td>
<td>Multi</td>
<td>ER</td>
<td>2012 or sooner</td>
<td>SLR</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>PRA_2008</td>
<td>Philippine Rheumatology Association</td>
<td>2008</td>
<td>Philippine</td>
<td>NG</td>
<td>Gout</td>
<td>Phy</td>
<td>NG</td>
<td>NG</td>
<td>NG</td>
<td>SLR</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>UT Austin_2009</td>
<td>University of Texas at Austin</td>
<td>2009</td>
<td>US</td>
<td>University of Texas at Austin</td>
<td>Adults with gout</td>
<td>Phy</td>
<td>NG</td>
<td>ER</td>
<td>NG</td>
<td>SLR</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>EULAR_2011</td>
<td>EULAR</td>
<td>2011</td>
<td>Multination</td>
<td>Pharmaceutical company</td>
<td>Gout</td>
<td>Phy</td>
<td>Multi</td>
<td>NG</td>
<td>NG</td>
<td>SLR</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>JSGNAM_2011</td>
<td>Japanese Society of Gout and Nucleic Acid Metabolism</td>
<td>2011</td>
<td>Japan</td>
<td>NG</td>
<td>Hyperuricemia or gout</td>
<td>NG</td>
<td>NG</td>
<td>ER</td>
<td>NG</td>
<td>SLR</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>ACR_2012</td>
<td>ACR</td>
<td>2012</td>
<td>US</td>
<td>ACR, NIAMS, NIH</td>
<td>Gout</td>
<td>Phy</td>
<td>Multi</td>
<td>NG</td>
<td>Intermittent</td>
<td>SLR</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>SER_2013</td>
<td>Spanish Society of Rheumatology</td>
<td>2013</td>
<td>Spain</td>
<td>Pharmaceutical company</td>
<td>Gout</td>
<td>Phy</td>
<td>Multi</td>
<td>ER</td>
<td>Four years</td>
<td>SLR</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>SIR_2013</td>
<td>Italian Society of Rheumatology</td>
<td>2013</td>
<td>Italy</td>
<td>NG</td>
<td>Gout</td>
<td>Phy</td>
<td>Multi</td>
<td>NG</td>
<td>NG</td>
<td>SLR</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>FMOH_2014</td>
<td>Federal Ministry of Health (Nigeria)</td>
<td>2014</td>
<td>Nigeria</td>
<td>NG</td>
<td>Gout</td>
<td>Phys, Pts in Nigeria</td>
<td>Multi</td>
<td>NG</td>
<td>NG</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>CRA_2016</td>
<td>Chinese Rheumatology Association</td>
<td>2016</td>
<td>China</td>
<td>NG</td>
<td>Gout in China</td>
<td>Phy</td>
<td>Multi</td>
<td>NG</td>
<td>NG</td>
<td>SLR</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>EULAR_2016</td>
<td>EULAR</td>
<td>2016</td>
<td>Europe</td>
<td>NG</td>
<td>Gout</td>
<td>Phy, Pts</td>
<td>Multi</td>
<td>ER</td>
<td>Intermittent</td>
<td>SLR</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>TRA_2016</td>
<td>Taiwan Rheumatology Association</td>
<td>2016</td>
<td>Taiwan, China</td>
<td>NG</td>
<td>Hyperuricemia or gout</td>
<td>Phy, Pts</td>
<td>Multi</td>
<td>NG</td>
<td>NG</td>
<td>NG</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>ACP_2017</td>
<td>ACP</td>
<td>2017</td>
<td>US</td>
<td>ACP</td>
<td>Acute and recurrent gout</td>
<td>Phy</td>
<td>NG</td>
<td>ER</td>
<td>Five years</td>
<td>SLR</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>BSR_2017</td>
<td>The British Society for Rheumatology</td>
<td>2017</td>
<td>UK</td>
<td>No specific funding</td>
<td>Gout in the UK</td>
<td>Phy</td>
<td>Multi</td>
<td>ER</td>
<td>Planned in 2020</td>
<td>SLR</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Continued
Table 1

