



HAL
open science

A multiverse analysis of meta-analyses assessing acupuncture efficacy for smoking cessation evidenced vibration of effects

Manele El Bahri, Xu Wang, Tiphaine Biaggi, Bruno Falissard, Florian Naudet, Caroline Barry

► To cite this version:

Manele El Bahri, Xu Wang, Tiphaine Biaggi, Bruno Falissard, Florian Naudet, et al.. A multiverse analysis of meta-analyses assessing acupuncture efficacy for smoking cessation evidenced vibration of effects. *Journal of Clinical Epidemiology*, 2022, 152, pp.140-150. 10.1016/j.jclinepi.2022.09.001 . hal-03799864

HAL Id: hal-03799864

<https://hal.science/hal-03799864>

Submitted on 22 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A multiverse analysis of meta-analyses assessing acupuncture efficacy for smoking cessation evidenced vibration of effects

Manele El Bahri¹, Xu Wang², Tiphaine Biaggi¹, Bruno Falissard², Florian Naudet^{x34}, Caroline Barry^{x1}

^x These authors contributed equally to this work.

¹ CESP, INSERM U1018, Univ. Paris-Sud, UVSQ, Paris, France

² CESP, INSERM U1018, Univ. Paris-Sud, UVSQ, Villejuif, France

³ Univ Rennes, CHU Rennes, Inserm, CIC 1414 [(Centre d'Investigation Clinique de Rennes)], F- 35000 Rennes, France

⁴ Univ Rennes, CHU Rennes, Inserm, Irset (Institut de recherche en santé, environnement et travail) - UMR_S 1085, F-35000 Rennes, France

Abstract

Objective: To explore the impact of methodological choices on the results of meta-analyses (MAs), with acupuncture for smoking cessation as a case study.

Study Design: After performing an umbrella review (using MEDLINE, the COCHRANE Library, the Wan Fang database, and the Chinese Journal Full-text Database/March 2018) of MAs exploring the use of acupuncture for smoking cessation, we extracted all randomized controlled trials (RCTs). Numerous MAs were performed according to every possible combination of various methodological choices (e.g., characteristics of the intervention and control procedures, outcome, publication status, language) to assess their vibration of effects (VoE); or more precisely the existence of a Janus effect, i.e., whether the 10th and 90th percentiles in the distribution of effect sizes were in opposite directions.

Results: After including 7 MAs and 39 RCTs, we performed 496,528 MAs. The effect size was negative at the 10st percentile (– 0.1, favoring controls) and positive at the 90th percentile (1.17, favoring acupuncture). In all, 104,491 MAs showed a statistically significant difference in favor of acupuncture, while 392,037 failed to demonstrate the efficacy of acupuncture (including 96 that showed a statistically significant difference in favor of the control).

Conclusions: The methodological choices made in performing pairwise meta-analyses can result in substantial VoE, occasionally leading to opposite results.

Study registration: <https://osf.io/a7ncu/>

Keywords: Meta-analysis, Vibration of effect, Acupuncture, Smoking cessation, Effect size, Complementary therapies

Running title: Vibration of effects in meta-analyses assessing acupuncture

Journal Pre-proof

What is new?

- Choice of selection criteria impacted effect estimates in meta-analyses of acupuncture for smoking cessation.
- Certain methodological choices may be associated with important increases in effect sizes.
- Variation of effects evaluations allow for exploring controversies related to conflicting meta-analyses.

Journal Pre-proof

INTRODUCTION

Despite the controversy surrounding them, acupuncture and related interventions are widely used in current clinical practice [1]. Although the idea that "acupoints", that is, specific locations on the energy meridians claimed to regulate the flow of Qi along these meridians and correct energy imbalance [2], [3] seems esoteric, some studies have suggested a biological basis for acupuncture. It has been reported that acupuncture activates the release of different kinds of neuropeptides (e.g., enkephalins, beta-endorphin, and endomorphins) [4] and that electro-acupuncture protects dopaminergic neurons [5], [6], but these results have not been replicated [7]–[9]. Regardless of the biology, the evidence base for its use is also tarnished by reproducibility issues. With thousands of randomized controlled trials (RCTs), the literature is undermined by the hundreds of contradictory systematic reviews and meta-analyses (SRMA) assessing acupuncture efficacy conducted in recent decades.

Acupuncture for smoking cessation is a relevant example of these redundant and divergent meta-analyses. At least 10 [10]–[19] SRMAs of acupuncture for smoking cessation have been published. Some did not find sufficient evidence of an effect, e.g., [17], while others did, e.g., [11]–[13]. Such irreproducibility is a major problem, especially when it concerns the gold standard in evidence-based medicine — SRMAs [20].

Examining the vibration of effects (VoE) is a heuristic approach proposed for a systematic and objective assessment of the potential of obtaining discordant results due to designs that allow for a large range of analytical choices [21]. In the context of SRMAs, VoE can be explored by computing meta-analyses under various analytical strategies. Multiple potential sources of heterogeneity across acupuncture trials can generate VoE, including, among others, different types of acupuncture, different treatment protocols, and different control groups [22].

In this case study, we aimed to explore the irreproducibility of meta-analyses of acupuncture for smoking cessation by a three-step procedure: (1) evaluating the extent of divergences in meta-analyses with an umbrella review; (2) performing multiple overlapping meta-analyses by varying numerous analytical choices; (3) investigating the impact of the different methodological choices on VoE.

METHODS

Design

Before the study began, we registered a protocol in the Open Science Framework (osf.io/vy9p7).

Data source and search strategy

To be included in the umbrella review, SRMAs had to explore specifically the efficacy of acupuncture for smoking cessation and had to include evidence from at least two controlled clinical trials.

Electronic searches took place on PubMed/Medline, the COCHRANE Library, the Wan Fang database, and the Chinese Journal Full-text Database (CNKI). The same algorithm was used for all databases without restrictions of time or language and is detailed in **Web appendix 1**.

All primary studies included in these systematic reviews were identified. To complete this strategy, we searched for randomized and “quasi-randomized” controlled trials (i.e., trials that used an inadequate approach to generating the random sequences, such as alternation or date of birth) with the same search strategy we used for SRMAs.

Among all these primary studies, those eligible for the VoE analysis were all the RCTs comparing acupuncture or related therapies (stimulation in areas of the body or the face described by the study’s authors as acupuncture points) with an inactive control (placebo acupuncture, no treatment, other inactive or minimal intervention such as lactose capsules, vitamins or general tobacco control messages), without language restrictions. Trials were included whether acupuncture was given alone or as an adjunct to other interventions, as long as the other interventions were given to all groups (e.g., acupuncture and nicotine replacement vs. nicotine replacement alone). Risk of bias of primary studies was assessed with the Cochrane Collaboration tool for assessing risk of bias [23].

Study selection and data extraction are detailed in **Web appendix 1**

Definition of the different analytical scenarios

To explore VoE, we combined various plausible methodological choices that could be selected and combined by meta-analysts exploring the efficacy of acupuncture. These methodological choices were based on different inclusion/exclusion criteria for selecting RCTs (e.g., language restrictions, geographic location, type and style of acupuncture, treatment dose and duration, type of control group, outcome reported, publication status, or risk of bias). Those choices were taken before the research started during a discussion among all authors of all potential sources of VoE they could envision, based on their own expertise and knowledge of previous meta-analyses. **Web appendix 2** details all the different methodological choices and the rationales behind the choice. All possible combinations of criteria were made to create, in theory, 12,150,000 scenarios for testing the efficacy of acupuncture.

