BMJ Open

Efficacy and safety of balneotherapy in rheumatology: a systematic review and meta-analysis

Ikram Aribi,[1] Mikail Nourredine,[1,2] Caroline Giroudon ,[3] Emmanuel Massy ,[4] Jean-Christophe Lega ,[1,2] Behrouz Kassai,[1,2,5] Guillaume Grenet [1,2]

To cite: Aribi I, Nourredine M, Giroudon C, *et al*. Efficacy and safety of balneotherapy in rheumatology: a systematic review and meta-analysis. *BMJ Open* 2025;**15**:e089597. doi:10.1136/bmjopen-2024-089597

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2024-089597).

Received 04 June 2024
Accepted 11 February 2025

[1]Laboratoire de biométrie et biologie évolutive, Université Lyon 1, UMR CNRS 5558, Lyon, France
[2]Service hospitalo-universitaire de pharmaco-toxicologie, pôle de santé publique, Hospices Civils de Lyon, Lyon, France
[3]Team aux chercheurs, service de la documentation centrale, Hospices Civils de Lyon, Lyon, France
[4]Service de rhumatologie, centre hospitalier Lyon Sud, Hospices Civils de Lyon, Lyon, France
[5]Centre d'investigation clinique, INSERM CIC 1407/UMR 5558 CNRS, groupement hospitalier Est, Hospices Civils de Lyon, Lyon, France

Correspondence to
Dr Guillaume Grenet;
guillaume.grenet@ubc.ca

ABSTRACT

Objective The efficacy of balneotherapy in rheumatology remains unclear. We aimed to estimate its benefits and risks in rheumatology.

Methods We conducted a systematic review of randomised trials assessing any European balneotherapy for a rheumatological indication in adults versus any control, on clinical outcomes. We searched PubMed, Cochrane Library, Embase and https://clinicaltrials.gov/ (up to 28 November 2023). We used the Cochrane risk of bias tool version 2, funnel plot and asymmetry tests. We used a random effects model with an inverse-variance weighting method for standardised mean difference (SMD) and risk ratio (RR). We used the Grading of Recommendations Assessment, Development and Evaluation approach for two primary outcomes, pain and quality of life (QoL) at 3 months, and two safety outcomes, withdrawal and any adverse event (AE).

Results We included 29 trials in mechanical disorders, 9 in inflammatory diseases and 4 in fibromyalgia. The synthesis suggested a decrease in pain of a very low level of certainty (SMD: −0.72 (95% CI (−1.00; −0.44)), very serious risk of bias and of inconsistency, publication bias strongly suspected); an increase in QoL of a very low level of certainty (SMD: 0.56 (95% CI (0.37; 0.75)), very serious risk of bias and serious risk of inconsistency); inconclusive results regarding the risk of withdrawal (RR: 0.75 (95% CI (0.46; 1.20)), very serious risk of bias and serious risk of imprecision) and of AE (RR: 0.80 (95% CI (0.43; 1.50)), serious risk of bias and of inconsistency and very serious risk of imprecision).

Conclusion The certainty of the effect of balneotherapy in rheumatology was very low.

PROSPERO registration number CRD42023448206.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ A systematic review of randomised trials assessing balneotherapy in rheumatology was conducted.
⇒ The primary outcomes were the pooled treatment effect of balneotherapy for pain and quality of life at 3 months after the intervention.
⇒ Two sensitivity analyses, restricted to specific comparators (placebo-like and standard of care), assessed the robustness of the findings.
⇒ Two subgroup analyses explored potential heterogeneity in the treatment effect (according to the underlying indication for the intervention and the type of balneotherapy).
⇒ Grading of Recommendations Assessment, Development and Evaluation was used to assess the certainty of the evidence.

INTRODUCTION

Rational

Baths have been used to treat various orthopaedic conditions since ancient times.[1] In European countries, the term 'balneotherapy' is typically used to describe bathing therapy based on natural mineral or thermal waters.[1] Balneotherapy is mostly prescribed to patients with any form of arthritis.[1] However, the mechanism of action of balneotherapy in rheumatic diseases remains unclear. It could be a combination of mechanical, thermal and chemical effects[2]; for instance, hydrostatic pressure may contribute to alleviating symptoms in rheumatic diseases such as osteoarthritis,[3] balneotherapy might also help dissipate algogenic chemicals in inflammatory diseases such as rheumatoid arthritis[1] and could decrease oxidative stress in fibromyalgia.[4] Trials assessing balneotherapy in osteoarthritis have suggested a potential effect on pain but using an outdated grading system of the evidence,[5] and for low back pain, some benefits have been suggested but were mostly based on trials at high risk of bias (ROB).[6] For inflammatory disorders, previous meta-analyses reported contradictory results, from a lack of evidence[1] to a potential benefit[7] in patients with rheumatoid arthritis. In patients with spondyloarthritis, pain and quality of life (QoL) might be improved, but the validity of these findings was low.[8] Consequently, health insurance systems question the reimbursement of balneotherapy, as recently highlighted in France.[9] The cost-effectiveness of balneotherapy has also been debated in other European countries such as Spain,[10] Italy[11] and the Netherlands.[12] Indeed, the annual health insurance expenditures can exceed

€10 million as reported in Hungary.[13] The previous systematic reviews were restricted to specific indications, each including fewer than 10 trials[1 5 7 8]; or included at least 1 non-randomised trial while not reporting pooled estimates of the treatment effect.[6] The power of these analyses and the generalisability of their findings are therefore limited. Moreover, at least 3 new randomised trials totalling almost 500 participants recently assessed balneotherapy in rheumatology.[14–16] In this context, there was a need for updating and broadening the evidence synthesis of the potential effect of balneotherapy in any rheumatological indication.

Objective

The main objective was to estimate the benefits and risks of balneotherapy for its indications in rheumatology.

METHODS

The review methods (the review question, the search strategy, the inclusion/exclusion criteria, the tool for ROB assessment, the synthesis plan and the plan for investigating causes of heterogeneity) were established prior to conducting the review. The completed Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist[17] is available in online supplemental material S1. We followed international guidance on conducting evidence synthesis.[18]

Patient and public involvement

Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

Eligibility criteria

The PICO(S) was:

Patient: Adults only, with a rheumatological indication for balneotherapy. The two previous Cochrane reviews were limited to less than 10 trials each.[1 5] Thus, we aimed to assess the intervention in a broader population, to increase the power of the analysis. Moreover, such broad orientation allows matching less precise health insurance terminology (as in France, eg).

Intervention: Any balneotherapy undertaken in Europe, of any duration >10 days, based on any natural mineral water, mud, steam and any adjuvant treatment (including adjuvant physiotherapy). We followed the large definition of balneotherapy, in line with a previous Cochrane review[1] and the previous definition of what is balneotherapy.[19] Balneotherapy undertaken in Europe was first defined as interventions conducted in countries represented in the European spa association (ESPA);[20] if the intervention was conducted in a European country that did not appear among the countries represented in the ESPA, the decision to include the study was made on a case-by-case basis, and after a consensus was reached. We limited our review to Europe as many geographical factors could impact the effect of balneotherapy.[19] Therefore, the larger the geographical area considered, the greater the risk of heterogeneity. We believe that the European area provides a balance between sufficient power and not too much heterogeneity. Moreover, the term 'balneotherapy' is typically used in European countries.[1] Finally, balneotherapy was mostly developed in Europe.[2]

Control: Any control (standard of care (SOC) without balneotherapy, 'pseudo-balneotherapy' <11 days, no treatment, etc), as in the previous Cochrane reviews that compared the balneotherapy with 'another intervention or with no intervention'.[1 5]

Outcome: All clinical outcomes that were the primary outcome of the trial, including clinical scales and QoL, validated at least by a national learnt society; trials assessing balneotherapy on non-clinical outcomes were excluded.

Study type: Randomised controlled trials (RCT), assessing superiority or non-inferiority, and of multi or single-centre design. The review was limited to randomised trials to limit the ROB.

The other eligibility criteria were: (1) time frame/years considered: no time restriction; language: English reports only (non-English language would need significant supplementary workforce for low to no impact on treatment estimate).[21]

Publication status: any.