<table>
<thead>
<tr>
<th>Document</th>
<th>Issuing organisation</th>
<th>Year of publication</th>
<th>Country</th>
<th>Funding body</th>
<th>Target population</th>
<th>Target audience</th>
<th>Guideline development group</th>
<th>Guideline review</th>
<th>Guideline update</th>
<th>Evidence base</th>
<th>LOE</th>
<th>SOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consensus statements</td>
<td>Chinese College of Cardiovascular Physicians</td>
<td>2012</td>
<td>China</td>
<td>NG</td>
<td>Asymptomatic hyperuricemia with CVD</td>
<td>NG</td>
<td>NG</td>
<td>NG</td>
<td>NG</td>
<td>CS</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>3e_2013[36]</td>
<td>3e Initiative</td>
<td>2013</td>
<td>Multination</td>
<td>Pharmaceutical company</td>
<td>Gout</td>
<td>NG</td>
<td>Rheu</td>
<td>NG</td>
<td>NG</td>
<td>SLR</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>CSE_2013[37]</td>
<td>Chinese Society of Endocrinology</td>
<td>2013</td>
<td>China</td>
<td>NG</td>
<td>Hyperuricemia or gout</td>
<td>NG</td>
<td>NG</td>
<td>NG</td>
<td>NG</td>
<td>CS</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>3e_PT_2014[40]</td>
<td>Portuguese 3e Initiative</td>
<td>2014</td>
<td>Portugal</td>
<td>NG</td>
<td>Gout in Portuguese</td>
<td>NG</td>
<td>Rheu</td>
<td>NG</td>
<td>NG</td>
<td>SLR</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>3e_AU_NZ_2015[43]</td>
<td>Australian and New Zealand 3e Initiative</td>
<td>2015</td>
<td>Multination</td>
<td>NG</td>
<td>Gout</td>
<td>NG</td>
<td>Rheu</td>
<td>NG</td>
<td>NG</td>
<td>SLR</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>ACR_EULAR_2015[42]</td>
<td>ACR/EULAR</td>
<td>2015</td>
<td>Multination</td>
<td>ACR, EULAR</td>
<td>Gout</td>
<td>NG</td>
<td>ER</td>
<td>Intermittent</td>
<td>SLR</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>T2T_2016[49]</td>
<td>NG</td>
<td>2016</td>
<td>Multination</td>
<td>Pharmaceutical company</td>
<td>Gout</td>
<td>NG</td>
<td>Rheu</td>
<td>ER</td>
<td>NG</td>
<td>SLR</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

ACP, American College of Physicians; ACR, American College of Rheumatology; AMM, Academy of Medicine of Malaysia; ASCR, American Society of Clinical Rheumatologists; CS, consensus statement; CVD, cardiovascular diseases; ER, external review; EULAR, European League Against Rheumatism; LOE, level of evidence; MOH, Ministry of Health Malaysia; MSR, Malaysian Society of Rheumatology; Multi, multidisciplinary development group; NG, not given; NIAMS, National Institute of Arthritis and Musculoskeletal and Skin Diseases; NIH, National Institutes of Health; Phy, physicians; Pt, patients; Rheu, rheumatologists; SLR, systematic literature review; SOR, strength of recommendation; 3e, evidence, expertise, exchange.

Note: Document abbreviations may vary and should be checked in the original source.
among which 11 explicitly stated the involvement of a methodologist. Twelve documents provided information on their target audience, among which only three considered patients as one of the target audiences. Eighteen documents conducted systematic literature review as part of their development process, among which 17 documents reported the level of evidence supporting recommendations and graded the strength of recommendations. Ten documents clearly stated being externally reviewed. Five provided a clear time of update plan. Twelve documents provided information on their funding body, among which six were fully or partially funded by the pharmaceutical industry and the rest did not clearly declare their funding body.