Descriptive analysis

The meta-analyses performed under the different analytical scenarios have been summarized descriptively, including all information regarding: (1) the methodological choices used to perform the meta-analyses, and (2) the number of studies included in the respective meta-analyses.

Assessment of VoE

The meta-analysis used a random effects model. The treatment effect sizes (ES) were expressed as standardized mean differences (Cohen's d), as the different outcomes were assessed differently within the different studies. For binary outcomes, odds ratios (and their variance) were converted into Cohen's d with the formulas in Bornstein et al. [24]. For each scenario, we extracted the ES (with its 95% confidence interval), the corresponding P -value, the I^2 index, and the P -value of the Cochran Q test for heterogeneity.

To represent both the magnitude of the effects in all meta-analyses and their conclusions by a simple visual representation, we created a heat map to represent VoE. Effect size was plotted on the x-axis, and the logarithm of the P -value of the test for the difference between the acupuncture and control groups on the y-axis. As a principal outcome, we inspected the heatmap: a substantial VoE is demonstrated by the presence of a "Janus effect" regarding effect size, that is, an effect in opposite directions in the 90th compared with the 10th percentile of meta-analyses. Similarly, a heat map was drawn to show VoE for heterogeneity. The I^2 index was plotted on the x-axis, and the logarithm of the P -value of the Q-test on the y-axis.

Exploration of methodological choices associated with VoE

Two exploratory analyses were conducted to assess the methodological choices associated with VoE. A multiple linear regression was used to examine the association between each methodological choice and the change in effect size, while a binomial logistic regression model tested the association between each methodological choice and the conclusions of the meta-analyses (acupuncture is effective versus ineffective).

Change to the initial protocol

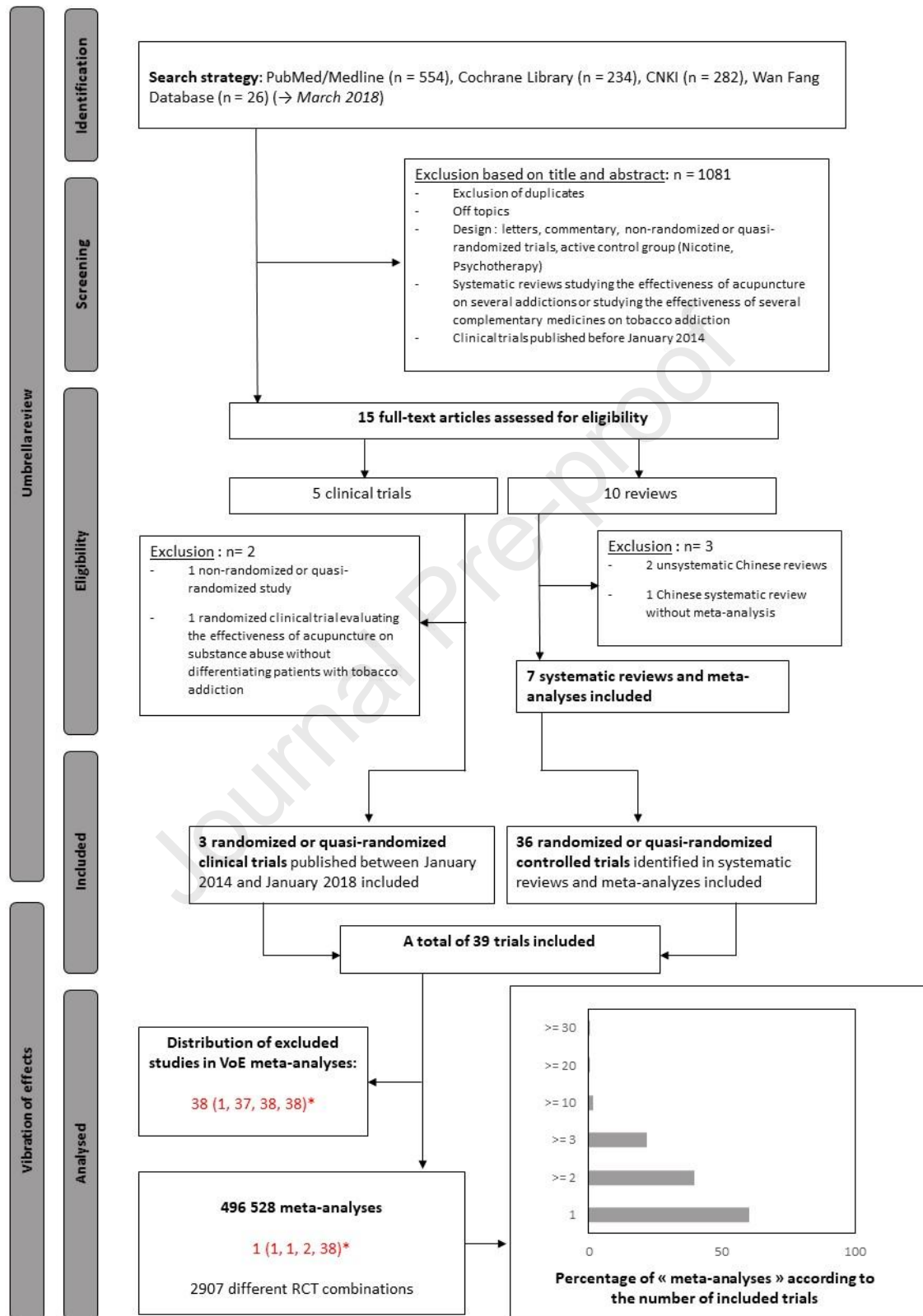
We initially planned to consider an additional methodological criterion — whether the studies included were single- or multicenter designs. Because of the inadequate reporting of this characteristic, with no study explicitly described as a multicenter design, this criterion was finally not used.

RESULTS

Umbrella review

This literature search took place on March 30, 2018. We identified seven SRMAs published between 1997 and 2015 that qualified for inclusion in the umbrella review. Six were published in English [10], [12], [15]–[18] and one in Chinese [25]. Characteristics of the meta-analyses included in the overview of reviews are summarized in **Web appendix 3**. All together, these meta-analyses reached conclusions that ranged from “not superior to sham acupuncture” [15] to “a significant effect” [12]. The median number of primary studies included in these SRMAs was 20 (range, 9-38). Among these primary studies, 36 met our inclusion criteria. After identifying three additional primary studies, published between January 2014 and January 2018, via systematic searches, we finally included 39 primary studies. **Figure 1** is the flowchart detailing the study selection process.

Figure 1: Flow diagram

*Distributions are presented as Median (Minimum, 1st quartile, 3rd quartile, Maximum)

Primary study characteristics

The main characteristics of the 39 primary studies included are set forth in **Web appendix 4 and summarized in table 1**. Among them, 30 clinical trials were published in English [26]–[55], 4 in Chinese [56]–[59], 4 more in French [60]–[63] and one in Italian [64].