Information sources

We undertook a comprehensive literature search using the main electronic databases PubMed [including MEDLINE], Embase [Elsevier] and Cochrane Library [Wiley]). These three databases were searched on 24 July 2023. Alerts were set up for all the databases queried and were stopped on 28 November 2023. We also searched for unpublished studies, reports and grey literature in reference lists, previous reviews on the same topic (review register: PROSPERO), trial register (clinicaltrials.gov), congress proceedings (International Society of Medical Hydrology and Climatology, World Federation of Hydrotherapy and Climatotherapy) and asking medical experts.

Search strategy

We used a combination of free-text and thesaurus terms for the concepts relevant to the topic. Searches were limited to documents published in English; no date restrictions were applied. The algorithms were developed with an information specialist (CG) and are available (online supplemental material S2).

Data management

We used the Covidence platform[22] for bibliographic records management and data extraction.

Selection process

Two reviewers (IA and GG) conducted the selection process using a standardised template implemented in the Covidence platform at each step (screening on title and abstract then selection on full text). This was done

independently, in duplicate. Consensus was searched for in case of disagreement by discussion among the authors.

Data collection process

Two reviewers (IA and GG) extracted the data using a standardised template implemented in the Covidence platform. Data extraction was checked. Consensus was searched for in case of disagreement by discussion among the authors. Study characteristics, population and setting characteristics, and outcome measures were extracted.

Data items

For the treatment effect on continuous outcomes, the point estimate according to the available data was extracted: postintervention mean-value and mean-change versus baseline, with its standard deviation (SD, calculated from the CI if this was reported instead of the SD), at the available time points. If the data were not available in a table but in a figure, we extracted the estimate from the figure. If a trial reported data at 6 months and 9 months but not 12 months after intervention, the 9-month data were used as proxy for 12 months. For the treatment effect on dichotomous outcomes, the number of events and the number of randomised participants in each arm were extracted.

Outcomes and prioritisation

Following the white paper of the French Society of Pharmacology and Therapeutics (Société Française de Pharmacologie et de Thérapeutique), the present meta-analysis should be considered a retrospective study.[23] Therefore, the analyses and results are exploratory only. As we expected heterogeneity in the outcomes reported in RCTs, the review focused on outcomes that are both (1) clinically relevant to the patient and (2) expected to be usually available. Moreover, the review included patients suffering from different diseases. Therefore, it was important that the outcomes allowed providing a treatment effect estimate independently of the background physiopathology. In this view, the two primary efficacy outcomes were pain intensity and QoL. We used the pain assessment scale as reported in the RCT (10 points or 100 points). The scale reported in the included trial to assess the QoL was used, and when several QoL measurements were reported, the less disease-specific measures, such as the generic Medical Outcome Study Short Form-36,[24] were prioritised in order to limit potential heterogeneity due to the underlying disease. A 3-month follow-up was defined *a priori* as the primary efficacy outcome for both pain and QoL.

The secondary efficacy outcomes were pain intensity and QoL at 6 months, 12 months and after the intervention (defined as: 'immediately' after or 'shortly' after (≤1 month) or 'during' or 'at the time' of the intervention, according to the available data).

The safety outcomes were withdrawal (due to adverse events (AEs) or serious AEs (SAEs) or for any reason according to the available data), SAEs and AEs.

ROB in individual studies

We used the ROB 2 tool[25] to assess the ROB of included studies. The ROB of each trial was assessed independently in duplicate by two reviewers (among IA, BK and GG). Consensus was searched for in case of disagreement by discussion among the authors. The ROB was assessed for the primary outcomes that are both continuous variables with potential missing data and both subjective outcomes. The ROB assessment was conducted to assess the effect of assignment to the interventions at baseline (intention to treat analysis).

Data synthesis

The review is limited to aggregated data. For the pain and the QoL outcomes, different scales were available. The measures of pain are pointed in the same direction: the lower the better for the patient (ie, less pain). For the QoL, some scales indicated a better health status by a higher score (the higher the better), while others indicated a better health status with a lower score (the lower the better). We multiplied the postintervention mean value of the lower the better QoL measures by −1 to ensure that all the QoL measures point in the same direction.[26] In the meta-analysis, pain outcomes were therefore in the direction that lower is better and QoL outcomes that higher is better. Because of the different scales, standardised mean difference (SMD) was used for pooling the estimates. As the mean change and postintervention mean value should not be combined when using SMD,[27] a distinct synthesis for mean change on one hand, and for postintervention mean value on the other hand is provided. However, for exploratory purposes only, we also reported a combined estimate as it has also been reported that combining these measures might not change the results.[28]

Summary measure

The inverse-variance weighting method was used to provide a pooled estimate of the balneotherapy effect (point estimate and its 95% CI) for each outcome. A random-effects model was used to allow the true population effect size to differ among studies. A restricted maximum likelihood estimator for τ^2 was used. To compute the summary effect's CIs, both the conventional method for random effects and the Hartung-Knapp modification[29] were used, and we kept the most conservative method (*ie.*, the method that produced the largest CI).[30] I^2 with its 95% CI was used to assess the heterogeneity of effect sizes. Statistical analysis was conducted using the R software[31] (V.4.3.1), in particular the package meta[32] (V.6.5-0). The statistical analysis was conducted in December 2023, that is, after the protocol registration and its amendments.

Additional analyses

A sensitivity analysis and subgroup analyses for the primary outcomes were planned *a priori*. We amended

the protocol as described in the section 'Registration and protocol'.

Sensitivity analysis

Differences in comparators might increase heterogeneity. To mitigate this risk, we conducted sensitivity analyses to assess the robustness of the results: restricted to placebo-like trials and restricted to the more frequent type of control arm.

Subgroup analyses

Subgroup analyses were conducted to explore potential heterogeneity in the treatment effect: (1) the potential impact of the type of balneotherapy, classified as bath, mud pack, bath plus mud pack and other, adapted from a previous definition[19] and (2) the potential impact of the underlying disease, classified by the main indication as mechanic for mechanical disorders, inflammatory for inflammatory and autoimmune diseases, and fibromyalgia for fibromyalgia. We did not use narrower disease categories to maintain the number of trials, and therefore, power, for subgroup analyses.

Reporting bias

Reporting bias was investigated using standard analyses (funnel plot,[33] Egger's[34] and Begg's[35] tests) for the two primary outcomes (pain and QoL at 3 months), and for the two safety outcomes that have been assessed using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach[36] (see below).

Confidence in cumulative evidence

Estimates of the strength of the evidence were provided following the GRADE approach[36] for the two primary outcomes (pain and QoL at 3 months) and for two safety outcomes (withdrawal and AE).

RESULTS

Study selection

We identified 2395 records from the bibliographic search. After removing duplicates and screening on title and abstract, 104 full-text records were excluded, mostly because of the intervention. We were unable to include three records because of a lack of information regarding their design despite contacting or trying to contact the authors.[37–39] In total, 42 studies[14–16 40–78] were included in the review (figure 1).

Among the included trials, seven were not two-arm parallel design. Among these, (1) four trials were three-arm parallel design, for two of which one arm was not included in the review (a short wave diathermy arm[43] and a mud arm that was not a thermal mud),[77] and for the other two, two arms were combined in one arm (same balneotherapy but at 36°C for one arm and at 38°C for the second intervention arm,[63] and two balneotherapy arms but in two different centres);[76] (2) one trial was a four-arm parallel design, of which two arms were excluded because they were neither balneotherapy nor control arms[65] and (3) two trials were cross-over trials, of which one trial reported only the first period[72] and one trial reported usable data for the first period and non-usable data of the second period.[73] All included trials were analysed as a two-arm parallel design.

Study characteristics

The oldest trial was published in 1989[74] and the most recent in 2023.[15] Most of the trials were conducted in Italy (Italy: k=13, Hungary: k=12, France: k=6, Germany: k=2, Spain: k=2, Portugal: k=2, Germany and Austria: k=1, Austria: k=1, Lithuania, Romania and The Netherlands: k=1 each). The main indications for balneotherapy were mechanical disorders (number of trials, k=29), inflammatory diseases (k=9), fibromyalgia (k=4). The intervention type was classified as: bath (k=21), bath plus mud pack (k=13), mud pack (k=3), other (k=5) (classification: see table 1; details: see online supplemental material S3). The control arms were classified as: SOC (k=17), placebo-like (k=13), other (k=9), waiting list (k=3). The duration of follow-up ranged from 1.5 to 52 weeks, the proportion of women from 6.7 to 98.0%, age from 40.6 to 75.5 years, and body mass index ranged from 24.8 to 29.9 kg/m^2 (table 1).

