Supplemental material 4: Detailed methods for consensus process

Systematic literature review: data collection

The systematic literature review addressed the domains of interest (see S2), covering the disease definition, name, organization of care, diagnostics and treatment. A search was conducted for Embase, Emcare, Web of Science, and Cochrane (initial search February 2021) and Pubmed (initial search May 11th 2023, last update April 18th 2024) The search string (see below) aimed to retrieve literature on the full spectrum of CNO, also including pediatric studies. In total, n=1385 papers were retrieved, excluding duplicates. Per domain of interest, papers were selected by searching title, abstract, and full text for key words (specified in the summary of evidence). Relevance of identified papers was assessed by reading full text by ATL. For several themes, the search was limited at the discretion of ATL, EMW and OMD to recent (systematic) reviews or studies with larger patient numbers only, as specified in the summary of evidence. The summary of evidence focused mainly on literature for adult CNO. When evidence from pediatric studies was considered relevant , this is explicitly mentioned in the text. In addition, the current European Alliance of Associations for Rheumatology (EULAR) recommendations for axial spondylarthritis (axSpA) and psoriatic arthritis (PsA) were consulted as these patients populations bear some clinical resemblance to adult CNO. Relevant considerations are embedded in the summary of evidence, with explicit mentioning that they were conceived for axSpA or PsA as appropriate. Lastly, for several domains that surpass the domain of CNO in specific (e.g. on differential diagnoses), supportive evidence was added manually. Current clinical practice standards were derived from a physician survey study that has been published previously (1).

Systematic literature review: full search string

(("Hyperostosis, Sternocostoclavicular"[majr] OR "sterno-costo-clavicular hyperostosis"[ti] OR "sternocostoclavicular hyperostosis"[ti] OR "sterno costoclavicular hyperostosis"[ti] OR "sternocosto clavicular hyperostosis"[ti] OR ("SCCH"[ti] AND "hyperostosis"[ti]) OR "SAPHO"[ti] OR "Acquired Hyperostosis Syndrome"[majr] OR "acquired hyperostosis syndrome"[ti] OR "Acquired Hyperostosis"[ti] OR "Acute pseudoseptic arthritis and palmoplantar pustulosis"[ti] OR "Aseptic osteomyelitis"[ti] OR "Bilateral clavicular osteomyelitis"[ti] OR "CRMO"[ti] OR "Chronic mandibular osteomyelitis"[ti] OR "chronic multifocal osteomyelitis"[ti] OR "Chronic multifocal symmetrical osteomyelitis"[ti] OR "chronic non bacterial osteomyelitis"[ti] OR "chronic non hematogenous osteomyelitis"[ti] OR "chronic recurrent multifocal osteomyelitis"[ti] OR "Chronic recurrent osteomyelitis"[ti] OR "Chronic sclerosing osteitis"[ti] OR "Chronic symmetric osteomyelitis"[ti] OR "Clavicular hyperostosis"[ti] OR "Clavicular periosteal new bone formation"[ti] OR "Condensing osteitis of the clavicle"[ti] OR "Diffuse sclerosing osteomyelitis"[ti] OR "Hyperostosis Syndrome"[ti] OR "Intersternocostoclavicular ossification"[ti] OR "Multifocal chronic osteomyelitis"[ti] OR "Multifocal sterile osteomyelitis"[ti] OR "Musculoskeletal syndromes associated with acne"[ti] OR

"non bacterial osteitis"[ti] OR "non bacterial osteomyelitis"[ti] OR "nonbacterial osteitis"[ti] OR "nonbacterial osteomyelitis"[ti] OR "Non-infectious osteitis"[ti] OR "Osteomyelitis of the bilateral clavicles"[ti] OR "pustulotic arthritis"[ti] OR "Pustulotic arthro-osteopathy"[ti] OR "Recurrent hyperostosis of the clavicle"[ti] OR "Sclerosis and hyperostosis of the manubrium sterni"[ti] OR "Spondylarthropathy with hidradenitis suppurativa and acne conglobata"[ti] OR "sternocostoclavicular osteoarthritis"[ti] OR "sternocostoclavicular pain"[ti] OR "sternocostoclavicular syndrome"[ti] OR "Subacute and chronic symmetrical osteomyelitis"[ti] OR "pustulotic arthro-osteitis"[ti] OR "pustulotic arthroosteitis"[ti] OR "sternocostoclavicular arthro-osteitis"[ti] OR "sternocostoclavicular arthroosteitis"[ti] OR "inter-sterno-costo-clavicular ossification"[ti] OR "intersterno costoclavicular ossification"[ti] OR "intersternocostoclavicular ossification"[ti] OR ("arthro-osteitis"[ti] AND "pustulosis"[ti])) AND english[la]