Appraisal of guidelines and consensus statements

Standardised AGREE II domain scores for each guidance document were shown as Figure 2 and were provided in value as online supplementary table 5. Scores for each AGREE II item were provided in mean as online supplementary table 6 and in detail as online supplementary table 7. The overall quality of guidance documents, as assessed by AGREE II, varied both between documents across domains and within documents between domains. The document with the highest domain scores was the gout management guideline published by the BSR in 2017, with five domains scoring above the upper quartile, followed by the guidelines published by the ACP in 2017 and the 2015 gout classification criteria by the ACR and the EULAR jointly, both with four domains scoring above the upper quartile. Guidelines did not always score higher than consensus statements. No tendency of improvement in the quality score over time was observed (online supplementary figure 1).

The AGREE II instrument evaluated guidelines and consensus statements in six domains, from the development, dissemination to implementation. The scope and purpose (domain 1) of a document clarifies its clinical questions. Proper involvement of stakeholders (domain 2) balances individuals’ biases. The rigour of development domain (domain 3) is most concerned by clinicians and ensures the validity of development methodology. Clearly presented recommendations (domain 4) conveyed precise and accessible information from the development group to clinicians. Good performances in the applicability domain (domain...
5) and the editorial independence domain (domain 6) guarantee the usefulness and the independence of documents.

Guidance documents received the highest scores for the scope and purpose domain (median 85.4%, range 66.7%–100.0%) and the clarity of presentation domain (median 79.2%, range 48.6%–98.6%), and the lowest scores for the applicability domain (median 10.9%, range 0.0%–66.7%) and the editorial independence domain (median 28.1%, range 0.0%–83.3%). The worst scored item was the monitoring or auditing criteria item (mean score 1.2, range 1.0–4.0), followed by the implementation advice or tools item (mean 1.7, range 1.0–4.8), the external review item (mean 2.1, range 1.0–6.0) and the updating procedure item (mean 2.1, range 1.0–6.5).

The ICC was 0.896. Group discussion modified 365/2208 (16.53%) of individual scores.

Synthesis of recommendations

Included guidance documents addressed four major themes: diagnosis of gout and hyperuricemia, treatment for hyperuricemia, treatment for acute gout attack and treatment for tophi. Figure 3 showed key recommendations and their inconsistencies.

Approaches to diagnostic strategies for gout and hyperuricemia

Thirteen guidance documents 17–20 22 36 38 40–43 46 49 51 covered the diagnosis of gout and 11 17 22 37 38 45–51 covered that of hyperuricemia. Online supplementary table 8 showed key recommendations. The identification of MSU crystals in synovial fluid or tophi was a gold standard for definite diagnosis, as recommended by all included documents. In the absence of MSU crystals, three aspects were commonly evaluated for gout diagnosis, namely the clinical manifestation, considered by all documents; the laboratory result, considered by all but one document 49; and the imaging result, considered by all but four documents 17 19 20 49 51.

Guidance documents differed when recommending the cut-off serum uric acid (SUA) level to diagnose hyperuricemia. For the patient population in general, four documents 38 47 48 51 recommended 7.0 mg/dL (or 420 μmol/L) as the cut-off, while two 17 45 preferred 6.8 mg/dL. Five documents 22 37 46 49 50 provided gender-specific cut-offs, recommending 6.0 mg/dL (or ~360 μmol/L) in female and 7.0 mg/dL (or ~420 μmol/L) in male. Asymptomatic hyperuricemia was defined in seven documents, 36 38 46–50 among which six 36 38 46–48 50 excluded patients with gout and two 36 48 excluded patients with tophi when making the diagnosis. Attitudes were inconsistent for whether or not patients with renal diseases can be diagnosed with asymptomatic hyperuricemia. Patients with renal diseases were not eligible for the diagnosis in the Japanese 36 and the Philippine 36 guidelines, but patients with pre-existing renal or cardiovascular diseases can receive this diagnosis in the 3 Initiative document. 36