ROB in the included trials

No included trial was at low risk of overall bias, 34 (81%) were at high risk of overall bias, and 8 (19%) had some concerns. In particular, deviations from intended interventions were the domain with the greatest frequency of high ROB (66.7% of the trials), while the randomisation process had the lowest frequency of high ROB (11.9% of the trials) (online supplemental material S4a). Among the 13 trials using a placebo-like design,[14 40 41 45 53 57 58 62 67–69 74 75] the overall ROB was high for 8 (61.5%)[41 45 53 57 62 67–69] (online supplemental material S4b). Trial sponsors were not systematically reported (online supplemental material S4c).

Results of individual studies

The point estimate (expressed as an SMD) of individual studies ranged from −2.43 to 0.22 for pain at 3 months (figure 2); and from 2.32 to 0.03 for QoL at 3 months (figure 3).

Results of syntheses

Primary outcomes

Pain at 3 months

Among the 21 trials reporting this outcome (2163 patients), 16 were at high ROB, 5 with some concerns. Regarding pain reported as the mean change from baseline, the intervention was associated with an SMD of −0.31 (95% CI (−0.54; −0.08)), with substantial heterogeneity (I^2=54%). Regarding pain reported as postintervention mean value, the intervention was associated with an SMD of −0.89 (95% CI (−1.25; −0.53)), with considerable heterogeneity (I^2=85%; figure 2).

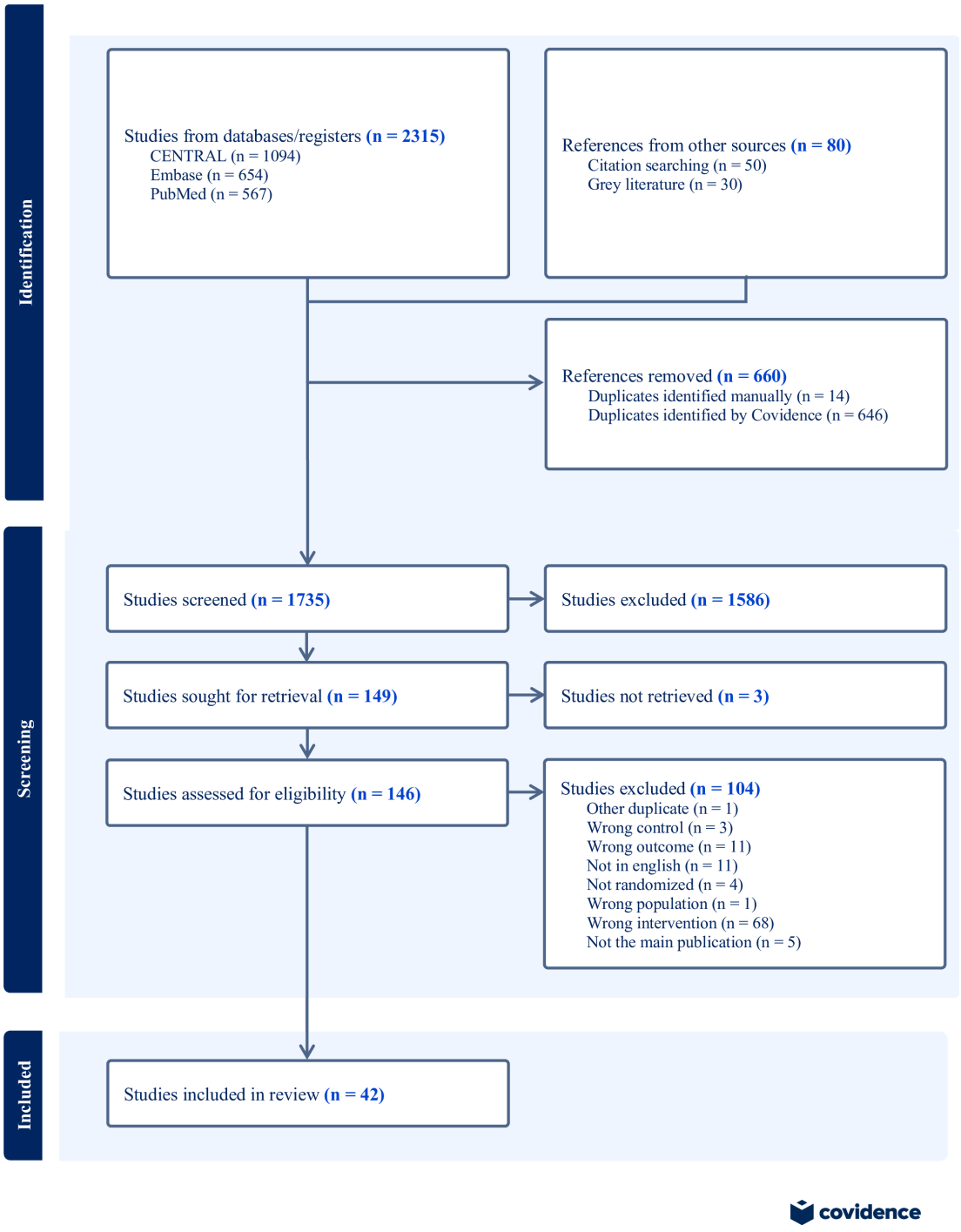

Figure 1 Flow chart of the systematic review, following PRISMA guidance (extracted from Covidence). PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

QoL at 3 months

Among the 18 trials reporting this outcome (1194 patients), 15 were at high ROB, 3 with some concerns. Regarding QoL reported as mean change from baseline, the intervention was associated with an SMD of 0.33 (95% CI (0.09; 0.56)), with low heterogeneity (I^2=12%). Regarding QoL reported as postintervention mean value, the intervention was associated with an SMD of 0.64 (95% CI (0.39; 0.88)), with substantial heterogeneity (I^2=62%; figure 3).

Secondary outcomes

Efficacy outcomes

Pain at 6 months was reported in 12 trials (2340 patients). The SMD was −0.26 (95% CI (−0.37; −0.16); I^2=0%) when reported as mean change from baseline, and −0.52 (95% CI (−0.71; −0.34); I^2=17%) when reported as postintervention mean value (online supplemental material S5).

Pain at 12 months was reported in 5 trials (1086 patients). The SMD was −0.11 (95% CI (−0.24; 0.02);