Systematic literature review: methods for appraising level of evidence

An appraisal of the quality of evidence was evaluated using the GRADE approach as outlined in the Cochrane Handbook for Systematic Reviews of Interventions (2). GRADE was performed by ATL and checked with EMW and OMD.

- For each statement or outcome in this Summary of Evidence, level of evidence is rated as “very low”, “low”, “moderate”, or “high”.
- Level of evidence may differ per clinical context. For example, for a certain laboratory marker, we may have “high” level of evidence on how often or how much it is increased, but we have no studies (“very low”) level of evidence on whether the evaluation of this marker improves diagnostic certainty or patient outcomes. Similarly, we may have a fair estimate on the prevalence of asymptomatic lesions (“moderate”), yet no studies on whether screening for them has clinical consequence or improves patient outcomes (“very low”). In such cases, the level of evidence is based on clinical utility to improve patient outcomes. This is specified in the summary of evidence where relevant.
- Evidence deriving from observational studies pertaining to clinical features in CNO/SAPHO was initially categorized as “high” level of evidence as for prevalence estimates randomization does not increase the validity.
- Evidence deriving from randomized controlled trials (RCTs) was initially categorized as “high” level of evidence.
- Evidence deriving from non-randomized studies of intervention (NRSIs) was initially categorized as “low” level of evidence
- Level of evidence was lowered in the presence of the following factors:

    - Significant risk of bias due to e.g.: confouding (NRSIs), absence of allocation concealment or absence of blinding (RCTs), loss to follow-up, selective reporting of outcomes.
    - Inconsistency in findings across studies (heterogeneity)
    - Indirectness (poor applicability of study results to the population of adult CNO/SAPHO)
    - Imprecision (large uncertainty in effect estimates, or few/only one study)
- Level of evidence was decreased by one level for each factor that prompted "serious concerns", or by two levels for each factor that prompted "very serious" concerns.
    - If only one/very few studies were available for the question at hand, level of evidence was marked as "very low" on the base of very serious imprecision.
    - Due to the general scarcity of literature in adult CNO/SAPHO, the presence of publication bias, which traditionally also lowers the level of evidence, could not be properly assessed.
- Level of evidence was increased one level for each of the following factors:
    - Particularly large effect sizes
    - Presence of dose-response relationships
    - Plausible residual opposing confounding

Expert panel constitution

To assemble a diverse expert panel, several strategies were employed (see also **figure A**). Firstly, individuals who had participated in the physician survey study were invited. These individuals had been approached via European Alliance of Associations for Rheumatology (EULAR), Rare Immunodeficiency, Autoinflammatory and Autoimmune Diseases Network (ERN RITA), European Reference Network on Connective Tissue and Musculoskeletal Diseases (ERN ReCONNET), European Reference Network on Rare Bone Disorders (ERN BOND), European Society of Endocrinology (ESE), European Calcified Tissue Society (ECTS), Rare Bone Disease Action Group, South East Asia and Pacific Area League Against Rheumatism (APLAR), Japan College of Rheumatology (JCR), American Society for Bone and Mineral Research (ASBMR), International Osteoporosis Foundation (IOF), International Federation of Musculoskeletal Research Societies (IFMRS)) (n=32) (1). Simultaneously, invitations were sent to first and last authors of CNO-related publications from the past 5 years, if not participated in the survey study already (n=46, total number of invited participants n=78). Moreover, the consensus initiative was disseminated via the aforementioned networks, with the addition of Childhood Arthritis and Rheumatology Research Alliance (CARRA), the American College of Rheumatology (ACR), Paediatric Rheumatology European Society (PRES), European Association of Nuclear Medicine (EANM) and the professional networks of EMW. This led to spontaneous applications from associated experts (n=21 in total). Recognizing the scarcity of expertise in CNO, the

initiative remained open to participants even after its formal commencement. ATL, OMD, and OB fulfilled the tasks of research fellow, supervising methodologist and chair, and minutes secretary respectively and were not eligible to formally vote.