Table 1. Distribution of primary studies and overlapping meta-analyses according to each possible methodological choice

<i>PICOS Category</i>	<i>Inclusion / exclusion criteria for selecting RCTs</i>	<i>Number (%) of primary studies fulfilling the criteria (n=39)¹</i>	<i>Number (%) of overlapping meta-analyses (n=496,528)²</i>
<i>Type of acupuncture (5)</i>	Only Ear-acupuncture, ear-acupressure, and auriculotherapy	24 (62)	91,776 (18)
	Only Electro-acupuncture	10 (26)	55,056 (11)
	Only Needle acupuncture with skin perforation	22 (56)	96,816 (19)
	Only noninvasive acupuncture without skin perforation	18 (46)	95,912 (19)
	Inclusion of all studies regardless of the type of acupuncture provided	39 (100)	156,968 (32)
<i>Intensity of treatment (3)</i>	Number of sessions ≤3	14 (36)	110,288 (22)
	Number of sessions >3	24 (62)	166,288 (33)
	No criteria on the number of sessions	39 (100)	219,952 (44)
<i>Treatment duration (3)</i>	Treatment duration ≤3 weeks	23 (59)	121,408 (24)
	Treatment duration >3 weeks	15 (38)	155,584 (31)
	No criteria on the treatment duration	39 (100)	219,536 (44)
<i>Treatment standardization (5)</i>	Standardized acupuncture	36 (92)	238,416 (48)
	Semi-standardized acupuncture	3 (8)	17,472 (4)
	Individualized acupuncture (all acupoints are selected, according to the Chinese diagnosis)	0 (0)	-
	Exclusion of Individualized acupuncture	39 (100)	-
	No criteria on the level of standardization	39 (100)	240,640 (48)
<i>Treatment (5)</i> <i>Acupuncture administered:</i>	Alone	39 (100)	221,512 (45)
	As an adjunct to another treatment	2 (6)	20,736 (4)
	As an adjunct to nicotine	1 (3)	6,144 (1)
	As an adjunct to nonpharmacological treatment (e.g. hypnotherapy)	1 (3)	26,624 (5)
	All comparisons whether acupuncture is administered alone or as an adjunct	39 (100)	221,512 (45)
<i>Comparator-control groups(5)</i>	Nonspecific acupuncture (at “incorrect” acupoints)	22 (56)	92,464 (19)
	Sham acupuncture at the same points	8 (21)	60,064 (12)
	All types of placebo acupuncture	32 (82)	127,872 (26)
	Only nonplacebo inactive comparators	8 (21)	65,360 (13)
	All inactive controls	39 (100)	150,768 (30)
<i>Outcome (5)</i>	Fagerström score (continuous)	5 (13)	24,848 (5)
	Cigarette consumption (continuous)	13 (33)	58,192 (12)

	Smoking cessation, or where this outcome is not reported, reduction in Cigarette consumption (binary)	38 (97)	170,328 (34)
	Smoking cessation (binary)		
	- Only smoking cessation biologically validated (eg: expired air CO concentration, serum cotinine)	14 (36)	75,248 (15)
	- All studies with a measure of smoking cessation (validated or not)	36 (92)	167,912 (34)
<i>Outcome-measurement term (3)</i>	Short term: <= 6 months (measure nearest 6 months)	37 (95)	209,536 (42)
	Long term > 6 months (last visit greater than 6 months)	10 (26)	77,696 (16)
	Last visit	39 (100)	209,296 (42)
<i>Studies -Sample size (2)</i>	Population analyzed >= 100	19 (49)	200,560 (40)
	No restriction on the sample size	39 (100)	295,968 (60)
<i>Language (3)</i>	In English only	30 (77)	217,832 (44)
	In Chinese only	4 (10)	32,512 (7)
	All languages	39 (100)	246,184 (50)
<i>Country (4)</i>	Only studies conducted in Asia (including China)	12 (31)	81,952 (17)
	Exclusion of studies conducted in Asia	27 (69)	123,856 (25)
	Exclusion of studies conducted in China	34 (87)	129,840 (26)
	All studies regardless of Country	39 (100)	160,880 (32)
<i>Risk of bias (3)</i>	Exclusion of quasi-randomized studies	35 (90)	185,440 (37)
	Exclusion of quasi-randomized studies, and of studies without a ITT analysis	18 (46)	110,736 (22)
	Inclusion of all studies regardless of randomization or ITT analysis	39 (100)	200,352 (40)
<i>Publication (2)</i>	Exclusion of unpublished studies	37 (95)	243,632 (49)
	Inclusion of published and un published data	39 (100)	252,896 (51)

Total of possible combination : 12,150,000

PICOS stands for: Patients; Intervention; Comparator; Outcomes; Study design.

¹ A trial could fulfill several criteria of the same category. For example, an acupuncture technique could correspond to auriculotherapy and involve electrical stimulation; trials could have both a short-term measure and a long-term measure, assess several outcomes, or could include more than two arms

² Number of meta-analyses (including at least 1 RCT) obtained with the corresponding methodological choice.

Twenty-four studies used auriculotherapy alone, six used facial acupuncture, and eight a mixture of points, located both on the ear and elsewhere on the body. One trial failed to specify the location of the stimulated points [32]. Different modes of stimulation were also used: manual acupuncture (n = 13), electro-acupuncture (n = 10), laser therapy (n = 3), acupressure (n = 12), and acupuncture in situ with indwelling needles (n = 8). All but three trials used a standardized therapeutic approach with exactly the same points for all patients. In the remaining studies, acupuncture treatment was semi-standardized, allowing the addition of personalized points, chosen according to the Chinese

diagnosis. Three trials [30], [42], [57] amounted to two parallel comparisons, as each included two different intervention groups, each with its own control group. Only 10 studies reported long-term measures after more than 6 months. Finally, most trials reported “smoking cessation” as the outcome (n = 36).

Risk of bias across the primary studies is presented in **Web Appendix 4**. The absence of details in reports prevented an assessment of the risk of bias in many trials. Four were quasi-randomized. Only five [33], [36], [46], [47], [51] studies were judged at low risk of bias for both randomization and allocation concealment. For adequacy of blinding, we assessed only participant blinding. There were 32 trials judged to be at low risk of bias where the comparator was designed to appear identical to the intervention: sham acupuncture at the same points used in the treatment group (eg, placebo needles mimicking skin penetration, sham electrostimulation without current) or nonspecific acupuncture stimulating active acupoints that are not considered relevant to the condition being treated or stimulating an area that is not a recognized "acupoint"). Subjects were not blinded in the remaining seven trials.

Vibration of effects across meta-analyses

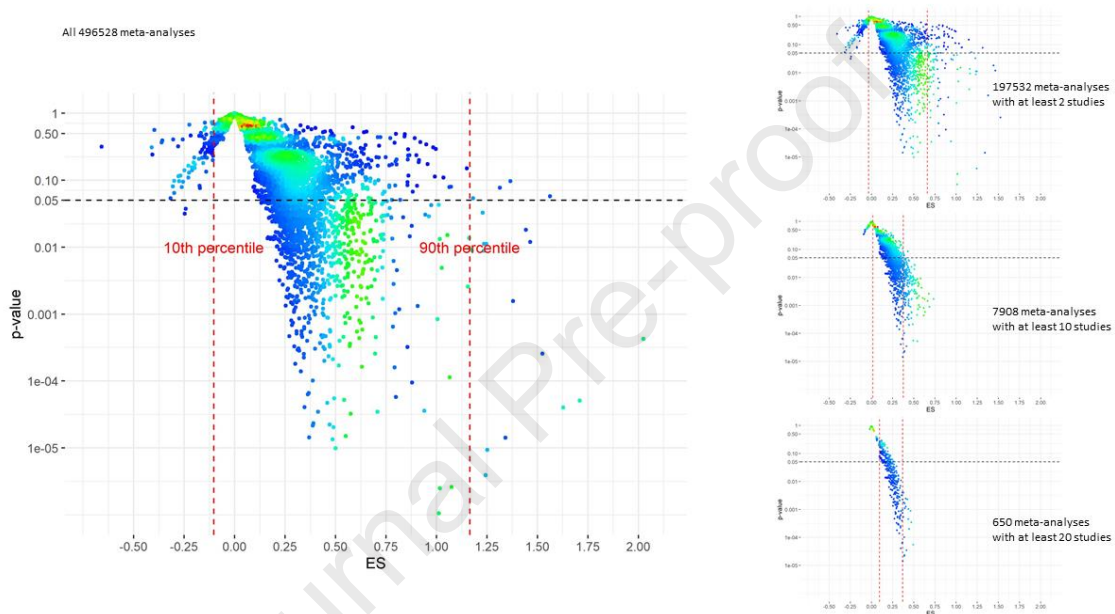
Of the 12,150,000 possible combinations of criteria, 496,528 (4.09%) matched some of the included trials, so that meta-analyses were possible. These 496,528 overlapping meta-analyses produced 2,907 different RCT combinations. **Table 1** presents the distribution of the meta-analyses according to each methodological choice and **Figure 1** describes the distribution of the number of trials included by meta-analysis. About 60.22% (n = 298,996) of these meta-analyses included only one trial, 39.78% (n = 197,532) at least two, 21.77% (n = 108,080) at least three, 1.59% (n = 7,908) at least 10, 0.13% (n = 650) at least 20, and 0.0001% (n=72) at least 30.