Table 1 Characteristics of included trials

Study identifier	Main indication	Intervention type	Control type	Single or multicentre	Follow-up	Women	Age	BMI
Annegret 2013[40]	Mechanics	Bath	Placebo-like	Single	36	60.0	58.3 (11.1)	27.5 (4.8)
Bálint 2007[41]	Mechanics	Bath	Placebo-like	Single	12	NA	NA	NA
Benini 2021[42]	Mechanics	Mud pack	Other	Single	36	91.1	65.8 (8.1)	
Cantarini 2007[43]	Mechanics	Bath-mud pack	SOC	Single	12	63.5	64.0 (8.6)	
Cantista 2020[44]	Mechanics	Bath	SOC	Single	12	66.7	75.5	27.8
Caporali 2010[45]	Inflammatory	Bath-mud pack	Placebo-like	Single	24	79.0	58.8 (7.8)	24.8 (3.8)
Chary-Valckenhaere 2013[46]	Mechanics	Bath-mud pack	WL	Multi	28	52.2	57.5 (9.7)	27.2 (5.1)
Ciprian 2013[47]	Inflammatory	Bath-mud pack	SOC	Single	24	6.7	46.7 (10.9)	NA
Cozzi 2015[48]	Inflammatory	Bath-mud pack	SOC	Single	6.4	69.0	53.2 (12.6)	NA
Espejo Antúnez 2013[49]	Inflammatory	Bath-mud pack	SOC	Single	1.6	72.0	71.1 (7.3)	29.1 (4.5)
Fioravanti 2007[54]	Fibromyalgia	Bath-mud pack	SOC	Multi	16	98.0	47.4 (10.0)	NA
Fioravanti 2010[52]	Mechanics	Bath-mud pack	SOC	Single	36	75.0	70.2 (5.0)	26.4 (4.1)
Fioravanti 2012[51]	Mechanics	Bath	SOC	Single	12	50.0	70.9 (7.4)	27.0 (3.5)
Fioravanti 2014[55]	Mechanics	Bath-mud pack	SOC	Single	48	86.7	70.8 (9.1)	25.4 (3.0)
Fioravanti 2015[50]	Mechanics	Bath-mud pack	SOC	Single	48	72.0	69.1 (10.1)	28.3 (4.1)
Fioravanti 2018[53]	Fibromyalgia	Bath	Placebo-like	Single	24	95.0	56.0 (7.7)	25.5 (4.7)
Forestier 2010[56]	Mechanics	Bath-mud pack	SOC	Multi	24	47.5	63.7 (9.8)	29.9 (5.3)
Franke 2000[57]	Inflammatory	Bath	Placebo-like	Single	24	76.7	58.4 (10.9)	25.7 (3.7)
Franke 2007[58]	Inflammatory	Bath	Placebo-like	Single	48	66.0	56.2 (11.4)	27.1 (4.5)
Fritsch 2022[59]	Inflammatory	Bath	SOC	Multi	10	93.0	54.3 (11.0)	NA
Gaisberger 2021[60]	Mechanics	Other	Other	Single	24	53.6	67.4 (4.4)	27.2 (4.6)
Gyarmati 2017[61]	Mechanics	Mud pack	Other	Single	16	95.6	64.5	NA
Hanzel 2018[62]	Mechanics	Bath	Placebo-like	Single	12	66.0	66.7 (4.8)	26.9 (3.2)
Horváth 2012[63]	Mechanics	Bath	Other	Single	13	81.0	63.2 (4.6)	28.4 (4.3)
Ionescu 2017[64]	Mechanics	Bath-mud pack	Other	Mono	1.5	63.1	55.7 (10.2)	28.3 (4.5)
Konrad 1994[65]	Mechanics	Bath	SOC	Multi	52	55.9	40.6 (8.7)	NA
Kovács 2002[68]	Mechanics	Bath	Placebo-like	Single	12	70.7	NA	NA
Kovács 2012[67]	Mechanics	Bath	Placebo-like	Single	24	91.5	59.5 (26.5)	29.6 (4.1)
Kovács 2016[66]	Mechanics	Bath	Other	Single	12	NA	59.9 (7.6)	NA
Kulisch 2014[69]	Mechanics	Bath	Placebo-like	Single	15	77.9	65.6 (7.1)	NA
Maindet 2021[16]	Fibromyalgia	Other	WL	Multi	24	90.8	49.8 (8.8)	27.2 (6.2)
NCT03289078 2017[70]	Mechanics	Other	SOC	Single	NA	NA	NA	NA
NCT05352477 2022[14]	Mechanics	Bath	Placebo-like	Single	12	NA	NA	NA
NCT05819437 2023[15]	Mechanics	Bath	SOC	Single	24	NA	NA	NA

Continued

Table 1 Continued

Study identifier	Main indication	Intervention type	Control type	Single or multicentre	Follow-up	Women	Age	BMI
Nguyen 1997[71]	Mechanics	Other	SOC	Single	24	81.4	63.5 (6.5)	NA
Peluso 2016[72]	Mechanics	Bath-mud pack	Other	Single	24	55.6	65.4 (7.5)	26.9 (8.5)
Pérez-Fernández 2019[73]	Fibromyalgia	Bath	SOC	Single	12	96.0	52.9 (9.9)	28.1 (4.6)
Santos 2016[78]	Inflammatory	Bath	WL	Single	12	86.4	58.4 (10)	NA
Szucs 1989[74]	Mechanics	Bath	Placebo-like	Single	3	NA	NA	NA
Tefner 2013[75]	Mechanics	Mud pack	Placebo-like	Single	12	84.9	63.5 (9.3)	NA
van Tubergen 2001[76]	Inflammatory	Other	Other	Multi	40	27.5	48.3 (9.7)	NA
Varzaityte 2020[77]	Mechanics	Bath	Other	Single	4	87.0	64.6 (11.4)	29.4 (4.3)

Follow-up in weeks; Women in %; age: in years, mean (SD); BMI in $kg.m^{-2}$, mean (SD); mechanics: mechanical disorders.
BMI, body mass index; NA, not available; SOC, standard of care; WL, waiting list.

I^2=0%) when reported as mean change from baseline and −0.23 (95% CI (−0.55; 0.09); I^2=0%) when reported as postintervention mean value (online supplemental material S6).

Pain after the intervention was reported in 26 trials (2567 patients). The SMD was −0.14 (95% CI (−0.77; 0.49); I^2=91%) when reported as mean change from baseline, and −0.62 (95% CI (−0.84; −0.40); I^2=69%) when reported as postintervention mean value (online supplemental material S7).

QoL at 6 months was reported in 11 trials (1659 patients). The SMD was 0.38 (95% CI (0.25; 0.51); I^2=0%) when as mean change from baseline, and 0.46 (95% CI (0.24; 0.68); I^2=48%) when reported as postintervention mean value (online supplemental material S8).

QoL at 12 months was reported in three trials (314 patients). The SMD was 0.07 (95% CI (−0.18; 0.33); I^2=0%) when reported as mean change from baseline, and 0.47 (95% CI (−0.05; 0.98); only one trial) when reported as postintervention mean value (online supplemental material S9).

QoL after the intervention was reported in 21 trials (1631 patients). The SMD was 0.34 (95% CI (0.09; 0.59); I^2=54%) when reported as mean change from baseline, and 0.34 (95% CI (−0.06; 0.73); I^2=86%) when reported as postintervention mean value (online supplemental material S10).

Safety outcomes

Among 13 trials (2062 patients, 123 events) that reported the risk of withdrawal, 10 were at high ROB and 3 with some concerns. The risk ratio (RR) of withdrawal was inconclusive (0.75, 95% CI (0.46; 1.20); I^2=12%, (online supplemental material S11).

The risk of SAE was reported in two trials (406 patients, 29 events). The RR was inconclusive (1.01, 95% CI (0.36; 2.85); I^2=21% (online supplemental material S12).

From 5 trials (1123 patients, 87 events) reporting the risk of AE, 2 were at high ROB and 3 with some concerns. The RR was inconclusive (0.80, 95% CI (0.43; 1.50); I^2=44% (online supplemental material S13).

Sensitivity analyses of the primary outcomes

Restricted to placebo-like design

Regarding pain at 3 months (9 trials, 1264 patients), the SMD was −0.21 (95% CI (−0.41; −0.01); I^2=45%) when reported as mean change from baseline, and −0.49 (95% CI (−0.86; −0.11); I^2=61%) when reported as postintervention mean value (online supplemental material S14).

Regarding QoL at 3 months (7 trials, 525 patients), the SMD was 0.28 (95% CI (−0.18; 0.73); I^2=62%) when reported as mean change from baseline, and 0.47 (95% CI (0.24; 0.69); I^2=0%) when reported as postintervention mean value (online supplemental material S15).