Approaches to treatment for hyperuricemia

Twenty-two guidance documents 14–17 19–22 36–41 43–52 covered the treatment for hyperuricemia and online supplementary table 9 summarised key recommendations. All but three documents 19 20 41 52 explicitly provided target levels for long-term SUA control, most of which recommended 6.0 mg/dL (or 360 μmol/L), except the South African guideline 31 which recommended 5.0 mg/dL (300 μmol/L). Two documents 16 22 recommended a lower limit of 3.0 mg/dL (or 180 μmol/L) for long-term SUA management. Among these two documents, only the 2016 EULAR guideline 16 explained the reason for providing a lower limit was that low SUA might increase the risk of neurodegenerative diseases, but the level of evidence and the grade of recommendation were both low.

All but six guidance documents 36 39 40 43 44 52 provided explicit indications for long-term ULT. The most commonly recommended indications were recurrent attacks, 14–17 19–22 41 45 48–51 tophi, 14–17, 19–22, 38 41 45–51 urate nephrolithiasis, 14–17 19–22, 37 48–51 arthropathy 16 17 21 22 38 41 45–51 and comorbidities. 14–16 19–22 37 47 49–51 The definition of recurrent attacks varied from at least once per year 17 to at least three times per year, 49 while the majority of documents 14–16 19–21 41 recommended twice per year as the cut-off.

Regarding the timing to initiate ULT, the documents did not agree on whether to start pharmacological ULT after an acute attack 14–17 19–22 36–41 43–52 or during an attack. 14–15 37 When recommending to start ULT after an attack, the preferred time to wait since the attack resolved varied from 2 weeks 37 48 to 6 weeks. 32 All guidance documents based their recommendation for this question on expert opinions, due to insufficient evidence. The explanations provided for starting ULT after an attack were that ULT was better discussed when a patient was not painful, 21 and that ULT initiation could prolong or worsen the acute attack. 32 Two documents 16 30 explicitly presented the currently conflicting views and insufficient evidence and stated consequently no recommendation for this issue.

When pharmacological ULT options were explicitly provided, allopurinol was recommended by all guidance documents 14–17 21 36 40 43 45 46 48–50 to be the first-line drug, while febuxostat was recommended by three documents 14 15 17 46 to be the first-line and by six documents 16 21 36 40 43 45 to be the second-line. However, recommendations on the dosage of allopurinol varied largely. The maximum daily allopurinol dose recommended varied from 300 mg 33 600 mg 22 37 47 800 mg. 14 15 17 38 45 to 900 mg 21 45 46 and the daily starting dose recommended in patients with normal renal function varied from 50 mg 19 20 22 47 48 51 to 200 mg. 21 As for patients with impaired renal function, the cut-off renal function to initiate dose adjustment was provided diversely as creatinine clearance 20–140 mL/min 37 45 46 49 51 or estimated glomerular filtration rate (eGFR) 130 mL/min/1.73 m 2. 21 One document preferred to depend allopurinol dosage solely on eGFR.
Figure 3  Summary of key recommendations for the diagnosis and treatment of gout and hyperuricemia. 3e, evidence, expertise, exchange Initiative; 3e_AU_NZ, Australian and New Zealand 3e Initiative; 3e_PT, Portuguese 3e Initiative; ACP, American College of Physicians; ACR, American College of Rheumatology; AMM, Academy of Medicine of Malaysia; ASCR, American Society of Clinical Rheumatologists; BSR, British Society for Rheumatology; CCCP, Chinese College of Cardiovascular Physicians; CRA, Chinese Rheumatology Association; CRA_multi, Chinese Multidisciplinary Expert Task Force on Hyperuricemia and Its Related Diseases; CSE, Chinese Society of Endocrinology; EULAR, European League Against Rheumatism; FMOH, Federal Ministry of Health (Nigeria); JSGNAM, Japanese Society of Gout and Nucleic Acid Metabolism; MOH, Ministry of Health Malaysia; MSR, Malaysian Society of Rheumatology; NIAMS, National Institute of Arthritis and Musculoskeletal and Skin Diseases; NIH, National Institutes of Health; PRA, Philippine Rheumatology Association; SAMA, South African Medical Association; SER, Spanish Society of Rheumatology; SIR, Italian Society of Rheumatology; T2T, treat-to-target; TRA, Taiwan Rheumatology Association; UT-Austin, University of Texas at Austin.
by limiting the maximum daily dose to 1.5 mg/eGFR in patients with renal impairment.72 HLA-B*5801 gene screening prior to allopurinol use was recommended by five guidance documents.14 15 21 22 37 38