Among these 496,528 meta-analyses, ES ranged from -0.66 to 2.02, with a mean of 0.35 and a median of 0.13. The ES was negative at the 10th percentile (ES = -0.1) and positive at the 90th percentile (ES = 1.17). In all, 46,612 meta-analyses had an ES below the 10th percentile (0.2% of them with significant differences in favor of the control group) and 49,616 had an ES above the 90th percentile (88.4% with significant differences in favor of acupuncture). Among all meta-analyses, 104,491 showed a statistically significant difference in favor of acupuncture, while 392,037 failed to do so and 96 even showed a statistically significant difference in favor of the control. **Figure 2** presents the heat map of the VoE. Due to the high number of meta-analyses including only one study, we performed sensitivity analyses by measuring the VoE for meta-analyses with at least 2, 3, 10, and 20 trials (those of 2, 10, and 20 trials are shown on **Figure 2**). **Web appendix 5** describes the

distribution of the ESs for all meta-analyses and for those with at least 2, at least 3, at least 10, and at least 20 trials.

Among the meta-analyses of at least two trials, the sign of the ES at the 10th percentile was opposite to that of the ES at the 90th percentile (ES = -0.03 vs. ES = 0.66). This persisted for meta-analyses of at least three trials, although the ES at the 10th percentile was exceedingly small (ES = -0.0044 at the 10th percentile vs. ES = 0.51 at the 90th percentile).

Figure 2: Vibration of effects for the comparison of acupuncture with the control.



A negative ES favors the control, and a positive ES favors acupuncture. The points represent the meta-analyses. The colors represent the densities. The analysis is present for all meta-analyses and for meta-analyses including 2, 10, or 20 primary studies.

Assessment of heterogeneity

Heterogeneity⁷⁷ was assessed in the 197,532 meta-analyses with at least two studies and is presented graphically in **Web appendix 6**. The estimated I^2 values range from 0 to 97.24%, with a median of 0.79%, a 1st quartile of 0%, and a 3rd quartile of 62.35%.

Methodological choices associated with VoE

The methodological choices associated with an ES increase (**Web appendix 7 and 8**) greater than 0.2 were inclusion of studies in Chinese only (+0.31 vs. inclusion of all languages), focus on Asian studies (+0.31 vs. inclusion of all countries), exploration of semi-standardized approaches +0.90 vs. inclusion of all approaches), and use of nonplacebo inactive comparators for acupuncture (no treatment, other inactive or minimal intervention) vs. inclusion of all control groups (+0.29). The choice of the

Fagerström Test for Nicotine Dependence as an outcome was also associated with an ES increase (+0.22 vs. smoking cessation/cigarette consumption). No methodological choice was associated with a decrease in ES greater than 0.2. The logistic regression model predicting the meta-analysis conclusions produced similar results (**Web appendix 8**).

DISCUSSION

Statement of principal findings

In this case study, we performed 496,528 meta-analyses evaluating the effectiveness of acupuncture in smoking cessation and observed divergent results: 104,491 meta-analyses (809 different RCT combinations) concluded that acupuncture was effective, while 392,037 (2,354 different RCT combinations) could not reach that conclusion. None of these meta-analyses was exhaustive, that is, included all 39 RCTs we retrieved; the most complete included 38 RCTs. The methodological heterogeneity was large both among SMRAs and among the studies they included. Using the VoE framework [21], [65], we showed that various combinations of possible methodological choices yield very different conclusions and explain some of the discrepancies across meta-analyses with several practical implications.

A 2014 systematic review reported that 838/840 (99.8%) RCTs evaluating acupuncture published in Chinese journals had positive results [66]. The plausibility of such results is questionable. SRMAs must, of course, balance the principle of exhaustiveness with the possible risk of bias in any study. This situation presents a dilemma, as excluding Chinese studies may miss a significant part of the information. In addition, it is possible that Western acupuncturists may be less well trained than their Chinese colleagues.

Concerning the level of standardization, some authors recommend evaluating the personalized interventions in RCTs of acupuncture and of nonconventional medicine in general [67]–[69], because individualization and flexibility may be a major component of the treatment.

But beyond acupuncture, the comparator may be a source of VoE. It has also been pointed out that trials of acupuncture including both “placebos” of acupuncture and no treatment control groups report that the former is associated with larger effects than the latter [70]. It remains controversial whether sham/nonspecific acu punctures are physiologically active (such as via the release of neurotransmitters) [71], or if they induce an inordinately high placebo effect [72] due to the psychological impact of the treatment ritual, greater patient's positive beliefs and expectations [73], or other causes [74]. It has also been suggested that acupuncture may be nothing but a “theatrical placebo” [75]. In any case, meta-analysis outcomes partially depend on the magnitude of the placebo

response of the control intervention : greater placebo response favors a negative outcome, whereas a smaller placebo response favors a positive outcome [72]. That is, results of RCTs may reflect bias, and SRMAs then amplify it.

Strengths and weaknesses of the study

Our study has certain limitations. First it is a case study based on acupuncture for smoking cessation. Acupuncture is a unique therapy in some ways, with major sources of heterogeneity across studies, and thus perhaps especially prone to VoE. Then, although our meta-analyses are based on methodological choices that individually make sense, combining some of them may lead to implausible or impossible scenarios. Certainly, it seems unlikely that a meta-analysis would include only studies written in Chinese while excluding trials carried out in China or Asia, and indeed, most combinations resulted in zero ($n = 7,795,287$) or only one trial ($n = 195,423$). Nonetheless, the sensitivity analyses of meta-analyses including at least 2 or 3 trials also found substantial VoE. In the larger meta-analyses — those including at least 20 studies — VoE resulted in either positive or negative conclusions (in terms of statistical significance) but without change in the direction of the effect. On the other hand, the literature is populated by SRMAs that are either empty or include only one study. An examination of 400 Cochrane systematic reviews showed 14% of them were empty and 7% of the remaining 86% included only one trial [76].

In addition, many analytical choices must be made in VoE analyses — an exercise involving some arbitrary and difficult choices and thus sometimes some redundancy (e.g., language and country in our study). For transparency purposes, we preregistered these choices to avoid the potential of allowing findings to bias our choices. But many other choices might have been made and would have resulted in different patterns of VoE. To limit such variability, we suggest that in future VoE analyses, these choices be guided on a systematic review.