Restricted to SOC design

Regarding the pain at 3 months (6 trials, 360 patients), the SMD was −1.50 (95% CI (−2.07; −0.92); I^2=82%) when

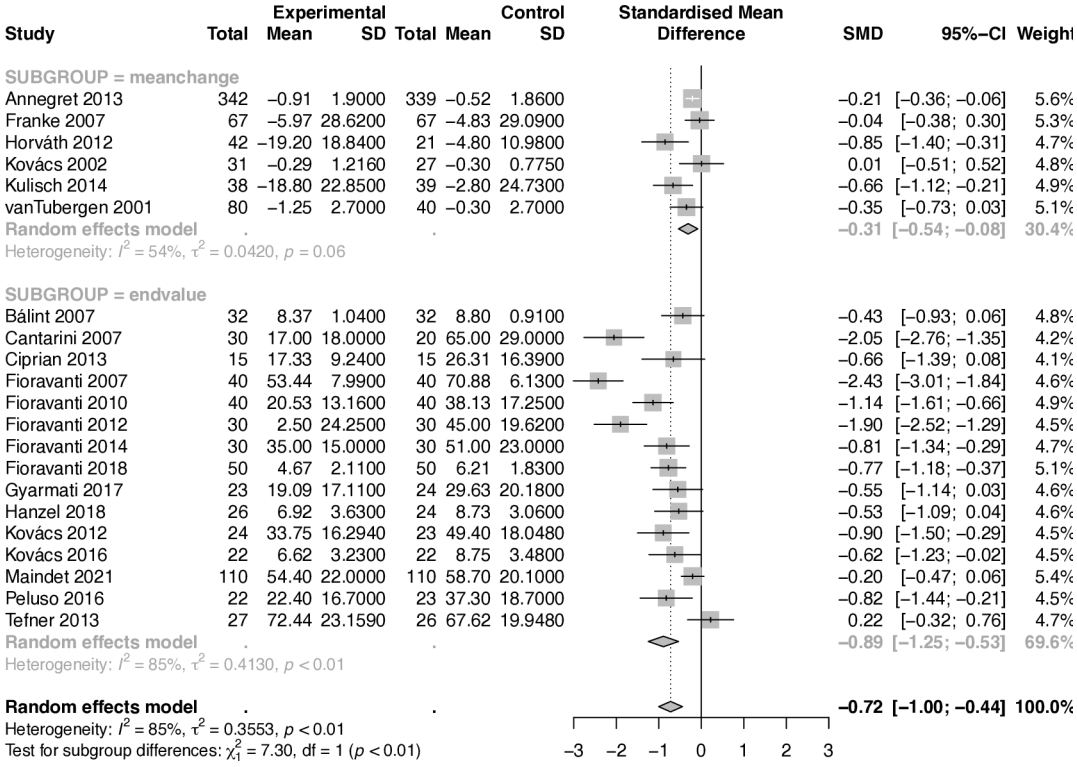

Figure 2 Forest plot of the effect of balneotherapy on pain, at 3 months. The data and the synthesis are provided for (1) the subset of trials that reported value as mean change ('SUBGROUP=meanchange', synthesis in bold grey), (2) the subset of trials that reported the end value ('SUBGROUP=endvalue', synthesis in bold grey), (3) overall (synthesis in bold black). SMD, standardised mean difference.

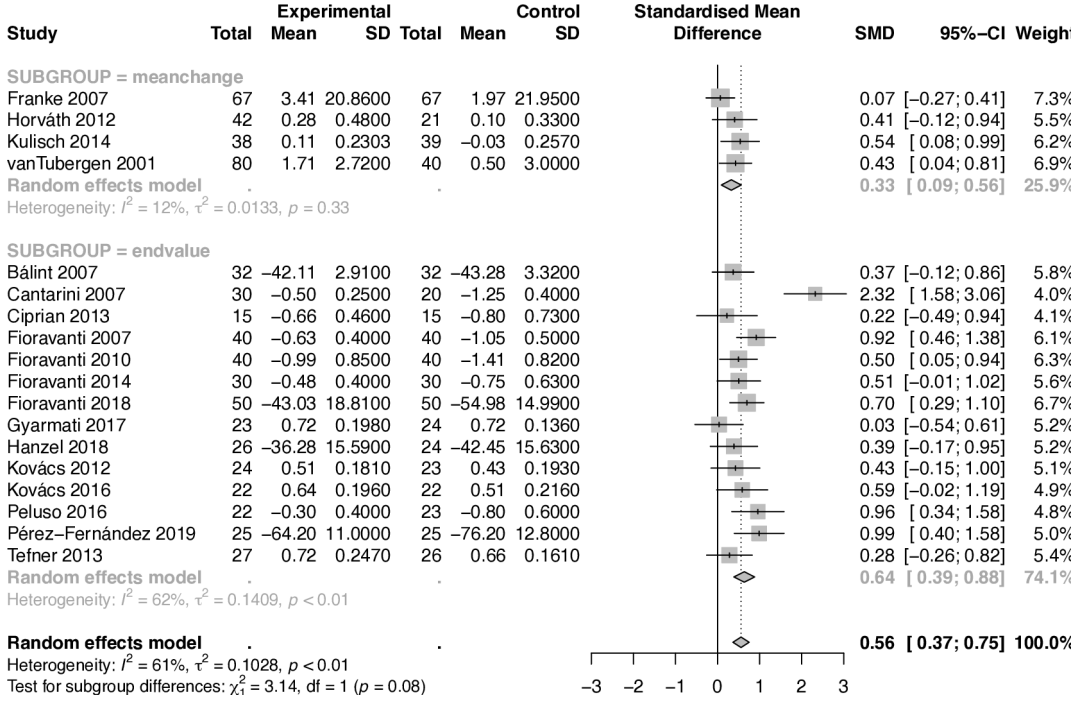

Figure 3 Forest plot of the effect of balneotherapy on quality of life, at 3 months. The data and the synthesis are provided for (1) the subset of trials that reported value as mean change ('SUBGROUP=meanchange', synthesis in bold grey), (2) the subset of trials that reported the end value ('SUBGROUP=endvalue', synthesis in bold grey), (3) overall (synthesis in bold black). SMD, standardised mean difference.

reported as postintervention mean value; no data were available for mean change from baseline (online supplemental material S16).

Regarding QoL at 3 months (6 trials, 350 patients), the SMD was 0.89 (95% CI (0.34; 1.44); I^2=78%) when reported as postintervention mean value; no data were usable for mean change from baseline (online supplemental material S17).

Subgroups analyses of the primary outcomes

Subgroups for pain at 3 months

Subgroup of interest was reported for 21 trials (2163 patients).

The intervention types bath, mud pack, bath plus mud pack and other were associated with an SMD of −0.59 (95% CI (−0.89; −0.30)), −0.16 (95% CI (−0.92; 0.60)), −1.32 (95% CI (−1.90; −0.73)) and −0.25 (95% CI (−0.47; −0.03)), respectively. The p value of the test for this subgroup effect was <0.01 (online supplemental material S18).

The intervention in mechanical disorder, inflammatory and fibromyalgia indications was associated with an SMD of −0.73 (95% CI (−1.02; −0.43)), −0.25 (95% CI (−0.55; 0.05)) and −1.11 (95% CI (−2.41; 0.18)), respectively. The p value of the test for this subgroup effect was 0.06 (online supplemental material S19).

Subgroups for QoL at 3 months

Subgroup of interest was reported for 18 trials (1194 patients).

The intervention types bath, mud pack, bath plus mud pack and other were associated with an SMD of 0.46 (95% CI (0.27; 0.66)), 0.17 (95% CI (−0.23; 0.56)), 0.88 (95% CI (0.33; 1.43)) and 0.43 (95% CI (0.04; 0.81)), respectively. The p value of the test for this subgroup effect was 0.22 (online supplemental material S20).

The intervention in mechanical disorder, inflammatory and fibromyalgia indications was associated with an SMD of 0.57 (95% CI (0.30; 0.85)), 0.23 (95% CI (−0.05; 0.50)) and 0.83 (95% CI (0.56; 1.10)), respectively. The p value of the test for this subgroup effect was <0.01 (online supplemental material S21).

Reporting bias

The visual inspection of the funnel plot and the p values of Egger and Begg's tests (0.004 and 0.009, respectively) were strongly in favour of a publication bias for the pain assessment at 3 months (online supplemental material S22).

The visual inspection of the funnel plot and the p values of Egger and Begg's tests (0.065 and 0.306, respectively) were inconclusive regarding the risk of publication bias for the QoL assessment at 3 months (online supplemental material S23), as for the risk of withdrawal (online supplemental material S24) and of AE (online supplemental material S25).