Formulation of the expert consensus recommendations

A schematic overview of the process for the development of the recommendations is presented in Figure 1. A two-survey Delphi process was started in May 2023. Both Delphi surveys were executed using the General Data Protection Regulation-compliant system of Calibrum (Surveylet); as only participant's opinions were assessed, no personal details of individual patients were gathered. Statements for the first Delphi were derived from the Summary of Evidence by ATL, EMW and OMD, and piloted by two collaborators who had previously participated in the preparatory physician survey study (GC and HGZ) (see **S5** for Delphi survey 1, including data overview). Experts were invited to provide anonymous ratings for each statement on a 9-point Likert scale, along with the opportunity for free-text commentary, such as suggestions for reformulations or additional content.

In total, 44 experts completed Delphi survey 1, of whom 36 were invited and 8 had volunteered via relevant networks (see above). Consensus on a specific statement was defined as a median score of at least 7/9 (indicating positive agreement with the statement), with an interquartile range (IQR) no larger than 25% of the total scale (indicating acceptable spread of scores). Group median scores, degree of spread, bipolarity assessments, and stakeholder group differences were analysed with SPSS Statistics version 25, IBM corp.

The results, plus a compilation of free-text comments, were made available before the first digital pre-meeting (June 2023, attended by n=39/44 survey completers). This meeting involved a structured discussion on the results of the first Delphi survey, with emphasis on dissent items.

Afterwards, a second Delphi survey was developed (see **S6** for the survey, including data overview). At this point, 2 experts withdrew from the project due to time constraints, and 13 more had self-applied via relevant networks. Of the total of 55 enrolled experts, n=43 completed the second Delphi survey in August 2023.

Following similar analysis strategies, the results of the second Delphi survey set the framework for a two-day in-person meeting held in October 2023, attended by 36 out of 55 panel members. This meeting incorporated in-depth discussions across all domains of interest, including a session featuring presentations from imaging experts (ANC, TD, FS, JT), a session with representatives from the Dutch CNO patient association, and a round-table session on a future research agenda.

Of note that after the second Delphi round, we transitioned from formal level of agreement metrics to a more open discussion format. The discussions were thematically organized around key topics such as disease definition, naming conventions, diagnostics, and treatment strategies. Each session began with an overview of the data from both Delphi rounds to inform the discussions, which were led by an independent methodologist. In most cases, the panel was able to reach unanimous consensus on the text presented in the final manuscript. However, in instances of differing opinions, votes were held to capture the majority viewpoint while addressing opposing perspectives as considerations within the recommendations. Minutes were kept to be able to reiterate panel member's individual viewpoints in follow-up meetings and during manuscript revisions.

Synthesizing all information from the Summary of Evidence, the two Delphi surveys, and the two meetings, a draft recommendations were prepared and circulated for feedback. A digital follow-up meeting was held in February 2024 to resolve remaining points of discussion. Subsequently, a revised version of the document was circulated and amended. Eventually, the final recommendations were also rated by the full panel on a 0-10 Likert scale, with 0 indicating no agreement and 10 indicating full agreement. Level of agreement metrics are displayed in the recommendation tables with mean score and standard deviation, as well as the proportion of the panel rating the recommendation 8/10 or higher. Authorship on the eventual manuscript was determined based on active participation, meeting attendance, and adherence to the International Committee of Medical Journal Editors (ICMJE) authorship guidelines.

References cited in this supplement:

1. Leerling AT, Clunie G, Koutrouba E, Dekkers OM, Appelman-Dijkstra NM, Winter EM. Diagnostic and therapeutic practices in adult chronic nonbacterial osteomyelitis (CNO). Orphanet J Rare Dis. 2023;18(1):206.
2. Schünemann HJ HJ, Vist GE, Glasziou P, Akl EA, Skoetz N, Guyatt GH. Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August 2023). Cochrane, 2023. Higgins JPT TJ, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editor: Cochrane; 2023.