The VoE framework may thus be considered an extreme form of sensitivity analyses that may sometimes lack clinical relevance. It is however based on an *a priori* determination of the methodological choices considered to be relevant and may be less arbitrary than the all-subset method, in which all possible meta-analyses of all possible subsets of studies are explored [77]. In addition, we performed a series of regressions to understand the sources of VoE. Although such an analysis must be interpreted cautiously because it is prone to confounding, it may help to understand the discrepancies found in multiple overlapping meta-analyses in a given field. Moreover, the VoE framework enables the *a posteriori* inspection of some combinations of interest. We were able to select two distinct plausible combinations of different criteria, both simple and plausible, that produced two contradictory meta-analyses (**Appendix 9**) — one negative (21 trials, $ES = 0.02$, $P =$

0.73) and the other positive (21 trials, ES = 0.44, $P = 0.001$). Future studies may use such divergent examples to explore how these overlapping meta-analyses have slipped past the safeguards of the peer review process. As the recent publication and retraction of a network meta-analysis comparing acupuncture with drug therapies in the treatment of chronic constipation suggests [78], highly biased SRMAs can appear in the peer-reviewed literature. This study, initially published in PLOS ONE, was subsequently retracted for various reasons including the imprecision of its search strategy, its vague inclusion and exclusion criteria, and errors in data extraction.

Strengths and weaknesses in relation to other studies, discussing important differences in results

The result we observed is in line with results observed for naltrexone in a previous case study exploring VoE in indirect comparisons of nalmefene and naltrexone in the reduction of alcohol consumption [65]. In that study, the VoE observed in indirect comparisons led to meta-analyses with opposite conclusions (in term of ES and statistical significance). The VoE observed in direct comparisons against placebo was smaller for nalmefene (a recent drug developed in two coherent research programs) than for naltrexone, which is an older drug with a myriad of pre- and post-approval RCTs conducted in very different settings. Unsurprisingly, these characteristics led to results very similar to those observed for acupuncture in smoking cessation, with substantial clinical and statistical heterogeneity between studies.

There is a large literature on the impact of various selection criteria used in meta-analyses such as population size [79], number of centers [80], and quality of trials included [81]–[85] on the meta-analysis results. By performing a multiverse analysis (i.e., performing all analyses across the whole set of alternatively processed data sets corresponding to a large set of reasonable scenarios) [86], VoE goes one step further and explores how much various major options in different combinations can potentiate each other and lead to totally contradictory conclusions. However, it implies predefining choices that may differ in different contexts. For example, in exploring VoE for drug studies, it might be easier to explore VoE based on the different sources from which the data from a study can be extracted because many documents are available (e.g., published articles, study reports, etc.); this is less true in acupuncture research. We therefore did not explore this important source of potential VoE [87]. Importantly, our quantification of the methodological choices possibly associated with VoE is exploratory and must be interpreted with caution. A major limitation of our approach lies in the fact that we performed a regression on a wide range of meta-analyses, some of which are redundant (and are combinations of the same set of studies). While point estimates are not expected to be biased, non-independence between those meta-analyses prevents any easy computation of confidence intervals, which we did not report for this reason. Solutions to derive such confidence

intervals (e.g., by weighting the results of the different studies or by analyzing only unique combinations of studies) require more development before they are implemented and adopted widely.

Meaning of the study: possible explanations and implications for clinicians and policymakers

As the massive production of meta-analyses has reached epidemic proportions in recent decades, with sometimes very contradictory results [88], new methods are needed to improve our understanding of the reproducibility issues within these studies. Our results point out once again that SRMA findings can be prone to certain vibration: a critical approach to appraising them is required [89].

Unanswered questions and future research

Our case study illustrates that in umbrella reviews the VoE framework may add value by mapping the possible uncertainty in a given field. Nonetheless, even bearing in mind the previous case study suggesting the value of the VoE framework [65], there is still insufficient evidence to adopt this tool widely as a definitive method for analyzing SRMAs. More case studies are needed, in the contexts, for example, of network meta-analyses and individual patient data meta-analyses. Exploring VoE systematically in a large set of SRMAs in different topics will provide a better sense of its utility.

Authors' contributions

FN, CB, and MEB conceived and designed the experiments. MEH, XW, TB extracted the experiments. MEH and CB analyzed the data. MEH and FN wrote the first draft of the manuscript. CB and BF contributed to the writing of the manuscript. MEH, XW, TB, BF, FN, and CB agreed with the results and conclusions of the manuscript. All authors read and approved the final manuscript.

Data sharing statement

Data and code are available on the Open Science Framework (<https://osf.io/tznsr/>).

Funding

No specific funding for this study.

REFERENCES

- [1] E. Ernst, 'Acupuncture--a critical analysis', *J. Intern. Med.*, vol. 259, no. 2, pp. 125–137, Feb. 2006, doi: 10.1111/j.1365-2796.2005.01584.x.
- [2] J.-E. Park *et al.*, 'A literature review of de qi in clinical studies', *Acupunct. Med.*, vol. 31, no. 2, pp. 132–142, Jun. 2013, doi: 10.1136/acupmed-2012-010279.
- [3] T. J. Kaptchuk, 'Acupuncture: theory, efficacy, and practice', *Ann. Intern. Med.*, vol. 136, no. 5, pp. 374–383, Mar. 2002, doi: 10.7326/0003-4819-136-5-200203050-00010.
- [4] J.-S. Han, 'Acupuncture and endorphins', *Neurosci. Lett.*, vol. 361, no. 1–3, pp. 258–261, May 2004, doi: 10.1016/j.neulet.2003.12.019.
- [5] H.-J. Park *et al.*, 'Acupuncture prevents 6-hydroxydopamine-induced neuronal death in the nigrostriatal dopaminergic system in the rat Parkinson's disease model', *Exp. Neurol.*, vol. 180, no. 1, pp. 93–98, Mar. 2003, doi: 10.1016/s0014-4886(02)00031-6.
- [6] X.-Y. Liu *et al.*, 'Electro-acupuncture stimulation protects dopaminergic neurons from inflammation-mediated damage in medial forebrain bundle-transected rats', *Exp. Neurol.*, vol. 189, no. 1, pp. 189–196, Sep. 2004, doi: 10.1016/j.expneurol.2004.05.028.
- [7] R. C. Chapman, C. Benedetti, Y. H. Colpitts, and R. Gerlach, 'Naloxone fails to reverse pain thresholds elevated by acupuncture: acupuncture analgesia reconsidered', *Pain*, vol. 16, no. 1, pp. 13–31, May 1983, doi: 10.1016/0304-3959(83)90082-9.
- [8] D. F. Bossut and D. J. Mayer, 'Electroacupuncture analgesia in rats: naltrexone antagonism is dependent on previous exposure', *Brain Res.*, vol. 549, no. 1, pp. 47–51, May 1991, doi: 10.1016/0006-8993(91)90597-o.
- [9] I. K. Gareus, M. Lacour, A.-C. Schulte, and J. Hennig, 'Is there a BOLD response of the visual cortex on stimulation of the vision-related acupoint GB 37?', *J. Magn. Reson. Imaging JMRI*, vol. 15, no. 3, pp. 227–232, Mar. 2002, doi: 10.1002/jmri.10059.
- [10] R. Ashenden, C. A. Silagy, M. Lodge, and G. Fowler, 'A meta-analysis of the effectiveness of acupuncture in smoking cessation', *Drug Alcohol Rev.*, vol. 16, no. 1, pp. 33–40, Mar. 1997, doi: 10.1080/09595239700186311.
- [11] P. Castera, J. Nguyen, J. Gerlier, and S. Robert, 'Is acupuncture beneficial in tobacco cessation, and is its effect specific? A meta-analysis', *Acupunct. Moxibustion*, vol. 1, no. 4, pp. 76–85, 2002.
- [12] H. M. Cheng, Y. C. Chung, H. H. Chen, Y. H. Chang, and M. L. Yeh, 'Systematic review and meta-analysis of the effects of acupoint stimulation on smoking cessation', *Am J Chin Med*, vol. 40, no. 3, pp. 429–442, 2012, doi: 10.1142/S0192415X12500334.
- [13] M. Tahiri, S. Mottillo, L. Joseph, L. Pilote, and M. J. Eisenberg, 'Alternative smoking cessation aids: a meta-analysis of randomized controlled trials', *Am. J. Med.*, vol. 125, no. 6, pp. 576–584, Jun. 2012, doi: 10.1016/j.amjmed.2011.09.028.
- [14] G. Ter Riet, J. Kleijnen, and P. Knipschild, 'A meta-analysis of studies into the effect of acupuncture on addiction', *Br. J. Gen. Pract. J. R. Coll. Gen. Pract.*, vol. 40, no. 338, pp. 379–382, Sep. 1990.
- [15] A. R. White, K. L. Resch, and E. Ernst, 'A meta-analysis of acupuncture techniques for smoking cessation', *Tob. Control*, vol. 8, no. 4, pp. 393–397, 1999, doi: 10.1136/tc.8.4.393.
- [16] A. White and R. Moody, 'The effects of auricular acupuncture on smoking cessation may not depend on the point chosen--an exploratory meta-analysis', *Acupunct. Med. J. Br. Med. Acupunct. Soc.*, vol. 24, no. 4, pp. 149–156, Dec. 2006, doi: 10.1136/aim.24.4.149.
- [17] A. R. White, H. Rampes, J. P. Liu, L. F. Stead, and J. Campbell, 'Acupuncture and related interventions for smoking cessation', *Cochrane Database Syst. Rev.*, no. 1, p. CD000009, Jan. 2014, doi: 10.1002/14651858.CD000009.pub4.
- [18] Y. M. Di, B. H. May, A. L. Zhang, I. W. Zhou, C. Worsnop, and C. C. L. Xue, 'A meta-analysis of ear-acupuncture, ear-acupressure and auriculotherapy for cigarette smoking cessation', *Drug Alcohol Depend.*, vol. 142, pp. 14–23, Sep. 2014, doi: 10.1016/j.drugalcdep.2014.07.002.