Certainty of evidence

Regarding the pain at 3 months, the inconsistency, indirectness and imprecision were assessed as 'very serious', 'not serious' and 'not serious', respectively. The overall certainty in the estimate was 'very low'. Regarding the QoL at 3 months, the inconsistency, indirectness and imprecision were assessed as 'serious', 'not serious' and 'not serious', respectively. The overall certainty in the estimate was 'very low'. For clarity, we reported the combined estimates of mean change and postintervention mean values for pain and QoL. Regarding the risk of withdrawal, the inconsistency, indirectness and imprecision were assessed as 'not serious', 'not serious' and 'serious', respectively. The overall certainty in the estimate was 'very low'. Regarding the risk of AE, the inconsistency, indirectness and imprecision were assessed as 'serious', 'not serious' and 'very serious', respectively. The overall certainty in the estimate was 'very low' (table 2).

DISCUSSION

The results of the present study meet the aim of the study. They indicate that most of the available trials assessing the effect of balneotherapy in rheumatology are at high ROB. Overall, the suggested decrease in pain and the suggested increase in QoL appeared to be of very low level of certainty, that is, the review does not support a benefit of balneotherapy in rheumatology. In addition, the assessment of the safety of balneotherapy was inconclusive, and therefore, there is no reliable evidence of a favourable risk-benefit ratio of balneotherapy in rheumatology.

General interpretation of the results in the context of other evidence

Previous reviews suggested a lack of evidence but were limited to specific indications and included fewer than 10 trials.[1 5 8] Our study confirms the lack of evidence for a broader landscape in rheumatology, based on a much bigger sample size (42 included trials). The previous large review, which included 26 studies, did not report pooled estimates of the treatment effect and did not exclusively include randomised trials[6] as was the case herein. It is also of note that, although Fernandez-Gonzalez *et al*[7] concluded to positive effects of balneotherapy, they included only 7 studies and did not report an assessment of the certainty of the evidence, whereas the present review included 42 studies and provided a certainty assessment following GRADE. Finally, one more strength of our review is the pooled estimates and the certainty assessments for safety outcomes also, not provided in the previous reviews.[1 5–8] The present review also found the absence of a study at low ROB, which might be related to the specific challenges when assessing such complex interventions.[7 9] The findings also provide evidence for potential publication bias in the field of balneotherapy that could not have been assessed in the two previous Cochrane reviews because of the low number of trials included.[1 5] The results of the present study also underscore the difficulty in accurately estimating the treatment effect of complex interventions such as balneotherapy, from the supposed specific effect of thermal water/mud

Table 2 Summary of finding including GRADE assessment

Certainty assessment							No of patients		Effect			
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Balneotherapy	Any control	Relative (95% CI)	Absolute (95% CI)	Certainty	Importance
Pain intensity (follow-up: mean 3 months; assessed with: SMD)												
21	Randomised trials	Very serious*	Very serious†‡§	Not serious	Not serious¶	Publication bias strongly suspected**	1121	1042	–	SMD **0.72 SD lower** (1 lower to 0.44 lower)	⨁◯◯◯ Very low	IMPORTANT
Quality of life (follow-up: mean 3 months; assessed with: SMD)												
18	Randomised trials	Very serious*	Serious† §††	Not serious	Not serious¶	None‡‡	633	561	–	SMD **0.56 SD higher** (0.37 higher to 0.75 higher)	⨁◯◯◯ Very low	IMPORTANT
Withdrawal (follow-up: mean 3 months; assessed with: RR)												
13	Randomised trials	Very serious*	Not serious† §§	Not serious	Serious¶¶	None‡‡	53/1069 (5.0%)	70/993 (7.0%)	**RR 0.75** (0.46 to 1.20)	**18 fewer per 1 000** (from 38 fewer to 14 more)	⨁◯◯◯ Very low	IMPORTANT
Adverse event (follow-up: mean 3 months; assessed with: RR)												
5	Randomised trials	Serious***	Serious† ††	Not serious	Very serious¶¶†††	None‡‡	38/564 (6.7%)	49/559 (8.8%)	**RR 0.80** (0.43 to 1.50)	**18 fewer per 1 000** (from 50 fewer to 44 more)	⨁◯◯◯ Very low	IMPORTANT

Summary of certainty assessment of balneotherapy versus any control, following the GRADE approach (number of studies, study design, risk of bias of the studies, inconsistency of the results, indirectness of evidence, imprecision of the estimate and overall certainty).
*Most of the trials: high risk of bias.
†Heterogeneity in populations (indication for the intervention), intervention (type of intervention), control arm and in measure of the outcome.
‡I^2 statistic suggested substantial to considerable heterogeneity.
§At least one subgroup effect was significant (nominal $p<0.05$).
¶>400 patients analysed, CI not wide, excluding the null effect.
**Funnel plot in favour, Egger and Begg's test significant (nominal $p<0.05$).
††I^2 statistic suggested low to substantial or moderate heterogeneity.
‡‡Funnel plot, Egger and Begg's tests: inconclusives.
§§I^2 statistic suggested low heterogeneity.
¶¶Number of events is probably not sufficient, the 95% CI overlaps no effect and fails to exclude important benefit or important harm.
***Most of the trials: some concerns.
†††Number of participants is probably not sufficient.
GRADE, Grading of Recommendations Assessment, Development and Evaluation; No, number; RR, risk ratio; SD, standard deviation; SMD, standardised mean difference.

to the non-specific effect of the adjuvant care such as resting and massage. This is supported by the sensitivity analyses that found that the effect estimate was smaller when balneotherapy was compared with placebo-like rather than SOC as the control. This smaller effect with a stronger comparator highlights the potential impact of the adjuvant care associated with balneotherapy and of a potential placebo effect.

Limitations of the evidence included in the review

The included trials were mostly at high ROB. Moreover, there is substantial heterogeneity in outcome measures concerning scale, time points and their reporting, limiting the amount of available data for synthesis (eg, data not available)[1 5 14 70] or reported in a way that precluded their use in the synthesis.[42 78] Finally, most of the comparators were non-active intervention (SOC or placebo-like comparator), limiting the comparability of balneotherapy to other specific interventions.

Limitations of the review processes

For pooling effect estimates, we used the SMD at the end of follow-up or mean change from baseline and end-intervention between groups. The translation of these changes in SMDs to clinical practice seems difficult. Moreover, we combined different indications (mechanical disorders, inflammatory, fibromyalgia) of balneotherapy in rheumatology. This approach aligns with the prevailing categorisations in the current funding of balneotherapy by the national health insurance in France, which is based on broad medical orientations such as 'rheumatology', 'phlebology', 'respiratory tract', etc. It is noteworthy that this amalgamation introduces heterogeneity, particularly since the underlying indication appears to impact the treatment effect, as evidenced in the subgroup analyses. The review was also limited by the strong publication bias and limited to direct comparisons. Finally, exploring the cost-effectiveness of balneotherapy was beyond the scope of this paper.

Implications of the results for clinical practice, policy and future research

Additional randomised trials with a low ROB are deemed essential to furnish dependable estimates of the impact of balneotherapy in rheumatology. The use of tap water as a control demonstrated the feasibility, although complexity, of conducting randomised, double-blinded, placebo-controlled trials. Furthermore, spin, misleading reporting and inadequate interpretation of the results were common. The registration of the protocols and the statistical analysis plan, guidelines for reporting[80] including enhanced reporting of safety outcomes will improve the quality of evidence in this field.

CONCLUSION

The present report updates the overall assessment of the potential benefit of European balneotherapy in rheumatology. The overall level of evidence regarding its potential benefit and risk appeared to be very low certainty.

Acknowledgements We thank Philip Robinson (DRS Hospices Civils de Lyon) for help in manuscript preparation.

Contributors CG designed the search strategy. IA and GG conducted the bibliographic selection and data extraction. IA, BK and GG assessed the risk of bias of included studies. MN and GG conducted the statistical analyses. IA and GG assessed the certainty of evidence. GG wrote the first draft of the manuscript. CG, EM, J-CL and BK revised the manuscript. Guarantor of the review: GG.

Funding This review was funded by the French association for thermal research (AFRETh, http://www.afreth.org/).

Competing interests GG and IA received a grant from the French association for thermal research—'Association Française pour la Recherche Thermale' (AFRETh, http://www.afreth.org/). BK, MN, CG, EM and J-CL had no competing interest.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information. The data extracted from included studies and data used for all analyses are available in the current report.