For patients with asymptomatic hyperuricemia, 14 guidance documents14 15 17 21 36 39 40 43 47 49 51 52 commented on the option of pharmacological ULT, among which, 5 17 21 38 51 52 explicitly recommended no treatment under any circumstances. Three documents47–49 recommended pharmacological treatments in asymptomatic hyperuricemia patients with comorbidities47 48 or with very high SUA levels,40 47–49 but their cut-off SUA level to indicate ULT varied from 8.0 mg/dL47 48 to 13.0 mg/dL49. We also found that the Portuguese consensus40 was incoherent by itself stating that no pharmacological ULT was recommended as a general principle, but also stating that pharmacological ULT was recommended for patients with SUA higher than 9 mg/dL. No direct evidence was provided by any document to support pharmacological treatment for asymptomatic hyperuricemia, and such recommendations were only made in concern of the onset of gout49 and the risk of cardiovascular events.47 48

**Approaches to treatment for acute gout attack**

Twenty-one guidance documents14–17 19–22 36–41 43–46 48–52 covered the treatment for acute gout attack and online supplementary table 10 summarised their key recommendations. Non-steroidal anti-inflammatory drugs (NSAIDs) was recommended by all but three documents19 20 39 41 as the first-line pharmacological treatment, while colchicine by 11 documents.14–17 21 22 36 37 40 43 45 48 Colchicine was recommended to be given in a fixed dose by three documents38 40 48 and in a loading dose followed by different doses by six documents.14–17 19 20 22 38 51 52 Seven documents21 36 41 43 45 49 50 only recommended the total daily dose for colchicine, regardless of the regimen, and their doses recommended varied from 1 mg21 49 50 to 2.4 mg.52 Surprisingly, one document43 recommended 1.8 g colchicine in 24 hours without any further explanation, which was likely a typo. Systemic steroids were recommended by all but 3 documents,37 39 44 among which 614–17 19 20 36 43 recommended them as the first-line option and 1021 22 38 41 45 46 48 50–52 recommended them when NSAIDs and colchicine were contraindicated or intolerant. Intra-articular steroids injection was recommended by 14 documents,14–17 21 22 36 38 40 43 45 46 48 50–52 among which 54–16 21 36 43 clearly recommended it as the first-line option.

**Approaches to treatment for tophi**

Twenty-one guidance documents14–17 19–22 36–41 43–46 48–52 covered treatment for tophi and online supplementary table 11 showed their key recommendations. Surgery was recommended by nine documents,22 36 38 40 43 48 49 51 among which five39 36 38 43 49 explicitly presented its indications, most commonly nerve compression22 36 38 43 and infection.36 38 43 The risk for surgery was only discussed by one document39 and only the risk of delayed wound healing was stated. Long-term ULT was recommended by all but two documents,44 52 but the drugs used for pharmacological treatment was only explicitly recommended by eight of them.15–17 21 37 43 46 51

**DISCUSSION**

**Principal findings and interpretations**

This systematic review, including 16 guidelines and 8 consensus statements, found generally low methodological quality and inconsistent recommendations from guidance documents covering the diagnosis and management of gout and hyperuricemia. During revision of our work, the English version of two documents, from the Chinese Multidisciplinary Expert Task Force on Hyperuricemia and Related Diseases54 and the Taiwan Rheumatologist Association,55 respectively, were released. Despite increase in the number of guidance documents published between 2003 and 2017, the quality of documents in all domains did not seem to improve with time. To date, this is the first systematic appraisal for the quality of guidelines and consensus statements pertaining to both gout and hyperuricemia.