- [19] J. Schwartz, 'Evaluation of acupuncture as a treatment for smoking', *Am. J. Acupunct.*, vol. 16, no. 2, pp. 135–142, 1988.
- [20] K. C. Siontis, T. Hernandez-Boussard, and J. P. A. Ioannidis, 'Overlapping meta-analyses on the same topic: survey of published studies', *BMJ*, vol. 347, p. f4501, Jul. 2013, doi: 10.1136/bmj.f4501.
- [21] C. J. Patel, B. Burford, and J. P. A. Ioannidis, 'Assessment of vibration of effects due to model specification can demonstrate the instability of observational associations', *J. Clin. Epidemiol.*, vol. 68, no. 9, pp. 1046–1058, Sep. 2015, doi: 10.1016/j.jclinepi.2015.05.029.
- [22] S. Birch, 'Controlling for non-specific effects of acupuncture in clinical trials', *Clin. Acupunct. Orient. Med.*, vol. 2–3, no. 4, pp. 59–70, 2003, doi: 10.1016/S1461-1449(03)00025-2.
- [23] J. P. T. Higgins *et al.*, 'The Cochrane Collaboration's tool for assessing risk of bias in randomised trials', *BMJ*, vol. 343, p. d5928, Oct. 2011, doi: 10.1136/bmj.d5928.
- [24] M. Borenstein, L. V. Hedges, J. P. Higgins, and H. R. Rothstein, *Introduction to meta-analysis*. John Wiley & Sons, 2021.
- [25] Z. Liu, Y. Wang, Y. Wu, and J. Yang, '[Condition and effectiveness evaluation of acupuncture for smoking cessation]', *Zhongguo Zhen Jiu Chin. Acupunct. Moxibustion*, vol. 35, no. 8, pp. 851–857, Aug. 2015.
- [26] D. Antoniou *et al.*, 'Aural electrical stimulation for smoking cessation - a double blind, placebo controlled study [Abstract].', *Eur. Respir. J.*, vol. 26, no. Suppl. 49, p. 2492, 2005.
- [27] A. Aycicegi-Dinn and W. Dinn, 'Efficacy of an alternative smoking cessation treatment', *J. Addict. Dis.*, vol. 30, no. 4, pp. 368–381, Oct. 2011, doi: 10.1080/10550887.2011.609803.
- [28] Y. Cai, C. Zhao, S. U. Wong, L. Zhang, and S. K. Lim, 'Laser acupuncture for adolescent smokers—a randomized double-blind controlled trial', *Am. J. Chin. Med.*, vol. 25, no. 3–4, pp. 443–449, 2000.
- [29] F. Clavel, S. Benhamou, A. Company-Huertas, and R. Flamant, 'Helping people to stop smoking: randomised comparison of groups being treated with acupuncture and nicotine gum with control group', *Br. Med. J. Clin. Res. Ed*, vol. 291, no. 6508, pp. 1538–1539, Nov. 1985, doi: 10.1136/bmj.291.6508.1538.
- [30] F. Clavel-Chapelon, C. Paoletti, and S. Benhamou, 'A randomised 2 x 2 factorial design to evaluate different smoking cessation methods', *Rev. Epidemiol. Sante Publique*, vol. 40, no. 3, pp. 187–190, 1992.
- [31] J. A. Cottraux, R. Harf, J. P. Boissel, J. Schbath, M. Bouvard, and J. Gillet, 'Smoking cessation with behaviour therapy of acupuncture—a controlled study', *Behav. Res. Ther.*, vol. 21, no. 4, pp. 417–424, 1983, doi: 10.1016/0005-7967(83)90011-6.
- [32] G. Docherty, D. Gordon, and L. McAlpine, 'Laser and NRT smoking cessation programmes in areas of high social deprivation [Abstract]', 2003, vol. 58(Suppl 3), p. iii43.
- [33] D. J. Fritz, R. M. Carney, B. Steinmeyer, G. Ditson, N. Hill, and J. Zee-Cheng, 'The efficacy of auriculotherapy for smoking cessation: a randomized, placebo-controlled trial', *J. Am. Board Fam. Med. JABFM*, vol. 26, no. 1, pp. 61–70, Feb. 2013, doi: 10.3122/jabfm.2013.01.120157.
- [34] A. J. Georgiou, C. P. Spencer, G. K. Davies, and J. Stamp, 'Electrical stimulation therapy in the treatment of cigarette smoking', *J. Subst. Abuse*, vol. 10, no. 3, pp. 265–274, 1998, doi: 10.1016/s0899-3289(99)00005-x.
- [35] V. Gilbey and B. Neumann, 'Auricular acupuncture for smoking withdrawal', *Am. J. Acupunct.*, vol. 5, no. 3, pp. 239–247, 1977.
- [36] J. Gillams, G. T. Lewith, and D. Machin, 'Acupuncture and group therapy in stopping smoking', *The Practitioner*, vol. 228, no. 1389, pp. 341–344, Mar. 1984.
- [37] D. He, J. E. Berg, and A. T. Høstmark, 'Effects of acupuncture on smoking cessation or reduction for motivated smokers', *Prev. Med.*, vol. 26, no. 2, pp. 208–214, Apr. 1997, doi: 10.1006/pmed.1996.0125.
- [38] C. M. Kerr, P. B. Lowe, and N. I. Spielholz, 'Low level laser for the stimulation of acupoints for smoking cessation: a double blind, placebo controlled randomised trial and semi structured interviews', *J. Chin. Med.*, vol. 86, pp. 46–51, 2008.