ORCID iDs
Caroline Giroudon http://orcid.org/0000-0001-7947-9437
Emmanuel Massy http://orcid.org/0000-0002-4842-8618
Jean-Christophe Lega http://orcid.org/0000-0002-9398-2968
Guillaume Grenet http://orcid.org/0000-0002-5237-2581

REFERENCES

1 Verhagen AP, Bierma-Zeinstra SMA, Boers M, *et al*. Balneotherapy (or spa therapy) for rheumatoid arthritis. *Cochrane Database Syst Rev* 2015;2015:CD000518.
2 Sukenik S, Flusser D, Abu-Shakra M. The role of spa therapy in various rheumatic diseases. *Rheum Dis Clin North Am* 1999;25:883–97.
3 Fioravanti A, Karagülle M, Bender T, *et al*. Balneotherapy in osteoarthritis: Facts, fiction and gaps in knowledge. *Eur J Integr Med* 2017;9:148–50.
4 Çetinkaya FN, Koçak FA, Kurt EE, *et al*. The Effects of Balneotherapy on Oxidant/Antioxidant Status in Patients With Fibromyalgia: An Observational Study. *Arch Rheumatol* 2020;35:506–14.
5 Verhagen AP, Bierma-Zeinstra SM, Boers M, *et al*. Balneotherapy for osteoarthritis. *Cochrane Database Syst Rev* 2007;CD006864.
6 Forestier R, Fioravanti A, Bender T, *et al*. Crenobalneotherapy for low back pain: systematic review of clinical trials. *Int J Biometeorol* 2022;66:13–23.
7 Fernandez-Gonzalez M, Fernandez-Lao C, Martin-Martin L, *et al*. Therapeutic Benefits of Balneotherapy on Quality of Life of Patients with Rheumatoid Arthritis: A Systematic Review. *Int J Environ Res Public Health* 2021;18:13216.

8 Forestier R, Bugnard A, Thomas T. Balneotherapy in spondyloarthropathy: A systematic review. *Therapie* 2022;77:723–30.
9 PLFSS 2023: les députés s'attaquent au remboursement des cures thermales par l'Assurance maladie. egora.fr; 2023. Available: https://www.egora.fr/actus-pro/assurance-maladie-mutuelles/76912-plfss-2023-les-deputes-s-attaquent-au-remboursement-des
10 Torres-Pruñonosa J, Raya JM, Crespo-Sogas P, *et al*. The economic and social value of spa tourism: The case of balneotherapy in Maresme, Spain. *PLoS ONE* 2022;17:e0262428.
11 Fioravanti A, Valenti M, Altobelli E, *et al*. Clinical efficacy and cost-effectiveness evidence of spa therapy in osteoarthritis. The results of 'Naiade' Italian Project. *Panminerva Med* 2003;45:211–7.
12 Van Tubergen A, Boonen A, Landewé R, *et al*. Cost effectiveness of combined spa-exercise therapy in ankylosing spondylitis: a randomized controlled trial. *Arthritis Rheum* 2002;47:459–67.
13 Varga V, Boncz I, Sebestyén A, *et al*. Utilization indicators of balneotherapy in Hungary. *Orv Hetil* 2019;160:22–8.
14 NCT05352477. Balneotherapy With Acquabios Thermal Water in Knee Osteoarthritis, 2022. Available: https://www.cochranelibrary.com/central/doi/10.1002/central/CN-02398609/full
15 NCT05819437. Knee osteoarthritis and balneotherapy. 2023. Available: https://www.cochranelibrary.com/central/doi/10.1002/central/CN-02553909/full
16 Maindet C, Maire A, Vermorel C, *et al*. Spa Therapy for the Treatment of Fibromyalgia: An Open. *Randomized Multicenter Trial J Pain* 2021;22:940–51.
17 Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.
18 Higgins JPT, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester: Wiley-Blackwell, 2008:649.
19 Gutenbrunner C, Bender T, Cantista P, *et al*. A proposal for a worldwide definition of health resort medicine, balneology, medical hydrology and climatology. *Int J Biometeorol* 2010;54:495–507.
20 European spas association. 2023. Available: https://europeanspas.eu
21 Nussbaumer-Streit B, Klerings I, Dobrescu AI, *et al*. Excluding non-English publications from evidence-syntheses did not change conclusions: a meta-epidemiological study. *J Clin Epidemiol* 2020;118:42–54.
22 Covidence - Better systematic review management, 25 July 2023. Available: https://www.covidence.org/
23 Livre blanc - gt méthodologie. 2023. Available: https://sfpt-fr.org/livreblancmethodo/index.htm
24 Veehof MM, ten Klooster PM, Taal E, *et al*. Comparison of internal and external responsiveness of the generic Medical Outcome Study Short Form-36 (SF-36) with disease-specific measures in rheumatoid arthritis. *J Rheumatol* 2008;35:610–7.
25 Sterne JAC, Savović J, Page MJ, *et al*. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898.
26 Chapter 6: choosing effect measures and computing estimates of effect. 2024. Available: https://training.cochrane.org/handbook/current/chapter-06
27 Chapter 10: analysing data and undertaking meta-analyses. n.d. Available: https://training.cochrane.org/handbook/current/chapter-10
28 da Costa BR, Nüesch E, Rutjes AW, *et al*. Combining follow-up and change data is valid in meta-analyses of continuous outcomes: a meta-epidemiological study. *J Clin Epidemiol* 2013;66:847–55.
29 IntHout J, Ioannidis JPA, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Med Res Methodol* 2014;14:25.
30 Jackson D, Law M, Rücker G, *et al*. The Hartung-Knapp modification for random-effects meta-analysis: A useful refinement but are there any residual concerns? *Stat Med* 2017;36:3923–34.
31 Ihaka R, Gentleman R. R: A Language for Data Analysis and Graphics. *J Comput Graph Stat* 1996;5:299–314.
32 Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health* 2019;22:153–60.
33 Sterne JAC, Sutton AJ, Ioannidis JPA, *et al*. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011;343.
34 Egger M, Davey Smith G, Schneider M, *et al*. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629–34.
35 Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994;50:1088–101.
36 Guyatt GH, Oxman AD, Vist GE, *et al*. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924–6.
37 Grujic Z, Tomic T, Jevtic N, *et al*. The effect of cooled-down peloid and water in patients with active rheumatoid arthritis. *Osteoporos Int* 2017;28.
38 Marcu F, Lazar L, Mușiu G. The impact of balneo-physical-kinetic therapy in bearing joints osteoarthritis. *Arch Balk Med Union* 2009;44:315–20.
39 NCT01544647. Spa therapy in knee osteoarthritis (oa): nancy-thermal. 2012. Available: https://www.cochranelibrary.com/central/doi/10.1002/central/CN-01535887/full
40 Annegret F, Thomas F. Long-term benefits of radon spa therapy in rheumatic diseases: results of the randomised, multi-centre IMuRa trial. *Rheumatol Int* 2013;33:2839–50.
41 Bálint GP, Buchanan WW, Adám A, *et al*. The effect of the thermal mineral water of Nagybaracska on patients with knee joint osteoarthritis--a double blind study. *Clin Rheumatol* 2007;26:890–4.
42 Benini C, Rubino G, Paolazzi G, *et al*. Efficacy of mud plus bath therapy as compared to bath therapy in osteoarthritis of hands and knees: a pilot single-blinded randomized controlled trial. *Reumatismo* 2021;73.
43 Cantarini L, Leo G, Giannitti C, *et al*. Therapeutic effect of spa therapy and short wave therapy in knee osteoarthritis: a randomized, single blind, controlled trial. *Rheumatol Int* 2007;27:523–9.
44 Cantista P, Maraver F. Balneotherapy for knee osteoarthritis in S. Jorge: a randomized controlled trial. *Int J Biometeorol* 2020;64:1027–38.
45 Caporali R, Bellometti S, Rossi S, *et al*. Safety of Mud-Bath Applications in Moderately Active Rheumatoid Arthritis. *J Musculoskelet Pain* 2010;18:81–7.
46 Chary-Valckenhaere I, Loeuille D, Kohler F, *et al*. Spa therapy in the treatment of chronic shoulder pain due to rotater cuff tendinopathy: Rotatherm, a large randomized multicentre trial. *Ann Phys Rehabil Med* 2012;55:e352.
47 Ciprian L, Lo Nigro A, Rizzo M, *et al*. The effects of combined spa therapy and rehabilitation on patients with ankylosing spondylitis being treated with TNF inhibitors. *Rheumatol Int* 2013;33:241–5.
48 Cozzi F, Raffeiner B, Beltrame V, *et al*. Effects of mud-bath therapy in psoriatic arthritis patients treated with TNF inhibitors. Clinical evaluation and assessment of synovial inflammation by contrast-enhanced ultrasound (CEUS). *Joint Bone Spine* 2015;82:104–8.
49 Espejo Antúnez L, Caro Puértolas B, Ibáñez Burgos B, *et al*. Effects of Mud Therapy on Perceived Pain and Quality of Life Related to Health in Patients With Knee Osteoarthritis. *Reumatología Clínica (English Edition)* 2013;9:156–60.
50 Fioravanti A, Bacaro G, Giannitti C, *et al*. One-year follow-up of mud-bath therapy in patients with bilateral knee osteoarthritis: a randomized, single-blind controlled trial. *Int J Biometeorol* 2015;59:1333–43.
51 Fioravanti A, Giannitti C, Bellisai B, *et al*. Efficacy of balneotherapy on pain, function and quality of life in patients with osteoarthritis of the knee. *Int J Biometeorol* 2012;56:583–90.
52 Fioravanti A, Iacoponi F, Bellisai B, *et al*. Short- and long-term effects of spa therapy in knee osteoarthritis. *Am J Phys Med Rehabil* 2010;89:125–32.
53 Fioravanti A, Manica P, Bortolotti R, *et al*. Is balneotherapy effective for fibromyalgia? Results from a 6-month double-blind randomized clinical trial. *Clin Rheumatol* 2018;37:2203–12.
54 Fioravanti A, Perpignano G, Tirri G, *et al*. Effects of mud-bath treatment on fibromyalgia patients: a randomized clinical trial. *Rheumatol Int* 2007;27:1157–61.
55 Fioravanti A, Tenti S, Giannitti C, *et al*. Short- and long-term effects of mud-bath treatment on hand osteoarthritis: a randomized clinical trial. *Int J Biometeorol* 2014;58:79–86.
56 Forestier R, Desfour H, Tessier J-M, *et al*. Spa therapy in the treatment of knee osteoarthritis: a large randomised multicentre trial. *Ann Rheum Dis* 2010;69:660–5.
57 Franke A, Reiner L, Pratzel HG, *et al*. Long-term efficacy of radon spa therapy in rheumatoid arthritis—a randomized, sham-controlled study and follow-up. *Rheumatology (Sunnyvale)* 2000;39:894–902.
58 Franke A, Reiner L, Resch KL. Long-term benefit of radon spa therapy in the rehabilitation of rheumatoid arthritis: a randomised, double-blinded trial. *Rheumatol Int* 2007;27:703–13.
59 Fritsch K, Nagy G, Szekanecz Z, *et al*. Balneotherapy, a Complementary Non-pharmacological Approach for Non-Inflammatory Complaints in Systemic Lupus Erythematosus: A Pilot Study. *In Vivo* 2022;36:3010–7.
60 Gaisberger M, Fuchs J, Riedl M, *et al*. Endogenous anandamide and self-reported pain are significantly reduced after a 2-week multimodal treatment with and without radon therapy in patients with knee osteoarthritis: a pilot study. *Int J Biometeorol* 2021;65:1151–60.
61 Gyarmati N, Kulisch Á, Németh A, *et al*. Evaluation of the Effect of Hévíz Mud in Patients with Hand Osteoarthritis: A Randomized, Controlled, Single-Blind Follow-Up Study. *Isr Med Assoc J* 2017;19:177–82.