**Comparison with existing research**

Guidance documents assessed in our study performed well in the domains of scope and purpose (domain 1) and clarity of presentation (domain 4), but poorly in the domain of applicability (domain 5). These results were consistent with two previous reviews,56 57 one of which systematically assessed the quality of all guidelines for gout and the other assessed three documents released, respectively, by the 3e Initiative36 the ACR14 15 and the EULAR.18 58 Our study systematically included both guidelines and consensus statements in the field of both gout and hyperuricemia, and the diverse performance by different AGREE II domains was shared across both types of document.

This distribution of AGREE II domain scores has been observed by many previous guideline appraisal studies, in which documents scored higher in the scope and purpose domain and the clarity of presentation domain, and lower in the applicability domain and the editorial independence domain. This domain score distribution was not only shared by guidance documents for endocrinology diseases, such as diabetes59 60 and thyroid disorders,31 61 and rheumatology diseases, such as rheumatoid arthritis72 73 and systemic lupus erythematosus,64 but also shared by documents for diseases in other clinical specialities.33 65–67 Despite generally low and varied scores in the applicability domain, guidance documents for gout and hyperuricemia performed obviously poorer comparing with documents for other conditions,31–33 59–61 63–67 suggesting that improving the usefulness of guidance being more challenging in gout and hyperuricemia. One major impediment to good applicability of guidance document is the time and cost to perform economic evaluations and pilot studies, and a stable and long-term task force of guideline development is required.
to conduct these evaluations and studies. Although forming such a task force is practically difficult in some regions and countries, guidance documents were suggested to at least inform audience the need to consider these issues.60 Low scores in the editorial independence domain often resulted from lacking of detailed information on the influence of funding body and on the conflict of interests. We found that 50% of documents declaring funding sources were supported by the pharmaceutical industry, calling for awareness of the potential influence of pharmaceutical industry on the synthesis of clinical guidance and for the need of promoting transparency in financial declaration.

Clinical implications and future research
Guidance documents were concordant and recommended a target for SUA ≤6.0 mg/dL (or 360 μmol/L) for long-term control, to consider recurrent attacks as one of the indications for ULT (although the definitions for recurrent attacks differed), to consider allopurinol as the first-line ULT and NSAIDs as the first-line drug in acute attack, and to consider long-term ULT in patient with tophi. Despite these similarities, recommendations differed in the majority of items and these discrepancies might come from several sources, including ethnic difference, quality of documents and lack of evidence.

Ethical and social differences are important reasons why recommendations may vary between guidelines and consensus, and such diversity is to be encouraged, in order to best meet the needs of local populations. One example was that Asian guidance documents were more likely to recommend HLA-B*5801 gene screening before prescribing allopurinol.22 37 38 HLA-B*5801 gene screening was promoted because the risk of hypersensitivity reactions associated with allopurinol is significantly increased in individuals carrying the variant allele HLA-B*5801. Studies suggested that the frequency of this variant allele are higher in Han Chinese, Korean and Tai people than that in the Caucasian population,14 15 21 and that that HLA-B*5801 gene screening prior to allopurinol initiation is cost-effective for Asians but not Caucasians.60 61 These findings are consistent with the preferences of Asian documents. Providing ethnicity-specific recommendations or explicitly specifying the ethnicity of target audience helps clarify this source of inconsistency and improves the precision of recommendations.