- [39] Y. Lamontagne, L. Annable, and M. A. Gagnon, 'Acupuncture for smokers: lack of long-term therapeutic effect in a controlled study', *Can. Med. Assoc. J.*, vol. 122, no. 7, pp. 787–790, Apr. 1980.
- [40] J. Leung, 'Smoking cessation by auricular acupuncture and behavioral therapy.', *Psychol. Int. J. Psychol. Orient*, vol. 34, pp. 177–187, 1991.
- [41] G. P. Martin and P. M. Waite, 'The efficacy of acupuncture as an aid to stopping smoking', *N. Z. Med. J.*, vol. 93, no. 686, pp. 421–423, Jun. 1981.
- [42] L. Parker and M. Mok, 'The use of acupuncture for smoking withdrawal.', *Am. J. Acupunct.*, vol. 5, pp. 363–366, 1977.
- [43] R. P. Steiner, D. L. Hay, and A. W. Davis, 'Acupuncture therapy for the treatment of tobacco smoking addiction', *Am. J. Chin. Med.*, vol. 10, no. 1–4, pp. 107–121, 1982, doi: 10.1142/S0192415X82000178.
- [44] Z. Tian and Y. Chu, 'Treating smoking addiction with the ear point seed pressing method', *J. Chin. Med.*, vol. 52, pp. 5–6, 1996.
- [45] N. R. Waite and J. B. Clough, 'A single-blind, placebo-controlled trial of a simple acupuncture treatment in the cessation of smoking', *Br. J. Gen. Pract. J. R. Coll. Gen. Pract.*, vol. 48, no. 433, pp. 1487–1490, Aug. 1998.
- [46] A. R. White, K. L. Resch, and E. Ernst, 'Randomized trial of acupuncture for nicotine withdrawal symptoms', *Arch. Intern. Med.*, vol. 158, no. 20, pp. 2251–2255, Nov. 1998, doi: 10.1001/archinte.158.20.2251.
- [47] A. R. White, R. C. Moody, and J. L. Campbell, 'Acupressure for smoking cessation--a pilot study', *BMC Complement. Altern. Med.*, vol. 7, p. 8, Mar. 2007, doi: 10.1186/1472-6882-7-8.
- [48] Y.-K. Wing, A. Lee, E. L. Y. Wong, P. Leung, L. Zhang, and E. S. Y. Pang, 'Auricular Acupressure for Smoking Cessation: A Pilot Randomized Controlled Trial', *Med. Acupunct.*, vol. 22, no. 4, pp. 265–271, Dec. 2010, doi: 10.1089/acu.2010.0763.
- [49] T.-P. Wu, F.-P. Chen, J.-Y. Liu, M.-H. Lin, and S.-J. Hwang, 'A randomized controlled clinical trial of auricular acupuncture in smoking cessation', *J. Chin. Med. Assoc. J. CMA*, vol. 70, no. 8, pp. 331–338, Aug. 2007, doi: 10.1016/S1726-4901(08)70014-5.
- [50] M.-L. Yeh, C.-Y. Chang, N.-F. Chu, and H.-H. Chen, 'A six-week acupoint stimulation intervention for quitting smoking', *Am. J. Chin. Med.*, vol. 37, no. 5, pp. 829–836, 2009, doi: 10.1142/S0192415X09007314.
- [51] A. L. Zhang, Y. M. Di, C. Worsnop, B. H. May, C. Da Costa, and C. C. L. Xue, 'Ear acupressure for smoking cessation: a randomised controlled trial', *Evid.-Based Complement. Altern. Med. ECAM*, vol. 2013, p. 637073, 2013, doi: 10.1155/2013/637073.
- [52] R. de P. Silva, E. de C. L. Chaves, S. C. Pillon, A. M. Silva, D. da S. Moreira, and D. H. Lunes, 'Contributions of auriculotherapy in smoking cessation: a pilot study', *Rev. Esc. Enferm. U P*, vol. 48, no. 5, pp. 883–890, Oct. 2014, doi: 10.1590/s0080-6234201400005000015.
- [53] H.-C. Kang, K.-K. Shin, K.-K. Kim, and B.-B. Youn, 'The effects of the acupuncture treatment for smoking cessation in high school student smokers', *Yonsei Med. J.*, vol. 46, no. 2, pp. 206–212, Apr. 2005, doi: 10.3349/ymj.2005.46.2.206.
- [54] M. Bilici, S. Güven, S. Köşker, A. Şafak, and Ü. B. Semiz, 'Electroacupuncture Therapy in Nicotine Dependence: A Double Blind, Sham-Controlled Study', *Noro Psikiyatri Arsivi*, vol. 53, no. 1, pp. 28–32, Mar. 2016, doi: 10.5152/npa.2015.9887.
- [55] S. Lee and H. Park, 'The Effects of Auricular Acupressure on Smoking Cessation for Male College Students', *West. J. Nurs. Res.*, vol. 39, no. 3, pp. 374–387, Mar. 2017, doi: 10.1177/0193945916660080.
- [56] Y. Li, T. Shen, L. Cao, W. Zhang, and Q. Wu, 'Clinical research of the seed of cowherb auricular pressing in treating 70 cases of quitting smoking', *Jilin J. Tradit. Chin. Med.*, vol. 29, no. 6, pp. 505–506, 2009.
- [57] L. Li, J. Sun, D. Xu, L. Shen, X. Chen, and A. Yao, 'The approach of ear acupuncture combining to psychological intervention with the health belief mode in community tobacco control.', *Zhejiang Prev. Med.*, vol. 23, no. 9, pp. 23–25, 2011.