62 Hanzel A, Horvát K, Molics B, *et al*. Clinical improvement of patients with osteoarthritis using thermal mineral water at Szigetvár Spa-results of a randomised double-blind controlled study. *Int J Biometeorol* 2018;62:253–9.
63 Horváth K, Kulisch Á, Németh A, *et al*. Evaluation of the effect of balneotherapy in patients with osteoarthritis of the hands: a randomized controlled single-blind follow-up study. *Clin Rehabil* 2012;26:431–41.
64 Ionescu EV, Tica I, Oprea C, *et al*. Adiponectin correlation with bioclinical benefits of using natural therapeutic factors in knee osteoarthritis. *Acta Endocrinol (Buchar)* 2017;13:308–13.
65 Konrad K. The Way Forward for Balneotherapy. *Rheumatology (Sunnyvale)* 1994;33:301.
66 Kovács C, Bozsik Á, Pecze M, *et al*. Effects of sulfur bath on hip osteoarthritis: a randomized, controlled, single-blind, follow-up trial: a pilot study. *Int J Biometeorol* 2016;60:1675–80.
67 Kovács C, Pecze M, Tihanyi Á, *et al*. The effect of sulphurous water in patients with osteoarthritis of hand. Double-blind, randomized, controlled follow-up study. *Clin Rheumatol* 2012;31:1437–42.
68 Kovács I, Bender T. The therapeutic effects of Cserkeszölö thermal water in osteoarthritis of the knee: a double blind, controlled, follow-up study. *Rheumatol Int* 2002;21:218–21.
69 Kulisch Á, Benkö Á, Bergmann A, *et al*. Evaluation of the effect of Lake Hévíz thermal mineral water in patients with osteoarthritis of the knee: a randomized, controlled, single-blind, follow-up study. *Eur J Phys Rehabil Med* 2014;50:373–81.
70 NCT03289078. Evaluation of functional improvement after spa therapy in fontcaude center. 2017. Available: https://www.cochranelibrary.com/central/doi/10.1002/central/CN-01564249/full
71 Nguyen M, Revel M, Dougados M. Prolonged effects of 3 week therapy in a spa resort on lumbar spine, knee and hip osteoarthritis: follow-up after 6 months. A randomized controlled trial. *Br J Rheumatol* 1997;36:77–81.
72 Peluso R, Caso F, Costa L, *et al*. Mud-bath therapy and oral glucosamine sulfate in patients with knee osteoarthritis: a randomized, controlled, crossover study. *Clin Exp Rheumatol* 2016;34:618–24.
73 Pérez-Fernández MR, Calvo-Ayuso N, Martínez-Reglero C, *et al*. Efficacy of baths with mineral-medicinal water in patients with fibromyalgia: a randomized clinical trial. *Int J Biometeorol* 2019;63:1161–70.
74 Szucs L, Ratko I, Lesko T, *et al*. Double-blind trial on the effectiveness of the Puspokladany thermal water on arthrosis of the knee-joints. *J R Soc Health* 1989;109:7–9.
75 Tefner IK, Gaál R, Koroknai A, *et al*. The effect of Neydharting mud-pack therapy on knee osteoarthritis: a randomized, controlled, double-blind follow-up pilot study. *Rheumatol Int* 2013;33:2569–76.
76 van Tubergen A, Landewé R, van der Heijde D, *et al*. Combined spa-exercise therapy is effective in patients with ankylosing spondylitis: a randomized controlled trial. *Arthritis Rheum* 2001;45:430–8.
77 Varzaityte L, Kubilius R, Rapoliene L, *et al*. The effect of balneotherapy and peloid therapy on changes in the functional state of patients with knee joint osteoarthritis: a randomized, controlled, single-blind pilot study. *Int J Biometeorol* 2020;64:955–64.
78 Santos I, Cantista P, Vasconcelos C, *et al*. Balneotherapy and Rheumatoid Arthritis: A Randomized Control Trial. *Isr Med Assoc J* 2016;18:474–8.
79 Boutron I, Tubach F, Giraudeau B, *et al*. Methodological differences in clinical trials evaluating nonpharmacological and pharmacological treatments of hip and knee osteoarthritis. *JAMA* 2003;290:1062–70.
80 EQUATOR network. Enhancing the quality and transparency of health research. 2024 Available: https://www.equator-network.org/