However, it is worrying that low methodological quality of guidance documents may also lead to discrepant recommendations and consequent variability in application. Our study suggested that comparing to high-quality documents,14 15 17 21 36–40 43 47–49 51 52 and low-quality ones,22 37 38 44 47 52 56 were more likely to provide ambiguous prioritisation of both (1) ULT drugs for hyperuricemia and (2) steroid options for acute attack. A quick notice was that when making this rough summary, we considered a document to be high-quality when it scored above the upper quartile in at least three out of the six AGREE II domains, and to be low-quality when it scored below the lower quartile in at least three out of the six AGREE II domains. Among all AGREE II domains, those pertaining to stakeholder involvement, rigour of development, applicability and editorial independence could be improved by standardising developing processes, which consequently improved the reliability of recommendations. These results reinforced that it is better for clinicians to refer to high-quality guidance documents instead of the low-quality ones. However, when high-quality documents are unavailable in local language, referring to low-quality local documents might mislead clinical practice in the region. Selecting appropriate guidance documents to follow in clinical practice is thus more challenging for non-English speaking countries, including China.57 Moreover, the oldest document included in our study was the South African Medical Association guideline, published in 2003, and no guidance document in either English or Chinese was released in South African in the last 16 years. This finding suggested that some old documents might still affect regional practice. Efforts to timely update or declare the withdrawal of existing guidance documents are also critical for clinical practice.

Guidance documents are considered as the starting point to identify evidence gaps and to prioritise research questions.70 Evidence gaps were discussed in the recommendations of both (1) treatment for asymptomatic hyperuricemia, by 11 14 15 36 37 39 43 out of 14 documents,14 15 17 21 36–40 43 47–49 51 52 and (2) timing to initiate ULT, by 21 39 out of 14 documents.14–17 21 22 36–49 48 49 51 52 Although the rest of documents provided explicit recommendations, they based their recommendations either on indirect evidence or expert opinions. As for gout and hyperuricemia, evidence synthesis is warranted for the effects of pharmacological ULT in patients with asymptomatic hyperuricemia and for the optimal timing to initiate ULT in patients with the acute attack.

Strengths and limitations
Strengths of our review included a systematic approach to identify guidance documents pertaining to the diagnosis and management of hyperuricemia and gout. Both guidelines and consensus statements were evaluated and compared. We used the AGREE II instrument, an international, validated and rigorously developed tool, to assess the quality of document development and we tailored the AGREE II instrument to point-by-point scoring criteria (online supplementary file 1) to improve the objectivity and reproducibility of our study. We summarised all key recommendations, and compared and visualised the inconsistencies among them, providing a concise but informative overview for clinicians and researchers.

Our study also has limitations. First, we only included documents published in English or Chinese, which could lead to a risk of neglecting essential documents from regions not using English or Chinese as the first language. We attempted to mitigate this risk by tailoring our search strategy to identify the English versions of guidance documents published from these regions. Second, unconscious bias from a subjective rating of documents was inevitable. We avoided inviting coauthors of guidance documents as reviewers to prevent subconscious competing interest,
and conducted two rounds of group discussions to minimise subjective bias. Third, the AGREE II instrument itself has weaknesses, although it was the most commonly used tool to assess the quality of guidance documents. The AGREE system assigns equal weight to all six domains, regardless of their relative importance. Although better methods of guideline development and greater transparency of reporting are associated with more reliable recommendations, they do not guarantee better patient outcomes. Hence, the quality scores assessed by the AGREE II should be interpreted with caution, especially when used to indicate which guidelines to follow in clinical practice. Moreover, the subjective interpretation of scoring criteria impeded the replicability of AGREE II studies and direct comparison of quality scores in guidance documents provided by different reviews. Fourth, our literature search was over 12 months old when the study was ready to publish, affecting the timeliness of our study. However, we eventually decided not to update the literature at a late stage of the study, because of the infeasibility of bringing together all reviewers with another round of centralised training and appraisal, and the risk of inconsistent scoring criteria for each reviewer after a long time since their previous scoring. Moreover, a quick review of publications in PubMed, using the same search strategy (online supplementary table 1) and limiting the publication date from 1 September 2016 to 21 January 2019, did not find any new relevant documents, reassuring us of the timeliness of our study.