- [58] Y. Fang, '[Clinical study on auricular acupuncture for treatment of smoking addiction][Chinese]', *Shanghai J. Acupunct. Moxibustion*, vol. 2, pp. 30–31, 1983.
- [59] Y. Huang, 'The Clinical Observation of Nicotine Dependence Treated with Smoke Three-Needles Combine Auricular Pressure Therapy', *Masters Diss. Guangzhou Univ. Chin. Med.*, 2012.
- [60] J. C. Lacroix and F. Besançon, '[Tobacco withdrawal. Efficacy of acupuncture in a comparative trial]', *Ann. Med. Interne (Paris)*, vol. 128, no. 4, pp. 405–408, Apr. 1977.
- [61] G. Lagrue, J. L. Poupy, A. Grillot, and J. C. Ansquer, '[Antismoking acupuncture. Short-term results of a double-blind comparative study]', *Nouv. Presse Med.*, vol. 9, no. 13, p. 966, Mar. 1980.
- [62] A. Vandevenne, M. Rempp, G. Burghard, Y. KUNTZMANN, and F. JUNG, 'Study of the specific contribution of acupuncture to tobacco detoxication [Etude de l'action spécifique de l'acupuncture dans la cure de sevrage tabagique]', *Sem Hôp. Paris*, vol. 61, no. 29, pp. 2155–2160, 1985.
- [63] J. Vibes, 'Clinical trial of the role of acupuncture in the fight against tobacco addiction [Essai thérapeutique sur le rôle de l'acupuncture dans la lutte contre le tabagisme].', *Acupuncture*, vol. 51, pp. 13–20, 1977.
- [64] A. Circo, A. Tosto, S. Raciti, R. Cardillo, M. Gulizia, and M. Oliveri, 'First results of an anti-smoke outpatient unit: Comparison among three methods [Primi risultati di un ambulatorio antifumo. Confronto fra tre metodice].', *Riv. Cardiol. Prev. E Riabil.*, vol. 3, no. 2, pp. 147–151, 1985.
- [65] C. Palpacuer, K. Hammas, R. Duprez, B. Laviolle, J. P. A. Ioannidis, and F. Naudet, 'Vibration of effects from diverse inclusion/exclusion criteria and analytical choices: 9216 different ways to perform an indirect comparison meta-analysis', *BMC Med.*, vol. 17, no. 1, p. 174, Sep. 2019, doi: 10.1186/s12916-019-1409-3.
- [66] Y. Wang, L. Wang, Q. Chai, and J. Liu, 'Positive Results in Randomized Controlled Trials on Acupuncture Published in Chinese Journals: A Systematic Literature Review', *J. Altern. Complement. Med.*, vol. 20, no. 5, pp. A129–A129, May 2014, doi: 10.1089/acm.2014.5346.abstract.
- [67] J. Gueguen, C. Hill, and C. Barry, 'Complementary Medicines', in *Wiley StatsRef: Statistics Reference Online*, American Cancer Society, 2015, pp. 1–13. doi: 10.1002/9781118445112.stat05556.pub2.
- [68] J. Gueguen, 'Evaluation des médecines complémentaires: quels compléments aux essais contrôlés randomisés et aux méta-analyses?', Université Paris Saclay (COMUE), 2017.
- [69] H. MacPherson, 'Out of the laboratory and into the clinic: acupuncture research in the real world', *Clin. Acupunct. Orient. Med.*, vol. 1, no. 2, pp. 97–100, 2000.
- [70] K. Linde, K. Niemann, A. Schneider, and K. Meissner, 'How large are the nonspecific effects of acupuncture? A meta-analysis of randomized controlled trials', *BMC Med.*, vol. 8, p. 75, Nov. 2010, doi: 10.1186/1741-7015-8-75.
- [71] I. Lund and T. Lundeberg, 'Are minimal, superficial or sham acupuncture procedures acceptable as inert placebo controls?', *Acupunct. Med. J. Br. Med. Acupunct. Soc.*, vol. 24, no. 1, pp. 13–15, Mar. 2006, doi: 10.1136/aim.24.1.13.
- [72] T. J. Kaptchuk, P. Goldman, D. A. Stone, and W. B. Stason, 'Do medical devices have enhanced placebo effects?', *J. Clin. Epidemiol.*, vol. 53, no. 8, pp. 786–792, Aug. 2000, doi: 10.1016/s0895-4356(00)00206-7.
- [73] Y.-C. Zheng, T.-T. Yuan, and T. Liu, 'Is acupuncture a placebo therapy?', *Complement. Ther. Med.*, vol. 22, no. 4, pp. 724–730, Aug. 2014, doi: 10.1016/j.ctim.2014.05.005.
- [74] C. Paterson and P. Dieppe, 'Characteristic and incidental (placebo) effects in complex interventions such as acupuncture', *BMJ*, vol. 330, no. 7501, pp. 1202–1205, May 2005, doi: 10.1136/bmj.330.7501.1202.
- [75] D. Colquhoun and S. P. Novella, 'Acupuncture is theatrical placebo', *Anesth. Analg.*, vol. 116, no. 6, pp. 1360–1363, Jun. 2013, doi: 10.1213/ANE.0b013e31828f2d5e.

- [76] C. Maynié-François *et al.*, 'Quality of evidence of the efficacy of therapeutic interventions on patient-important outcomes in Cochrane's systematic reviews' abstracts: A survey', *Therapie*, pp. S0040-5957(20)30153-0, Sep. 2020, doi: 10.1016/j.therap.2020.09.004.
- [77] I. Olkin, I. J. Dahabreh, and T. A. Trikalinos, 'GOSH - a graphical display of study heterogeneity', *Res. Synth. Methods*, vol. 3, no. 3, pp. 214-223, Sep. 2012, doi: 10.1002/jrsm.1053.
- [78] 'Retraction: Comparison of acupuncture and other drugs for chronic constipation: A network meta-analysis', *PLoS ONE*, vol. 13, no. 7, p. e0201274, Jul. 2018, doi: 10.1371/journal.pone.0201274.
- [79] A. Dechartres, D. G. Altman, L. Trinquart, I. Boutron, and P. Ravaud, 'Association between analytic strategy and estimates of treatment outcomes in meta-analyses', *JAMA*, vol. 312, no. 6, pp. 623-630, Aug. 2014, doi: 10.1001/jama.2014.8166.
- [80] A. Dechartres, I. Boutron, L. Trinquart, P. Charles, and P. Ravaud, 'Single-center trials show larger treatment effects than multicenter trials: evidence from a meta-epidemiologic study', *Ann. Intern. Med.*, vol. 155, no. 1, pp. 39-51, Jul. 2011, doi: 10.7326/0003-4819-155-1-201107050-00006.
- [81] J. Savović *et al.*, 'Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials', *Ann. Intern. Med.*, vol. 157, no. 6, pp. 429-438, Sep. 2012, doi: 10.7326/0003-4819-157-6-201209180-00537.
- [82] L. Wood *et al.*, 'Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study', *BMJ*, vol. 336, no. 7644, pp. 601-605, Mar. 2008, doi: 10.1136/bmj.39465.451748.AD.
- [83] P. Jüni, D. G. Altman, and M. Egger, 'Systematic reviews in health care: Assessing the quality of controlled clinical trials', *BMJ*, vol. 323, no. 7303, pp. 42-46, Jul. 2001, doi: 10.1136/bmj.323.7303.42.
- [84] J. Pildal, A. Hróbjartsson, K. J. Jørgensen, J. Hilden, D. G. Altman, and P. C. Gøtzsche, 'Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials', *Int. J. Epidemiol.*, vol. 36, no. 4, pp. 847-857, Aug. 2007, doi: 10.1093/ije/dym087.
- [85] K. F. Schulz, I. Chalmers, R. J. Hayes, and D. G. Altman, 'Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials', *JAMA*, vol. 273, no. 5, pp. 408-412, Feb. 1995, doi: 10.1001/jama.273.5.408.
- [86] S. Steegen, F. Tuerlinckx, A. Gelman, and W. Vanpaemel, 'Increasing Transparency Through a Multiverse Analysis', *Perspect. Psychol. Sci. J. Assoc. Psychol. Sci.*, vol. 11, no. 5, pp. 702-712, Sep. 2016, doi: 10.1177/1745691616658637.
- [87] E. Mayo-Wilson *et al.*, 'Cherry-picking by trialists and meta-analysts can drive conclusions about intervention efficacy', *J. Clin. Epidemiol.*, vol. 91, pp. 95-110, Nov. 2017, doi: 10.1016/j.jclinepi.2017.07.014.
- [88] J. P. A. Ioannidis, 'The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses', *Milbank Q.*, vol. 94, no. 3, pp. 485-514, Sep. 2016, doi: 10.1111/1468-0009.12210.
- [89] J. de Vrieze, 'The metawars', *Science*, vol. 361, no. 6408, pp. 1184-1188, Sep. 2018, doi: 10.1126/science.361.6408.1184.

Authors' contributions

FN, CB, and MEB conceived and designed the experiments. MEH, XW, TB extracted the experiments. MEH and CB analyzed the data. MEH and FN wrote the first draft of the manuscript. CB and BF contributed to the writing of the manuscript. MEH, XW, TB, BF, FN, and CB agreed with the results and conclusions of the manuscript. All authors read and approved the final manuscript.

Journal Pre-proof

What is new?

- Choice of selection criteria impacted effect estimates in meta-analyses of acupuncture for smoking cessation.
- Certain methodological choices may be associated with important increases in effect sizes.
- Variation of effects evaluations allow for exploring controversies related to conflicting meta-analyses.

Journal Pre-proof

Declarations of interest: none

Journal Pre-proof