CONCLUSIONS
The methodological quality needs to be improved in the current guidelines and consensuses on the diagnosis and management of gout and hyperuricemia, as assessed by the AGREE II. Inconsistent recommendations are common, even in some key aspects. Promoting standard methods for guidance documents development, and synthesising high-quality clinical evidence to fill in evidence gaps, are warranted to improve the quality of guidance documents.

Author affiliations
1Department of Endocrinology and Metabolism, West China Hospital, Sichuan University, Chengdu, China
2Institute of Health Informatics, University College London, London, UK
3Department of Gastroenterology, West China Hospital, Sichuan University, Chengdu, China
4Department of Oto-Rhino-Laryngology, West China Hospital, Sichuan University, Chengdu, China
5Jockey Club School of Public Health and Primary Care, Prince of Wales Hospital, The Chinese University of Hong Kong, Hong Kong, China
6The Second Clinical College, Nanjing University of Chinese Medicine, Nanjing, China
7Chinese Evidence-Based Medicine Center, West China Hospital, Sichuan University, Chengdu, China
8Farr Institute of Health Informatics Research, University College London, London, UK
9Health Data Research UK London, University College London, London, UK
10The National Institute for Health Research University College London Hospitals Biomedical Research Centre, University College London, London, UK
11Evidence-Based Medicine Center, School of Basic Medical Sciences, Lanzhou University, Lanzhou, China
12Division of Population Health and Genomics, Ninewells Hospital and School of Medicine, University of Dundee, Dundee, United Kingdom

Author's contributions
HT and SL conceived this study.QL, JS-WK and SL designed the inclusion/exclusion criteria and the searching resource and strategy. QL, JS-WK, HC, LL and XS designed the appraisal strategy of each included guideline and consensus. QL and XL searched literature search and extracted data. QL, XL, JW, HL and SL assessed the quality of each document. QL analysed and visualised the outcomes. S-CC, AS, YC, ZA, XS and HH provided critical review. QL, XL and SL drafted the manuscript. All authors discussed actively in the protocol of the study.

Funding
This research received no specific funding from any bodies in the public, commercial or not-for-profit sectors. AS is supported by the National Institute for Health Research (NIHR) University College London Hospitals Biomedical Research Centre and a THS Institute postdoctoral fellowship. HH is a NIHR Senior Investigator. His work is supported by: (1) Health Data Research UK, which is funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directors, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation and Wellcome Trust. (2) The BigData@Heart Consortium, funded by the Innovative Medicines Initiative-2 Joint Undertaking under grant agreement No 116074. This Joint Undertaking receives support from the European Union’s Horizon 2020 research and innovation programme and EFPA; it is chaired, by DE Grobbee and SD Anker, partnering with 20 academic and industry partners and ESC. (3) The NIHR University College London Hospitals Biomedical Research Centre. SL was supported by grants from the National Natural Science Foundation of China (grant number 81400611 and 21534008), National Basic Research Program of China (grant number 2015CB942800), the Scientific Research Project of Health and Family Planning Commission of Sichuan Province (grant number 130029, 150149, 17PJ063 and 17PJ445), Cholesterol Fund by China Cardiovascular Foundation and China Heart House and the International Visiting Program for Excellent Young Scholars of Sichuan University.

Competing interests
None declared.

Patient consent for publication
Not obtained.

Provenance and peer review
Not commissioned; externally peer reviewed.

Data availability statement
All data in this paper were obtained from published studies. No additional data are available from the authors.

Open access
This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES


